Improving Consensus Development for Health Technology Assessment: An International Perspective

Council on Health Care Technology
ISBN: 0-309-55564-7, 176 pages, 6 x 9, (1990)

Improving Consensus Development for Health Technology Assessment: An International Perspective

Clifford Goodman and Sharon R. Baratz, editors

Council on Health Care Technology
Institute of Medicine

National Academy Press
Washington, D.C. 1990

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 advisor to the federal government and its own initiative in identifying issues of medical care, research, and education.

The COUNCIL ON HEALTH CARE TECHNOLOGY was established in 1986 by the Institute of Medicine as a public-private entity to address issues of health care technology and technology assessment. The council is committed to the well-being of patients as the fundamental purpose of technology assessment. In pursuing that goal, the council draws on the services of experts in medicine, health policy, science, engineering, and industry.

The Council on Health Care Technology sponsored a workshop on International Consensus Development for Medical Technology Assessment in London on June 7, 1989. The workshop and these proceedings were supported in part by the National Center for Health Services Research and Health Care Technology Assessment, grant number HS 05526. The opinions and conclusions expressed here are those of the authors and do not necessarily represent the views of the National Academy of Sciences or any of its constituent parts, the U.S. Department of Health and Human Services, or the organizations with which the authors are affiliated.

Library of Congress Catalog Card Number: 90-60936
International Standard Book Number: 0-309-04239-9
Additional copies of this report are available from: National Academy Press 2101 Constitution Avenue, NW Washington, DC 20418

Printed in the United States of America

S127

ACKNOWLEDGMENTS

Among the various authors, reactors, and other participants, representatives of 11 countries contributed to this effort. The editors wish to thank the authors of the consensus program profiles and the authors and presenters of the five main workshop papers, whose names are given with their respective contributions to this publication. The profiles were done well in advance of the workshop and formed much of the basis for the papers presented there. The five main papers constituted largely original work and provided the necessary multinational perspective.

Important insight was provided by the reactors designated for the main presentations, including Bjørn Backe, Anton Casparie, Johan Calltorp, Richard Chrzanowski, Gil Hill, Dominique Jolly, Egon Jonsson, Jari Kankaanpää, Anna-Liisa Kauppila, Niek Klazinga, Jonathan Lomas, Duncan Neuhauser, and J.J.E. van Everdingen. Bryan Jennet gave a thoughtful and stimulating welcome to the workshop. Itzhak Jacoby was instrumental in helping to plan the workshop agenda. Based on the workshop deliberations, the writing group of 10 people from five countries drafted, reviewed, and edited multiple versions of the recommendations included here for improving consensus development efforts.

We gratefully acknowledge the assistance of the King's Fund Centre for hosting the workshop and wish to thank in particular Barbara Stocking, Jackie Spiby, and Maria Said. Holly Dawkins and Evan

son Joseph of the Institute of Medicine staff were most helpful with logistical arrangements and related support. Partial support for this project was provided by grant HS 05526 from the National Center for Health Services Research and Health Care Technology Assessment.

COUNCIL ON HEALTH CARE TECHNOLOGY

PREFACE

Recently, in many countries, the interests of different groups concerned with health care have focused on the use of medical technologies—their safety, efficacy, and effectiveness; their cost-effectiveness and cost-benefit; their impacts on quality of care; and their social, legal, and ethical implications. The sum of these varied interests comprises the field of health care technology assessment.

The Council on Health Care Technology was created in the United States to promote the development and application of technology assessment in health care and the review of health care technologies for their appropriate use. The council was established as a public private enterprise at the Institute of Medicine, a component of the National Academy of Sciences, through the Health Promotion and Disease Prevention Amendments of 1984 (P.L. 98-551, later amended by P.L. 99-117). In 1987 the U.S. Congress extended support for the council as a public-private venture for an additional three years (by P.L. 100-177).

The goals and objectives of the council, as stated in the report of its first two years of operations, are "to promote the development and application of technology assessment in medicine and to review medical technologies for their appropriate use. The council is guided in its efforts by the belief that the fundamental purpose of technology assessment is to improve patient well-being and the quality of

care." In pursuing these goals, the council seeks to improve the use of medical technology by developing and evaluating the measurement criteria and the methods used for assessment, to promote education and training in assessment methods, and to provide technical assistance in the use of data from published assessments.

The council has conducted its activities through several panels and committees. Members of these groups reflect a broad set of interested constituencies—physicians and other health professionals, patients and their families, payers for care, biomedical and health services researchers, manufacturers of health-related products, managers and administrators throughout the health care system, and public policymakers.

The Methods Panel of the Council on Health Care Technology has worked toward the improvement of the methods, techniques, and procedures of technology assessment. The panel objectives included strengthening of the ability of health care institutions to acquire primary data for the assessment of medical technology, increasing the number of assessments of medical technology that are based upon primary data, strengthening of the methods that provide alternatives to randomized controlled clinical trials, and development of technology assessment methods, following the development of methods for health quality assessment and assurance. The Methods Panel conducted a variety of projects in pursuit of these goals.

Through the council, the Methods Panel promoted the examination of group judgment methodologies by way of a series of workshops. In June of 1989 the council organized a one-day workshop on International Consensus Development Conferences in conjunction with the annual meeting of the International Society for Technology Assessment in Health Care. The workshop allowed participants to review and consolidate findings on alternative approaches to consensus development efforts, develop recommendations or guidelines for conducting these efforts, and identify research needs for resolving methodologic questions. This report records the proceedings and findings of that workshop.

WILLIAM N. HUBBARD, JR., CHAIR
JEREMIAH A. BARONDESS, CO-CHAIR

CONTENTS

Program Profiles

INTRODUCTION

Group judgment methods are perhaps the most widely used means of assessment of medical technologies in many countries. The consensus development conference is a relatively inexpensive and rapid mechanism for the consideration and evaluation of different attributes of a medical technology including, for example, safety, efficacy, and efficiency, among many others. The current concept of the consensus development program originated at the National Institutes of Health (NIH) in the United States. The first conference was held in the United States in 1977 as part of an effort to improve the translation of NIH biomedical research findings for use in clinical practice. Since that time, the methods of conference organization and conduct, as well as dissemination strategies for the results, have evolved in the attempt to refine the U.S. process. Many countries have initiated their own consensus development programs. Although the idea of consensus development is common to all programs, the consensus development process and dissemination mechanisms differ across countries. Individuals in a variety of countries have remodeled the NIH consensus development conference methodology to adapt the technological assessment to the particular national context of the program.

The Methods Panel of the Council on Health Care Technology initiated a project to compare and contrast international consensus development programs in order to share the cumulative insight and

experience gained in the different countries. The goal of the workshop was to examine existing programs and to formulate suggestions for improvements in the use of the group judgment methodology. Individuals in the field of consensus development from eleven countries gathered for one day to work toward improved consensus development conferences and mechanisms to translate these findings into better patient care.

Before the workshop, individuals developed written profiles of nine consensus development programs in eight countries, including Canada (two programs), Denmark, Finland, The Netherlands, Norway, Sweden, the United Kingdom, and the United States. For each profile, authors considered the national context of the program, the scope of the program, the format and conduct of the consensus conferences, the documentation and use of evidence in the process, and the dissemination and impact of consensus recommendations. Workshop participants received copies of these profiles in preparation for the workshop. The final versions of the profiles are included in the first part of this report.

Workshop discussion was based upon five major presentations. Using material provided in the program profiles and other resources, each main presenter provided an examination of one particular aspect of consensus development across the different countries. Speakers presented papers on the role and sponsorship of consensus development programs in national health care systems, the topic and scope of the programs, the documentation and use of evidence in the consensus development processes, the format and conduct of consensus development processes, and the dissemination and impact of consensus development exercises.

At the end of the workshop, a working group met to formulate recommendations for improving consensus development for assessing health technologies. The working group attempted to address the issues raised in the discussions and to record the solutions developed by participants in the workshop. The recommendations may require adaptation to the different national contexts of the programs in order to improve the quality and impact of international consensus development programs. These recommendations represent the consolidated view of the writing group and are not necessarily the views of the National Academy of Sciences or any of its constituent parts, the U.S. Department of Health and Human Services, or the organizations with which the authors are affiliated. These recommen

dations may serve as a useful guide for improving consensus development programs or developing new programs.

The five major presentation papers that address different aspects of the consensus development programs are included in the first section of this report. The program profiles are provided in alphabetical order by country in the second section. The recommendations of the working group follow these sections. An international consensus development bibliography is provided to facilitate further research on consensus development. The articles contained in the bibliography are specifically oriented toward consensus development; therefore, some of the references used in the papers in this report that do not directly address consensus development have been omitted. The bibliography includes additional articles on consensus development that were not cited by any of the report authors. A list of authors is provided in the final section of the report.

ELEMENTS OF THE CONSENSUS DEVELOPMENT PROCESS

Sponsorship and Role of Consensus Development Programs within National Health Care Systems

Itzhak Jacoby

In 1977, under the leadership of Donald Fredrickson, the U.S. National Institutes of Health (NIH) began sponsoring consensus development conferences. This new program was NIH's innovative response to the widely perceived need for assessment of the safety and efficacy of technologies. In the U.S. Congress and elsewhere, debates were raging about health technology assessment in general and the role of the federal government in particular, including the impact of NIH's research effort on clinical practice. Senator Edward Kennedy asked in a 1976 speech:

> Shouldn't some institution in our society have an ongoing function of reviewing not just new knowledge that might be transferred into clinical practice but also old knowledge that underpins current procedures involving risks, high costs, or simply great inconvenience to millions of patients in order to determine what needs to be changed, updated, or further researched? (Perry, 1988).

Further impetus for developing the program came from rapidly escalating health care costs, which were linked by the public and policymakers to the uncontrolled diffusion of expensive, but not necessarily cost-effective, technologies. The program also fulfilled Dr. Frederickson's desire for NIH to bolster information transfer at the interface between biomedical science and clinical practice.

With more than 75 consensus conferences to its credit, the NIH program has served as a model for the United States and other countries in the development of health technology assessment. Similar

programs that have emerged in several European countries and Canada differ markedly from the NIH model in their goals, scope of consideration, and mode of disseminating results. This divergence has resulted from differences among the countries' health care delivery systems and the nature of the organizations that sponsor the conferences. These differences in the context of technology assessment play a critical role in determining the scope, format, and impact of each country's consensus development program.

The consensus development process is designed for conflict resolution. It borrows from several other conflict resolution modes: the judicial process, collegiate peer review, democratic debate, and collective bargaining. As pointed out by Kosecoff et al. (1987), for the process to have impact and meaning, topics must involve clinical practice-related controversies and a gap between knowledge and practice. The process has its greatest effect in the resolution of controversies about technologies or issues on the cutting edge of medical practice, and therefore, it should be reserved for this purpose.

Broad participation of multiple factions is necessary in sponsoring, planning, and implementing the conferences in order to increase the likelihood of proper selection of topics, encourage development of conclusions applicable to practice decisions and policy-making, and produce a significant impact on health care delivery. Participants should include the biomedical research community, those in medical practice, payer organizations, and organizations concerned with the administration and delivery of services. To meet the needs of these factions, consensus conferences should aim to resolve conflicts about safety, efficacy, cost and cost-effectiveness, indications and contraindications, societal acceptance, and other ethical issues. Any compromise in this agenda, especially given the urgent need for consensus results and the high cost of these conferences, can only be considered a missed opportunity. The context within which health technology assessment—and consensus development in particular—takes place in each country has determined the extent to which this ideal is realized.

THE U.S. EXPERIENCE IN CONSENSUS DEVELOPMENT

The focus of the NIH Consensus Development Program is narrowly defined, in line with the agency's mission as a biomedical

research institution. The program limits assessment to issues of safety and efficacy, excluding social, economic, and ethical considerations. Topics are proposed by one of NIH's many component institutes when sufficient data are available to conduct such an assessment and when some element of controversy exists. Participants in the process are broadly recognized leaders in their respective fields. Technical and research experts present information to an appropriate panel of physician specialists, informed public representatives, and others with specialized expertise. This panel reviews the scientific data presented to develop the consensus statement. Representatives of physician specialty groups, third-party payers, or health administrators are not usually participants in the planning or execution of the conference.

From the beginning, the NIH process was supposed to seek a "technical consensus" on the clinical significance of new findings, the adequacy of validation for efficacy and safety, and the need for further research. NIH recognized that further consensus development on the broader health care delivery issues should complement its technical consensus before a particular clinical modality is recommended for adoption. This second process was then referred to as *interface consensus* and was to deal not only with safety and efficacy, but also with cost; cost-effectiveness; and legal, ethical, and other societal issues (Perry, 1988).

To provide this broader assessment, the U.S. Congress established the National Center for Health Care Technology (NCHCT). The center's responsibilities included consensus development on the full range of clinical practice-related issues. The one conference that the center sponsored as a follow-up to the NIH consensus conference on coronary artery bypass surgery complemented the NIH conclusions with considerations of broad societal issues. The NCHCT succumbed to budgetary and political pressures in 1981 after only two years of operation. While it existed, the NIH added the word *technical* to the title of its consensus development program to distinguish NIH's responsibilities from those of the center. Although the word *technical* has since been dropped, NIH-sponsored consensus conferences still conform to the restricted model.

The broad goal of the NIH program is to facilitate the appropriate and timely application of biomedical research findings to clinical practice. In evaluating the extent to which the program has attained that goal, analysis both of conferences that have produced an impact

and those that have not point to the advantages of broad sponsorship. For the most part, NIH relies upon topic selection criteria concerning the adequacy of the evidence for resolving controversies and the thoroughness with which the clinical implications of scientific or technologic developments have been explored and discussed in the literature. Although these criteria are reasonable in the context of scientific peer review, they do not address the needs of the intended recipients of technology transfer: physicians, hospitals, and other providers of health care, as well as their patients. To increase the impact of consensus development conference results, those who choose topics should systematically consider the current state of practice along with the current state of science in determining the potential usefulness of a conference.

The NIH consensus development conference on the treatment of primary breast cancer illustrated the importance of considering the state of practice when planning a conference. This conference failed to produce a change in practice because there was widespread preconference conformity with the major consensus recommendation (Kanouse et al., 1989).

An inability to effect change may also result from the panel's avoidance of the areas of greatest controversy in generating recommendations. The U.S. conference on coronary artery bypass surgery is a case in point. In the consensus statement from that conference, the panel confined its recommendations to the relatively clear cases at the extremes, where there was less controversy, instead of providing unequivocal recommendations regarding the most common clinical scenarios leading to bypass surgery, where clinicians might welcome some assistance in making choices (Kanouse et al., 1989). In general, sole custodianship of any process encourages a natural tendency for a group to avoid controversy and to safely promote the view of the sponsor, which, in this case, was the biomedical research community. Greater involvement of physician specialty societies and payer organizations, including health insurance groups and agencies such as the Health Care Financing Administration that reimburse providers and beneficiaries with government funds, might help to produce the most urgently needed consensus development information.

When participation in NIH consensus development conferences extends beyond the biomedical research community, the benefits are clear. A recent evaluation of the NIH Consensus Development Pro

gram singled out the 1980 conference on cesarean childbirth as a success in both information transfer and practice impact. Since both the obstetrical community and the lay public viewed sharply rising cesarean section rates with alarm, NIH's sponsorship of this conference was spurred by a suggestion from the American College of Obstetricians and Gynecologists. The conference thereby capitalized on the existing physician awareness and concern that formed a basis for motivating change in hospital policies and physician practices. The conference was oriented to this well-defined interest and performed a unique service in summarizing the accumulated knowledge on this topic, pointing out the practice implications of this information, and developing clear recommendations about changing the management of certain deliveries. All of these activities had never before occurred with the clinician as the intended audience. Finally, NIH arranged for publication of the consensus statement in two major obstetrical journals, and the American College of Obstetricians and Gynecologists took an active interest in helping to disseminate the findings (Kanouse et al., 1989).

Given these elements—a scientifically based and clinically relevant message delivered for the first time to a receptive audience that recognized the need for change—it is perhaps not surprising that the conference on cesarean sections was especially successful in changing physicians' and hospitals' practices (Kanouse et al., 1989). Such a combination of favorable conditions may be the exception rather than the rule for consensus development topics. The likelihood of identifying motivation for change in a controversial area may be enhanced by seeking out the active participation of the potential users of consensus statements.

An analysis of the topics selected for the NIH Consensus Development Program over the last eight years would suggest a shift away from controversial subjects accompanied by the production of conclusions that are not only less controversial but also more general and less helpful to the practicing health care professional community or patients. Two exceptions, in addition to the one on cesarean sections, are the conferences on liver transplantation and the use of ultrasound in prenatal care. In both cases, outside pressure was brought to bear on NIH to hold the conferences.

To achieve the greatest impact with consensus development conferences in the United States, NIH should contribute, along with other participants from the broader health care community, to a con

sensus process administered by an organization with a much broader mission. Topics should be selected that represent a significant challenge from the perspective of clinicians and the public at large. The influence, resources, and insight of the biomedical research community, combined with those of physician specialty societies, third-party payers, and individuals involved in health care policy and administration, would constitute a formidable force for change in health care.

One additional change in the current procedure should also be considered. Occasionally, a consensus development panel has produced a statement with assertions that are not supported by evidence or a statement that does not adequately answer one or more of the questions posed. It should be routine for all consensus statements to be reviewed by an independent group, as are papers in peer-reviewed journals and reports from the National Academy of Sciences. This peer review should result in the production of more valid statements.

CONSENSUS DEVELOPMENT OUTSIDE THE UNITED STATES

As the NIH Consensus Development Program in the United States matured, international interest in the process began to emerge. In late 1981, Sweden's consensus development program was established. The program was the result of numerous discussions and visits by representatives of the Swedish Medical Research Council, a sister agency of NIH, and the Swedish Planning and Rationalization Institute for the Health and Social Services (Spri), which is responsible for health planning in Sweden. Although a great deal of discussion has occurred concerning the similarities between the U.S. and Swedish programs, the cosponsorship of the Swedish program by a health planning organization has received little attention. Spri's orientation has had a distinctive impact on the consensus development program's activities. This arrangement was particularly important considering the decentralization of the national health service delivery system in Sweden, where most decisions are made by regional councils.

The first Swedish consensus development conference addressed total hip-joint replacement. It followed a conference on the same topic in the United States. The Swedish conference, in line with the sponsors' interests, added a question on the cost implications of the procedure to questions shared with the U.S. conference. Other top

ics have been selected primarily on the basis of their importance to health service delivery and the practice of medicine, in addition to the willingness of the scientific community to deal with the issue. An examination of the impact of the Swedish conference on hip-joint replacement indicated an increase in the number of procedures immediately following the conference (Calltorp, 1988).

In the United Kingdom, the consensus development program began in 1984. Unlike other programs, this endeavor stemmed not from legislative or governmental impetus but from individual interest and belief in the process. The program found a base in The King's Fund Centre for Service Development, an academic and philanthropic institution prominent in the British medical community for training health service managers. The organizers of the consensus development program nevertheless sought the participation of government agencies responsible for the National Health Service and participation of physicians' groups. These organizations, however, while supporting the purpose of the program, chose to participate as observers and adopt a "wait and see" attitude.

After lengthy negotiations, the first topic chosen was coronary artery bypass grafting. At the time, only slightly more than 10 procedures per 100,000 population were performed in the United Kingdom, as opposed to more than 80 procedures per 100,000 population in the United States and approximately 40 per 100,000 in other European countries. Since most health services utilization experts agreed that the European rate was appropriate, it was not surprising that the consensus process recommended raising the rate for the United Kingdom. In recognition of Britain's limited health care resources, the consensus statement provided criteria for clinicians to use in the selection of patients who could benefit most from the medical process for receipt of treatment. Compared with the U.S. conference on this topic in 1981, the King's Fund conference produced very detailed clinical advice as to indications and contraindications for the coronary artery bypass procedure.

In addition, as discussed in a detailed comparison of the U.K. and U.S. conferences on bypass surgery (Stocking, 1985), the questions considered by the panel went beyond scientific issues to encompass the costs and implications of increased use of coronary artery bypass grafting on the British National Health Service. In line with these interests, and especially because of the capped budget of the service, the U.K. panel consisted of people who were not experts on the

procedure; half of the panel members were not clinicians. Subsequent conferences in the United Kingdom have dealt both with narrow issues in medicine as well as broad, societal issues. Increased emphasis has been placed on the impact on the health care delivery system of adopting new technologies, as well as social, economic, and cultural implications, in an effort to make the consensus statements as useful as possible in the public policy arena. This broad scope was feasible because the sponsoring institution was an academic center rather than an active participant in the health delivery system.

In The Netherlands, a technology assessment program has been initiated by the National Organization for Quality Assurance in Hospitals (Centraal Begeleidingsorgaan voor de Intercollegiale Toetsing), a physician organization. The program began as a physician-initiated quality assurance program, similar to the U.S. professional standards review organizations and, later, the professional review organizations. Only after its emergence was it recognized as an analog to the consensus development program. This program is significantly different from the previous two examples. Each assessment is a long process, involving many internal review cycles. The main feature that should be highlighted is that, since the consensus development process is sponsored by a body of physicians, there is little opportunity for contributions by other participants in health care delivery. Because it is sponsored by physicians, the program has a high potential for making a significant impact on practice behavior. In fact, the conclusions reached by this consensus development program serve as the basis for the quality assessment program in place in Dutch hospitals (Klazinga et al., 1987).

In France, the Institut National de la Santé et de la Recherche Médicale (INSERM), a sister organization of NIH, attempted to establish a consensus development program in 1987. The effort was unsuccessful because of the research organization's inability to obtain the active participation of the country's social security administration and physician organizations in sponsoring consensus conferences. This experience emphasizes the importance, in some quarters, of bringing the major participants in health services delivery into the assessment process in order to ensure the usefulness of results.

The Israel Ministry of Health has sponsored jointly with the World Health Organization (WHO) and the U.S. NIH a consensus confer

ence on magnetic resonance imaging and positron emission tomography (PET). Israel's need for this conference developed from a controversy that arose between the philanthropic organization Hadassah, the Ministry of Health, and the Sick Fund, which provides care to the majority of Israelis, over the uses of the two technologies. Lacking the expertise to resolve the issue, Israel called upon WHO, which in turn called upon NIH. This broad sponsorship resulted in the acquisition of the needed expertise. The conference concluded that, although PET has great scientific importance, it was clearly not ready for clinical applications. As is characteristic of a small country with a pluralistic health care system, the Israeli conference covered practical aspects of economic and human resources in addition to assessment of the science. The conclusions drawn at the conference had a direct impact on the provision of imaging services in Israel.

Interestingly, the United States had begun to plan a conference on PET, but when it became clear to the planners that the recommendation was going to be negative for clinical applications, the conference was canceled. In April 1989, the *Wall Street Journal* carried an article entitled "Debate Grows Over Clinics Pushing Costly PET Scans," which examined the movement of PET into clinical settings with questionable and costly applications (*Wall Street Journal*, 1989). Israel's definitive action has enabled that country to deal effectively with PET before it emerged as a problem, as now appears to be the case in the United States.

CONCLUSIONS

This paper described five models for consensus development sponsorship and examined them in the context of the health service delivery systems of the respective countries. Other consensus development programs worldwide can be identified as belonging to one of these five models (see Table 1). It is critical for program sponsors to consciously consider the model they intend to apply in light of their desired objectives.

The goal of consensus development programs should be to resolve conflicts in as broad a range of factors as is relevant, including safety, efficacy, cost and cost-effectiveness, indications and contraindications, acceptability, and other societal and ethical issues. The likelihood of realizing this goal will be improved by expanding the

sponsorship of the process to include all major participants in the biomedical research and health care communities. These sponsors should make every effort to select topics that represent a significant challenge from the perspective of clinicians and the public at large. In planning conferences, sponsors should consider both the current state of science and the current state of clinical practice to assure that an impact on clinical practice is possible. In particular, preference should be given to topics for which motivation to change policy and clinical practice already exists among clinicians. Finally, consensus development programs should adopt a procedure to review consensus statements before they are published in final form.

TABLE 1 Sponsorship and Scope of Consensus Conferences in Five Countries

Country	Sponsors	Scope
United States	National Institutes of Health	Safety and efficacy
The Netherlands	National Organization for Quality Assurance in Hospitals	Quality assurance
Sweden	Swedish Medical Research Council and Swedish Planning and Rationalization Institute for the Health and Social Services	Safety, efficacy, and cost-effectiveness
United Kingdom	The King's Fund College	Broad societal
Israel	All partners in national health, plus the World Health Organization and the U.S. National Institutes of Health	Broad consensus for application

By marshaling the resources of the research, practice, health financing, and health policy communities, the consensus development process can contribute significantly to any nation's health agenda.

References

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.

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 NIH consensus development program. Santa Monica, Calif.: RAND Corporation.

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

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.

Perry, S. 1988. Consensus development: An historical note. International Journal of Technology Assessment in Health Care 4:481-484.

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

Wall Street Journal. April 3, 1989. Debate grows over clinics pushing costly PET scans. B1.

Topic and Scope of Consensus Development Conferences: Criteria and Approach for Selection of Topics and Properties for Assessment

Tore Scherstén

During the past 10 to 15 years, unparalleled advances in biomedical and technical research have revolutionized the practice of medicine. Powerful new health care technologies for prevention, diagnosis, and treatment have emerged, for example, vaccines, computed tomography (CT) scanners, magnetic resonance imaging, and organ and cell transplantation procedures. New biotechnologies such as the technique of recombinant DNA, gene cloning and protein production, monoclonal antibody formation, and microchemical instrumentation all open up novel, almost unlimited, opportunities for medicine. These advances, however, have created serious problems for the public and its representatives, for the patient, and for the medical community. The problems are not only scientific and medical but also of ethical, economic, social, and legal concern. Therefore, not surprisingly, the need and demand for comprehensive assessments of health care technologies have grown in the developed world. A technological assessment should provide decision makers with useful information and guidelines on policy questions that concern the application of medical technology. The recommendations should be based on the validation of safety and efficacy, of cost-effectiveness, and of social and ethical implications of the use of the technology—in short, the assessment should answer the question of whether the benefits from the use of a particular technology outweigh the costs in human and monetary terms.

SCOPE OF THE CONFERENCE

The consensus development conference, a rapid data synthesis method, has the potential to be a powerful instrument for the provision of balanced advice about a technology and for the definition of the need for further information and research. By and large, these are the main reasons for the great interest and rapid adoption of this synthesis method in the Western world.

The consensus development conference process does not generate new scientific data and therefore cannot solve or contribute to the solution of a true scientific conflict. The conference can disclose and define what is known and what is not known about a technology and can also contribute to new interpretations and evaluations of the available scientific data. Thus, the participants may formulate the best current judgment of a given technology in relation to the health care service requirements, the patients' desires, and the demands of the society. The aims and the expectations of a consensus development conference differ between countries because of variations in the cultural settings and in the health care systems.

In the United States, where the consensus development conference originated, the National Institutes of Health (NIH) initially proposed the process in 1977 as a means to facilitate the transfer of biomedical research results into clinical practice. NIH considered the consensus development conference to be a natural extension of the biomedical research community's obligation to assume more responsibility for the practical implications of research results. The primary goal of the conference process was to validate the safety and efficacy of a particular technology. The original intent of the scope of the program was to emphasize emerging technologies. However, with time the scope of the program has widened, and most of the technologies assessed through the consensus exercises are already in use (Perry, 1987, 1988).

The NIH emphasis on safety and efficacy was natural in view of the activities of the National Center for Health Care Technology (NCHCT). NCHCT was involved in assessment activities related to the delivery of health services, hospital management, and billing systems in the time period of 1977 to 1981. The NCHCT assessments focused on the economic and social implications of medical technologies (P.L. 95-623; U.S. Congress, 1978).

Seven European countries, including Denmark, Finland, The Neth

erlands, Norway, Sweden, Switzerland, and the United Kingdom, have started consensus development conference programs. These countries have modified the conference according to the cultural settings and health care systems of their societies. In general, the scope of the European consensus exercises is broader than that of the NIH conferences. European conferences usually include an economic evaluation of the technology and an attempt to address the ethical implications of technological use. In the Nordic countries, the conferences also examine the availability of a technology to ensure equal access for all patients in need of the technology. In these countries, the conferences create a base for societal debate and evaluation of the medical technology. The consensus development conference has also been used in policy-making and in the allocation of resources. The target groups of the conferences in these countries include doctors, politicians, administrators, and the public (Andreasen, 1988; Klazinga et al., 1987).

Questions about technologies may address the effects on patients, the patients' relatives, and organizations, including the potential additional personnel and equipment requirements (for example, the need for physicians, assistants, or technicians). The effects on human resources in terms of education and training of new specialists and the employment of persons in isolated areas, as well as the overall opportunity costs to society of technological use have also been examined. There are several examples in European countries where a conference has influenced the allocation of resources. In Sweden, for example, the resources for total hip-joint replacement were increased and reallocated after a consensus development conference. The diffusion of CT scanners was also enhanced after the conference on the management of stroke.

Only one direct comparison between the NIH conference and the European model has been performed (Rogers et al., 1982). NIH arranged a conference on hip-joint replacement, and a similar conference was held on the same topic in Sweden. The conferences provided the opportunity to compare consensus development in two cultures with different health care systems. The Swedish conference was convened to discuss the costs and the availability of the technology in Sweden as well as safety and efficacy. The conference in the United States focused on safety and efficacy but did not include considerations of costs or availability. Despite differences in the cultures and the focus of the conferences, the consensus statements

were remarkably similar in the evaluation of efficacy and safety. This comparison supports the possibility of the transfer of information on biomedical technologies between countries and the use of evaluations done in different countries independent of the original health care systems. In addition, the process may be used effectively in different cultural settings.

SELECTION OF TOPIC

In the United States as well as in other countries with consensus development conference programs, all types of medical technologies (i.e., for prevention, diagnosis, treatment, and rehabilitation) have been examined.

The initial intent of the NIH program was to emphasize new and emerging technologies, although most of the conferences in the United States and other countries have addressed established technologies already in widespread use. This selection seems natural in view of the fact that the majority of technologies in use have never been comprehensively evaluated for safety, efficacy, and social consequences.

The following criteria for the selection of topics are similar in all countries.

- The topic under consideration should be important from a quantitative and/or qualitative point of view; i.e., it should be of medical importance, affect a significant number of people, and/or be very costly.
- There must be an available base of scientific information about the technology.
- There must be a scientific debate about the use of the technology and a discrepancy between the available knowledge and its practical application in medicine.

In the Nordic countries the availability of the technology for the people (i.e., equity) has also been an important criterion for selection.

International experience with consensus development conferences has shown that the method is important for the evaluation of the quality of clinical practice and/or for the provision of information and guidelines to administrative and political decision makers. Data concerning the efficacy and safety of a medical technology are of

universal value and therefore can be transferred between health care systems. On the other hand, evaluation of factors of importance for education, health care system organization, and economics will have limited transferability between countries. Hence, there are good reasons for a country to develop a national consensus conference program.

References

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

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

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 4:481-485.

Rogers, E.M., J.K. Larsen, and C.U. Lowe. 1982. The consensus development process of medical technologies: A cross-cultural comparison of Sweden and the United States. Journal of the American Medical Association 248:1880-1882.

U.S. Congress. 1978. P.L. 95-623. Health Services Research, Health Statistics and Health Care Technology Act of 1978.

Documentation and Use of Evidence in the Consensus Conference Process*

Gérard Breart

The purpose of a consensus development conference, as defined by the U.S. Office of Medical Applications of Research (OMAR) of the National Institutes of Health (NIH), 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 and that will be useful to health professionals and the public at large (OMAR, NIH, 1988). The consensus statement should reflect the available scientific evidence.

Critics have highlighted the shortcomings of the consensus development process and considered the consensus statement to be more of a compromise of divergent viewpoints than an in-depth analysis of the existing research evidence (Ahrens, 1985; Levitt and Potish, 1988; Lomas, 1986; Oliver, 1985). Jacoby (1988) states that:

> The quality and quantity of data available for specific questions at these recent conferences still varied considerably. The separate elements of the consensus statements, however, did not always reflect this variability in how strongly conclusions were stated, resulting in inappropriately strong conclusions.

These two different opinions on the origin and nature of consensus statements allude to the fact that, in some circumstances, there

* The author acknowledges, with thanks, Sharon R. Baratz for revising the manuscript.

may be a gap between the available data (consensus process origin) and the consensus statement (consensus process result). The aim of this paper is to outline the elements that may explain this gap. This discontinuity between the available evidence and the final statement may be due to the consensus process itself or to the general decision-making process for clinical practice.

THE CONSENSUS PROCESS

In order to ensure that the final statement is based primarily on scientific evidence, the whole consensus process has to be oriented toward this objective. The following three conditions must be met:

1. Scientific evidence must exist and must be clearly and completely presented to the panel in an easily accessible and understandable form.
2. Members of the panel must be ready to base their decisions on scientific evidence. Panelists must not have a strong opinion on the topic before the conference. Panelists must be given criteria for the evaluation of scientific studies.
3. The environment (actual practices, outside pressures) must allow room for decisions based on scientific data.

These three conditions have to be kept in mind throughout the organization of the following stages of the consensus development process:

- preparation of the process
- review of the existing evidence
- selection of the panel
- selection of the speakers
- formulation of recommendations.

Preparation of the Process

At least two conditions must be satisfied for the consensus statement to be based on scientific evidence: (1) the existence of a reasonable body of research data and (2) variability in practice patterns concerning the proposed topic.

Given insufficient data, the consensus statement may only be a compromise of contradictory expert opinions. The relationship be

tween current practice for a technology or clinical condition and the consensus development conference is of particular importance. If few practitioners diverge from the standard procedure, it is likely that the consensus statement will reflect the general opinion, even in the absence of good evidence. On the other hand, if practices vary considerably, it will be clear to the panel that choices have to be made. The panel should be prepared and willing to make choices on a scientific basis. One such an example has been observed in Canada for the consensus conference on cesarean birth (Battista and Fletcher, 1988). The importance of topic selection implies that the preparation of a consensus development conference should include a survey on actual practice as opposed to opinion on practice. The objective of the survey would be to identify problem areas and the potential for evidence to aid in finding solutions. This "need assessment" phase was an integral part of the Canadian conference (Lomas, 1986).

Review of the Existing Evidence

To guarantee that the review of the literature will be complete and systematic, a specific individual should be assigned to this task. This staff person should prepare a summary of each main paper using a standard presentation that explicitly states the criteria for evaluation of the quality of the studies. This individual should also prepare a synthesis of the literature using the technique of meta-analysis whenever applicable.

The review should include published papers and unpublished results as well information on studies in progress. (Unpublished studies may tend to have negative findings more often than published ones do.) The review should consider articles related to efficacy or effectiveness of the therapeutic or preventive method evaluated and papers that address other criteria for assessments, for example:

- actual practices in the field and consequences for the health care system of any modification in practice
- direct or indirect adverse effects of the evaluated procedure.

The last point is particularly important as evaluations of "early detection procedures" by noninvasive technologies, generally regarded as safe, should include consideration of the consequences of false-positive or false-negative diagnoses.

Selection of the Panel

The panel should be as neutral as possible, and OMAR guidelines should be carefully followed:

A. Balanced representation on the consensus panel is crucial because a range of expertise on the panel is important to the panel's ability to deal with varied scientific material presented, and a diversified panel enhances the credibility of the consensus statement.
B. Panel members must be thoughtful, able to weigh evidence, and capable of collaborative work.
C. Panelists should have no vested interest in the technology being reviewed.
D. The size of panels has varied from 9 to 16 members; 12 or 13 is a reasonable working group.
E. The panel should contain balanced representation from various sectors of professional and community life and should not be professionally identified with advocacy or promotional positions with respect to the consensus topic (OMAR, NIH, 1988).

In addition, it may be useful to give the panel basic articles concerning the evaluation of the quality of data.

Selection of the Speakers

Speakers should be selected for their scientific expertise. Conference planning committees should provide precise information to speakers concerning the topic they have to address. Speakers who present reviews on the topic should be asked to include all of the opposing data and interpretations in their presentations. Speakers should receive information on the evaluation of the quality of published data. Speakers who present their own data should receive specific recommendations on the format for presenting the methodology of the research.

Formulation of Recommendations

The members of the panel are asked to base their recommendations on scientific evidence. It may be easier to comply with this rule if members of the panel are asked to explicitly support their recommendations by data from the literature (even if this evidence

is not included in the final statement) and if members of the panel use standard sentences as a grading scale for recommendations. In their study of the periodic health examination, Battista and Fletcher (1988) proposed the following scale.

A. There is good evidence to support the recommendation that the condition should be specifically considered in a periodic health examination.
B. There is fair evidence to support the recommendation that the condition should be specifically considered in a periodic health examination.
C. There is poor evidence regarding the consideration of the condition in a periodic health examination, and recommendations may be made on other grounds.
D. There is fair evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination.
E. There is good evidence to support the recommendation that the condition should be excluded from consideration in a periodic health examination.

SCIENTIFIC EVIDENCE AND CLINICAL PRACTICE

The discussion on how to increase the use of scientific evidence in a consensus development process can also be broadened to the discussion of the relation between scientific evidence and use of innovative procedures. Surveys conducted in France in the field of perinatology (i.e., on ultrasound examinations, electronic fetal monitoring, and prescription of beta-mimetics) give an example of this relation (Blondel et al., 1989; Breart, 1984; Ringa et al., 1986, 1989).

Of the doctors who provided antenatal care, 99 percent used routine ultrasound screening for every pregnant woman. When there was no complication, 2 percent of the respondents said that they prescribed only one ultrasound examination, 47 percent prescribed two, 45 percent prescribed three, and 6 percent prescribed four or more. Overall, 96 percent of the respondents considered that improvement in diagnosis and in pregnancy outcome were reasons for obstetric ultrasound (the latter was most frequently classified as the primary reason); 65 percent of the doctors mentioned demand by the mother and the safety of the procedure as reasons for ultrasound;

most of the time these reasons were ranked at the third or fourth level in the list of reasons (Blondel et al., 1989).

Only 1 of 455 people who attended deliveries did not use electronic fetal monitoring during labor. This screening was mostly applied to every woman: one percent of the respondents used it for high-risk women only. This pattern of practice was established before 1975 by 39 percent of the obstetricians, between 1975 and 1979 by 37 percent, and after 1980 by 24 percent. Most of the obstetricians gave earlier detection of fetal distress as the first reason for their using the procedure.

Overall, 99 percent of the respondents said that they used beta-mimetics for the prevention of preterm delivery. When asked under what circumstances they prescribed these drugs, 74 percent of the doctors said that they administered oral beta-mimetics prophylactically, 63 percent used intravenous or oral beta-mimetics when they detected signs of premature maturation of the cervix, 76 percent prescribed them for women with ruptured membranes, and 49 percent prescribed them when the cervix was dilated 4 cm or more. The principal reason given by 65 percent of the obstetricians for using beta-mimetics was the tocolytic effect of these drugs; 38 percent indicated delay of delivery as a primary concern. Two other reasons (better compliance with rest and their psychological effect) were less often quoted, and they were mostly the third and fourth reasons for using these drugs. Thirty-eight percent of the doctors had not changed their practice concerning the use of beta-mimetics since 1980, 11 percent had prescribed them more often, and 51 percent had prescribed them less often (Blondel et al., 1989).

Side effects of the treatment were more frequently reported as reasons for decreasing beta-mimetic use (90 percent) than was the lack of effectiveness of these drugs (74 percent), and the former was more frequently mentioned as the first reason. Improvement in neonatal care and the existence of other treatments to prevent preterm delivery were rarely given as the principal reason.

Upon analysis of the evidence drawn from the clinical trials (Breart, 1984; King et al., 1985; Ringa et al., 1986, 1989), it is clear that the actual practices have not taken into account the results of the randomized controlled trials that did not favor extensive use of any of these procedures or that were published after use of the procedures had spread.

However, these results do not indicate that use of any new technology is not based upon scientific evidence. A systematic review

of all articles published on ultrasound examination in four obstetrical journals read by French obstetricians was performed for the period from 1979 to 1984. Of the 182 papers identified, 137 were related to the diagnostic assessment of ultrasound, 14 concerned the effects on medical practice, and only 9 papers studied the effects of ultrasound scanning on health. Most of the 137 papers on diagnostic assessments of ultrasound came out in favor of the procedure. Many articles have been published on the topic; thus, the decision to use this technology seems to have been based on scientific publications. However, these publications addressed only the first level of evaluation, namely, the diagnostic value of the procedure, whereas an epidemiologist would have expected a more complete evaluation.

According to the data presented here, it can be said that the decision to practice routine ultrasound examination has been made on a theoretical basis, rather than on proof of its efficacy and effectiveness. The same holds true for the diffusion of beta-mimetics for a variety of clinical conditions; the propagation of the drugs was based on their ability to stop uterine contractions and not on the proof of their efficacy in reducing the preterm birth rate (the original purpose of these compounds). Similar results have also been observed for electronic fetal monitoring (Breart, 1984).

The surveys conducted in France also revealed that the decision to modify any given practice seems to be based on possible side effects rather than on doubts concerning its effectiveness. Therefore, if clinical practice is mainly based on "theoretical" consideration, as well as on side effects, this must be taken into consideration in the consensus development process. To take into account the first point, it may be useful to advise the members of the consensus development panel to clearly distinguish between theoretical considerations and actual proof. For this task, the panel should receive information about the theoretical basis for the effectiveness of the proposed intervention. This summary should include presentations of causal pathways in a manner similar to that proposed by Battista and Fletcher (1988) for preventive practices (Figure 1).

The presentation of the theoretical basis as well as the causal pathways should be accompanied by a list of questions concerning the efficacy and the effectiveness of the proposed interventions at each step. For the instance of early detection of high serum cholesterol levels, the ultimate purpose is to prevent the occurrence of coronary heart disease (CHD). The intermediate steps along the causal pathway are the ability of the detection procedure to identify

individuals with high cholesterol levels (step 1), the ability of dietary counseling and/or pharmacologic treatment to lower serum cholesterol levels (step 2), and the successful prevention of CHD resulting from the lowering of serum cholesterol levels (step 3). The most important link here (step 4), is the efficacy or effectiveness of dietary counseling and/or pharmacologic treatment. The utility of early detection depends upon how well these intermediate steps lead to prevention of CHD in asymptomatic individuals with high serum cholesterol levels (step 5).

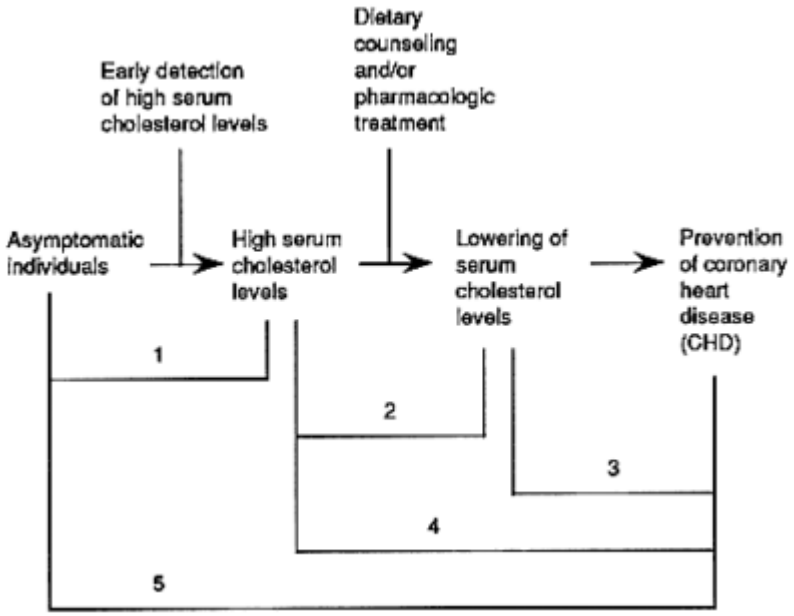

Figure 1 Early detection of high serum cholesterol levels.
SOURCE: Battista and Fletcher (1988).

A very careful search of published and unpublished results should be performed to address the issue of side effects. Since the theoretical beneficial effects of a given procedure are taken into consideration in the consensus development process, the theoretical side effects have to be considered with the same weight.

The problem of the use of scientific evidence to make decisions is not unique to consensus development, as it is a general problem for

decision making in clinical practice. Better use of scientific evidence in the consensus development process may come from understanding and consideration of the process of clinical decision making by practitioners.

References

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

Battista, R.N., and S.W. Fletcher. 1988. Making recommendations on preventive practices: Methodological issues. American Journal of Preventive Medicine 4(4)(Suppl.):53-67.

Blondel, B., V. Ringa, and G. Breart. 1989. The use of ultrasound examinations, intrapartum fetal heart rate monitoring and beta-mimetic drugs in France . British Journal of Obstetrics and Gynaecology 96:44-51.

Breart, G. 1984. Un problème de décision en perinatologie: l'Enregistrement continu du coeur foetal pendant le travail. Revue d'Épidémiologie et de Santé Publique 32:34-39.

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

King, J.F., M.J.N.C. Keirse, A. Grant, and I. Chalmers. 1985. Tocolysis. The case for and against. Pp. 199-200 in Preterm Labor and its Consequences, R.W. Beard and F. Sharp, eds. London: Royal College of Obstetricians and Gynaecologists.

Levitt, S.H., and R.A. Potish. 1988. The consensus statements on adjuvant chemotherapy in breast cancer: Fact or artifact. American Journal of Oncology 11:73-76.

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

Lomas, J., G.M. Anderson, M.W. Enkin, E. Vayda, R. Roberts, and B. MacKinnon. 1988. The role of evidence in the consensus exercise, results from a Canadian consensus exercise. Journal of the American Medical Association 259(20):3001-3005.

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 non-consensus conferences on coronary heart disease. Lancet i:1087-1089.

Ringa, V., B. Blondel, and G. Breart. 1986. Évaluation des traitements contre la prematurité. Pp. 191-198 in Control and Management of Parturition, C. Sureau, P. Blot, D. Cabrol, et al., eds. Paris: Institut National de la Santé et de la Recherche Médicale-John Libbey.

Ringa, V., B. Blondel, and G. Breart. 1989. Ultrasound in obstetrics: Do the published evaluative studies justify its routine use? International Journal of Epidemiology 18:489-497.

Format and Conduct of Consensus Development Conferences: A Multination Comparison*

Elizabeth A. McGlynn, Jacqueline Kosecoff, and Robert H. Brook

INTRODUCTION

The purpose of this paper is to compare the methods used in nine countries to organize and conduct consensus development conferences on scientific issues related to the delivery of medical care. Our comments focus primarily on the format and conduct of these conferences. However, we also briefly discuss the context, selection, and scope of topics; the role of evidence; and the dissemination of results. We confine our remarks on these latter topics to descriptive statements that illuminate issues relevant to a discussion of process.

We address the process by which consensus development conferences are conducted because the approach taken may influence the results, or the acceptability of the results, either intentionally or unintentionally. Thus, at each step of the process it is important to consider whether the mechanics of conducting the conference are enhancing or detracting from achieving its goals and objectives. To the extent that the mechanism itself interferes with the intended outcomes, changes in the approach are warranted.

* This paper is to be published in a forthcoming issue of the *International Journal of Technology Assessment in Health Care,* Cambridge University Press, New York.

METHODS

We set out to examine the process by which consensus development conferences are conducted in Canada, Denmark, Finland, The Netherlands, Norway, Sweden, Switzerland, the United Kingdom, and the United States. Although most of these countries began with the model used in the United States by the National Institutes of Health (NIH), many important variations have been introduced (Casparie and van Everdingen, 1985; Jennett, 1985; Perry, 1987, 1988; Stocking, 1985; Vang, 1986).

The 10 programs considered are:

1. The Canadian Task Force on the Periodic Health Examination (Canada I)
2. The Canadian Research Group (Canada II)
3. The Danish Medical Research Council and the Danish Hospital Institute
4. The Medical Research Council of the Academy of Finland
5. National Organization for Quality Assurance in Hospitals in The Netherlands
6. The Norwegian Institute for Hospital Research and the Norwegian National Research Council
7. The Swedish Planning and Rationalization Institute for the Health and Social Services
8. The Swiss Institute of Public Health
9. The King's Fund Forum in the United Kingdom
10. The U.S. National Institutes of Health Office of Medical Applications of Research (OMAR).

The consensus development conference is a complex entity, and researchers have undertaken scientific evaluations of the conferences and their impact in their respective countries (Calltorp, 1988; Johnsson, 1988; Kanouse et al., 1989; Lomas et al., 1988; Wortman et al., 1988). An international comparison is challenging because there may be as much or more variation within a particular country regarding how individual conferences are conducted as there is among countries (Andreasen, 1988). For example, more than 75 such conferences have been held by NIH in the United States, and although the model was basically the same, implementation certainly varied across the conferences. We have attempted to use the most recent data available to represent each country in order to capture any

changes that have taken place over time in the conduct of the conferences.

We used three sources of information to prepare this paper:

1. Profiles prepared by representatives from each country for the June 1989 workshop on international consensus development for medical technology assessment organized by the Institute of Medicine
2. Published articles and consensus statements related to such conferences
3. Comments received after the conference from representatives of each country.

We divided the consensus development process into four stages for our investigation:

1. Context of the consensus development process
2. Prepanel process
3. Panel composition
4. Consensus panel meeting.

The context within which consensus conferences are conducted encompasses the nature of the audience, the topics considered, and how the topics are selected. The prepanel process includes all activities that are required to stage the consensus conference, such as selecting the chair, panel members, and presenters as well as preparation of background information. Panel composition includes the type of panelists, their qualifications, and the process by which they are selected. The consensus panel meeting stage describes the activities at the actual conference, such as the use of public forums and private sessions, the type of information considered in arriving at consensus, and the group process by which consensus is actually achieved.

RESULTS

In this section, we describe how each of the countries described in this paper conducts consensus conferences. We have organized the section around the four major areas investigated: context, prepanel process, panel composition, and the consensus panel meeting. Within each area we addressed several questions. Tables 1 through 4 summarize our findings.

Context

We begin by examining the context in which consensus development conferences are conducted because the format of conferences may differ depending on the intended audience for the conference and what issues are addressed. The results are displayed in Table 1.

Audience

The potential audience for any of these conferences is fairly broad. In most cases, countries indicate that a variety of different groups are included in the target audience. We, however, tried to determine which audience was the primary target for each country in order to detect differences in the focus of planning and conference efforts. Six of the programs consider health professionals to be the audience for their consensus conferences (Canada I and II, Finland, The Netherlands, Switzerland, and the United States). Within those programs, most seem to be addressing physicians who are involved in direct patient care; however, academic-based physicians and research scientists are also targeted. Two countries (Norway and Sweden) consider national health authorities, such as health planners and political decision makers, as well as health professionals to be the primary audience for the conference. Denmark and the United Kingdom indicate that the public and health professionals are their primary audiences; most countries suggest that the public may be interested in the findings. Switzerland is the only country that included third-party payers in its audience list.

Issues

The scope of issues addressed in any one conference varies considerably, and most countries appear to have a fairly broad perspective. Countries typically indicate that they intend to address a variety of issues, including safety, effectiveness, efficacy, appropriateness, service requirements, economics, political and social concerns, and needs for future research. The evidence that is available in any particular area limits the issues that can be addressed. Most consensus development conferences address issues of effectiveness and efficacy (the strength of the clinical literature), while few address economic or cost issues (rarely addressed in clinical or epidemio

TABLE 1 Context in Which Consensus Development Conferences Are Conducted

Context	United States (NIH)	Canada I
Who is the audience for consensus product?	Health professionals; public	Practicing primary care physicians; medical educators; policymakers; general public
What issues does panel address?	Benefits; appropriate use of technology; future research	Appropriateness of procedures to prevent adverse outcomes; focus is on the condition rather than the procedure; effectiveness (benefit > harm?) or, if no information, efficacy
Who selects topics?	Institutes within NIH and OMAR	Task Force on Periodic Health Examination
No. of topics	75	1

Canada II	Denmark	Finland
Practicing physicians	Public; health planners; health professionals	Health care providers; researchers; public; providers of agricultural products (cholesterol)
Appropriateness benefit > risk	Effectiveness ; safety; adoption level; economic, organizational, and ethical implications	Safety; efficacy; effectiveness; service requirements; adoption level; economic and other (e.g., ethical) implications
Investigators from participating universities (McMaster, Univ. of Toronto, Univ. of British Columbia)	Subcommittee of Danish Research Council	Medical Research Council
2	6	3

Context	The Netherlands	Norway	Sweden
Who is the audience for consensus product?	Clinicians; hospital administrators; policymakers	National health authorities; hospital owners; health professionals	Politicians; administrators; educators; planners; public
What issues does panel address?	Primary: safety, efficacy, effectiveness; also: cost-effectiveness, service requirements	Effectiveness; efficacy; risks; costs; consequences: organizational, psychological/social, ethical	Safety; efficacy; effectiveness; cost/economic implications; minor role for ethical, legal, social implications
Who selects topics?	Scientific Council of CBO	Technology Assessment Committee	Representatives from Medical Research Council and Swedish Planning and Rationalization Institute of Health Services form a steering committee
No. of topics	22	2	9

Switzerland	United Kingdom	Degree of Variation among Countries
Physicians; health care administrators and decision makers third-party payers	National/local policymakers; clinicians; health professionals; public	Six focus on health professionals; one on public two on national politics or planners; United Kingdom considers all potential audience groups to be equally important
Safety; effectiveness; substitution versus replacement; appropriate uses; cost	Scientific merits of technology for different patients; cost; safety; effectiveness; legal, social, service, ethical implications	Scope varies: most have global perspectives; few seem likely to truly address efficacy; four examine appropriateness; one examines whether technology is a replacement or substitute
Planning committee	King's Fund steering group including clinicians, professional bodies, representatives of Dept. of Health	Most combine government and specialty society input (except United States and Canada II)
1	7	1-40+

logic studies). Four programs (Canada I and II, Switzerland, and the United States) address appropriateness of technology use.

Topic Selection

Most countries rely upon a combination of government and specialty societies to select topics for consensus development conferences. The exceptions are the United States and Canada II. The United States relies upon staff within the government-funded NIH to select topics, but does not formally involve relevant U.S. medical specialty societies (Office of Medical Applications of Research, National Institutes of Health, 1988). Canada II selects topics based on interests of the principal investigators and the availability of funding to explore various technologies; subsequent to topic selection, the researchers generally seek government funding and specialty society support. In most countries, topic selection comes from a standing committee responsible for technology assessment activities (e.g., Scientific Council of CBO in The Netherlands, Research Councils in Denmark and Norway).

Topics

The variation in topic selection within any country depends in part on how many consensus conferences have been conducted. In the United States and The Netherlands, which have conducted more than 75 and 27 conferences, respectively, there is considerable variation in the types of topics selected. In general, across all countries, four types of topics have been addressed: treatment of conditions (e.g., otitis media, schizophrenia), diagnostic procedures (e.g., detection of breast cancer, diagnosis of venous thrombosis), therapeutic procedures (e.g., adjuvant chemotherapy for breast cancer, coronary artery bypass surgery), and planning issues (e.g., need for asylum for the mentally ill, impact of routine HTL V-III [human T-lymphotropic virus type III] antibody testing on the safety of the blood supply).

Conferences have also considered technologies at varying stages of development, from new or emerging to established and outmoded. The stage of development is important both because it has implications for the availability of evidence and because the degree of con

troversy surrounding the technology may be related to its stage of development.

Prepanel Process

Although most countries spend between six months and one year planning for the consensus development conference, different types of activities are undertaken during that time, as shown in Table 2.

Responsibility for Planning

Three different types of groups have responsibility for planning the consensus conferences. In seven countries (Denmark, Finland, The Netherlands, Norway, Sweden, Switzerland, and the United Kingdom), planning is done by a special working group generally appointed by the committee responsible for selecting the topic. In two countries (Canada and the United States), staff are responsible for preparing for the conference. In the United States, employees of the relevant bureaus, divisions, or institutes within NIH conduct planning activities. In Canada II, planning is done by the academic-based research team. In one country (Canada I), planning is done by the consensus panel itself.

Review of Literature

Most of the conferences rely upon oral presentations of scientific evidence as the basis for making judgments about the technology under consideration. In most countries, information is sent ahead of time to panel members to prepare them for the conference. For example, in the United Kingdom, panel members receive introductory textbook-type information (necessary for the lay members of the panel) as well as a comprehensive set of readings, drawn from a computer search of all relevant literature, and abstracts prepared by the expert speakers who will present information at the actual meeting.

Canada I and II and Switzerland prepare a formal synthesis of the literature in advance of the panel meeting. For Canada I, which arrives at a consensus through several iterations of papers, the synthesis of the literature provides a starting point for developing the

TABLE 2 Types of Activities Undertaken during the Prepanel Process

Prepanel Process	United States (NIH)	Canada I
Time spent in preparation for panel meeting?	12-15 months	Ongoing process
Who is responsible for planning and preparation?	NIH staff: OMAR coordinator; bureau, institute, or division (BID) coordinator; representatives from sponsor agencies	Task force members
Extent and nature of literature review?	Bibliography and paper abstracts sent in advance; sometimes background reports prepared or expert summaries of state of science; BID coordinator determines scope of literature search	Original reports of clinical trials or epidemiologic studies; quality of evidence is graded; expert opinion used if no evidence is available
Is there a review of current clinical practice (e.g., efficacy, current uses of the technology)?	For some topics, current uses are summarized	Some efforts made to measure patient compliance; working on rules for using data on efficacy (e.g., decision analysis techniques)
Are patient outcome data used in planning?		Both intermediate and final outcomes are considered; disease prevention is main outcome; examining broader range of outcome measures (e.g., quality of life)
Are recommendations or specific questions prepared for panel in advance?	Conference questions developed by planning committee (four to six questions with one pertaining to future research); structure of questions allows answers to be drawn strictly from scientific literature	Part of background paper which is prepared by panel member and then revised over time

Canada II	Denmark	Finland
12 months	6-10 months	8-12 months
Project staff responsible for preparation of materials and meeting planning	Planning committee: head of expert group, head of questioning panel, head of secretariat, one or two topic specialists; appointed by DMRC subcommittee	Planning Committee is nominated by Medical Research Council
Synthesis of literature; all relevant research examined; mutually exclusive indications for use of technology developed	Abstracts of papers to be presented are sent in advance; no mention of formal literature review	Data are presented by speakers and include reviews, controlled trials, epidemiologic studies, and state of science; abstracts of papers are sent to panelists in advance
During needs assessment phase, presented to panel before conference by analysts; describe claims data	For some topics, current uses are summarized	Current clinical practice used only once in planning phase to develop preliminary responses to questions; efficacy considered throughout
To some extent in development of indications		
Indications for use of procedure are rated by panelists in advance of meeting; formal consensus questions prepared in advance based on needs assessment phase	Planning committee formulates main conference questions	Prepared responses to questions was attempted only once (for otitis media)

Prepanel Process	The Netherlands	Norway	Sweden
Time spent in preparation for panel meeting?	9-12 months	12 months	8-12 months
Who is responsible for planning and preparation?	All planning and preparation done by the working group in collaboration with CBO staff; 5-10 meetings	Technology Assessment Committee is established by National Research Council	Specially appointed working group
Extent and nature of literature review?	Syllabus is compiled; experts present evidence; few literature reviews have been done; evidence graded using Sackett method; background papers use research plus clinical experience	Randomized controlled trials (RCTs) and compilation of scientific evidence (abstracts of expert testimony)	Reviews, RCTs, epidemiologic studies, expert presentations; abstracts sent in advance
Is there a review of current clinical practice (e.g., efficacy, current uses of the technology)?		Ad hoc study before conference to document spread and use of technology, cost, and health effects	Experts asked to present data on actual state of practice
Are patient outcome data used in planning?	Efforts being made to introduce decision analysis techniques into conferences in the preparatory phase	Where available	Not explicitly
Are recommendations or specific questions prepared for panel in advance?	Draft consensus statement is prepared in advance and sent to conference participants	Yes	No

Switzerland	United Kingdom	Degree of Variation among Countries
14 months	9-12 months	Most seem to take about a year
Planning committee drawn from Swiss Public Health Institute, Federal Office of Social Insurances, appropriate medical associations, and Sickness Funds Association	Planning committee nominated by King's Fund	Two have staff; seven have working groups; one has a consensus panel
Synthesis of bibliographic references; full text of expert reports to panel four weeks prior to conference; experts given guidelines on how to synthesize literature.	Reading material sent to panel before the conference includes introductory texts, actual literature, and 1,000-word abstracts from speakers	Three have formal synthesis; three have oral testimony; one has bibliography; seven have abstracts sent in advance; one has reading material
Data on magnetic resonance imaging was collected for one year through a registry	Small surveys done where relevant; experts asked to present data	Six rely on expert testimony; one asks panelists to use experience; one uses review in planning conference; four conduct special surveys
Partially, although no validation or controlled trial data available	Where available	Six use outcomes to some degree (usually based on availability); The Netherlands working on incorporating patient outcome data through decision analysis techniques
Specific questions prepared 6 months before conference; draft of statement prepared 2 weeks before conference	Advance preparation of questions or statement specifically disallowed	Five have recommendations in advance; six have questions in advance; two have no advance preparation; one has patient indications

consensus statement. Canada I has the most formal approach to rating the relationship between the recommendations made and the scientific merit of the literature on which those recommendations are based (Battista and Fletcher, 1988). Those recommendations based on clinical trials receive the highest grade, whereas those recommendations based on the opinions of respected medical authorities alone are given a lower grade.

Canada II combines the NIH consensus development conference approach with a model developed for the RAND Corporation/University of California at Los Angeles (UCLA) Health Services Utilization Study (Chassin et al., 1987; Lomas et al., 1988). The latter approach uses a synthesis of the literature as a starting point for developing mutually exclusive categories of patients who might be candidates for the technology under evaluation. These "indications" for technology use are then rated by the panel, using a two-tier Delphi approach, on a nine-point scale ranging from extremely appropriate (nine points) to extremely inappropriate (one point); values in the midrange (four to six points) represent equivocal ratings of the use of the technology in those patients. Where no literature exists to inform the ratings of certain patient categories, expert opinion is used to expand the available information.

In some cases, U.S. consensus development panels have had a synthesis prepared, but this varies from conference to conference. In eight countries (Denmark, Finland, The Netherlands, Norway, Sweden, Switzerland, the United Kingdom, and the United States), abstracts of papers that will be presented at the conference are sent to panel members in advance.

In general, there appears to be very little systematic effort to survey the available literature and summarize the state of the science for the topic under consideration. Only infrequently is an attempt made to synthesize the literature and assess the scientific merit of the research. Further, in those countries that rely upon experts to present summaries or evaluations of scientific information, it is unclear how the experts who will present testimony are selected and whether the group of experts fairly represents the range and distribution of evidence in the literature. Most consensus statements do not refer to the literature upon which findings are based, which makes it difficult to determine whether and to what extent the literature has had an influence on the conference's conclusions and whether some literature is more influential than other. Most countries indicate that

they favor using results from randomized clinical trials and epidemiologic studies, but it is difficult to ascertain the extent to which this is true in actual practice.

Review of Current Clinical Practice

Most countries indicate that their consensus development conferences are intended to address issues related to the efficacy of certain medical technologies. Four countries (The Netherlands, Norway, Switzerland, and the United Kingdom) conduct special surveys to collect data on current uses of the technology. In Finland, current clinical practice experience was used once (for otitis media) in the planning phase to develop preliminary responses to questions; however, this practice is not the rule. Experts are relied upon in six countries (Canada II, Denmark, The Netherlands, Sweden, the United Kingdom, and the United States) to provide information on current practice. An evaluation of the U.S. experience, however, demonstrated that conference topics and subsequent recommendations addressed issues, such as the discontinuation of Halstead radical mastectomies, that were largely obsolete in clinical practice prior to the actual conference (Kosecoff et al., 1987). Canada I is working on decision-making rules for using data on efficacy and is considering how decision analysis might be used to expand the information available with which to make decisions about technologies. Canada II uses analyses of claims data for information on current uses. Canada II also uses the experience of individual panel members to represent what is occurring in clinical practices. In rating indications for the use of the technology, panel members are asked to draw on their own experience and make ratings based on how they would approach treatment for each type of patient seen in their own practice (Lomas et al., 1988). Sometimes the experiences of panel members also result in changes to the structure or form of the indications (Park et al., 1986).

Patient Outcome Data

Although many countries indicate that they assess the appropriateness and outcomes of use for a technology, the definitions of *appropriate* and *outcome* are not always explicit. Possible approaches to defining appropriateness include cost-benefit, cost-effectiveness,

and risk-benefit analyses. In each case, information about outcomes experienced by different types of patients is necessary for a formal evaluation of appropriateness. Only Canada I has made an explicit effort to include data on patient outcomes in the decision-making process; both intermediate and final outcomes are considered, but the prevention of disease is the primary final outcome. Canada I is also examining the introduction of broader outcome measures (e.g., quality of life) into its consensus development process (Battista and Fletcher, 1988). Canada II uses a risk-benefit definition of appropriateness for structuring and rating indications for use of the technology (Lomas et al., 1988). The Netherlands has indicated that efforts are under way to introduce decision analysis techniques into the conference planning process and, because the consensus development panel itself also does the planning, such efforts would certainly have the opportunity to affect panel deliberations (Klazinga et al., 1987). Other countries indicate that they use such data when they are available, but the extent to which this is done and the effect of such information on the consensus development process remains unclear.

Advance Preparation of Recommendations or Questions

Because of the time pressures surrounding the actual consensus development conference, we were interested in the extent to which recommendations or questions are prepared in advance in draft form as a starting point for panel deliberations. In five countries (Canada II, Denmark, Norway, Switzerland, and the United States), questions to be addressed by the panel are prepared in advance. This provides the focus for the conference and presumably affects the selection of experts who will present evidence at the conference. In Finland, prepanel drafting of answers to questions was attempted only once to help manage a particularly controversial debate regarding treatment of otitis media. In Canada II, as previously described, a formal set of indications is created. In four countries (Canada I, The Netherlands, Norway, and Switzerland), some form of draft consensus statement is prepared in advance and is made final during the conference. Canada I uses an entirely different approach from the rest of the countries; it relies on a permanent panel that is responsible for an ongoing process of writing papers evaluating the use of certain preventive measures and diagnostic procedures used in peri

odic health examinations (Battista and Fletcher, 1988). The panel writes and revises papers (the equivalent of consensus statements) based on deliberations at regularly scheduled panel meetings; as long as two years may be spent on the preparation of these statements. In The Netherlands, panelists receive the draft consensus statement in advance of the meeting. In two countries (Sweden and the United Kingdom), no advance preparation of questions or recommendations is undertaken; in the United Kingdom, such advance work is specifically disallowed.

Panel Composition

The acceptability of consensus development conference recommendations depends to some degree on the qualifications of panel members. We examined variations in the panel size, the process of selection, and qualifications of the panel chair and members. The results are displayed in Table 3.

Panel Size

Panels range in size from 9 to 18 members. Panel size is an interesting consideration given the implications for the group process required to accomplish the task at hand. Groups of 18 will probably function differently than groups of 9; the likelihood of everyone participating equally in the decision-making process declines as the number of panelists increases. On the other hand, larger groups enhance the range of viewpoints that can be taken into account in considering complex issues.

Panel Chair

In general, the planning committee selects the panel chair. The exception to this is in Sweden, where no chair is selected; leadership is shared by different persons drawn from the expert group (i.e., the leadership rotates depending on the issue under discussion). In most countries, the qualifications of panel chairs are fairly general, including the individual's stature as a scientist and leadership abilities. An evaluation of the process of selecting panel chairs in the United States suggests that this is done informally and is based on staff familiarity with individuals who are currently conducting re

TABLE 3 Consensus Development Conference Panel Composition

Panel Composition	United States (NIH)	Canada I
Number of panelists	9-16	
Who selects the panel chair?	Director of sponsoring BID with advice from the planning committee	
Background of the panel chair?	Stature in field as physician or scientist; skills in conducting meetings; no established position on technology under consideration	
Who selects panel members?	Planning committee	
Background of the panel members?	Researchers; health professionals; epidemiologists/ biostatisticians; public (ethicists, economists, lawyers, theologians, consumers); moved from balanced to neutral panels	Clinical epidemiologists from family practice, internal medicine, psychiatry, pediatrics, infectious disease, nursing. All have academic affiliations; balanced

Canada II	Denmark	Finland
10	12-16	16
Principal investigator	Subcommittee of Danish Medical Research Council	Planning committee
	Prominent medical doctor in specialty not related to technology being assessed	Distinguished researcher; skill in conducting meetings
	Planning committee	Planning committee
Seven physicians (five obstetricians, one general practitioner, one neonatologist); generalists and specialists; academic and community practice; users/nonusers of technology; regional representative; three nonphysicians (one epidemiologist, one lawyer, one consumer); balanced	Physicians, journalists, other academics, politicians, patient representatives; none are working with technology being assessed; neutral	Researchers; clinicians; other health care professionals; lay representatives (social scientists, editors)

Panel Composition	The Netherlands	Norway	Sweden
Number of panelists	10-15	10-12	16-18
Who selects the panel chair?	Scientific Council of CBO	National Research Council's technology assessment committee	Leadership is shared by different persons generally drawn from the expert group
Background of the panel chair?	Scientific expert	Member of the planning committee	
Who selects panel members?	Scientific Council of CBO in consultation with chair	Technology Assessment Committee	Steering committee
Background of the panel members?	Scientific association representatives, experts. May add: nursing, general practitioners, physiotherapists	Half medical experts; half nonmedical experts (health economists, ethicists, lawyers, journalists, health administrators, health politics); neutral	Half medical experts; half from health, care profession economists, policy, epidemiologists, administrators, patients

Switzerland	United Kingdom	Degree of Variation among Countries
14	12	9-18
Planning committee	King's Fund	In general, planning committee; Sweden does not have a chair
Skill in leading group decision process; strong knowledge in specific technology area	Creditable persons in different specialty (e.g., neurosurgeon for coronary artery bypass panel)	Scientific stature; leadership ability (not really clear other than this)
Planning committee	King's Fund and planning committee	Panel selected generally by planning committee
Researchers in field; health professionals (specialists and those in related fields); epidemiologists; health care administrators; third-party payers; balanced	Two specialists in topic; half medical (not technology users); half nonmedical (judicial model); neutral	All use a combination of scientists, physicians, and lay people; three balanced, four neutral, three unknown

search in the topic area (Wortman et al., 1988). In the United States, in addition to stature and leadership requirements, the panel chair must have no established position on the technology under consideration. In one conference in the United Kingdom, the panel chair was deliberately selected from a specialty that does not use the technology being evaluated (e.g., a neurosurgeon chaired the coronary artery bypass panel). Denmark also selects an individual from a nonuser physician specialty to chair the panel.

Panel Members

The panel members are also generally selected by the planning committee, and it appears that the selection process is not systematic, but rather relies upon the people with whom the planning committee is familiar. In small countries, the informal approach may produce the same list of potential members as a more formal process. In all countries, the panels are composed of both scientists and laypeople. Generally, the panel distribution is half scientist and half laypeople. In the Canada II panel on cesarean section, there were five obstetrician-gynecologists, one general practitioner, one neonatologist, one epidemiologist, one lawyer, and one consumer (or 70 percent physicians and 30 percent other). Nonmedical experts tend to include health economists and policymakers, epidemiologists, administrators, and patients. In the United Kingdom, among the medical panelists, two were specialists in the topic area and the other four were nonusers of the technology.

There are two approaches to the composition of panels: balanced and neutral. Balanced panels are designed to encompass the range of opinions in the field; experts representing different viewpoints are brought together on the panel itself. The balanced panel has the advantage of well-informed members who can engage in an exchange of opinion. The disadvantage of balanced panels is that it may be difficult to find meaningful middle ground. A neutral panel is composed of people who do not have a stated opinion on the technology being evaluated; the neutral panel is more like a jury who will weigh the evidence and come to a decision. The advantage of the neutral panel is that its judgments will presumably be based upon the evidence presented, rather than being influenced by preexisting opinions or experiences of panelists. The disadvantage, particularly for conferences concerned with complex technologies, is that it may be

problematic to educate the panel sufficiently to allow for informed judgments. Three programs endeavor to use balanced panels (Canada I and II and Switzerland). The rest of the countries have neutral panels or do not indicate their approach on this dimension. The United States has changed over time from balanced panels to neutral panels because of experiences with balanced panel meetings "degenerating into unconstructive debate" (Jacoby, 1985).

Consensus Panel Meeting

The processes for the conduct of panel meetings exhibit perhaps the least variation overall. Most countries have adopted the format used by NIH with little concrete variation (Table 4).

Product

In nine programs, the consensus panel is charged with producing a written statement of consensus on the issues addressed in the topic area. In Canada I the product is practice guidelines; in Canada II the product includes a consensus statement as well as ratings of indications for use of the procedure (e.g., cesarean section) in different types of patients. It does not appear that any country has detailed guidelines concerning the content and/or specificity of the consensus statement. The U.S. guidelines, for example, "encourage panelists to strive for statements that (1) recommend concrete, specific actions; (2) differentiate patients into subclasses when appropriate; and (3) offer didactic advice to the clinician on precise techniques that should be used" (Kanouse et al., 1989; Mullan and Jacoby, 1985). The level of detail in the statements we reviewed varies considerably (both between and within statements). Statements (with the exceptions of Canada I and II) tend to make little or no reference to the literature on which the findings are based, and the types of recommendations vary from general to specific.

Meeting Time

Most countries follow the U.S. model of conferences meeting times, which last two and a half days, during which the panel writes or modifies the consensus statement in an all-night session starting on the second day. The exceptions are Canada I (which relies upon

TABLE 4 Consensus Panel Meeting Processes

Consensus Panel Meeting	United States (NIH)	Canada I
What is the final product?	Consensus statement; press conference	Synthesis of scientific evidence; recommendations for practice policies (guidelines for practice)
How long does the panel meet?	Two and a half days (including nights)	Four to eight iterations over one to two years; one hour discussion per condition per meeting
Extent of public forums; number of audience members?	One to two days of evidence presented in public; audience: 200-700	None
Private panel sessions?	Meet night before public forum; begin drafting statement after first day of evidence	All work done in private meeting
How is consensus defined?		
Who makes the final decision?	Panel, but the audience may offer suggestions	Task force
Are formal votes taken?	No	No
What criteria are used for making decisions?	None	Value of scientific evidence based on predetermined criteria
How is disagreement handled?	In some cases a minority report was prepared (rarely done)	Not clear
Sensitivity to group process issues?	Not clear	Not clear

Canada II	Denmark	Finland
Indications rated for appropriateness, inappropriateness, and equivalent uses of technology and consensus statement	Consensus statement	Consensus statement
Two and a half days	Three days	Two and a half days (including nights)
Presentations on unpublished data from 11 experts	Public sessions on all three days; audience may ask questions; audience: 150	Two days of plenary sessions with medical community, expert witnesses, public; audience: 150-160
All work done in private meeting to produce an interim consensus statement	Expert group and panel meet night before; panel meets evening of first day; consensus statement prepared in closed session	One to two meetings before conference is held; meet night before plenary session
	Unanimity	
Panel	Questioning panel	Panel
Yes	No	No
Benefit > risk	None	None
Comparison of ratings indicates level of agreement	Public debate in newspapers and professional journals after dissemination of results	Not clear
Voting is done anonymously; opinions known before panel chair can ensure these are brought up	Not clear	Not clear

Consensus Panel Meeting	The Netherlands	Norway	Sweden
What is the final product?	Consensus statement	Consensus statement press conference	Consensus statement
How long does the panel meet?	One to two days (including nights)	Two and a half days (through night on last day if necessary)	Two and a half days
Extent of public forums; number of audience members?	Practitioners and public; audience: 150-1,000	Physicians, experts and public audience: ~200	One and a half days for expert presentations, questions, and discussion; audience: 200
Private panel sessions?	Experts meet several times ahead of the actual conference	Private sessions to write the consensus statement	Panel starts writing on afternoon of second day
How is consensus defined?	Chair formulates guidelines		Panel sets the rules
Who makes the final decision?	Audience is asked if they agree; final decision is with panel	Panel	Panel presents at press conference
Are formal votes taken?	No	No	No
What criteria are used for making decisions?	None	None	None
How is disagreement handled?	Points of disagreement are mentioned in consensus statement text	Not clear	Press conference
Sensitivity to group process issues?	Not clear	Not clear	Not clear

Switzerland	United Kingdom	Degree of Variation among Countries
Consensus statement; summary presentation in open forum; press conference	Consensus statement	Nine consensus statements; one practice guideline; one indication rated (plus consensus statement); many hold press conference to release statement
Two and a half days (no all night session)	Three days, 21 hours; no overnight session	Most are two and a half days and write through night Canada I is four to eight iterations over one to two years
One day of evidence; half day of discussion; audience: 150	Same as NIH; audience: 200	All except Canada I rely on public forum
Two half-day executive sessions; meet night before plenary session	Executive session for writing statement; some revisions made after public presentation	All write statements in executive session; Canada II has indications rated before and during meeting
No formal voting; comments on draft	No clear definitions	No clear definitions
Chairman	Panel	Panel makes decision in all but one case; public input in two; chairman decides in Switzerland
No	No	No, except Canada II
Value of scientific evidence; conformity with state of practice	None	Generally no explicit criteria used
Detailed discussion; if no consensus, disagreement mentioned in final document	Not clear	U.S., some minority reports; Denmark, public debate; most have press conferences; Canada II, ratings; Switzerland and The Netherlands, mentioned in consensus statement
Little experience handling problems in a nonhomogeneous group; panel did not question experts thoroughly	Not clear	

four to eight iterations of a paper over one to two years), Canada II (which relies upon rated indications from an iterative process using the RAND/UCLA modified Delphi method to form the basis for panel recommendations), and The Netherlands (which meets 5 to 10 times before the final meeting at which the consensus statement is completed).

Public Forums

All programs except Canada I rely upon public forums during which evidence is presented. The audience size ranges from 150 to 1,000, with most audiences being about 200. In most cases, the audience may ask questions or make comments, although the degree to which this is done varies. It is not clear whether or not the audience contributes substantially to the process; rather, it appears that the audience is in attendance primarily to observe the proceedings.

Private Panel Sessions

The bulk of work required to answer questions and to draft, revise, and complete the consensus statement is done in private sessions. Most panels meet the night before the conference begins to receive instructions about the process of the meeting itself and to hear the "charge" to the panel (i.e., questions to be answered). In many cases, the questions to be addressed have already been made public (Norway). In Finland and The Netherlands, the panel convenes prior to the actual conference. In Finland, one or two meetings are held; in The Netherlands, several meetings are held during the conference planning process. The familiarity of panelists with one another presumably affects the group process dynamics. This may be accomplished in some countries without preconference meetings because the panel members already know each other (Wortman et al., 1988).

Definition of Consensus

Little information exists on how each country defines consensus. The implied definition is unanimous agreement with the consensus

statement. Because unanimity is expected, there are few formal mechanisms available for dealing with disagreement (with the exception of Canada II). The Netherlands and Sweden indicate that the chair and panel (respectively) are responsible for formulating guidelines to define consensus, but whether that requirement introduces any more systematic consideration into the process is unclear. Most countries release consensus statements at press conferences and/or in journals, which presumably allows a forum for dissenters to present their opinions. In the United States, a minority report may be used if serious disagreement remains, although this option is rarely employed (Jacoby, 1985). Switzerland and The Netherlands mention any remaining sources of disagreement in the consensus statement.

Decision-Making Process

In all but one country, the panel makes the final consensus decision; in Switzerland, the chair is responsible for final decisions. Public input is solicited in two countries (The Netherlands and the United States). No formal votes are taken in the process of writing the consensus statement. There are also no formal criteria for making decisions. In Canada I, there is a formal rating of the evidence using criteria that are predetermined, and the consensus statements are tied to this rating of the quality of evidence.

Sensitivity to Group Process Issues

There is little available evidence on the extent to which conferences are sensitive to group process issues. The size of many of the consensus panels suggests that, in the absence of concerted efforts, many panelists may not voice their opinions during discussion. The lack of formal decision-making criteria and formal voting suggests that undercurrents of dissent may go unrecognized. The practice of writing consensus statements in an all-night session suggests that individuals may agree because they are simply too tired to continue disagreeing. The presentation of evidence only in oral form would seem to particularly disadvantage lay members or scientists in unrelated fields who are unfamiliar with the literature and/or the science of the technology. This is likely to result in less participation in the

discussion by the generalists, potentially resulting in decision making that is biased by particularly knowledgeable or vocal panel members.

DISCUSSION

Consensus development conferences are used in many countries to assess the uses of medical technologies. Most of the conferences are conducted using a variation of the model that was developed in the United States by NIH. In this paper we have described the processes by which such conferences are conducted and the extent of variation among several countries along a series of specific dimensions. In this section we discuss the major concerns that arose from our review.

Link between Goals and Inputs

Most countries indicate that they use consensus development conferences to address a variety of issues, including the safety, effectiveness, efficacy, appropriateness, and consequences (political, social, ethical) of medical care and technology. In most cases, however, there appears to be a gap between these goals and the information available upon which to base such judgments.

Few countries undertake a systematic examination of the clinical or health services research literature in order to construct a synthesis of what is known from published sources about the use of a technology. If a synthesis is done, it rarely relies upon a formal meta-analysis or any quantitative process of combining evidence.

There are several advantages to beginning the process of planning a consensus development conference by conducting a thorough literature synthesis. First, the synthesis provides all panelists with a common starting point for discussions. Few individuals, even those who work with a technology in clinical practice or research settings, are completely informed or have synthesized and can remember all findings from the literature. Panelists may tend to remember the studies that support their respective positions and forget or disregard information that runs counter to their beliefs. The common ground provided by the synthesis may be particularly important for lay panelists or scientists from different disciplines; in most countries, about half the panelists are in such categories.

Second, the synthesis and the information about the state of clinical practice as it exists in a community setting provide "the facts" from which to build a set of questions to be addressed by a conference. In one evaluation of the consensus development conference process in the United States, the authors found that at the beginning of the conference, panelists believed that the right questions were being asked about a technology, but that their opinions about the critical issues changed as evidence was presented during the course of the conference (Wortman et al., 1988). Other research found that conferences focused on areas of clinical medicine believed to be in need of change when actual clinical practice had already been corrected (Kosecoff et al., 1987).

A synthesis of the literature provides a means of identifying not only what is known but also what is not known from published sources (Jacoby, 1988). The missing pieces may be answered through presentations by experts based on research that has not yet been published or through a systematic evaluation of sources such as patient medical records, use of decision analysis techniques, or expert opinion. In fact, we found that many countries conduct ad hoc surveys prior to a conference to collect data on current uses of the technology being evaluated. In addition, a synthesis, if done carefully, can help combine conflicting information and focus the panel on the most critical concerns. For instance, the literature synthesis may reveal that studies supporting the use of a technology are based on data obtained from poorly conducted studies, while data that do not support technological use come from randomized controlled clinical trials.

Third, the synthesis provides a systematic way of identifying the experts in the field. Previous evaluations have suggested that the selection of experts is not necessarily systematic (Wortman et al., 1988). This is not meant to suggest that the informal network approach results in the selection of worse (or even different) experts than a systematic approach, but that the informal approach is less scientifically defensible and has the potential for introducing unintended biases into the consensus development process.

Finally, linking of each consensus statement to its supporting evidence would allow users of the consensus findings to understand the scientific bases for the recommendations and would make clear where judgments are based on clinical trials, epidemiological studies, observational studies, or expert opinion. This might also enable panels

to make statements with varying degrees of certainty based on the science available to back the statement, which could, in turn, encourage the discussion of controversial topics.

Formalizing the Group Process

Consensus development conferences appear to be run fairly informally with respect to the criteria for making decisions, definition of consensus, formal voting or polling of panel members, and handling disagreements. A more formal process might not change the outcomes substantially, but it would make the process more scientifically defensible and replicable. An increase in the formality of the process might also allow for the devotion of more discussion time to controversial issues, if provisions are made for a range of consensus definitions.

The group process might be enhanced if the discussion is divided into disagreements about underlying assumptions. For example, panelists could discuss the probabilities of positive and negative outcomes associated with the use of a technology versus the patient utilities associated with those outcomes. If a conference recommended doing procedure *x* in clinical situation *y*, the recommendation should be based on achieving certain specified outcomes. If both good and bad outcomes result, it would be useful to know what weights patients place on those outcomes (e.g., how much risk of death a person is willing to take to gain a specified improvement in functioning). Finally, for those conferences that deal with economic or management issues, it should be made clear how these considerations modified the conference findings regarding clinical policy (e.g., the use of the technology is clinically acceptable but the financial implications are unacceptable).

Although not stated explicitly, it appears that consensus is defined in most conferences as unanimity. It has been suggested that requiring unanimity may result in statements that represent the "lowest common denominator" of opinion. Allowing consensus to take on a meaning that is less restrictive than unanimity expands the range and type of issues that can be addressed. The most straightforward way of introducing this type of definition of consensus into the process is by taking formal votes throughout the development of the consensus statement. Voting also provides a mechanism that allows for disagreement without necessarily endangering the overall process.

Alternatively, polling could be used (as it is in juries where unanimity is required) to ascertain how close the group is getting to full agreement. This type of polling throughout the process provides a means for ensuring that everyone can be heard. (The chair can ask those who disagree with a statement to state the source of their differences).

Voting or polling can be done in two ways. The first is anonymous (i.e., using secret ballots), where individuals are not identified with particular positions. The RAND/UCLA and Canada II conferences used this approach, allowing those who disagree some protection from undue pressure to change a position. The second approach is public voting, which may be more intimidating from the standpoint of a group process, but which may allow the group to focus on unresolved problems and force dissenters to defend their positions.

Further, voting can be done simply (i.e., yes or no) or on some scale that reflects the level of agreement or disagreement. Canada II, for example, uses the method developed at RAND/UCLA that asks panelists to rate indications for the use of a procedure on a nine point scale ranging from "extremely appropriate" to "extremely inappropriate." Another approach is to have panelists rate their agreement with a particular statement or recommendation on, for example, a five-point scale ranging from "strongly agree" to "strongly disagree." Obviously, yes/no votes provide a method that is simpler and perhaps easier to implement, to increase the certainty of adopting or rejecting recommendations. The advantage of using scales is that they allow panelists somewhat more latitude in agreeing or disagreeing with recommendations. Further, this information can be used to suggest a hierarchy or degree of certitude of those recommendations that should be implemented immediately versus those recommendations that require further study or that do not have strong support from the panel. The RAND Corporation has developed software for quickly entering the votes (through secret ballot) and processing the ratings; thus, technology is available to allow for reasonable implementation of the more complex approach.

Writing the Consensus Statement

Most countries appear to use the U.S. model when the panel writes the statement in an overnight session and presents the results the next morning. Although this has the benefit of producing a statement very promptly, it would seem to raise questions about the qual

ity of the product. As sleep deprivation and clear thinking are rarely compatible, the danger arises that the consensus reflects the opinions of those with stamina rather than representing a thoughtful consideration and deliberation based on the evidence. Although the introduction of an additional day or two into the process raises implications concerning the availability of experts and increased costs, the gains in the quality of the consensus findings may be well worth the additional time.

There may be alternative approaches for conserving the amount of available meeting time. Along with the literature review prepared in advance, some countries begin drafting the consensus statement before the meeting, giving panelists the opportunity to consider the issues and focus on those areas of particular disagreement or controversy. The Canada II method developed at RAND/UCLA, for example, requires that panelists rate indications for the appropriate use of a procedure in advance using a modified Delphi method. Analysis of the first set of ratings indicates areas of potential disagreement, which are then given more time during the panel meeting. By shifting some of the responsibility onto panelists to prepare for the conference, the use of the standard meeting time may be more effectively targeted to those activities for which the group process and interactions are important. One alternative is seen in The Netherlands, where CBO staff help to complete the statements at a meeting that occurs one to three months after the consensus meeting development conference. This provides time for writing and reflection before concluding the process.

CONCLUSIONS

The processes by which consensus development conferences are conducted can affect the value and validity of the final product. Although the NIH model has been used extensively in other countries, it has never been shown to produce replicable results or to be preferable to other models. Further, for all approaches, there is always the potential for improvement in the process. The continued use of consensus development conferences to achieve agreement on the state of the science for a particular technology suggests a need to enhance the inputs to the process and the methods by which issues and experts are identified. An improvement in the inputs to the process has several advantages and seems likely to enhance the final

product. Formalizing the process by which consensus conferences are conducted would enhance the credibility of the statements by making allowances for varying degrees of consensus and providing an outlet for disagreement that does not detract from the overall endeavor. Such formal mechanisms as voting or polling, whether done anonymously or not, may allow for more controversial issues to be considered. Finally, the current proclivity toward writing the consensus statement in an all-night session seems to be unnecessarily grueling and may have a negative effect on the final product. By improving inputs at the planning stages, it should be possible to develop schedules that allow for consideration of serious issues on a more humane time frame.

References

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

Battista, R.N., and S.W. Fletcher. 1988. Making recommendations on preventive practices: Methodological issues. American Journal of Preventive Medicine 4(4)(Suppl.):53-67.

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.J.E. van Everdingen. 1985. Consensus development conferences in the Netherlands. International Journal of Technology Assessment in Health Care 1:905-912.

Chassin, M.R., J. Kosecoff, R.E. Park, C.M. Winslow, K.L. Kahn, N.J. Merrick, J. Keesey, A. Fink, D.H. Solomon, and R.H. Brook. 1987. Does inappropriate use explain geographic variation in the use of health care services? A study of three procedures. Journal of the American Medical Association 258(18):2533-2537.

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):309.

Jennett, B. 1985. First consensus development conference in United Kingdom: On coronary artery bypass grafting, commentary by the chairman 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.

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 NIH consensus development program. Santa Monica, Calif.:RAND Corporation.

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

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 and physician practice. Journal of the American Medical Association 258:2708-2713.

Lomas, J., G.A. Anderson, M. Enkin, E. Vayda, R. Roberts, and B. MacKinnon. 1988. The role of evidence in the consensus process, results from a Canadian consensus exercise. Journal of the American Medical Association 259(20):3001-3005.

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

Office of Medical Applications of Research, National Institutes of Health. 1988. Guidelines for the selection and management of consensus development conferences. Bethesda, Md. Pamphlet.

Park, R.E., C.M. Winslow, K.L. Kahn, N.J. Merrick, J. Keesey, A. Fink, D.H. Solomon, and R.H. Brook. 1986. Physician ratings of appropriate indications for six medical and surgical procedures. American Journal of Public Health 76(7):766-772.

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 4:481-484.

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.

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

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)(Fall):469-498.

Dissemination and Impact of Consensus Development Statements

Arnold D. Kaluzny

Consensus development conferences were initiated in the United States at the National Institutes of Health (NIH) in 1977 as a method for developing a statement of accepted standards for clinical practice either with regard to the use of a particular technology or the treatment of a particular disease or syndrome. Other countries have organized conferences, although their formats and purposes have changed over time and between countries. The purpose of this paper is to consider the main elements of the dissemination activities in various countries' programs, to consider different methods of diffusion, and to assess their impacts in various settings. The analysis presented is based on program profiles prepared by each country for the International Workshop on Consensus Development for Medical Technology Assessment (1989). A comparative review of these profiles provides a unique opportunity to analyze international efforts in the dissemination of consensus recommendations and to suggest possible areas for further research.

ELEMENTS OF DISSEMINATION ACTIVITIES

While the elements of the dissemination process for any particular program vary by country, it is possible to compare these vis-à-vis barriers to diffusion identified by existing theory and research. Although the list could be quite extensive, attention will be given to

four factors with varying degrees of tractability that may directly influence dissemination activities: attributes of the technology under review and the resulting consensus statement, adopter characteristics, environmental constraints and incentives, and communication channels.

Attributes

The technology selected for review and the technological attributes considered are important determinants of dissemination. Unfortunately, the literature on the diffusion of medical technology is often characterized by inconsistent findings and thus is often discounted in our overall effort to better manage the review process and the dissemination of state-of-the-art technology. In part, this inconsistency is accounted for by what some researchers (Downs and Mohr, 1976) have termed a unitary approach to developing innovation theory. This means that all innovations, regardless of type and/or specific attributes, are considered as equal and subject to the same theory. Increasingly, it is recognized that both the rate and speed of adoption and diffusion are a function of the interaction of the type and attributes of the innovation with various adopter characteristics (Fennell and Warnecke, 1988). Moreover, there is some evidence to suggest that both diffusion and adoption of various types of programs or technologies are not totally random. There appears to be a predictable order that the adopting unit follows as it tend to implement and/or adopt a particular activity (Fennell, 1984). Certain types of technologies may be linked in such a way that the implementation of one tends to facilitate implementation of another. Thus, the selection of topics for consensus and the range of criteria upon which judgments are made are extremely important in the design of a dissemination strategy.

A review of consensus development conferences indicates that they take a fairly inclusive view of the types of technology subject to the consensus development process. While all programs specify criteria for inclusion, the technologies reviewed tend to be fairly eclectic and not subject to any apparent a priori strategy that would facilitate dissemination. Among the programs considered here, the one exception to this generalization is the Canadian Task Force on the Periodic Health Examination. The Canadian Task Force limits the focus of their evaluation and dissemination efforts to preventive

services offered to asymptomatic individuals within a primary care setting. This implicit programmatic theme may facilitate dissemination and the subsequent adoption of consensus recommendations by physicians.

Different countries assess a range of attributes for particular technologies in the consensus development programs. As documented in the innovation literature (Fineberg, 1985; Scott, in press), these attributes influence diffusion of technology as well as the ability to have an impact on physician practice. For example, technologies that have a favorable cost-benefit ratio, are compatible with ongoing practice patterns, and are consistent with reimbursement policies will be disseminated more readily to potential adopters. Table 1 presents the various attributes upon which technologies are assessed according to three major groupings. The U.S. program and the Canadian Research Group focus explicitly and almost exclusively on the effectiveness of technology. The Canadian Task Force and The Netherlands emphasize effectiveness, with secondary consideration given to the psychosocial, economic, ethical, and legal implications. The programs in Finland, Norway, Denmark, the United Kingdom, and Sweden appear to give equal weight to effectiveness as well as other attributes, including cost and service requirements.

Technology and its associated attributes are important factors in the diffusion process, yet an equally critical element concerns the attributes of the consensus statement itself. How clear is the statement? Is it prescriptive or discursive, and does it provide concrete and specific actions or guidelines that physicians can follow in clinical practice? While there has been no systematic study of how the attributes of the consensus statement influence physician decision making, one U.S. study of the NIH consensus development process attempted to assess the attributes of the consensus statement (Kanouse et al., 1987). The researchers selected 24 statements for analysis and found three dimensions for statement classification: discursive, didactic, and scholarly. Discursive statements tend to be long and abstract and contain few recommendations. Consensus statements characterized as didactic offer clinicians practical and detailed guidance, while scholarly statements offer up-to-date descriptions of the scientific evidence bearing on a topic and devote more attention to detail than most statements. The critical issue is obviously the relationship of these and other attributes to an actual change in physician knowledge and behavior.

TABLE 1 Technological Attributes by Country Program

Attributes	Canada (TFPHE[a])	Canada (RG[b])	Denmark	Finland	The Netherlands	Norway	Sweden	United Kingdom	United States
Effectiveness	1	1	1	1	1	1	1	1	1
Safety	2	0	1	1	2	1	1	1	0
Other[c]	2	0	1	1	2	1	1	1	0

Note: 0 = no explicit consideration; 1 = primary consideration; 2 = secondary consideration.

[a] TFPHE = Task Force for the Periodic Health Examination.

[b] RG = Research Group.

[c] Other = cost, resource requirements, acceptability, ethical, legal, or related attributes.

TABLE 2 Evaluation Efforts by Country Program

	Canada (TFPHE[a])	Canada (RG[b])	Denmark	Finland	The Netherlands	Norway	Sweden	United Kingdom	United States
Evaluation	no	yes	no	no	yes	yes	yes	yes	yes
Formal assessment of:									
Changes in practice patterns		X				X			X
Interventions for dissemination		X							

[a] TFPHE = Task Force for the Periodic Health Examination.

[b] RG = Research Group.

Adopter Characteristics—Behavior and Structure

Physician characteristics and the context of their practice are critical to any dissemination effort. Table 2 indicates that several countries have conducted evaluations of their programs and that the level of awareness of both the conferences and their content vary by country and within country by physician characteristics. For example, evaluation results suggest that there are significant differences in the level of awareness between U.S. and Swedish physicians. In Sweden, a national sample of physicians revealed a significant level of awareness among targeted groups. Awareness of a single consensus conference was very high, ranging from 86 to 94 percent in all target groups, with approximately 7 to 10 percent of the respondents indicating that the consensus statement evoked some changes in clinical practice (Johnsson, 1988). In a national sample of U.S. physicians, 41 percent reported general awareness of the program but only 18 percent said they were somewhat or very familiar with it. Awareness varied by specialty, with oncologists indicating the greatest familiarity and family practitioners the least (Kanouse et al., 1987).

Awareness also appears to be a function of physicians' socio-demographic characteristics. Analysis of the U.S. data (Kanouse et al., 1987) revealed that physicians who had heard of the program were somewhat older, had practiced medicine about two years longer, were less likely to work in private group practice and more likely to work in a hospital, clinic, or other institutional setting; or were more likely to be the members of a medical school's teaching staff and to report that they had responsibility for training students, residents, and interns. Moreover, their information habits and preferences were quite different. Physicians who were aware of the program reported spending more time reading journals such as the *Journal of the American Medical Association* (where many NIH consensus statements are published) and tended to talk informally with their colleagues about medical topics. They also reported spending 30 percent more time attending continuing medical education courses and tended to receive patient referrals from a larger number of physicians than did those who were less aware of program recommendations.

What is the relationship of conference recommendations to actual clinical practice? Three countries report evaluation efforts to link conference recommendations with actual clinical practice patterns. Of these three studies, only the Norwegian results suggest that the

consensus conference recommendations affected practice patterns. However, as the evaluation suggests, and as discussed under the section dealing with environmental constraints, this may be more a function of environmental factors than of a consensus statement. The researchers used a preconference, postconference survey design to gather data from the country-wide registration of ultrasound procedures. The analysis revealed a reduction in the average number of diagnostic ultrasound examinations per woman following the promulgation of the consensus conference recommendations on the use of ultrasound in pregnancy.

Data from two remaining studies indicate the contrary. A recent Canadian analysis of practice guidelines for use of cesarean sections on the attitudes and behavior of physicians reveals that there was substantial awareness and agreement on the proposed guidelines as well as a self-reported change in practice (Lomas et al., 1989). However, the actual analysis of hospital discharge data suggested that the consensus guidelines produced little or no change in actual practice. A similar analysis comparing self-reported practice with actual practice was conducted as part of a U.S. evaluation (Kanouse et al., 1987). Here, changes were measured in several hospital-based procedures that were subject to consensus conference recommendations, providing a comparison between physician self-reported and actual practice vis-à-vis exposure to particular consensus recommendations (Kosecoff et al., 1987). Results showed that the conferences largely failed to stimulate change in physicians' practices, despite moderate success in reaching appropriate target audiences. While this analysis involved physicians and hospitals in only one geographic region of the United States, and thus is subject to severe limits of generalization, the data did reveal that physicians' preferred practice patterns bear a strong relationship to what they actually do. The link, however, between consensus development conferences and actual practices was quite disappointing. For example, in the analysis of a breast cancer conference, relationships between physician awareness of the conference and their compliance with the recommendations reflected preexisting differences, not a program effect.

Environmental Constraints and Incentives

The context in which dissemination occurs can facilitate or inhibit the process. The literature on diffusion points to a series of organi

zations at the federal, state, and local levels that affect this process (Fineberg, 1985). For example, in Norway, the endorsement by the Health Directorate of the consensus conference statement on ultrasound screening of pregnant women as the national guideline affected the use of ultrasound screening in clinical practice (Backe and Nafstad, 1989). Equally important, but often overlooked, may be subseries of "supply-side factors" (Robertson and Gatignon, 1987). Sources of innovative technology other than consensus statements may influence physician behavior and are important in determining the amount of persuasive information being transmitted to potential adopters. For example, the aggressive marketing strategies of private corporations that produce medical technologies and the clinical policies of hospitals and clinics that provide care influence physician behavior.

The extent to which environmental constraints and incentives affect dissemination or are considered in formulating a dissemination plan varies according to the structure and function of the program. The very composition of a consensus development panel is often intended to address some of these constraints by including relevant administrators and policymakers in the consensus development process. For example, the Swedish program considers social, organizational, and economic aspects of technologies. Thus, the conferences involve experts in medicine, health economics, epidemiology, and health policy, as well as administrators and concerned patient groups. The resulting statements go beyond addressing the safety and efficacy of the technology. The statements are directed toward a much broader audience, with dissemination targeted to concerned physicians, politicians, and administrators.

Evaluation studies have focused both on the awareness of consensus development conferences by health administrators and politicians and their awareness of consensus results. Among one sample of Swedish health administrators and politicians, awareness of consensus conferences was high (Calltorp, 1988). Eighty-nine percent indicated that they knew about the conference. Ninety-nine percent of the administrators were aware of the consensus conference, and 85 percent of the politicians knew about the consensus conference. When queried about their awareness of the outcomes (i.e., the content of the consensus statements), results were equally impressive. Eighty-three percent of the respondents were aware of the outcome of one or more conferences. Administrators ranked highest (96 percent of the administrators were aware of one or more statements).

Eighty-four percent of the politicians (full-or part-time) were aware of one or more of the statements. The latter ranged from a low of 49 percent who were aware of diagnostic imaging of liver tumors and 67 percent who indicated awareness of the recommendations on sight-improving surgery.

Communication Channels

The channels used to learn about new medical technologies have been a frequent subject of investigation (Eisenberg, 1986; Fineberg, 1985). Consensus development conferences have used a variety of channels, giving primary importance to professional journals, both general and specialty, and direct mailings of consensus statements. Table 3 lists the various channels used by countries' programs as part of their dissemination efforts. Few programs use direct mail or continuing education. The majority of programs disseminate various forms of publications to health care providers, managers, and policymakers, whereas fewer programs target the general public. The evaluation of specific channels has been limited. Two exceptions include the RAND Corporation study (Kanouse et al., 1987) of the U.S. consensus development program and the Canadian Research Group proposal. The latter is a potential series of quasi experiments to evaluate alternative dissemination strategies (Lomas, 1989). Specifically, the Canadian Research Group is conducting a large randomized controlled trial to evaluate two dissemination strategies; one strategy uses local "educational influentials" and the other strategy involves the use of chart audit and feedback.

The U.S. evaluation suggests that consensus recommendations are more likely to reach specialists than generalist physicians (Kanouse et al., 1987). The study was not able to establish a direct link between awareness and publication in a generalist type channel, such as the *Journal of the American Medical Association* (JAMA), which publishes most of the NIH consensus statements. The analysis revealed that reading JAMA had no predictive value after other factors such as specialty and type of practice were taken into account. Reading the *New England Journal of Medicine* and various specialty journals, however, and being part of a well-defined network of clinicians was associated with greater awareness of the consensus development program. Similar patterns existed with respect to actual knowledge of relevant conferences, with higher levels of awareness being recorded by physicians reading specialty journals.

TABLE 3 Communication Channels by Country Program

Channel	Canada (TFPHE[a])	Canada (RG[b])	Denmark	Finland	The Netherlands	Norway	Sweden	United Kingdom	United States
Publications									
General professional		X	X	X	X	X	X	X	X
Specific professional	X	X	X	X	X	X	X	X	X
Health policy/ management					X		X		X
General public			X	X	X	X	X	X	X
Direct mail			X		X			X	X
Continuing education			X				X		

[a] TFPHE = Task Force for the Periodic Health Examination.

[b] RG = Research Group.

The U.S. evaluation also assessed continuing medical education (CME) programs and direct mailings as alternative communications channels (Kanouse et al., 1987). Dissemination through CME programs was considered an important source in that more than two-thirds of respondents in the physician survey regarded conferences, meetings, and CME programs as "very important" information sources both for first hearing about new medical procedures and for deciding whether to use them. Nearly half of the physicians surveyed had learned of the Consensus Development Program through direct mailing of different reports by the National Institutes of Health. A substudy to examine direct mailings to five relevant specialties in the metropolitan St. Louis, Missouri, area revealed that significantly more physicians who received such mailings were aware of the conference and its recommendations than was a comparative sample of physicians who did not receive such direct mailings (Jacoby and Clarke, 1986).

AREAS OF FUTURE RESEARCH

The development of consensus development conferences, many with different functions and structures operating within a variety of different health care systems, provides an opportunity for future collaborative research and evaluation. Variations in function, structure, and setting provide a unique opportunity to set up a series of natural experiments that can increase the overall understanding of the basic diffusion process as well as the effectiveness of alternative dissemination strategies. This understanding is critical for (1) enhancing the effectiveness of consensus development conferences vis-à-vis their stated goals and objectives and (2) providing a data base from which to evaluate the consensus conference approach to dissemination compared with other methodologies and programmatic initiatives that can influence physician practice patterns (e.g., standards and protocol participation). Listed below are several research areas that capitalize on a systematic and comparative assessment of consensus development conferences.

- *Evaluation of specific dissemination methods.* The effectiveness of specific dissemination strategies is an empirical question and is contingent on several covariants. One approach would be to set up a series of experiments or quasi experiments to evaluate alterna

tive strategies beyond those traditionally used such as professional journals and newsletters. For example, the Canadian Research Group randomized trial evaluates two implementation strategies: use of local influentials versus use of chart audit and feedback (Lomas, 1989). Obviously, this requires replication in other settings along with the opportunity to evaluate other dissemination strategies such as the use of "clinical alerts" by relevant governmental agencies, computerized physician query systems, and the use of follow-up meetings for selected clinicians that are related to a particular consensus statement.

- *Role of organizational intermediaries.* An underlying assumption of most dissemination efforts is that the individual physician is the unit of analysis and should be the target of the dissemination effort. In reality, however, health care organizations are involved in the delivery and financing of the technology and represent important constituent groups. Thus, they are critical actors in the adoption process. This is clearly recognized by several of the European programs, in that their panels include administrators and policymakers, yet even within this context, the explicit target is the physician.

  The explicit targeting of organizations and their decision-making processes requires an understanding of organizational behavior, which is quite different from individual adoption processes. Consideration needs to be given to a range of factors including structural characteristics, the role of coalitions within organizations, and the idea of secondary choices (i.e., the implementation decision by the organization and the subsequent adoption decision by the physician).
- *Interaction of attributes, adopters, and environmental characteristics.* A central issue is the kind of diffusion and adoption process that may occur with different types of technology, different types of consensus statements, and different environmental conditions. The opportunity to explore each, as well as their interactions, is present in a cross-cultural assessment of dissemination practice. There is a unique opportunity to employ a natural experiment using a series of tracer technologies for consensus development programs and dissemination efforts in a variety of countries. This approach would clarify the nature of the interaction between the attributes of the technology being diffused, the characteristics of the adopting unit and the political and policy context of various countries. For example, one would expect different diffusion patterns in competitive and regulatory environments. The ability to monitor these dif

fusion patterns over time and to compare different delivery systems permits insight as to how larger cultural factors and values influence practice patterns (Waitzkin, 1983). This type of evaluation would help us move away from a "unitary approach to diffusion" and toward building and testing theory that would have greater relevance to the actual design of dissemination strategies.

- *Dissemination and changes in actual clinical practice.* While evaluations of dissemination strategies have focused primarily on awareness of the consensus development program and/or specific program recommendations, the real issue is whether dissemination and the resulting awareness actually change physician practice patterns. Few evaluations have attempted to assess such changes; and evaluations that examine the relationships between dissemination efforts, levels of awareness, and attitudes toward specific conferences with physician behavior are needed. This information would provide a critical link to targeting areas of practice requiring consensus recommendations since it is precisely these areas of practice that determine physician readiness to comply with recommended changes.
- *Revising assumptions about the role of dissemination.* A good share of our diffusion models, or at least the assumptions underlying them, are borrowed from a simpler time and a simpler problem. Clearly, the dissemination of information to change physician practices is far more complex than efforts to change the buying habits of the general public. A close examination of clinical practice patterns and decision-making processes reveals a level of intractability that is not easily influenced by fairly simple dissemination practices. For example, a qualitative analysis comparing decision-making practices of physicians in the United Kingdom and the United States in terms of adoption of formed versus dynamic (unformed) technologies revealed that physician decision-making processes are greatly shaped by the perspectives of local clinical practices and are not easily influenced by more formal dissemination channels. Ann Greer suggests: "there are no magic signatories or formats which will cause knowledge to jump off the page and into practice" (Greer, 1988).

The complexity of the decision-making process clearly suggests that consideration needs to be given to other strategies beyond simple dissemination and that in the future the relative cost-effectiveness of these strategies must be given greater attention. This is not to sug

gest that consensus development conferences and concomitant dissemination strategies be abandoned, but they need to be supplemented with private and other public sector initiatives consistent with the underlying process of clinical practice. One approach may be to couple the involvement of clinicians in ongoing research (e.g., clinical trials) with the systematic selection of topics for consensus development and subsequent diffusion to the practicing community. For example, within the United States, the National Cancer Institute has instituted a Community Clinical Oncology Program that allows local practitioners to participate in clinical trials research. Combining this type of involvement with the diffusion of consensus statements may be worthy of consideration. Another possibility is to capitalize on larger ongoing initiatives to target dissemination efforts. For example, in the United States the Joint Commission on Accreditation of Healthcare Organizations has launched a program to monitor selected organizational and clinical outcome indicators as part of the overall hospital accreditation process. The availability of this information may help to target dissemination efforts to those clinical areas amenable to and/or requiring change. Finally, the changing of practice patterns requires a new recognition of the role of patients and administrators in the clinical decision-making process. As reflected in many of the European consensus development conferences, consumers and/or their representatives, along with administrators and various policymakers, are important contributors in the evaluation of given technologies. Consumer involvement and influence in changing physician practice patterns is not well understood and is worthy of investigation and evaluation.

SUMMARY

The consensus development programs and the resulting consensus statements represent one approach to influencing clinical practice. The development of this approach and its adaptation in both structure and function under different countries' initiatives provide an important opportunity to assess its utility under a variety of conditions. Evaluations of programs' effectiveness in changing clinical practice patterns must take into account attributes of the technologies considered, as well as the consensus statements themselves and a variety of organizational and environmental factors that influence

the dissemination process. The consensus statement should be recognized as only one influence on clinical practice and should be considered an integral part of a broader strategy of dissemination and technology transfer.

References

Backe, B., and P. Nafstad. 1989. Development in practice of diagnostic ultrasound in Norway after the consensus conference recommending routine screening. Paper presented at the Fifth Annual Meeting of the International Society of Technology Assessment in Healthcare, London.

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.

Downs, G., and L. Mohr. 1976. Conceptual issues in the study of innovation. Administrative Sciences Quarterly 21(4)(December): 700-714.

Eisenberg, J.M. 1986. Doctors' Decisions and the Cost of Medical Care: The Reasons for Doctors' Practice Patterns and Ways to Change Them. Ann Arbor, Mich.: Health Administration Press.

Fennell, M. 1984. Synergy, influence and information in the adoption of administrative innovations. Academy of Management Journal 27:113-129.

Fennell, M., and R. Warnecke. 1988. The Diffusion of Medical Innovations. New York: Plenum Press.

Fineberg, H.V. 1985. Effects of clinical evaluation on the diffusion of medical technology. Pp. 176-207 in Assessing Medical Technologies, Institute of Medicine, National Academy of Sciences. Washington, D.C.: National Academy Press.

Greer, A. 1988. The state of the art versus the state of the science. International Journal of Technology Assessment in Health Care 4:5-26.

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

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.

Kanouse, D.E., R.H. Brook, J.D. Winkler, J. Kosecoff, S.H. Berry, G.M. Carter, J. P. Kahan, L. McKloskey, W.H. Rogers, C.M. Winslow, G.M. Anderson, L. Brodsley, A. Fink, and L. Meredith. 1987. Changing medical practice through technology assessment: an evaluation of the NIH consensus development program. Santa Monica, Calif.: RAND Corporation .

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.

Lomas, J. 1989. Profile for the consensus development program in Canada. Paper prepared for the Institute of Medicine Workshop on Consensus Development for Medical Technology Assessment, London.

Lomas, J., M.W. Enkin, G.M. Anderson, K. Domnick-Pierre, E. Vayda, and W. Hannah. 1989. The role of consensus statements in changing physician's awareness and attitudes, knowledge and behavior. Paper presented at the Fifth Annual Meeting of the International Society of Technology Assessment and Healthcare, London.

Robertson, T., and H. Gatignon. 1987. The diffusion of high technology innovations: A marketing perspective. In New Technology as Organizational Innovation: The Development and Diffusion of Microelectronics, J. Pennings and A. Buitendam, eds. New York: Ballinger.

Scott, W.R. In press. Innovations in medical care organizations. Medical Care Review, Summer.

Waitzkin, H. 1983. The Second Sickness. New York: Free Press.

PROGRAM PROFILES

Profile of a Consensus Development Program in Canada: The Canadian Task Force on the Periodic Health Examination

Renaldo N. Battista

NATIONAL CONTEXT

The Canadian Task Force on the Periodic Health Examination was created in 1976 by the Conference of Deputy Ministers of Health, with a mandate to examine the scientific merit of the annual checkup and to determine how periodic health examinations might enhance or protect the health of the population. The work of the task force was and continues to be funded by the National Department of Health and Welfare (federal government). However, the membership is composed of individuals who are not directly linked to the government and who have academic affiliations.

The task force reviews the scientific evidence regarding the effectiveness of preventive services that are to be offered to asymptomatic individuals within primary health care settings. The sum of the evidence about whether a clinical procedure can be expected to be of benefit to the population as a whole is synthesized, and practice policy recommendations are formulated.

The recommendations were originally targeted to practicing primary care physicians; but we now recognize the importance of addressing medical educators, policymakers, and the general public as well. The recommendations of the task force do not carry formal legal or legislative weight; they are simply practice guidelines to be used as a reference by health care providers, policymakers, and

medical educators. They may be used as guidelines for reimbursement of medical services, but the impact of making preventive services explicitly reimbursable has been minimal.

The Canadian Task Force on the Periodic Health Examination has an annual budget of approximately $100,000 (Canadian) for consensus development activities. For the past three years, the government has provided half of this amount as a supplemental fund for a full-time research coordinator; the rest of the funds are used to conduct and support three to four meetings each year. The process is relatively inexpensive, partly because panel members are asked to use their own resources to produce documents. The task force assumes responsibility for the purchase and dissemination of reprints.

SCOPE OF THE PROGRAM

The task force assesses the appropriateness of applying specific clinical procedures for the prevention of adverse health outcomes. The emphasis is on clinical procedures that can be carried out in the primary care setting; they may be physical examinations, laboratory tests, vaccinations, or counseling activities. In the case of primary prevention, where the goal is to prevent the initiation of the disease process, the technology applied might be a vaccine or an intervention such as counseling and patient education. In the case of secondary prevention, the goal is to detect a disease process before it becomes symptomatic in order to apply therapy when it will have a greater impact on the progress of the disease or to detect a risk factor for disease. In the case of secondary prevention, we would evaluate not only the early detection procedure (physical examination or laboratory test) but also the diagnostic and treatment interventions.

The focus of the task force is on preventable conditions rather than the clinical procedures themselves. For instance, rather than focusing on the efficacy of the digital rectal examination, we examine the evidence on the preventability of prostate cancer by all procedures reviewed in the literature, including per-rectal ultrasound, laboratory tests, and the digital rectal examination. In choosing which procedures to consider for review, however, we choose techniques that can be expected to have an impact on clinical practice in the present or in the very near future—that is, new or established technologies.

We are primarily interested in the effectiveness of the application of a clinical procedure; that is, does this procedure net more benefit than harm in those to whom it is offered? When there is no evidence as to its effectiveness in the real clinical situation, we then examine its efficacy, (i.e., its ideal benefit-to-harm ratio). However, this is not the sole basis for our recommendations; we also consider safety; acceptability to the patient and provider; cost-effectiveness; and ethical, psychological, or legal implications.

Only conditions that have a potential for prevention are considered. Whether the condition carries a considerable burden of suffering is another important criterion for its choice; this burden may be quantitative (e.g., mortality, morbidity, years of life lost) or qualitative (e.g., perception by the public as being of concern). We may also choose to consider a condition with respect to a given technology when it is associated with high costs or when it has other impacts on clinical practice. An example of this would be the evaluation of intrauterine electronic fetal monitoring during labor (Battista and Fletcher, 1988; Canadian Task Force on the Periodic Health Examination, 1979).

FORMAT AND CONDUCT OF THE PROCESS

The task force is composed of a stable panel of members who are well versed in the clinical epidemiologic approach and who possess expertise in their own fields of pediatrics, family medicine, psychiatry, infectious diseases, geriatrics, and nursing. Various conditions will be under consideration at any one time, and each member will be responsible for reviewing the scientific evidence of one condition. Our process is iterative in that members undertake an overview of the literature, prepare a background paper, present it to the task force for discussion, modify the paper or literature review based on the feedback, present a new draft at the next meeting, and so on, until the entire task force is satisfied with the extent of the review and comes to a consensus on the implications of the evidence for clinical practice. Depending on the complexity of the condition, arriving at consensus requires from four to eight iterations spread over one to two years. Background papers are distributed to the members before the meetings, and approximately one hour is allocated to a discussion of each condition at each meeting.

Formulation of recommendations for practice are reached by con

sensus. The recommendations are primarily based on the value of the scientific evidence according to predetermined criteria. These criteria are primarily based on the value of the scientific evidence according to the quality of study design, study execution, and reporting. In cases where the scientific evidence is clear and the task force is reassured that all pertinent studies have been considered, consensus is easily reached. When the evidence is poor or equivocal, the recommendation is made on other grounds such as cost, ethics, or safety considerations. Since this evidence rests primarily on the opinions of experts, the achievement of consensus is more difficult and may require several iterations. After consensus is achieved, the task force's position is often sent to an expert in the field for peer review before it is published in English in the Canadian Medical Association Journal (Canadian Task Force on the Periodic Health Examination, 1979, 1984, 1986, 1988, 1989) and in French in L'Union Médicale du Canada. These medical journals have a combined total circulation of 65,000, reaching almost all of the physicians in Canada.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

As mentioned previously, each member prepares a background paper that is circulated to other members before the meetings. Only original reports of clinical trials and epidemiologic studies are considered; these are identified from computerized MEDLINE searches, key citations, and consultations with experts in the field. The quality of the evidence from each study is always given priority in the consensus development process; expert opinion is only invoked in the absence of evidence.

The background paper is condensed to include only the key information and key references pertaining to the decision; this concise statement appears in the medical journals. Currently, the task force is moving toward formal documentation of its decision-making process so that each recommendation can be traced. This will probably not be published, but will be made available to interested parties.

DISSEMINATION AND IMPACT

Task force meetings are held on a regular basis three times per year. Only members and others who act as consultants participate in

these meetings, so wider notice is not given. It is a closed process; neither the agenda of the conditions being discussed nor the interim statements are made available to the general public, although representatives from the Department of Health and Welfare are notified of all developments.

The intended impact of the task force recommendations is to change clinical practice and enhance provider knowledge (Battista and Mickalide, in press). The recommendations are published in the Canadian Medical Association Journal and L'Union Médicale du Canada, and reprints are sent to those on a growing list of people who have expressed interest in the task force's work. Since the task force released its first report in 1979, there have been significant changes in clinical practice that are compatible with our recommendations, but it is difficult to make a causal link with the task force recommendations. One measure of impact is the request for reprints of our reports; more than 40,000 requests for the 1979 report have been received from people from all over the world. The work of the periodic health examination task force is widely quoted in the lay press, and in academic circles it has been used in curriculum design and as a standard for preventive behavior. The consensus process of the task force has not been evaluated formally or informally. However, a study of general practitioners that was conducted in Quebec and New Brunswick three to four years after the first task force report was presented documented the varying level of integration of some preventive activities into clinical practice and witnessed the need for improvement (Battista, 1983; Battista and Spitzer, 1983; Battista et al., 1985).

References

Battista, R.N. 1983. Adult cancer prevention in primary care: Patterns of practice in Quebec. American Journal of Public Health 73(9):1036-1039.

Battista, R.N., and S.W. Fletcher. 1988. Making recommendations on preventive practices; methodological issues. American Journal of Preventive Medicine 4 (4)(Suppl.):53-67.

Battista, R.N., and A.D. Mickalide. In press. Integration of preventive services into primary care--a conceptual framework for implementation. In Preventing Disease: Beyond the Rhetoric, R. Goldbloom and R.S. Lawrence, eds. New York: Springer-Verlag.

Battista, R.N., and W.O. Spitzer. 1983. Adult cancer prevention in primary care: Contrasts among primary care settings in Quebec. American Journal of Public Health 73(9):1040-1041.

Battista, R.N., C.S. Palmer, B.M. Marchand, and W.O. Spitzer. 1985. Patterns of preventive practice in New Brunswick. Canadian Medical Association Journal 132:1013-1015.

Canadian Task Force on the Periodic Health Examination. 1979. The periodic health examination. Canadian Medical Association Journal 121:1193-1254.

Canadian Task Force on the Periodic Health Examination. 1984. The periodic health examination: 2. 1984 Update. Canadian Medical Association Journal 130:1278-1285.

Canadian Task Force on the Periodic Health Examination. 1986. The periodic health examination: 2. 1985 Update. Canadian Medical Association Journal 135:724-727.

Canadian Task Force on the Periodic Health Examination. 1988. The periodic health examination: 2. 1987 Update. Canadian Medical Association Journal 138:618-626.

Canadian Task Force on the Periodic Health Examination. 1989. The periodic health examination: 2. 1989 Update. Canadian Medical Association Journal 141:209-216.

Profile of a Consensus Development Program in Canada: The Canadian Research Group

Jonathan Lomas

NATIONAL CONTEXT

The Canadian Research Group consensus development program has no particular national context, as it is a group of interested researchers who engage in opportunistic consensus development to undertake research on the consensus development process. The group consists of individuals from McMaster University and the University of Toronto in Ontario and the University of British Columbia in Vancouver. The program obtains funding through peer review grant applications and not in connection with any ongoing program funds. The cost of conducting the consensus development exercises is approximately $70,000-$100,000 (Canadian).

The underlying objective of our consensus development program is to translate existing research evidence into clinical practice. The principal audience is, therefore, practicing clinicians. It may be, however, that consumers and/or funding agents in the system are useful targets. The ultimate goal of the program is to change clinical practice where it has been demonstrated to be inappropriate.

SCOPE OF THE PROGRAM

The scope of the program is quite limited, as it is primarily a vehicle for research. We have been opportunistic in choosing our

topics; thus, we have selected topics for which (1) there is some evidence of a lack of congruence between the research evidence and clinical practice and (2) we have obtained some support for the project from a specialty society or a group of physicians. To date, we have conducted two consensus development conferences: the appropriate use of cesarean section (in conjunction with the Society of Obstetricians and Gynecologists of Canada) and priority-setting among candidates for coronary artery bypass surgery (with an interested group of cardiovascular surgeons and cardiologists).

The conferences address the effectiveness of clinical services almost exclusively. The programs do not consider cost-effectiveness or service requirements. In some cases, ethical, legal, and social implications come to play a role in the development of consensus.

FORMAT AND CONDUCT OF THE PROCESS

The entire consensus development process takes 6 to 15 months, depending on the urgency and the availability of staff and research funds. During the first stage, the researchers identify priority areas by analyzing existing administrative data sets in health care. The assessment areas which, at first pass, do not appear to be congruent with existing research evidence are highlighted. During the next stage, the researchers identify a panel of individuals, most of whom are clinicians but not all of whom are from the specific clinical area under investigation. An attempt is made to include epidemiologists on the panel. The two most important elements of the process are (1) the preparation of a comprehensive background paper on existing research evidence for the area under discussion, with clear consideration of the methodologic quality of the evidence, and (2) the completion by panelists of indications questionnaires. The indications questionnaires consist of representative scenarios for the clinical conditions under study. The panelists rate the scenarios according to their appropriateness for intervention (similar to the RAND Corporation technique). Speakers or witnesses do not necessarily make formal personal presentations, although the panel does meet on two occasions to consider the question under study. During the first meeting, the panel develops an interim idea for the statement followed by dissemination and discussion of the resulting document. The second and concluding meeting, which may be done through the mail, is held to complete the statement.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

Staff are hired specifically to undertake comprehensive literature reviews, with substantial attention given to the methodologic quality of the various studies. The results of the literature review are presented in a format that focuses on the high-methodologic-quality studies as opposed to the low-methodologic-quality studies. The thrust of the consensus development exercise is to make scientific evidence outweigh expert opinion when data are available.

DISSEMINATION AND IMPACT

The Canadian Research Group is particularly interested in the dissemination and impact of the consensus development statement. Dissemination has occurred by publication in journals and some targeted mailing to specialists who will be affected by the statement. The active attempts to implement the statement are, however, potentially more important than this passive dissemination process. The research group is currently conducting a large randomized controlled trial to evaluate different ways to assess the value of two alternative implementation strategies for consensus recommendations in community hospitals. One strategy involves local educational influentials (individuals identified by their local colleagues); the other strategy focuses on the use of chart audit and data feedback. The results of this evaluation are not yet available but should be available in 1990.

In the meantime, the research group has evaluated the impact of the first consensus development statement (cesarean section) on the basis of its passive dissemination. The statement had an impact on clinicians' attitudes and self-reported practices. The impact on their knowledge is less marked, and the impact on their actual practice, as measured by hospital discharge data, is minimal, although statistically significant (Lomas et al., 1989).

References

Lomas, J., G.M. Anderson, K. Domnick-Pierre, E. Vayda, M.W. Enkin, and W.J. Hannah. 1989. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. New England Journal of Medicine 321:1306-1311.

Profile of the Consensus Development Program in Denmark: The Danish Medical Research Council and The Danish Hospital Institute

Torben Jørgensen

NATIONAL CONTEXT

The Danish consensus development program began in 1983 with the first conference on early detection of breast cancer. The idea of developing consensus by way of a public conference originates from the National Institutes of Health in the United States. A subcommittee under the Danish Medical Research Council (DMRC) initiated the conference program in Denmark. The subcommittee of the DMRC has cooperated with the Danish Hospital Institute to plan and implement six consensus conferences to date. The Danish Hospital Institute is a nonprofit institution, supported in part by the government and the county councils. The Medical Research Council and the Danish Hospital Institute have been the main sponsors of the program, although additional public and private funds have also supported the program. At present there are two consensus development conferences per year.

The main goals of the consensus development program are to inform the public about the state of the art of important health problems and alternative treatments for these problems, and to provide an information base for health professionals, administrators, and politicians involved in decision making for health care planning and in the formulation of research agendas. The first two conferences had the supplemental goal of investigating whether this imported

consensus methodology for technology assessment was appropriate for Danish society in the absence of a long-standing tradition of public hearings such as those seen in the United States.

The costs for one consensus development conference, including conference planning and staffing, the printing and dissemination of the report, and overhead costs, amount to 1989 DK 300,000-350,000 (approximately U.S. $50,000). This does not include fees for the planning group, questioning panel, or expert group, as their respective contributions are provided free of charge.

SCOPE OF THE PROGRAM

In Denmark, six consensus development conferences have occurred on medical topics. The next conference, Reduction of Cancer Mortality by 15 Percent before the Year 2000, is currently being planned. The previous conferences were as follows:

November 1983	Early Detection of Breast Cancer
December 1985	Prevention and Treatment of Dental Caries
September 1986	Cholesterol and Ischemic Heart Disease
April 1987	Secretory Otistis Media (Glue Ear)
October 1988	Physical Training and Health
January 1989	Senile Dementia.

A wide variety of technologies and health problems have been chosen as topics for consensus development conferences in the relatively short period of the program's existence. Three of the conferences addressed specific technologies, although alternative technologies were considered for comparison. For example, in the case of Early Detection of Breast Cancer, x-ray mammography was compared with palpation and self-examination.

Preventive technologies were assessed in four of the six previous conferences. Preventive technologies will also be considered in the upcoming conference on reduction of cancer mortality. Prevention has not explicitly been cited as a principal goal of the Danish consensus conference program, although the conferences are intended to inform the public of important issues in health.

The technologies assessed are established, for the most part, but in some cases, new technologies have been assessed. The requests for topics to be assessed have come from the academic community of DMRC. The potential methods used to involve health plans in

the topic selection process are currently under consideration for future conferences.

Topic priorities are set by the subcommittee under DMRC, which includes the director of the Danish Hospital Institute. The subcommittee appoints the planning committee, which consists of the chair of the expert group and of the conference questioning panel, one or two members of the DMRC subcommittee, the head of the secretariat, and one or two specialists in the topic to be assessed.

FORMAT AND CONDUCT OF THE PROCESS

The consensus development process consists of three phases: (1) the planning phase (six to ten months), (2) the conference itself (three days), and (3) the publishing phase (one to two months). The last phase is addressed in the section on the dissemination and impact of the consensus statement.

The Planning Phase

During the first phase, the planning committee formulates the main questions of the conference and selects the expert group and the questioning panel. The expert group consists of 12-16 professionals, mainly medical doctors, but also includes nurses, sociologists, psychologists, and economists who are working with the technology or the health problem to be assessed. The questioning panel consists of doctors, academics, journalists, politicians, and patient representatives. Practical conference logistics are handled during this phase. Two weeks before the conference, the questioning panel receives abstracts of the papers to be presented.

The Conference

The evening before the first day of the conference, the questioning panel and the expert group meet for the first time and attend an oral presentation on the conference process. The two groups then meet separately to discuss strategies for the conference.

The chair of the questioning panel conducts the conference. The first day of the conference is devoted to the presentation of approximately 15 expert papers. Only brief questions for clarification are allowed after the presentation of each paper. The conference is

open to the public; about 150 individuals usually attend. Most of the attendees are medical professionals, although journalists from newspapers and sometimes the television networks may attend the conference.

In the evening, the questioning panel meets in closed session to prepare the questions for experts on the second day of the conference. During the three-to four-hour morning session, the panel asks questions of the expert group. Other participants are given a shorter amount of time to raise questions and to provide comments. In the afternoon, evening, and night of the second day, the questioning panel meets in a closed session to discuss the issues at hand and to formulate the consensus statement. There is usually full consensus among the members of the questioning panel.

The consensus statement is presented to the expert group and to the audience on the morning of the third day. The expert group may not be in complete agreement with the questioning panel. The only alternations made at this point are corrections to factual faults in the statement, if there are any. The consensus statement is the responsibility of the questioning panel. The conference concludes with the release of the statement to the press.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

The chair of the expert group is responsible for the selection of other expert presenters and for the agenda of the first day of the conference. The composition of the expert group and the program for the first day are discussed by the planning group during the preparation for the conference. The chair of the expert group is urged to come forward with experts with different opinions. If the expert group does not include individuals with a variety of opinions, experts in the audience may voice the selectively omitted opinions. The preferred situation is to include all points of view in the formal presentations made by the group of experts. Each expert in the group is responsible for documentation of his or her presentation; the evidence may include reports of clinical trials, epidemiologic studies, literature reviews, etc. Often, references are made to international experiences, but apart from a few Swedish and Norwegian experts, the experts are all Danish.

DISSEMINATION AND IMPACT

In the time immediately following the conference, there is usually public debate on the topic in the newspapers, followed by debate in Danish professional journals. The statement is printed and widely distributed free of charge to the appropriate clinical, administrative, and political decision makers, as well as to libraries and research institutions. The statement is translated into English for publication of a small number of copies. There is no scientific evidence for the impact of the consensus processes, and the program has not been formally evaluated.

There are two main project goals; the first is to inform the public. Each conference has been well covered by the press, and there has been public debate following the conference, but we do not know whether this has changed the behavior of citizens. The conference on Physical Training and Health, which was held in October 1988, probably had a greater impact than those of previous conferences. The Heart Association aided in the dissemination of an increased number of consensus statements for public education and in the production of an educational video based on the conference.

The second goal is to establish an information base for making decisions in health planning. Here again it is difficult to state the impact of the consensus development conference. Many other factors influence decision making and the dissemination of information to health planners, although we do believe that the consensus conferences have had some impact. For example, the statement from the conference on Early Detection of Breast Cancer, which was held in November 1983, could not recommend general mammography screening, as the panel did not find evidence of clinical efficacy. The experts present at this conference held the opposite opinion. Denmark has not introduced general mammography screening, even though Sweden and other countries often examined for comparisons have introduced such protocols. The experts at the conference were dissatisfied with the statement; this is often the case for conferences where the panel does not recommend the widespread diffusion of the medical technology in question.

Apparently, the use of consensus development conferences for the assessment of medical technologies has encouraged development of a number of similar conferences in Denmark for the evaluation of technologies in other fields. The Prime Minister asked for a consen

sus conference on water pollution, the results of which had an immediate impact on the Parliament's decision. The questioning panel has consisted of citizens (which were selected to represent the population) for some of these conferences. Prior to the consensus conference these individuals attend two weekend seminars to learn about the topic under consideration. The next consensus conference in this series will address mapping of the human genome.

Profile of the Consensus Development Program in Finland: The Medical Research Council of the Academy of Finland

Anna-Liisa Kauppila

NATIONAL CONTEXT

The Medical Research Council (MRC) of the Academy of Finland has initiated the organization of consensus development conferences. The Academy of Finland is the central government agency for science administration and science policy planning. The primary task of the Academy is to fund and promote basic research. The Academy attempts to increase the resources for scientific research, to improve the efficiency of Finnish scientific research, to coordinate research work across administrative boundaries, and to plan research and science policy in Finland. The MRC is one of the Academy's seven councils. The MRC members are leading researchers and represent the faculties and the departments of community health in the medical schools as well as other research institutes in Finland. The evaluation of scientific issues occurs at the Academy because of the increased demand for technology assessments and the increased desire for better interaction between the research community, the politicians, and Finnish citizens.

The purpose of the consensus development conference program of the MRC is to raise interest in developing the assessment of health care technology in Finland. The MRC is the primary agency for organizing and sponsoring consensus development conferences for medical technology assessment. The MRC seeks the cooperation of

the League of Hospitals and the Institute of Public Health and, depending on the theme of the conference, the support of any appropriate association or nonprofit corporation. Three consensus development conferences have been sponsored in Finland. The first conference, held in 1985, was on the treatment of acute otitis media; the second, in 1987, was on the treatment of schizophrenia; and the third, in April 1989, was on cholesterol and coronary heart disease.

The consensus development method is a readily available method of assessment for adoption and adaptation in any national health care system. The desire to start the program was due both to an interest in the consensus development method itself and the need to apply the method in a practical situation when there was controversy over the appropriate use of a particular medical technology. The experience of the first conference on the treatment of otitis media was encouraging. The MRC hopes to gain experience on the use of the consensus development conference and to find ways to adapt the method to the needs of the Finnish health care system.

The principal goals of the Finnish consensus development conferences have not yet been fully defined, as the program is relatively new. Different goals were established in planning each conference. At present, the main goals of the program include the accumulation of experience with the method, development and adaptation of the method, and evaluation of the impact of consensus development conferences. The main purpose of each conference is similar to that of conferences conducted by the National Institutes of Health (NIH) in the United States, that is, to evaluate publicly the scientific information on health care technologies and to arrive at a consensus statement that will be useful for health care providers, researchers, and the public. Individuals from a variety of disciplines participate in the attempt to reach consensus on different topics. Researchers, practitioners, health care providers, and planners in health care at the community and national levels are included in the evaluation of the appropriate use of a technology and the implications of this use for patients and society.

At present, the practical goals of the program are to:

- provide a setting for the evaluation and review of the scientific evidence in support of or in opposition to the use of a health or health-related technology
- disseminate information from researchers to clinical practitioners, health care providers, and consumers

- facilitate the diffusion, adoption and appropriate use of sound technologies by reaching the decision makers in health care or other appropriate sectors, for example national legislators
- identify gaps in the current state of knowledge for a technology in order to initiate research.

The National Board of Health sponsors the publication of the consensus statements and participates in the dissemination of the consensus report. Administrators of the National Board of Health have served as members of the panel in two conferences. The National Board of Health is very interested in the consensus development program, although it does not have an official role in organizing the conferences.

The primary intended users of the consensus statements vary according to the theme of the conference. For example, the conference on acute otitis media targeted clinical practitioners, whereas the conference on cholesterol and coronary heart disease addressed a wide range of interested parties, from the providers of agricultural products and nutritionists to clinical practitioners.

SCOPE OF THE PROGRAM

The Finnish program has no set priorities as to the kind of technologies to be assessed (e.g., drugs, devices, or procedures). The technologies may be new or established; and the technologies considered may be for prevention, diagnosis, treatment, or rehabilitation.

A working group of the MRC plans consensus development conferences, although the MRC makes the final decisions with regard to topic selection and the membership of the planning committee. The MRC chooses the conference topic from a list of suggestions compiled by the working group. The MRC officially solicits topic suggestions from a wide group of interested parties, including the National Board of Health; the Hospital League of Finland; the members of the MRC; and the staffs of university hospital clinics, medical schools, and the departments of community health.

The MRC has not strictly defined topic priorities, although the basic criteria for selection resemble those used by NIH, as follows.

1. The subject under consideration should be medically important and should address major national health problems.

2. There should be a scientific controversy or a gap between current knowledge and practice or variation in practice.
3. An adequately defined base of scientific information must be available on the topic.
4. The potential to obtain consensus must exist.
5. There should be public interest in the issues addressed.

The issues addressed in a consensus development conference usually include the safety, efficacy, effectiveness, service requirements, adoption level, and economic implications of the technology, as well as social, psychological, ethical, or legal considerations as appropriate. The Finnish MRC holds the multidisciplinary approach to the broad assessment of technologies to be of great importance for the consensus development conference process.

FORMAT AND CONDUCT OF THE PROCESS

The planning process for a consensus development conference takes approximately a year and a half. The solicitation of topics from a wide range of interested parties by the MRC marks the start of the planning process. This initial planning phase for topic selection is carried out by a small working group on consensus development conferences. This group invites specialists to join in the initial planning phase as necessary. Knowledgeable researchers prepare papers on two or three topics to explain the criteria for selection of the particular subject. The primary organizations involved in the consensus development program (the National Board of Health, the League of Hospitals, and the Institution of Public Health) come to agree that the topic is the most suitable for consideration before the conference occurs. The MRC selects the most popular suggestions for topics and then nominates the full conference planning committee.

The chairperson of the planning committee is usually a member of the MRC with experience in the consensus development process. Members of the MRC working group also participate in the planning committee. The MRC identifies outside experts to participate in the planning committee. These individuals are the primary advocates of controversial opinions on the topic to be assessed. Two or three critics of the technology serve on the planning committee. The potential chairperson for the conference and other nonbiased experts

in a variety of fields participate on the committee. A representative of the League of Hospitals also takes part in the planning process.

The planning committee carefully considers the selection of the chairperson for the conference. All members of the committee must agree on the choice of the chairperson. He or she is usually a distinguished researcher who possesses the personal skills to successfully lead the panel toward consensus. The conference chair is invited to join the planning committee if the individual selected is not already a member of the committee.

The planning committee formulates the consensus questions and the conference agenda. The committee devotes substantial effort to the definition of consensus development questions. The different viewpoints of the members of the planning committee are reflected in the conference agenda and in the selection of speakers. In order that all views be represented in the program, discussants may be chosen together with the speakers. A discussant reviews a speaker's paper before the conference and then provides criticisms or comments on the speaker's viewpoint. The planning committee is considered crucial to the consensus development conference process, as the expertise of the group allows for the identification of speakers and discussants and the definition of an appropriate agenda. The planning committee acts as the nucleus of the consensus development process.

All appropriate experts are represented as speakers, including basic researchers, clinicians, epidemiologists, health economists, psychologists, and representatives from a variety of other disciplines. Nutritionists and representatives for the agricultural production industry participated as speakers in the 1989 conference on cholesterol. Speakers present both their own data and overviews of other available scientific data for their 15-to 20-minute presentations.

The planning committee must fully agree as to the selection of panel members. The committee carefully selects a panel with the appropriate range of expertise for the particular conference. Researchers, clinicians, planners, administrators, and other health care providers are represented on the panel. Panelists should not be advocates of a particular position on the topic. Representatives of the public on the panel may be editors of scientific journals or experts from other areas, for example, social sciences or information sciences. Designated patient representatives have not been included as panel members, although individuals chosen for their appropriate

expertise in a specific field may at one time have been patients. Speakers and members of the audience present the patients' perspectives. The size of the panel has varied; in the consensus development conference on cholesterol there were 16 panel members.

The basic format for a conference is similar to the NIH model and occurs over a period of two and a half days. The Finnish format continues to evolve, as each conference has been slightly different. The conference is an open, widely announced public meeting. Members of the medical community and the public are invited. Between 100 and 160 people have attended consensus development conferences.

The panel receives copies of the speakers' presentations, including all tables and graphs, three to four weeks before the conference. For the most part, panels meet one or two times before the actual conference occurs. The panel also meets the night before the actual public conference. During the preconference meetings the panel members discuss as a group their methods of approaching the statement (e.g., how groups should be composed to focus on specific questions) and of considering the presentations by speakers. The panel has the opportunity to raise additional issues or concerns for an open discussion. The panel can request additional information on the topic. The panel members may elect to submit questions to the speakers, to address the issues raised during this open discussion or "ventilation session" in their presentations. The panel may produce an introduction to their statement for discussion the night before the conference.

The planning committee chair opens the conference with a brief introduction to the consensus development process. The chair of the speakers' group then provides background information on the technology in question; the presenter summarizes areas where no controversy exists and a brief review of the relevant epidemiologic and scientific data. The first two days of the conference are devoted to plenary sessions where the experts deliver their presentations. This is followed by questions from the panel and the audience in a public discussion. Nonbiased experts alternate as the chair for different parts of the consensus development conference; for example, different individuals chair sections of the public meeting that address particular questions. The chairs may be members of the planning committee or speakers. The panel drafts the consensus statement in the evenings of the first and second days of the conference. The

panel may work until the early morning hours of the third day. The development of consensus occurs throughout the planning process of the conference. The conference represents the nodal point for consensus development—either the panel succeeds in agreeing on the statement or no resolution of the controversy occurs.

The consensus statement is presented to the audience for discussion on the morning of the third day. During the final discussion of the statement, proposals for changes can be made that are based only on the evidence given by the speakers or by the audience during the two preceding days. The panel will decide on any alterations following this discussion. After this point, no changes can be made. The chair of the panel presents the statement to the press.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

The planning committee is responsible for gathering material for the panel and the conference. The panel receives background material two to three months before the conference. The background material consists of literature reviews, epidemiologic studies, reports of clinical trials, relevant journal articles, and any available summaries of the state of the science. The planning committee distributes speakers' papers to the panel one month before the conference. The panel can only consider the evidence presented in the public meeting for development of the statement. The consensus statement does not include references to the materials used at the conference.

DISSEMINATION AND IMPACT

Announcements of the conferences appear in professional journals and newsletters in the health care sector. Media representatives receive personal invitations to attend the consensus conference and the press conference on the last day of the meeting. Announcements are distributed by mail to a variety of individuals in health care, including the Hospital League, the Society of Health Care Assessment of Finland, medical schools, institutes, selected university departments, university clinics, central hospitals, health administrators, communal administrators, etc.

The consensus statement is published in the main Finnish medical

journal shortly after the conference. A press conference occurs at the end of the consensus development conference, and a press release is distributed. The statement and the speakers' papers are published as a single document through the cooperative endeavors of the Finnish Academy and the National Board of Health. The sponsoring organizations are responsible for the dissemination of the report to members of their societies. To date, the impact of the consensus development program has not been formally evaluated.

Profile of the Consensus Development Program in the Netherlands: National Organization for Quality Assurance in Hospitals (CBO)

Niek S. Klazinga, Anton F. Casparie, and J. J. E. van Everdingen

NATIONAL CONTEXT

The origin of the CBO (Centraal Begeleidingsorgaan voor de Intercollegiale Toetsing, or National Organization for Quality Assurance in Hospitals) consensus development process lies in the quality assurance activities of medical specialists in hospitals (Klazinga et al., 1988). In the early 1980s, it became evident that it was difficult to develop criteria for audit studies at the hospital level for several controversial medical topics. The development of guidelines for medical practice on a national level was needed. In 1981, the Scientific Council of CBO decided to start a consensus development program similar to the one conducted by the National Institutes of Health in the United States. CBO is an independent nonprofit organization founded in 1979 by the Dutch Specialists Association and the Association of Medical Directors. The Scientific Council of CBO represents all 34 scientific medical associations in The Netherlands. The purposes of the CBO consensus development program are (1) to establish guidelines on controversial medical issues and (2) to ensure the quality of care by promoting behavioral change among medical practitioners (Casparie and van Everdingen, 1985a, b).

The primary goal of the consensus development program is to develop guidelines for daily medical practice such that they not only

represent the state of the art but also are acceptable to the medical community. The program is run for and by the profession itself (scientific associations), but finds recognition among financiers, government authorities, and patients' organizations. In recent government policy papers, consensus development is mentioned as one of the profession's contributions to quality assurance. The clinicians (specialists, general practitioners, nurses, physiotherapists, etc.) are the intended primary users of the consensus statements; however, recommendations may also be of interest to hospital administrators and policymakers (Casparie et al., 1987).

There are usually four conferences each year with an average attendance of 300 individuals per conference. The total yearly cost of the program is approximately dFl. 400,000 (U.S. $200,000). This includes costs for personnel, materials, meeting arrangements, overhead, etc.

SCOPE OF THE PROGRAM

The program assesses selected clinical problems. The choice of technology to be assessed depends on the problem in question: Drugs, devices, medical or surgical procedures, support systems, and organizational or administrative systems may be examined. Technologies for discrete stages of intervention are discussed for different clinical conditions; for example, prevention of bedsores, diagnosis of deep venous thrombosis, and treatment for osteoporosis (van Everdingen and Casparie, 1988; van Everdingen et al., 1988). Table 1 lists the CBO consensus development programs that have taken place over the years. Some of the technologies assessed are new (used only by practitioners in university hospitals) or already established and/or widespread. Some have been labeled as obsolete during the process of consensus development.

The properties of a technology that are normally addressed are safety, efficacy, effectiveness, and, to a lesser extent, cost-effectiveness and service requirements. Suggestions for consensus development topics come from scientific associations, peer review committees in hospitals, or medical foundations such as The Netherlands Heart Foundation (consensus on cholesterol). The Scientific Council of CBO is responsible for final selection of topics.

TABLE 1 CBO Consensus Development Program

Year	Program
1982	Blood transfusion therapy
1983	Traumatic lesions of the back
	Mammography policy
1984	Severe brain damage
	Melanoma of the skin
	Thrombocyte transfusion policy
1985	Solitary thyroid nodules
	Prevention of bedsores
	Osteoporosis
	Foot problems of diabetic patients
1986	Diagnosis of deep venous thrombosis
	Nonscrotal testis
	Treatment of bedsores
	Drug addicts in prison
	Hypercholesterolemia
1987	Prevention of herpes neonatorum
	Hemophilia
	Follow-up colon polyps
	Cholesterol
	Suspect lymph nodules in the neck
	Diagnosis of atopic syndrome
	Total hip joint replacement
	Follow-up of colorectal cancer
1988	Diagnosis of dementia
	Sports and cardiac pathologies
1989	Prevention of deep venous thrombosis
	Prevention of hospital infections
1990	Diagnostics for lung carcinoma
	Hypertension
	Acute otitis media
	Nutrition and allergy
1991	Cerebrovascular accident
	Diabetic retinopathy
	Treatment of deep venous thrombosis

The following criteria are applied in the choice of topics:

- controversial in the literature and among practitioners
- relevant in terms of health benefit
- feasibility of consensus development
- relevant for medical practice
- sufficient scientific data available.

FORMAT AND CONDUCT OF THE PROCESS

The consensus development process requires about two years. First, a topic is selected by the Scientific Council of CBO, and then a suitable chair is selected for the working group. Formation of the complete working group takes about three months. All members of the working group are experts and official representatives of their respective scientific associations. Representatives from the fields of nursing, physiotherapy, and general practice are invited to participate as needed. During the following 12 months, the working group meets six to ten times to develop a syllabus with background information and to draft consensus statements.

The consensus development conference takes place after completion of the draft statements by the working group. All practitioners and interested persons can attend the conference; attendance has ranged from 150 to 1,000 individuals. The conference itself lasts one to two days. Two weeks in advance of the conference, participants receive the syllabus with background information and drafts of proposed consensus statements. All statements are defended by the working group as a whole. After each presentation, there is ample time for questions; later in the program the audience may comment on the draft consensus statements. Considerable time is allocated for discussion between members of the audience and members of the working group. At the end of the day, the chair summarizes the results and tries to formulate a definite set of consensus guidelines. The chair asks the audience explicitly whether they agree with the consensus text, and then he closes the meeting. One month after the consensus development conference, the working group meets for the last time to finalize the text of the consensus statement for publication.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

The information used during the consensus development process comes from the experts in the working group. Sometimes, formal literature searches are performed on a small scale, but in most cases the experts bring in literature as well as their own material. The working group may use reports of clinical trials, epidemiologic studies, and literature reviews for particular topics (Klazinga et al., 1987). The available evidence was weighed systematically on several occasions according to the method proposed by Sackett during the National Institutes of Health conference on the prevention of venous thrombosis and pulmonary embolism, March 24-26, 1986 (Sackett, 1986). On other occasions, the evidence collected from the literature and the experts' experience was combined. In the syllabus written by the members of the working group, reference is made to the data used during the consensus development process. The most important literature is also mentioned in the text of the final consensus statement.

DISSEMINATION AND IMPACT

CBO consensus development conferences are widely announced to reach physicians in The Netherlands via professional journals and direct mailing to the appropriate specialty groups. Announcements are posted in hospitals to attract attention. The final consensus statements are published in the *Nederlands Tijdschrift voor Geneeskunde* (a Dutch medical journal); and copies are sent to those who attended the conference as well as all medical staff, hospital administrators, and chairpersons of peer review committees. Scientific associations that cosponsor the consensus development conference sometimes send the consensus results to their members. The text can be purchased from CBO, and the CBO staff is also active in disseminating the consensus reports to hospital practitioners who are involved in quality assurance. In a few cases, follow-up meetings have been held for specialists who were unable to attend the consensus development conference.

The consensus statements are intended to change clinical practice by altering physician behavior. This is why CBO works for large-scale commitment and involvement of medical specialists in the

consensus development process. Although the consensus guidelines have no legislative bearing, they acquire status from the involvement of the different scientific associations and experts. The CBO consensus development program has been evaluated in terms of both process and effect (van Everdingen, 1988). Evaluations have addressed:

- the consensus meeting process and the activities in the working groups
- impact of awareness of consensus guidelines among medical practitioners
- impact of consensus guidelines by practitioners (formulation of protocols and criteria setting for audit studies in hospitals)
- effects of consensus guidelines on the behavior of medical specialists (on the hospital and national levels).

The evaluation activities of CBO are summarized in Table 2.

References

Casparie, A.F., and J.J.E. van Everdingen. 1985a. Consensus development in The Netherlands. European Newsletter on Quality Assurance 2:1, 5.

Casparie, A.F., and J.J.E. van Everdingen. 1985b. Consensus development conferences in The Netherlands. International Journal of Technology Assessment in Health Care 1(4):905-912.

Casparie, A.F., N.S. Klazinga, J.J.E. van Everdingen, and P.P. Touw. 1987. Healthcare providers resolve clinical controversies: The Dutch consensus approach. Australian Clinical Review 7(24):43-47.

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

Klazinga, N.S., A.F. Casparie, and J.J.E. van Everdingen. 1988. Quality assurance on the crossroad of medical decision making and management of the health care process in the hospital (abstract). Theoretical Surgery 3:42-43.

Sackett, D.L. 1986. Statistical issues in clinical trials. Pp. 36-40 in Program and Abstracts for the NIH Consensus Development Conference on the Prevention of Venous Thrombosis and Pulmonary Embolism. Bethesda, Md.: Office of Medical Applications of Research.

van Everdingen, J.J.E. 1988. Consensusontwikkeling in de Geneeskunde. Thesis (English summary). Utrecht, Bonn: Scheltema & Holkema.

van Everdingen, J.J.E., and A.F. Casparie. 1988. Consensus development conferences on osteoporosis (letter to the editor). British Medical Journal 296:61-62.

van Everdingen, J.J.E., N.S. Klazinga, and A.F. Casparie. 1988. Blood transfusion policy in Dutch hospitals. International Journal of Health Care Quality Assurance 1:16-19.

TABLE 2 Evaluation Activities in Relation to CBO's Consensus Development Program

Topics of Evaluation	Method	Scope[a]	Effect
Consensus meeting	Survey of participants	Standard procedure performed for 22 conferences; participants surveyed varied from 70 to 100	Permanent feedback on consensus procedures
Working group	Survey of all members of the working group	Standard procedure performed for 22 conferences; working group size varied from 10 to 15	Feedback on functioning working groups
Awareness of consensus guidelines on:			
Blood transfusion	Oral survey of 50% of the 150 clinical chemists in The Netherlands	*N* = 75	100% awareness
Deep venous thrombosis	Survey of general practitioners	*N* = 435/706	37% mention consensus as most important source of their own policy.
Use of consensus guidelines for:			
Audit activities	Collection of criteria for audit studies performed on consensus development topics	Different topics; different numbers of studies	Feedback on applicability of guidelines
Protocol development			
On bed sores (Consensus Conference 1985)	Surveys in hospitals and nursing homes	1984: *N* = 143/477; 1986: *N* = 89/143; 1989: *N* = 167/477	Used frequently
On blood transfusion	Oral survey of 50% of the 150 clinical chemists in The Netherlands	*N* = 75	
On melanoma of the skin	Survey of the regional cancer centers	*N* = 8/8	Three centers formulated guidelines and two of them used the consensus regulations

Topics of Evaluation	Method	Scope[a]	Effect
Effect on behavior			
Melanoma of the skin (consensus conference, 1984)	Analysis of histopathological reports from 19 laboratories with computerized files	1983: *N* = 324; 1984: *N* = 383; 1985: *N* = 384	Major compliance of pathologists on reporting guidelines; less, but growing, compliance of dermatologists with guidelines on incisional procedures
Blood transfusion	Analysis of annual reports of blood banks	*N* = 22/22	Growing compliance with consensus guidelines; frequent use for criteria setting
	Analysis of audit studies	*N* = 17	

[a]N indicates the number of respondents/number of people surveyed, as available.

Profile of the Consensus Development Program in Norway: The Norwegian Institute for Hospital Research and the National Research Council

Bjørn Backe

NATIONAL CONTEXT

The Norwegian Institute for Hospital Research (NIHR) initiated the first consensus conference, Use of Ultrasound Screening in Pregnancy, in 1986. This nonprofit, health services research institute mainly conducts health planning and contract research for health authorities and hospital owners. The NIHR arranged the first consensus development conference in cooperation with the Health Services Research Unit of the National Institute of Public Health. The NIHR sponsored the conference in conjunction with the Royal Norwegian Ministry of Health and Social Affairs.

In 1987, the Ministry of Health and Social Affairs asked the Norwegian National Research Council (NRC) to develop a program for consensus development conferences in Norway. The NRC recommended a three-year program with two consensus development conferences per year. At present, the Norwegian Institute for Hospital Research and the NRC work together to organize and finance the consensus development program. The first conference in this program, on mammography screening, occurred in February 1989. The second conference, on reduction of the population's cholesterol level, occurred in October 1989.

The purpose of the program is to improve the policy and practice for areas of concern in the health services by providing objective

information on controversial medical matters for political and administrative decision makers, medical professionals, patients, and general consumers.

The panel from the first conference on ultrasound in pregnancy recommended that one routine examination should be offered to all pregnant women (Norwegian Institute for Hospital Research, 1987). The Directorate of Health immediately endorsed the statement as the national guidelines for obstetrical use of ultrasound. Recent research indicates that the consensus development conference statement has had a measurable influence on the practice of ultrasound (Nafstad and Backe, 1989). Two national cross-sectional surveys of ultrasound practice have been made: a short time before and two years after the consensus conference. In this time period, there was a reduction in the total number of examinations performed and an increase in the proportion of pregnant women being examined (98 percent). There are roughly 52,000 deliveries per year in Norway. At present approximately 12,000 fewer examinations are performed per year in Norway compared with the 1986 rate for the procedure.

The second consensus panel noted that mammography screening should not yet be routinely offered, pending the results of ongoing research that is expected to provide further evidence toward resolution of the controversy. The Minister for Health and Social Affairs referred to the statement in a parliamentary debate shortly after publication. She agreed with the consensus statement and stated that the national guidelines (i.e., the treatment program for the use of mammography) would be published in the near future, based on the consensus development conference statement.

The national health authorities, the county hospital owners, and the health professionals have used the statements from the Norwegian consensus development conferences as a reliable form of advice. The two past conferences have been funded jointly by the Royal Norwegian Department for Health and Social Affairs and the NIHR. The total cost per conference is about NKr 382,580 (U.S. $55,851). This includes direct costs for the planning group, conference report, informational material, and staffing, including overhead costs.

The panelists and the expert presenters do not receive compensation for their participation. Funds for the conference are used to employ the leader of the conference planning group. The planning group leader may be one of the researchers at the NIHR who works

part time with consensus development conferences or, in some instances, may be employed on a short-term contract. The planning group leader is responsible for the financial aspects of the consensus development conference program. For the most part, the planning group leaders have been involved in research projects that addressed issues concerning the particular topic for the consensus development conference.

SCOPE OF THE PROGRAM

A medical technology must meet the following criteria before assessment by the consensus conference method.

- The technology must be of broad general interest.
- The ethical, social, and/or other consequences must be large.
- The number of patients must be large, or the costs great.
- The topic is within the medical-scientific framework (i.e., there must be enough scientific evidence for the questions to be answered on a scientific basis).
- There must be disagreement as to the utility of the technology (i.e., there must be a real dispute within the medical profession on how to interpret the available scientific results, and how to apply the technology in practice).
- The statement should have the ability to influence the diffusion of the technology and to alter current clinical practice.

The term medical technology broadly defines those instruments, practices, and procedures based upon medical/biological knowledge. The Norwegian program does not have explicit limitations as to the appropriate stage in the life cycle of a technology for assessment by a consensus development conference. The program will probably focus upon new and established technologies.

At present, NRC and NIHR are collecting suggestions for future conference topics. NRC established a subcommittee on technology assessment: the TA committee. Two health services research institutes, the Royal Norwegian Ministry for Health and Social Affairs and the Directorate of Health, are represented on the TA committee. The committee chair is a representative of the Medical Research Council.

The TA committee selects conference topics, drafts the questions, and appoints a planning group for each conference. As a rule, both

the chair of the consensus panel (appointed by the TA committee) and one or more of the topic experts in the field participate in the planning group. The planning group refines the questions for the panel, selects panelist and experts, and prepares the practical arrangements for the conference. The work of the planning group is subject to the approval of the TA committee.

The aspects of a technology addressed at a consensus development conference include effectiveness and efficacy (or practical, achievable health benefits); health risk; monetary cost; and organizational, ethical, social, and psychological consequences of increased or decreased use of a technology. Equal access to the technology (equity) is a particularly important consideration. The questions are formulated so that the statements can be used as a basis for treatment programs that serve as guidelines for the treatment of certain conditions or patient groups. The consensus development conference statement itself is insufficient for use as a treatment program, as treatment programs provide far more detail on patient care.

FORMAT AND CONDUCT OF THE PROCESS

The necessary preparation time for a consensus development conference, from the selection of a topic to the composition of the consensus statement, is approximately one year. The conference itself lasts two and a half days; the panel completes the statement during the second night of the conference.

The panel members are offered assistance from NIHR in the assembly of relevant literature. Panelists received reports of pertinent randomized controlled trials (RCTs) and other relevant literature before the two recent conferences. No formal protocol exists for the distribution of information to the panel before each conference. The chair of the panel and the planning group decide upon a procedure for materials procurement and dissemination for the individual conference.

The Norwegian program does not have a formal procedure for selection of the panelists. Approximately half of the panel is to come from the medical profession, including experts in epidemiology and statistics, community medicine, and general practice, as well as experts in other medical fields who are familiar with the technology in question but who are not experts on the subject. The nonmedical members of the panel are experts in health economics,

ethics, law, health journalism, health administration at various levels, and health politics.

As a rule, the panelist should not have expressed strong opinions on the consensus topic. They are invited as individuals, rather than as members or representatives of groups or organizations. They are experts in their fields, although they are not necessarily experts on the topic of the conference. They receive both written and oral information on the concept of consensus development. In general, the panel of representative, well-informed individuals seeks to achieve consensus through discussion and debate.

The conferences are public meetings with short scientific presentations by experts (15 minutes each) that are followed by discussions. During the general discussion period, the panel members question the experts. Questions and comments may also come from the audience. Approximately 200 people attended the first two conferences, most of whom were physicians.

On the morning of the third day, the panel presents the consensus statement. The panel prepares the final version of the statement without the assistance of any other individuals. The panel meets privately a number of times during the conference. The first closed session occurs the evening before the start of the public presentations.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

The first two Norwegian conferences focused upon technologies for health screening. In both cases, the question of utility rested heavily upon a small number of RCTs (four and three studies, respectively). If the RCTs were performed in Nordic countries, the principal author presented the material. In other cases, the conference planners commissioned an expert to present the material.

The experts must submit a written abstract of their oral presentation. Panelists receive the abstracts a minimum of four weeks before the conference. The working language throughout the conference is Norwegian.

As a rule, only published evidence has been considered. Documentation of both the state of knowledge and the state of practice are important elements of the consensus development conference process. Often studies must be performed before each conference to

estimate the actual diffusion, use, costs, and health effects of the relevant technology. Appropriate institutions and experts are commissioned to prepare reports in these cases.

Panelists do not adhere to any formal guidelines for weighting the presentations. The panel reaches consensus through discussion. The process is not rigidly structured. References are not available in the consensus statement, as they are not considered to be necessary. The panels do, however, provide conclusions and detail the reasoning behind the consensus statement.

DISSEMINATION AND IMPACT OF THE CONSENSUS STATEMENT

After each conference, the proceedings are published in a report, together with the consensus statement and a short review of the main reasons for holding the conference. The mechanisms for dissemination of the consensus findings include:

- press conference
- newsletter to the media
- distribution of the statement to administrative authorities
- publication of the statement in the *Journal of the Norwegian Medical Association* and distribution of the statement to other appropriate journals, and to radio, television, and other news media.

An evaluation of the impact of the first conference (which was held in 1986) is in progress as a part of a thesis on the diffusion of medical technologies. The previously mentioned study of the practice of ultrasound before and two years after the consensus development conference was presented at the June 1989 meeting of the International Society of Technology Assessment in Health Care (Backe and Nafstad, 1989). The sponsoring organizations are currently considering the feasibility of an evaluation of the ongoing three-year program.

References

Backe, B., and P. Nafstad. 1989. Development in practice of diagnostic ultrasound in Norway after the consensus conference recommending routine screening (abstract). Presented at the annual meeting of the International Society of Technology Assessment in Health Care, London.

Nafstad, P., and B. Backe. 1989. Prevalence and practice of obstetrical ultrasound in Norway (in Norwegian). Tidsskr Nor Laegeforen 109:2975-2978.

Norwegian Institute for Hospital Research. 1987. Ultrasound in pregnancy: Consensus statement, 1986. International Journal of Technology Assessment in Health Care 3(3):463-470.

Profile of the Consensus Development Program in Sweden: The Swedish Medical Research Council and the Swedish Planning and Rationalization Institute for the Health and Social Services

Stefan Håkansson and Ingemar Eckerlund

NATIONAL CONTEXT

Sweden was the first country outside of the United States to adopt the consensus development conference as a method for technology assessment (Calltorp and Smedby, 1989). The format of the Swedish conferences follows the U.S. National Institutes of Health model quite closely. Since 1982, 11 consensus development conferences have been arranged jointly by the Swedish Medical Research Council (MFR) and the Swedish Planning and Rationalization Institute of the Health and Social Services (Spri).

The goal of each conference is to reach consensus among specialists from different disciplines concerning the safety, benefit, and appropriate use of a particular medical technology. Practicing physicians, representatives of the lay public, health care administrators, and politicians often participate. Swedish consensus development conferences have three major purposes:

1. to review and assess the scientific base for a medical technology
2. to contribute to the dissemination of knowledge concerning the technology
3. to provide information concerning the appropriate use of the technology.

The cost of a Swedish consensus conference is about Skr 500,000 (about U.S. $75,000) in 1989 prices. This includes costs for planning, staffing, and report dissemination and overhead costs associated with consensus development activities. The costs are equally divided between the MFR and Spri.

SCOPE OF THE PROGRAM

Since 1982, the following consensus development conferences have been held:

- Total Hip Joint Replacement (1982)
- Treatment of Myocardial Infraction (1983)
- Treatment of Depressive Disorders (1984)
- Sight Improving Surgery (1984)
- Diagnostic Imaging of Liver Tumor (1985)
- Cerebral Hemorrhage and Stroke—Diagnosis and Treatment (1986)
- Urinary Incontinence in Adults—Diagnosis and Treatment (1986)
- Chronic Leg Ulcers—Diagnosis and Treatment (1988)
- Postoperative Wound Infections—Hygienic Routines in Hospital (1988)
- Preoperative Routines (1989)
- Venous Thrombosis—Diagnosis, Prevention and Treatment Indications (1989).

The types of technologies that have been addressed are primarily medical and surgical procedures. Table 1 details the types of technological interventions and the stages in the life cycle of technologies that have been assessed by way of consensus development.

Safety, efficacy, effectiveness, cost, and economic implications are technological properties or concerns that are usually addressed during the consensus development conferences. Ethical, legal, or social implications have hitherto played a relatively minor role.

The criteria for selection of a topic in the consensus development program include the following.

- The topic should be medically important and have the potential to eventually produce changes in medical practice (i.e., there is a gap between current knowledge and current medical practice).
- The topic should be clearly defined, and a corresponding scientific knowledge base must exist.

TABLE 1 Consensus Development Conference Topics According to Life Cycle Stages of Applicable Technologies Assessed

	Stages in the Life Cycle of Technologies			
Types of Technological Intervention	Emerging	New	Established	Obsolete
Prevention			Wound infections; Venous thrombosis; Incontinence	
Diagnosis		Liver tumor; Venous thrombosis; Preoperative routines	Liver tumor; Stroke; Incontinence; Venous thrombosis; Leg ulcers; Preoperative routines	Liver tumor; Preoperative routines; Venous thrombosis
Treatment	Sight improvement	Hip replacement; Acute myocardial infarction; Sight improvement; Venous thrombosis	Acute myocardial infarction; Wound infections; Stroke; Incontinence; Venous thrombosis; Leg ulcers	Venous thrombosis
Rehabilitation	Hip replacement			

- It must be possible to answer the questions on technical scientific grounds, and the result must not be biased by the subjective values of the panel.

FORMAT AND CONDUCT OF THE PROCESS

The planning and implementation of the consensus development conference process lasts from one to one and a half years. The conference itself lasts for two and a half days.

The MFR's initiative group on the assessment of medical technology considers topics for consensus development conferences. The proposals from this group go to a steering committee, having representation from the sponsoring organizations (MFR and Spri), which in turn selects topics, develops conference questions, and appoints participants to the panels and expert groups. Since 1985, both the Federation of the County Councils of Sweden and the Swedish Society of Medicine have been asked to suggest conference topics. The conferences are planned in detail by special working groups appointed for each conference. These working groups usually consist of the panel chair, the expert group, two representatives from each of MFR and Spri, one representative from the Swedish Society of Medicine, and one representative from the National Board of Health and Welfare.

The panels normally consist of approximately 16-18 members. About half of the panelists are medical experts, while the others usually represent health economics, epidemiology, health policy, administration, and concerned patient groups.

Although the format of the Swedish conferences follows the U.S. model quite closely, the scope is somewhat broader. In addition to the evaluation of the safety and efficacy of a technology, the Swedish conferences address health care organization, cost effectiveness, and social and ethical questions.

One and a half days are dedicated to expert presentations, including questions and discussion. The speakers present not only their own data but also general overviews of available data in order to present the actual state of the art. The panel begins to write the statement on the afternoon of the second day. On the third day the panel presents the statement to the experts and the audience (approximately 200 individuals) for discussion. The panel meets again

after this general session to decide on possible modifications. The conference ends with a press conference.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

The expert group assembles information for each consensus development conference. The presentations by the experts are based on various kinds of information. For example, reports of clinical trials, epidemiologic studies, and literature reviews have been used as evidence.

Panelists and participants receive paper abstracts about one month before the conference begins. The consensus statement's validity rests solely upon the scientific evidence presented to the panel for decision making. Information that has not been presented at the conference is not to be considered by the panel. Only presentations by the experts, distributed materials, and answers to the questions of the panelists are to be taken into account in the panel's statement.

There are no explicit rules for the development of consensus and the use of evidence during the executive sessions of the panel. The panels for each conference decide upon the definition of consensus. The chairperson leads the discussion and consideration of evidence in the format chosen by the panelists.

The final product of the consensus development process is the consensus statement. In general, consensus statements do not cite articles used as evidence. Each consensus statement contains a list of the members of the program committee, the expert speakers, and the panelists.

DISSEMINATION AND IMPACT

Spri publishes the consensus statements and distributes them free of charge to politicians, health care administrators, chief medical officers, and concerned health care personnel. The statements are also published in their entirety in the *Journal of the Swedish Medical Association*. The Swedish consensus statements regularly receive considerable publicity through the radio and television media. The press actively disseminates the key points of the statements.

A committee worked for two years to evaluate Swedish consensus development conferences and to trace the impact of consensus statements on the principal actors in the health care arena (Berfenstam et al., 1986; Calltorp, 1988; Johnsson, 1988). The studies were directed toward the evaluation of the impact of the conference on specialists in the actual specialties concerned and on the top administrators and politicians in the county councils.

Details of the evaluations were presented by Calltorp (1988) and Johnsson (1988). The main conclusions are as follows.

- The statements are very well known among all major groups concerned. The politicians and administrators tend to report a higher practical utility of the statements.
- The statements are regarded as good educational material, as they provide useful background reading for planning and decision making.
- The conferences and the statements could be influential in the priority setting process within county councils.

References

Berfenstam, R., G. Carlsson, and L. Werkö. 1986. Konsensuskonferensers Roll och Effekter. (The Role and Effects of Consensus Development Conferences.) Stockhom: Medicinska Forskningsradet, Spri.

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.

Calltorp, J., and B. Smedby. 1989. Technology assessment activities in Sweden. International Journal of Technology Assessment in Health Care 5:263-268.

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.

Profile of the Consensus Development Program in the United Kingdom: The King's Fund Forum

Jackie Spiby

NATIONAL CONTEXT

The consensus development program in the United Kingdom was initiated by Barbara Stocking and Bryan Jennett. These two individuals had observed the consensus development program in the United States and were eager to develop similar conferences in the United Kingdom. They approached various funding bodies to support the initiative and eventually obtained funding for the first conference from the King's Fund. This funding was then extended for three additional conferences and subsequently for another four, for a total of eight.

The King's Fund is an independent charitable organization concerned with the development of health services, management training, and policy analysis in the health care arena. The independence of the Fund allows the consensus development program to bring together a wide variety of individuals from different organizations to work together and speak freely on medical technology assessment. The King's Fund has a very extensive network within the health services and is generally well regarded.

The purpose of the consensus development program is to promote public debate about important controversial issues. The program aims to produce an authoritative independent consensus statement that can be used to initiate changes in health services and stimulate

research. The initial program was also established to experiment with and investigate the usefulness of consensus development conferences within the U.K. context.

The primary users of the statement vary according to the topic but include national and local policymakers, clinicians, hospital managers, health professionals, and consumers.

SCOPE OF THE PROGRAM

The King's Fund established the consensus development program specifically to investigate the applicability of the format of a consensus development conference for different topics in health care; therefore, the program has included clinical, public health, and social policy issues (Table 1).

The topics are chosen by the King's Fund Fora Steering Group, a multidisciplinary group of senior members in the health care field that oversees the whole program of conferences. This group is chaired by Bryan Jennett. Suggestions for topics are received from a variety of sources, and a short list is produced by the Steering Group.

The main criteria used for inclusion of a topic in the program are the following.

- It is an important public health issue.
- There is multidisciplinary involvement.
- There is real controversy.
- There are data available.
- There is public interest.
- It is timely.
- It is considered that the consensus development format is acceptable for investigation of the particular problem.
- It is considered that change can be identified and that there is potential for initiating change.

Specific technologies or clinical problems may be assessed in the program. Once the Steering Group chooses a topic, all relevant issues are included in the assessment, for example, legal, economic, political, social, and organizational issues. A technology may be considered at any stage in its life cycle.

TABLE 1 Consensus Development Conferences in the United Kingdom, 1984-1989

Title	Date	Special Features
Coronary Artery Bypass Surgery	November 1984	First U.K. conference
Breast Cancer Treatment	October 1986	Nonmedical chairman
Role of Asylum in Society	April 1987	Broader policy issues considered
Prenatal Screening	December 1987	Major ethical issues considered
Treatment of Stroke	June 1988	Return to more clinically based topic
Intensive Care	April 1989	In collaboration with the King's Fund Institute—expert panel producing statement prior to the conference
Cholesterol Measurement in the Prevention of Coronary Heart Disease	June 1989	Less than half of the panel was medically qualified

FORMAT AND CONDUCT OF THE PROCESS

The consensus development process for each conference lasts approximately one year from the Steering Group selects the topic to the publication of the final consensus statement. The actual conference itself lasts for three days.

Staff of the King's Fund Fora conduct a wide-ranging search to identify speakers and other contributors. They are informed in this process by a planning group that participates in the development of each conference. Each conference has a unique planning group. The search includes a computer search; interviews; literature review; and discussions with leading experts, policy institutes, the Department of Health, etc.

The King's Fund staff and the planning group choose the 12-person panel to provide a broad range of expertise. No more than six members are medically trained. Any panel would generally include clinicians, an epidemiologist, a statistician, an economist, a consumer, a nurse, and a manager. However, the final balance of the professionals is dependent on the actual subject. The panel is

nonexpert on the topic of the conference. However, one or two unbiased experts are included with the intent that they ensure that the panel does not misinterpret the information.

The relationship between the speaker and the data varies depending on the particular topics. Speakers may be asked to present their own data, but often speakers present an overview. On occasion, the Fora program has commissioned specific work, especially economic analysis for a conference.

The general format of the conference follows the style of the U.S. National Institutes of Health quite closely. The first day and a half consist of expert presentations and audience discussion. Sufficient time is provided for audience discussion, and the chairman is expected to ensure that the public is allowed to participate. Considerable effort is made to ensure that the audience is composed of professionals from different disciplines. Space in the audience is provided for members of the public. On the second day of the conference, during an open session, approximately ten members of the audience provide three-minute presentations on the topic. The panel then retires for the second afternoon and the morning of the third day to produced their statement. Each panel has managed to get some sleep on the night of the second day. The statement is then taken back to the audience for general discussion on the afternoon of the third day. The panel finalizes the statement during the end of the afternoon of the third day. Approximately 50 percent of the audience return to discuss the statement, and a lively debate usually occurs. The consensus statement cannot be modified once the panel completes the review during the conference.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

The King's Fund Fora staff assembles the information for the consensus development process. Initially, the staff conducts computer searches and wide-ranging literature reviews, along with interviews of experts and other knowledgeable people, to identify major issues and controversies. In view of the variety of issues considered by the conference, the staff uses many different types of literature in the search process (e.g., national reports, clinical trials, epidemiologic studies, literature reviews, editorials, and books).

Before the conference, the panel members receive a comprehen

sive selection of readings from the literature. The reading packet has included one or two introductory texts for those members who lack specific knowledge of the topic. Panel members also receive a 1,000-word abstract on the conference presentations from each speaker. There is no prior weighting of the evidence by the conference staff.

The panel is instructed to place considerable primary emphasis on the scientific evidence presented. Expert opinion should serve to support the available scientific evidence. Inevitably, during the course of the conference, the panel receives conflicting expert opinions. The comments of consumers from the audience are often important to the panel. The conference organizers try to provide an objective consumer view based on available scientific research.

The consensus statements do not cite references. Citations are not relevant for the statement, as it is produced in a format of approximately 3,000 words. References and identification of the evidence used can be found in the abstract book produced for each conference.

DISSEMINATION AND IMPACT

The announcement of the conference is made both in the relevant journals and by direct mailings. Evidence from the field of market research and from an unpublished study on the conference indicates that direct mailing is much more successful in increasing attendance.

A press conference is held on the afternoon of the third day of the conference to make direct contact with members of the press. Several of the conferences have succeeded in attracting attention from the radio and television media. Within a week, the statement is published and copies are sent to the press, relevant organizations, every district and regional health authority, and relevant clinicians. The statement is then available on request. The statement is usually published in the *British Medical Journal* the week following the conference. Editorials and articles in other relevant journals are also produced.

The intended impact of the statement depends on the content of the statement. The program aims to influence national and local policymakers, clinicians, other health professionals, and consumers. The statement does not always influence all of these groups. The statements are used by organizations as appropriate for their pur

poses. The King's Fund uses the statement within the Acute Services Program to develop standards and create health service development.

The Fora program has been evaluated formally and informally (Table 2). The results of these evaluations provide an ongoing form of assessment and are continually used to modify and enhance the program. Table 2 provides evidence of the impact of the consensus development process. The King's Fund continues to review and evaluate the consensus development program.

TABLE 2 Evaluation Activities in Relation to King's Fund Fora Program

Program Aspect	Evaluation Activity
Consensus meeting procedure	After all conferences the panel, speakers, and audience are surveyed as part of the regular monitoring and development of the program.
Role of consumers	Following the breast cancer conference, the role of the consumer in a consensus development conference and the accessibility of the statement to the public were reviewed.
International comparison	Ph.D. thesis comparing U.S. and U.K. conferences.
Review of impact	Survey of managers and clinicians following Coronary Artery Bypass Surgery conference to assess awareness of the statement and resultant change.
	Survey of nurses and consumers following Breast Cancer Treatment conference to review use made of statement.
	Survey of regional health authorities to consider impact of genetic screening statement on influencing regional policy.
	First phase of a study to review the impact of the statement of the Impact of Stroke conference has been undertaken to identify how close services mirror the recommendations. Follow-up is to take place two years postconference.
	Contact established with a small group of individuals who identified themselves as wishing to utilize the statement of the Impact of Stroke conference to initiate change. The progress that has been achieved is being monitored.

Profile of the Consensus Development Program in the United States: The National Institutes of Health Office of Medical Applications of Research*

Sharon R. Baratz

NATIONAL CONTEXT

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. Following the initiation of the Consensus Development Program in 1977, the Office of Medical Applications of Research (OMAR) was formally established in the Office of the Director of the National Institutes of Health (NIH) in 1978. OMAR is the focal point for activities to improve the assessment and translation of results from NIH-supported biomedical research into knowledge that can be applied safely and effectively in the practice of medicine and public health. OMAR is part of NIH, the primary government-sponsored biomedical research facility in the United States.

The principal vehicle for OMAR's efforts in the systematic assessment of biomedical technologies is the Consensus Development Program. Each consensus development conference is cosponsored by OMAR and one or more of the NIH bureaus, institutes, or divisions (BIDs). Other federal agencies with biomedical components

* The author acknowledges Jerry Elliott of OMAR, NIH, for reviewing the manuscript.

may participate in the sponsorship of particular conferences, depending on the topic (OMAR, NIH, 1988). The approximate cost per assessment for the OMAR conferences is $116,000 (Elliott, 1989). The purpose of the 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)" (OMAR, NIH, 1988).

The three main goals of the consensus development program are to:

- provide a setting for the evaluation and review of the scientific soundness of health technologies for a particular clinical condition or for a particular health-related technology, with emphasis on safety and efficacy
- aid in the diffusion of knowledge of advances in biomedical technology, through dissemination of the findings from the consensus development process to physicians and consumers
- to facilitate the diffusion, adoption, and appropriate use of technologies found to be sound.

SCOPE OF THE PROGRAM

A broad variety of technologies have been topics of consensus development conferences, including medical and dental drugs, devices, procedures, facilities, and support systems used in prevention, diagnosis, and treatment. Conferences tend to focus on a technology, for example, electroconvulsive therapy (1985) and magnetic resonance imaging (1987), or on a particular clinical problem and the alternative technologies applied for prevention, diagnosis, treatment, or rehabilitation of these, for example, travelers' diarrhea (1985) and adult urinary incontinence (1988).

Although the consensus development program was to have originally focused on emerging technologies, most of the conferences have addressed technologies already in clinical use, especially new or widely used technologies. This is largely because evaluative information regarding many emerging technologies is insufficient for the level of validity sought for consensus development conferences and because many technologies already in widespread use have

not been carefully scrutinized for safety and efficacy (Perry and Kalberer, 1980). The Consensus Development Program at OMAR is primarily concerned with the safety, efficacy, and clinical application of technologies. Conferences do not usually directly address social, ethical, legal, economic, or political issues surrounding technologies.

For the most part, the 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). The following criteria are currently in use by OMAR 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 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.

FORMAT AND CONDUCT OF THE PROCESS

Once a topic is chosen, a senior OMAR staff person works with a designated BID coordinator and other BID staff to organize the conference. OMAR's focus is on the consensus development process while the initiating BID's contribution concerns the scientific information required for the conference topic (Elliott, 1989). A planning committee of OMAR staff, BID staff, the conference chairperson,

and outside experts (who are not federal employees) begin to organize the conference. The program planning committee has four major functions:

1. to draft consensus questions (usually four to six questions)
2. to draft the conference program
3. to recommend conference speakers
4. to recommend consensus panel members (OMAR, NIH, 1988).

The chairperson of the consensus development conference is selected for his or her stature as a distinguished physician or scientist and for personal skills in chairing the open symposium portion of the conference and in leading the consensus panel. The size of the panels has varied from 8 to 16 individuals; most have had 10 to 12 members. OMAR seeks balanced representation from various sectors of professional and community life. The planning committee usually decides upon the appropriate areas of expertise needed for panelists of a particular conference. OMAR holds that all panelists, including the chair, should have no vested interest in the assessment topic and should be able to weigh evidence and to collaborate. 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 groups 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). The planning committee also selects speakers on the basis of their expertise and their ability to present evidence on the safety, efficacy, effectiveness, and service requirements, as appropriate, of the technology in question (Goodman, 1988). OMAR recommends that the planning committee suggest speakers to present opposing data and interpretations where controversy exists (OMAR, NIH, 1988).

The planning and implementation of a consensus development conference usually involves 12-14 months of work after topic selection. Observers have noted that the consensus development conference borrows aspects from the scientific meeting, the judicial process, and the town meeting (Jacoby and Rose, 1986; Mullan and Jacoby, 1985). Consensus development conferences are open meet

ings to which members of the public and the medical community are invited. The conference begins on a Sunday night with a closed session of the conference panel. During this session, panelists meet with OMAR staff to clarify questions about the consensus development conference. Several months before the conference, panelists divide into subcommittees to focus on particular consensus questions so that different panel members are responsible for writing specific portions of the consensus statement. All panelists remain responsible for the statement as a whole and are to follow all presentations and deliberations in the consensus development process. Panelists may meet as subcommittees on Sunday night. The following one and a half days are devoted to the plenary session for the expert presentations (15-30 minutes each, interspersed with open discussions involving speakers, panelists, and questions from the audience).

In the evenings, the panel convenes to draft consensus answers to the predetermined questions, considering the expert opinions of the conference speakers and other views expressed at the meeting. The panelists and chairperson for each conference decide on the rules for consensus in the executive sessions. The consensus view of the panel is not necessarily that of all panelists. If a panel cannot achieve full agreement on a particular point, the consensus statement may identify opposing or alternative opinions and/or majority-minority viewpoints. Few conferences have produced minority statements.

The chair reads the consensus statement 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 revise the statement based on comments received during this session. The conference concludes with a press conference.

DOCUMENTATION AND USE OF EVIDENCE IN CONSENSUS DEVELOPMENT

Panelists receive abstracts of the speakers' presentations at least one month in advance of the actual conference date, 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 staff of the National Library of Medicine 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. Speakers and members of the audience receive conference materials, including speakers' presentation summaries, a conference agenda, and logistical information, at the first public meeting of the panel. The consensus statement's validity rests in part upon the scientific evidence presented to the panel for decision making.

In general, the evidence presented prior to and during the conference is not formally weighted for integration in the consensus development process. A few conferences have employed decision-assist models to help the panel explore the implications of the data presented by speakers. The extent of the panelists' dependence upon such models in decision making has varied among the conferences (Jacoby and Pauker, 1986). Decision analysis may be used to provide a model of the sequence of potential strategies and outcomes for the questions at hand. The data available are used to structure the pathways in terms of probable occurrence and to compute utilities for the alternative outcomes. Strategies may be tested under different assumptions of risk and utility values. The decision-assist models are intended to "help structure complex alternatives in a rational way [and to allow] the incorporation of expertise and information from a variety of expert consultants without abdicating the decision to any one of them" (McNeil and Pauker, 1984).

There are no explicit rules for the consideration of evidence by the panel. The panels for each conference decide upon the definition of consensus. The chairperson directs and leads the discussion and consideration of evidence in a format acceptable to the panel.

The final product of the consensus development process is the consensus statement. Consensus statements do not cite articles used as evidence. Each consensus statement contains a list of the members of the program planning committee, the expert speakers, and the panelists.

DISSEMINATION AND IMPACT

The OMAR Director of Communications and the BID Information Officer develop a plan to announce the conference and to disseminate the consensus statements. Conferences usually receive

considerable attention from the medical media and general media at the time of their occurrence. The plan may include a combination of the following.

- 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.
- The *Journal of the American Medical Association* routinely publishes most of the consensus statements. Consensus statements are also published by specialty journals in the topic area.
- OMAR places notices in numerous professional journals to announce future conferences as well as the availability of consensus statements.
- The publication of the consensus statement along with selected papers from a CDC [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.
- Summary videotapes and audiotapes of the conference may also be prepared and distributed.
- A summary of the statement is also prepared and sent to appropriate specialty journals (OMAR, NIH, 1988).

Two consensus development conferences, on prostate cancer (1987) and urinary incontinence (1989), were televised live via satellite throughout the nation.

The consensus development program at NIH has undergone formal evaluation. The process continues to evolve in an attempt to improve health care practices in the United States. The procedures for conference planning, formulation of questions, and report dissemination have become more standardized, as have the formats for conducting the conferences and the final consensus statements.

Winkler et al. (1986) studied the dissemination of consensus development conference information in the popular press and concluded that the reports of the NIH conferences appear to be factual and balanced. Direct mailing seems to augment the success of information transfer to targeted groups. A variety of sources is needed to inform practitioners and the public (Jacoby and Clark, 1986).

OMAR conducted a survey to measure the effectiveness of two conferences (computed tomography scan of the brain, 1981; hipjoint replacement, 1982), as evidenced by the extent to which physicians were aware of the conferences and the conclusions reached at each one. The surveys determined that awareness varied greatly among different specialties. The study concluded that there was much room for improvement in information dissemination and that it would be fruitful to examine physicians' information-seeking habits, such as examining the role of opinion leaders, so as to better design strategies that more effectively disseminate conference results (Jacoby, 1983).

In another study by the RAND Corporation to evaluate the assessment program, Kanouse et al. (1987) investigated the impact of the consensus development conferences on physician awareness and behavior. They found that efforts to reach the practicing community can be improved. The consensus development program is more successful at reaching specialists than generalists. In addition, physicians who frequently participate in continuing medical education are more likely to have heard of consensus development conference recommendations. The physicians surveyed knew more about the content of specific consensus findings than about the program as a whole and the consensus development conference process. The physicians found information on clinical practice, in summary form accompanied by evidence, to be most useful.

The group from the RAND Corporation also investigated medical records in the state of Washington to determine whether the quality of care improved with respect to 12 recommendations by four consensus development conference panels. The results indicated that the conferences did not affect clinical practice in the manner intended. In some instances, physicians had adopted the consensus recommendations on patterns of care before the conference or no change was shown in comparisons of patterns before and after the conference or other constraints determined clinicians' practices, such as a lack of available resources (Kosecoff et al., 1987).

In early 1990, the Institute of Medicine (IOM) completed an evaluation of the NIH Consensus Development Program, conducted at the request of NIH. Although not charged with evaluating the program's dissemination activities and impact, the IOM did address program purpose and scope, role and placement of the program within NIH, aspects of the consensus development process, and financial support for the program (Institute of Medicine, 1990).

References

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

Institute of Medicine. 1990. Consensus Development at the NIH: Improving the Program . 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., and S.M. 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. Pauker. 1986. Technology assessment in health care: Group process and decision theory. Israeli 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.

Kanouse, D.E., R.H. Brook, J.D. Winkler, J. Kosecoff, S.H. Berry, G.M. Carter, J.P. Kahan, L. McKloskey, W.H. Rogers, C.M. Winslow, G.M. Anderson, L. Brodsley, A. Fink, and L. Meredith. 1987. Changing medical practice through technology assessment: An evaluation of the NIH consensus development program. Santa Monica, Calif.: RAND Corporation.

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.

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: The maturation of consensus conferences. Journal of the American Medical Association 254(8):1068-1072.

Office of Medical Applications of Research, National Institutes of Health. 1988. Guidelines for the selection and management of consensus development conferences. Bethesda, Md. Pamphlet.

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.

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.

RECOMMENDATIONS FOR STRENGTHENING CONSENSUS DEVELOPMENT FOR ASSESSING HEALTH TECHNOLOGIES

The following recommendations concern strengthening the use of consensus development for assessing health technologies. These recommendations were drafted by a writing group following the International Workshop on Consensus Development for Medical Technology Assessment, held in June 1989 at the King's Fund Centre in London, and coordinated by the Council on Health Care Technology, Institute of Medicine. The writing group, which included ten persons from five countries, drafted the recommendations based on the deliberations of the workshop. Members of the writing group included Gérard Breart, Clifford Goodman, Itzhak Jacoby, Egon Jonsson, Arnold Kaluzny, Pedro Koch, Jacqueline Kosecoff, Tore Scherstén, Jackie Spiby, and Caroline Weill.

The recommendations do not necessarily represent the views of the National Academy of Sciences or any of its constituent parts, the U.S. Department of Health and Human Services, or the organizations with which the authors are affiliated.

For the purpose of these recommendations, a consensus development program refers to an organizational entity that coordinates a series of consensus development conferences. The program may be a unit of, or sponsored by, one or more organizations. The conferences themselves are normally part of a broader process involving preconference planning and preparation and postconference activity such as dissemination of consensus statements. Each consensus

development conference normally addresses one or more related technologies by assessing a set of technological properties, concerns, impacts, or other aspects pertinent to the use and effects of these. Health technologies refer to the drugs, devices, procedures, and organizational and administrative systems used in providing or facilitating health care.

For the purpose of these recommendations, a consensus development process involves a group of experts or other representatives (i.e., a panel) that assesses a technology and formulates or prepares a set of findings that constitutes a consensus statement. The process of consensus development is based upon evidence provided to the panel in the form of literature and other documents, expert testimony, or other means. Although other group judgment processes may be conducted by remote panelists, consensus process panelists must have the opportunity for direct, face-to-face interaction in formulating and reviewing the panel's findings. These processes may be facilitated by staff, and a process could involve more than one panel, for example, a second group to provide an independent review of the main panel's work before a statement is adopted as policy.

RECOMMENDATIONS

1. Consensus development programs should be sponsored by organizations that have the ability to implement or effectively disseminate consensus findings.
2. Programs should adopt the goal of bringing about changes in health and medical practice and related policies of national health authorities, industry, payers, academic institutions, and other agents. Sponsors and panelists should be cognizant of the intended audience for the consensus findings and the intended means for disseminating the findings. The consensus program should identify the ways in which the program in general, and each conference in particular, are intended to effect change.
3. The consensus development program should describe specifically the scope of its concern related to technology assessment. This may include such technological aspects (i.e., properties, concerns, or impacts) as safety; efficacy, effectiveness, or patient out

comes; cost, cost-effectiveness, and related economic concerns; patient values or utilities; social, ethical, legal, and political concerns; guidelines for appropriateness, medical necessity, or quality; manner of application in research settings and in general or routine practice; and comparison of these among alternative technologies for specific clinical problems. These may vary from one conference to another. Programs are encouraged to assess all salient aspects of a technology; however, to the extent that a program's scope is necessarily more narrowly defined, consensus statements should note which important aspects are not addressed, and should note why these ought to be addressed by others.

4. Programs should seek to conduct assessments and provide recommendations in a timely fashion. On one hand, this requires consideration of the time needed to effectively plan, conduct, and report on conferences. On the other hand, consideration must be given to the pace at which technologies are developed and diffused and the need to provide timely guidance regarding their use.
5. The procedures and criteria for selecting conference topics and panel members should be documented.
6. The questions to be addressed by the consensus panel should be specific and manageable, that is, commensurate with the available evidence, the time available for the process, and other resources.
7. Panelists should represent the relevant health and medical professionals, methodologists such as epidemiologists and biostatisticians, economists, administrators, patient or other consumer representatives, and others who can provide relevant perspectives. Panel chairpersons and consensus program staff should be recognized as objective with regard to consensus topics and skilled in group processes.
8. Consensus development programs should seek the best available scientific evidence concerning the safety, efficacy, effectiveness, and other pertinent aspects of the technologies to be assessed. Clinical information should be made available to panelists about the use of procedures not only in research settings but in general or routine practice as well. Data from general practice are often not available in the literature and may have to be gathered by such means as review of patient records. For aspects for which evidence is not commonly available (e.g., social, ethical, or legal issues or values, preferences, or other patient, family, or community perspec

tives), programs may have to make special efforts to acquire or elicit reliable information.

9. Prior to a consensus development conference, programs should provide an ordered and categorized compilation or synthesis of research reports and related evidence concerning the technological aspects at issue. Sources may include reports of laboratory or bench studies, randomized clinical trials, epidemiologic or other observational studies, patient record audits, patient surveys, qualitative literature reviews, and individual expert opinion. Published literature as well as relevant unpublished surveys, trials, and data should be included. Source documents should be clearly cited and should be graded or otherwise rated for the strength or magnitude of their findings and for the rigor or quality of their methodologic approach. Copies of original documents should be available to panelists as needed.

   When resources and time permit, a meta-analysis of applicable data should be provided. Meta-analysis is a statistical method for obtaining quantitative answers to specific questions from multiple reports of primary studies on a particular subject. Using data 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.

10. All panelists should be involved in interpreting evidence, even though they will have varying levels of expertise for doing so. Consensus development programs should make available basic guidance or training concerning the use of evidence when needed to ensure that all panelists have at least a fundamental understanding of the role of evidence in the process.

11. The conduct of consensus development processes should be structured and documented. In particular, processes should provide for the following.

    a. In advance of the formal consensus conference, panelists should develop an organized compilation of points to be addressed at the conference, rather than relying on the conference alone to formulate these.

    b. An operational definition of consensus should be specified (e.g., full agreement, majority agreement) as well as how to present less than full agreement in the panel's findings (e.g., by citing minority opinions). This should be made known to panelists and other participants before a consensus conference is undertaken. Consen

sus may be determined by, for example, particular voting requirements or other defined rating mechanisms.

c. Evidence, including meta-analysis or other ordered information syntheses, expert testimony, etc., should be presented in a form that is comprehensible to all panelists, including those who may not be quantitative experts or specialists. Evidence should be presented in a manner that is consistent with the order and nature of the issues or questions to be considered and in a timely manner that allows sufficient opportunity for panelists' thorough review.

d. The consensus development conferences should be structured so that participants have the opportunity to contribute equally to the process. This is important in instances in which panelists represent multiple disciplines and may not be familiar with others' perspectives on consensus topics. In particular, a structured process is called for in the voting or other delineated means used in rendering a final set of consensus findings.

e. The duration and spacing of consensus conference sessions should be such that deliberations are not unduly affected by panelists' experiencing fatigue or lack of attention.

12. Each consensus statement should include the following:

a. a general description of the consensus development process or approach used;

b. notations regarding the strength of agreement or assurance of the panel concerning its findings;

c. a description of the reasoning used by the panel and the evidential basis for the consensus findings; statements should include summary descriptions of the conference source documents (i.e., at least a bibliography, preferably accompanied with annotations concerning topic, methodology, and findings of cited studies); and

d. recommendations for research needed to resolve those issues concerning which panels could not reach agreement and to otherwise advance understanding of the topics.

This documentation of process, evidence, and needed research should enhance internal consistency of the statement, enable users to follow the reasoning of the panelists, convey the context and applicability of the findings, and provide the basis for reassessment in light of new developments.

13. The consensus development program should provide for timely review, for example, in the form of peer review or an oversight committee established by the sponsoring organization, of the con

sensus development conference findings by experts who are not panelists for the particular conference that generated the findings. The purpose of this review would be to ensure that the questions posed to the panel were adequately addressed and that the findings were reasonably supported and not inconsistent or otherwise erroneous with respect to the evidence.

14. Programs should consider alternative modes of information dissemination to increase the effectiveness of inducing change. These efforts may include appropriate marketing approaches such as development of information products that best convey consensus findings to specific target audiences of medical professionals, patients, payers, industry, and others and effective means of calling attention to the need for more definitive research to resolve issues for which well-founded consensus could not be achieved.
15. Consensus development programs should provide for monitoring new developments that may overturn or significantly alter the available evidence pertaining to a technology that has been assessed in previous consensus development conferences, so that the program may call for a reassessment as appropriate. Such developments may include technological advances, reports of new clinical trials, and changes in the way technologies are applied in practice that call into question their safety or effectiveness.
16. The program should provide for formal evaluation of the impacts of the program and specific conferences, such as by acquiring data that would measure changes in practice behavior or contributions to policy initiatives, or demonstrate other impacts. This should include evaluation by parties that are independent of the program.
17. In order to improve consensus development processes, studies should be conducted that answer the following questions.

    a. How effective are consensus development processes in changing health and medical practice behavior?
    b. If consensus development processes are effective, what conditions or factors (e.g., identity of sponsoring organization, timing of conferences with respect to technologies' diffusion, documentation of panel's reasoning, dissemination strategies) contribute to their effectiveness?
    c. If consensus development processes are not effective in changing health and medical practice behavior, do they confirm prevailing practices?
    d. Can decision support resources, such as decision analysis or

particular decision-making techniques, strengthen consensus development processes?

e. Do consensus findings reflect increased understanding and convergence of the opinions of panelists or agreement only on the "least common denominator" (i.e., least controversial and most commonly accepted issues)?

INTERNATIONAL CONSENSUS DEVELOPMENT BIBLIOGRAPHY

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

Andreasen, P.B. 1987. La technologie médicale dans un petit pays au système de protection sociale avancé, le cas du Danemark. In Systèmes de santé, pouvoirs publics et financeurs: Qui controle quoi? La Documentation Française 149-154.

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.

Backe, B., and P. Nafstad. 1989. Development in practice of diagnostic ultrasound in Norway after the consensus conference recommending routine screening (abstract). Paper presented at the Fifth Annual Meeting of the International Society of Technology Assessment in Health Care, London.

Battista, R.N., and S.W. Fletcher. 1987. Making recommendations on preventive practices: Methodologic issues. Paper prepared for the International Symposium on Preventive Services in Primary Care: Issues and Strategies, October 4-7, Quebec, Canada.

Battista, R.N., and S.W. Fletcher. 1988. Making recommendations on preventive practices: Methodological issues. American Journal of Preventive Medicine 4(4)(Suppl.):53-67.

Battista, R.N., and R.S. Lawrence. 1988. Implementing preventive services--Introduction. American Journal of Preventive Medicine 4(4)(Suppl.):1-2.

Burke, J.F., H.S. Jordan, C.M. Boyle, and E. Vanner. 1981. An impact study of the consensus development conference on supportive therapy in burn care. Report 0-341-132/3718. Boston, Mass.: Health Research Institute, Inc., and Washington, D.C.: U.S. Government Printing Office.

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.

Canadian Task Force on the Periodic Health Examination. 1979. The periodic health examination. Canadian Medical Association Journal 121:1193-1254.

Canadian Task Force on the Periodic Health Examination. 1984. The periodic health examination: 2. 1984 Update. Canadian Medical Association Journal 130:1278-1285.

Canadian Task Force on the Periodic Health Examination. 1986. The periodic health examination: 2. 1985 Update. Canadian Medical Association Journal 135:724-727.

Canadian Task Force on the Periodic Health Examination. 1988. The periodic health examination: 2. 1987 Update. Canadian Medical Association Journal 138:618-626.

Canadian Task Force on the Periodic Health Examination. 1989. The periodic health examination: 2. 1989 Update. Canadian Medical Association Journal 141:209-216.

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

Casparie, A.F., and J.J.E. van Everdingen. 1985. Consensus development in The Netherlands. European Newsletter on Quality Assurance 2:1, 5.

Casparie, A.F., N.S. Klazinga, J.J.E. van Everdingen, and P.P. Touw. 1987. Health-care providers resolve clinical controversies: The Dutch consensus approach. Australian Clinical Review 7(24):43-47.

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

Forsythe, M. 1985. Personal view. British Medical Journal 290:63.

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

Greer, A. 1988. The state of the art versus the state of the science. International Journal of Technology Assessment in Health Care 4:5-26.

Henig, R.M. 1985. The process and provocations of consensus. Medical World News 14 (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.

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

Institute of Medicine. 1990. Consensus Development at the NIH: Improving the Program. 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., and S.M. 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. Pauker. 1986. Technology assessment in health care: Group process and decision theory. Israeli 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.

Jennett, B. 1985. First consensus development conference in the United Kingdom: On coronary artery bypass grafting, commentary by the chairman 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, 1979-83. 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 NIH consensus development program. Santa Monica, Calif.: RAND Corporation.

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

Klazinga, N.S., A.F. Casparie, and J.J.E. van Everdingen. 1988. Quality assurance on the crossroad of medical decisionmaking and management of the health care process in the hospital (abstract) . Theoretical Surgery 3:42-43.

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., and R.A. Potish. 1988. The consensus statements on adjuvant chemotherapy in breast cancer: Fact or artifact. American Journal of Oncology 11:73-76.

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

Lomas, J., G.M. Anderson, M.W. Enkin, E. Vayda, R. Roberts, and B. MacKinnon. 1988. The role of evidence in the consensus exercise, results from a Canadian consensus exercise. Journal of the American Medical Association 259(20):3001-3005.

Lomas, J., G.M. Anderson, K. Domnick-Pierre, E. Vayda, M.W. Enkin, and W.J. Hannah. 1989. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. New England Journal of Medicine 321:1306-1311.

Lomas J., M.W. Enkin, G.M. Anderson, K. Domnick-Pierre, E. Vayda, and W. Hannah. 1989. The role of consensus statement in changing physician's awareness and attitudes, knowledge and behavior. Paper presented at the Fifth Annual Meeting of the International Society of Technology Assessment and Health Care, London.

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: The maturation of consensus conferences. Journal of the American Medical Association 254(8):1068-1072.

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 non-consensus conferences on coronary heart disease. Lancet i:1087-1089.

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 4:481-484.

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.

Rogers, E.M., J.K. Larsen, and C.U. Lowe. 1982. The consensus development process of medical technologies: A cross-cultural comparison of Sweden and the United States. Journal of the American Medical Association 248:1880-1882.

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.

Stocking, B. 1986. Strategies for technology assessment and implementation in some European countries. International Journal of Technology Assessment in Health Care 2:19-26.

van Everdingen, J.J.E., and A.F. Casparie. 1988. Consensus development conferences on osteoporosis (letter to the editor). British Medical Journal 296:61-62.

van Everdingen, J.J.E. 1988. Consensusontwikkeling in de Geneeskunde. Thesis (English summary). Utrecht, Bonn: Scheltema & Holkema.

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. Ann Arbor, Mich.: 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)(Fall):469-498.

LIST OF AUTHORS

Bjørn Backe
Researcher
Norwegian Institute of Hospital Research
Trondheim-NTH Norway

Sharon R. Baratz
Research Associate
Council on Health Care Technology
Institute of Medicine
National Academy of Sciences
Washington, DC USA

Renaldo N. Battista, M.D., Sc.D.
Director
Division of Clinical Epidemiology
Department of Medicine
The Montreal General Hospital
Montreal, Quebec Canada

Gérard Breart, M.D.
Director of Research Unit No. 149
Epidemiological Research on Mother and Child
Paris France

Robert H. Brook, M.D., Sc.D.
Professor of Medicine and Public Health
Center for Health Sciences
University of California, Los Angeles
Los Angeles, CA USA
and
Deputy Director
Health Program
The RAND Corporation
Santa Monica, CA USA

Anton F. Casparie, M.D., Ph.D.
Professor of Socio-Medical Science
Department of Health Policy and Management
Erasmus University, and
Chair of the Scientific Council
CBO (National Organization for Quality Assurance in Hospitals)
Utrecht The Netherlands

Ingemar Eckerlund, M.Sc.
The Swedish Planning and Rationalization Institute for the Health and Social Services
Stockholm Sweden

Clifford Goodman, Ph.D.
Director
Council on Health Care Technology
Institute of Medicine
National Academy of Sciences
Washington, DC USA

Stefan Håkansson, Ph.D.
The Swedish Planning and Rationalization Institute for the Health and Social Services
Stockholm Sweden

Itzhak Jacoby, M.D.
Director
Division of Health Services Administration
Department of Preventive Medicine and Biometry
Uniformed Services University of the Health Sciences
Bethesda, MD USA

Egon Jonsson, Ph.D.
Professor
The Karolinska Institute and
Director
Swedish Council on Technology Assessment in Health Care
Stockholm Sweden

Torben Jørgensen
Department Head
Danish Hospital Institute
Copenhagen Denmark

Arnold D. Kaluzny, Ph.D.
Department of Health Policy and Administration
School of Public Health
and
Research Associate
Health Services Research Center
University of North Carolina at Chapel Hill
Chapel Hill, NC USA

Anna-Liisa Kauppila, M.Sc.
Planner
The Academy of Finland
Helsinki Finland

Niek Klazinga, M.D.
Staff Member
CBO (National Organization for Quality Assurance in Hospitals)
Utrecht The Netherlands

Pedro Koch, M.D.
Medical Service
Swiss Federal Office of Social Insurance
Bern Switzerland

Jacqueline Kosecoff, Ph.D.
Executive Vice President
Value Health Sciences, Inc.
Santa Monica, CA USA
and
Adjunct Professor of Medicine and Public Health
Center for Health Sciences
University of California, Los Angeles
Los Angeles, CA USA

Jonathan Lomas
Associate Professor
Department of Clinical Epidemiology and Biostatistics
Health Sciences Center
McMaster University
Hamilton, Ontario Canada

Elizabeth A. McGlynn, Ph.D.
Health Policy Analyst
The RAND Corporation
Santa Monica, CA USA

Tore Scherstén, M.D.
Professor of Surgery
Department of Surgery
Sahlgrenska Hospital
Göteborg School of Medicine
Göteborg Sweden

Jackie Spiby, M.B., M.F.C.M.
Consultant in Public Health Medicine
North West Thames Regional Health Authority
London United Kingdom

J.J.E. van Everdingen, M.D., Ph.D.
Staff Member
CBO (National Organization for Quality Assurance in Hospitals)
Utrecht The Netherlands

Caroline Weill
Professor
National School of Public Health
Hôpital National
Saint Maurice France