Appendix F: Synthesizing and Assessing the Evidence and Determining Practice Policies

Syntheses of effectiveness research on clinical preventive services and clinical practice policies have been issued by a number of different organizations, including professional societies, government agencies, third-party payers, and private researchers. The specialty societies that have issued specific recommendations on prevention include the American College of Physicians, the American Academy of Pediatrics, the American Academy of Family Physicians, the American College of Obstetricians and Gynecologists, the American College of Radiology, and the American Medical Association, Other private organizations include the Rand Corporation, the American Cancer Society, the American Heart Association, and the Institute of Medicine, National Academy of Sciences.

Several United States government agencies have organized external panels to synthesize the evidence on preventive medicine, or completed their own reviews of the evidence, often with input from outside experts. The National Institutes of Health (NIH) in the Public Health Service (PHS) in the Department of Health and Human Services (USDHHS), the National Cancer Institute in the NIH, the National Heart, Lung, and Blood Institute in the NIH, the Congressional Office of Technology Assessment (OTA), the Agency for Health Care Policy and Research, the Centers for Disease Control and Prevention (CDC), and the Office of Disease Prevention and Health Promotion (ODPHP), all have been involved in efforts to synthesize and evaluate effectiveness information.

Although the process of synthesizing the evidence on clinical practice is currently characterized by a diversity of decentralized efforts, there has been tremendous growth in interest in the methods used to synthesize and evaluate the evidence and, in general, these methods are becoming more rigorous and sophisticated. The Institute of Medicine has provisionally identified several attributes of good practice guidelines (see table F-l—Provisional Documentation Checklist for Practice Guidelines).

To assess the state of knowledge about the effectiveness of clinical preventive services, OTA looked to those organizations whose methods most reflected the criteria outlined by the Institute of Medicine. The methods employed by three different organizations, which generally took a relatively rigorous and systematic approach to reviewing the evidence on the effectiveness of preventive services, are described below. These organizations are the Canadian Task Force on the Periodic Health Examination (CTFPHE), the US. Preventive Services Task Force (USPSTF), and the Immunization Practices Advisory Committee (ACIP) of the CDC.

The Canadian Task Force on the Periodic Health Examination

The Canadian Task Force on the Periodic Health Examination (CTFPHE) was established in 1976 to recommend periodic health assessments for Canadian residents (29). The landmark contribution of CTFPHE was their use of a rigorous set of criteria to evaluate the evidence for or against the effectiveness and efficacy of any preventive intervention (83). The explicit criteria used by CTFPHE to rate the evidence on

Table F-l—Institute of Medicine Provisional Documentation Checklist for Practice Guidelines

Attribute	Item
Validity	Projected health outcomes if guidelines are followed. Information required to evaluate outcomes. Projected rests if guidelines are followed, information required to evaluate costs. Description of data, methods, and assumptions used to make projections, Explicit description of the relationship between the scientific evidence and the guidelines and explanations for any differences between the guidelines and the evidence. Explanations for any important differences between the guidelines in question and those developed by others. Thorough literature review describing scientific research including sponsors, settings, methodologies, findings, and qualifications. Description of methodology for evacuating the scientific literature and the results. Explicit assessment of the quality, consistency, clarity, and strength of the scientific evidence. Description of methodology for using expert or group judgment as a basis for evacuating scientific evidence or, in the absence of evidence, reaching a consensus based on expert opinion. Explicit description of the strength of expert consensus. Description of procedures, participants, and findings of review by experts and others not involved in the original development process. Description of methods, settings, and results of any protests of the guidelines,
Reliability/ reproducibility y	Description of methods and results of testing (1) the reliability of the development method and (2) the reproducibility of the clinical decisions reached by users of the guidelines.
Clinical applicability	Specifications by age, sex, race, clinical diagnosis, and other factors of the populations to which a set of guidelines apply. Description and analysis of the scientific literature or expert consensus that forms the basis for statements about the age, sex, and other factors of the populations to which a set of guidelines apply.
Clinical flexibility	Description and analysis of the scientific literature or expert consensus that forms the basis for statements about major foreseeable exceptions to applications of the guidelines. Listing of the basic information to be provided to patients and the kinds of patient preferences that may be appropriately considered. Listing of the data needed to document exceptions based on clinical circumstances, patient preferences, or delivery system characteristics.
Clarity	Methods and results of any testing of readability, logic, or understanding.
Multidisciplinary process	Description of the parties involved in developing the guidelines, their credentials and interests, and the methods used to solicit their views or to arrive at group judgments. Description of the procedures used to subject guidelines to review and criticism by experts not involved in the original development process, with summary of results.
Scheduled review	Timetable and method for the scheduled review. Description of the basis for arriving at the timetable or specific date.

SOURCE: Institute of Medicine, ***Clinical Practice Guidelines, Directions fora New Program,*** Field, M.J. and Lohr, K.N. (eds), (Washington, DC: National Academy Press, 1990).

effectiveness are shown in table F-2, ranked from the most to least credible.[1]

Each CTFPHE recommendation was assigned a letter grade, indicating the quality of the evidence which supported the recommendation (e.g., "A' indicated good evidence supporting the inclusion of a service, "C" indicated the evidence was poor, and "E" indicated there was good evidence that the service should be excluded). In their initial 1979 report, the Canadian Task Force issued recommendations for preventive services related to 78 potentially preventable conditions. Since their first report, CTFPHE has issued a number of updates and additional evaluations; for example, in 1993 CTFPHE issued an update on cholesterol screening (39). CTFPHE is in the process of updating the majority of their recommendations made since the original 1979 report and these will be published in mid-1994 (82).

Table F-2—Qualit y of Evidence Criteria Used by t he U.S. Preventive Services Task Force and the Canadian Task Force

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

SOURCE: U.S. Preventive Services Task Force, Guide *to Clinical Preventive Services.' An Assessment of the Effectiveness of 169 Interventions* (Baltimore: Williams and Wilkins, 1989).

The US. Preventive Services Task Force

In 1984, the Office of Disease Prevention and Health Promotion (ODPHP), in the U.S. Department of Health and Human Services, recommended the formation of the U.S. Preventive Services Task Force (USPSTF), a non-Federal, multidisciplinary panel of prevention experts (83). A 20-member panel was established in 1985 and in 1989 USPSTF published guidelines for the use of 169 preventive interventions,[2] USPSTF is working with CTFPHE to update their recommendations, which are scheduled for release in 1994.

The USPSTF's 1989 recommendations were based on a comprehensive literature search and the methods used to evaluate each study were systematic and explicit. To be considered effective by the USPSTF, screening tests, such as those used in cancer screening, had to be able to detect the target condition earlier than would have been the case without screening and with sufficient accuracy to avoid producing large numbers of false-positive and false-negative results (where accuracy refers to the test sensitivity, specificity, and positive predictive value) (see box F-1 for definitions of these terms). In addition, the test had to be reliable, that is, able to produce the same results when repeated. Even if a test accurately and reliably detected the disease at an early stage, it was not considered effective unless its use led to a better clinical outcome than would have occurred otherwise. That is, the interventions which followed a positive diagnosis for a condition had to be effective in preventing or delaying progress of the disease.

The USPSTF also used an explicit approach for evaluating the quality of the scientific evidence concerning the effectiveness of an intervention, and they placed the greatest confidence in evidence from randomized clinical trials (see table F-2). When there were no well-designed studies that supported an intervention, the USPSTF would recommend interventions that had demonstrated consistent benefits in a large number of studies of weaker design.

In making recommendations, the USPSTF evaluated the degree of efficacy of an intervention, the burden of illness, and the potential for negative consequences associated with its widespread, routine

[1] Note that table F-2 shows the criteria now used by the USPSTF and the CTFPHE. They arc a slightly revised version of the original criteria used by the CTFPHE in 1979. Specifically, category II-1, ''evidence obtained from well-designed controlled trials without randomization,'' was absent in the original criteria.

[2] ODPHP provides staff support for USPSTF, including background research on specific topics (232).

Box F-l—Important Concepts for Determining the Efficacy of a Screening Test

Sensitivity: The proportion of persons with a condition who correctly test positive when screened.

Specificity: The proportion of persons without a condition who correctly test "negative" when screened.

False Positives: A person without the disease who tests positive for the disease.

False Negatives: A person with the disease who tests negative for the disease.

Positive Predictive **Value: The** proportions of people correctly labeled diseased by the test. The positive productive value increases as the prevalence of the target condition in the screened population increases.

Accuracy: The USPSTF uses the term accuracy to refer to the performance of a test in terms of its sensitivity, specificity, positive predictive value, and negative predictive value.

Reliability: The ability of a test to obtain the same result when repeated.

Incidence: The number of new occurrences of the event in a specified time for a given population.

Prevalence: The ratio of the total number of all individuals who have an attribute or disease at a particular time, or during a particular period, to the population at risk for having the attribute or diseases.

SOURCES: U.S. **Preventive** Services **Task Force,** *Guide to Clinical Preventive **Services*** **(Baltimore,** *MD:* Williams and Wilkins, **1989); Maxcy-Rosenau, Last, JM. ed. Public Health and Preventive Medicine, 12th Edition (Norwalk, CT: Appleton-Century-Crofts,** 1986).

use. These negative effects may have included discomfort and physical injury, invasiveness, inconvenience, a longer period of morbidity due to early detection, overtreatment of borderline abnormalities, and anxiety from being falsely, or correctly, labeled as having the condition. For some preventive services no recommendation was made because the evidence was inadequate to decide for or against the procedure. In these cases, clinicians were advised to use their judgment to guide the application of the intervention.

Finally, interventions were often recommended for selected high-risk groups even though there was no evidence of greater effectiveness in these individuals than in the general population. The USPSTF argued that this policy was based on the recognition that the absence of evidence of effectiveness does not rule out effectiveness and if, in fact, the intervention is effective, individuals at increased risk of developing the disease are most likely to benefit.

There are several potential limitations to the USPSTF's methods. In choosing which target conditions to evaluate, the USPSTF considered both the burden of suffering from the target condition and the potential for effectiveness, but not the magnitude of the reduction in morbidity and mortality. Ideally, decisions relating to the widespread promotion of a preventive intervention may depend not only on whether the intervention is effective, but the expected magnitude of the effect, For example, the USPSTF assessed the effectiveness of cervical cancer screening, but not how many years of life would be saved if every women was routinely screened for cervical cancer.

A second limitation of the USPSTF recommendations is that they focus on interventions performed by physicians. For example, smoking education programs were not evaluated, with the exception of physician advice about smoking cessation. Other types of health education programs, such as labor and delivery and sex education classes, were not considered. In addition, preventive dental services were given little consideration, except as something which physicians should encourage. Similarly, the USPSTF's report does not explicitly evaluate the role of nonphysician providers. Nurses, social workers, physician assistants, and other health care providers may be able to provide many of the services described as appropriate by the USPSTF with equal effectiveness, and probably at lower cost, than can primary care physicians (e.g., advice regarding smoking cessation, blood pressure measurement, cholesterol measurement).

The Immunization Practices Advisory Committee

The Immunization Practices Advisory Committee (ACIP), an advisory group established by the CDC, issues recommendations on the use of new and existing vaccines. Recommendations typically describe the populations which should receive the vaccine, a schedule for vaccinations, and vaccine precautions and contraindications.

The ACIP meets several times during a year to review the evidence about the benefits and risks of vaccines and then issues its recommendations, ACIP members are selected from nominations made by professional and academic societies and represent experts in relevant disciplines (e.g., epidemiology, microbiology, public health, immunology, and public health practice). Representatives of the Food and Drug Administration (FDA) and the NIH act as ex-officio members, and the ACIP has liaison representatives from professional and governmental organizations.[3] Draft policy statements and background information are prepared by the CDC staff prior to the meetings. An attempt is made to gather all relevant background material, including both published and unpublished studies, such as unpublished studies from the vaccine manufacturer and the FDA.

The vaccines evaluated by the ACIP are licensed by the FDA, which does its own assessment of vaccine efficacy. The Center for Biologics Evaluation and Research of the FDA grants licensure for use of vaccines based upon demonstration of safety and efficacy. The approval process is complex and typically involves several sequential phases of evaluation, including initial testing of the vaccine in a small number of persons to determine its safety and immunogenicity; administration of the vaccine to a larger number of persons to obtain further data on adverse effects and the immune response; and controlled field trials with sufficient study subjects to develop reasonable estimates of safety and efficacy (104). The efficacy of a vaccine is usually measured in terms of protection against clinical disease (104). Although the FDA has primary responsibility for determining the safety and efficacy of vaccines, they do not issue recommendations concerning vaccine use, although they do provide input into the recommendations issued by the ACIP.

[3] These organizations include the American Academy of Family Physicians, American Academy of Pediatrics, the American College of Physicians, the American Hospital Association, the American Medical Association the Canadian National Advisory Committee on Immunization, the Department of Defense, and the National Vaccine Program.