Developing an Evidence-Based *Guide to Community Preventive Services*—Methods

Peter A. Briss, MD, Stephanie Zaza, MD, MPH, Marguerite Pappaioanou, DVM, PhD, Jonathan Fielding, MD, MPH, MBA, Linda Wright-De Agüero, PhD, MPH, Benedict I. Truman, MD, MPH, David P. Hopkins, MD, MPH, Patricia Dolan Mullen, DrPH, Robert S. Thompson, MD, Steven H. Woolf, MD, MPH, Vilma G. Carande-Kulis, MS, PhD, Laurie Anderson, PhD, MPH, Alan R. Hinman, MD, MPH, David V. McQueen, ScD, Steven M. Teutsch, MD, MPH, Jeffrey R. Harris, MD, MPH, The Task Force on Community Preventive Services

Abstract: Systematic reviews and evidence-based recommendations are increasingly important for decision making in health and medicine. Over the past 20 years, information on the science of synthesizing research results has exploded. However, some approaches to systematic reviews of the effectiveness of clinical preventive services and medical care may be less appropriate for evaluating population-based interventions. Furthermore, methods for linking evidence to recommendations are less well developed than methods for synthesizing evidence.

The *Guide to Community Preventive Services: Systematic Reviews and Evidence-Based Recommendations* (the *Guide*) will evaluate and make recommendations on population-based and public health interventions. This paper provides an overview of the *Guide*'s process to systematically review evidence and translate that evidence into recommendations.

The *Guide* reviews evidence on effectiveness, the applicability of effectiveness data, (i.e., the extent to which available effectiveness data is thought to apply to additional populations and settings), the intervention's other effects (i.e., important side effects), economic impact, and barriers to implementation of interventions.

The steps for obtaining and evaluating evidence into recommendations involve: (1) forming multidisciplinary chapter development teams, (2) developing a conceptual approach to organizing, grouping, selecting and evaluating the interventions in each chapter; (3) selecting interventions to be evaluated; (4) searching for and retrieving evidence; (5) assessing the quality of and summarizing the body of evidence of effectiveness; (6) translating the body of evidence of effectiveness into recommendations; (7) considering information on evidence other than effectiveness; and (8) identifying and summarizing research gaps.

Systematic reviews of and evidence-based recommendations for population-health interventions are challenging and methods will continue to evolve. However, using an evidence-based approach to identify and recommend effective interventions directed at specific public health goals may reduce errors in how information is collected and interpreted, identify important gaps in current knowledge thus guiding further research, and enhance the *Guide* users' ability to assess whether recommendations are valid and prudent from their own perspectives. Over time, all of these advantages could help to increase agreement regarding appropriate community health strategies and help to increase their implementation.

Medical Subject Headings (MeSH): community health services, decision making; evidence-based medicine; systematic reviews; methods; population-based interventions; practice guidelines; preventive health services; public health practice; task force (Am J Prev Med 2000;18(1S):35–43)

Introduction

The *Guide to Community Preventive Services: Systematic Reviews and Evidence-Based Recommendations* (the *Guide*) is an initiative of the U.S. Department of Health and Human Services and is being developed by a 15-member, independent, nonfederal Task Force on Community Preventive Services (the Task Force) in cooperation with many public and private sector partners.[1] The Task Force is supported by staff of the Centers for Disease Control and Prevention (CDC) and others who are developing, disseminating, and implementing the *Guide*. The *Guide* will make specific recommendations on selected interventions defined as activities that prevent disease or injury or that promote health in a group of people. Preventive interventions for individuals are considered in the *Guide*'s sister publication, the *Guide to Clinical Preventive Services*,[2] and are not included in the *Guide to Community Preventive Services*. Interventions recommended in the *Guide* will address 15 major topic areas (i.e., chapters) selected by the Task Force.[3] Chapters are organized into the following three major sections: (1) changing risk behaviors (e.g., reducing tobacco product use or increasing levels of physical activity); (2) reducing specific diseases, injuries, or impairments (e.g., reducing the occurrence of vaccine-preventable diseases or cancer); and (3) addressing ecosystem and environmental challenges (e.g., reducing health disparities attributable to differences in socioeconomic status). *Guide* reviews[4,5] and recommendations[6,7] are expected to be released as they are completed over the next two years; they will also be collected and published in book form.

Systematic reviews and evidence-based recommendations are playing an increasingly important role in decision-making about health-related issues. Over the past 20 years, information on the science of synthesizing research results has exploded.[8,9] Much of the available work on research synthesis on health related topics has, however, focused on clinical preventive services (i.e., preventive practices applied to target conditions among asymptomatic individuals)[2] and medical care.[9–11] Furthermore, experience in linking evidence to practice recommendations exists[2,12] but is more limited. Therefore, the Task Force is working to refine available approaches to systematic reviews and evidence-based recommendations for population-based and public health interventions. This paper provides an overview of the process being used to review evidence and to translate that evidence into recommendations provided in the *Guide*. An example of a review and its resulting recommendations that were developed using these methods is shown elsewhere in this issue.[5,7]

Methods for Developing Reviews and Recommendations

The Task Force determined that recommendations in the *Guide* should be based on systematic reviews of evidence aimed at showing the relationship of the intervention to particular outcomes and an explicit process for translating the evidence into recommendations. In the *Guide*, the term *evidence* includes: (1) information that is appropriate for answering questions about an intervention's effectiveness; (2) the applicability of effectiveness data (i.e., the extent to which available effectiveness data is thought to apply to additional populations and settings); (3) the intervention's other effects (i.e., side effects, including important outcomes of the intervention not already included in the assessment of effectiveness whether they are harms or benefits, intended or not intended, and health or non-health outcomes); (4) economic impact; and (5) barriers that have been observed when implementing interventions. *Guide* recommendations are primarily based on evidence of effectiveness.

For the purposes of the *Guide*, evidence is derived generally from observation or experiment. Acceptable methods for gathering and evaluating evidence varies, based on the issue addressed. For example, the *Guide*'s process uses data from comparative studies—those that compare outcomes among a group exposed to the intervention versus outcomes in a concurrent or historical group that was not exposed or was less exposed—to answer questions about whether interventions are effective. However, it may use noncomparative studies to describe barriers and to collect economic information. Additional illustrations of approaches to evaluating different types of evidence are explored in

From the Division of Prevention Research and Analytic Methods, Epidemiology Program Office (Briss, Zaza, Wright-De Agüero, Truman, Hopkins, Carande-Kulis, Anderson, Harris), Centers for Disease Control and Prevention (CDC), Atlanta, Georgia; Office of Global Health (Pappaioanou), CDC, Atlanta, Georgia; Los Angeles Department of Health Services (Fielding), Los Angeles, California; University of Texas-Houston School of Public Health (Dolan Mullen), Houston, Texas; Department of Preventive Care, Group Health Cooperative of Puget Sound (Thompson), Seattle, Washington; Medical College of Virginia (Woolf), Fairfax, Virginia; Task Force for Child Survival and Development (Hinman), Atlanta, Georgia; National Center for Chronic Disease Prevention and Health Promotion (McQueen), CDC, Atlanta, Georgia; Merck & Co., Inc. (Teutsch), West Point, Pennsylvania

The names and affiliations of the Task Force members are listed on page v of this supplement and at http://www.thecommunityguide.org

Address correspondence and reprint requests to: Peter A. Briss, MD, Senior Scientist and Development Coordinator, Community Preventive Services Guide Development Activity, Epidemiology Program Office, MS K73, Centers for Disease Control and Prevention, 4770 Buford Highway MS-K73, Atlanta, GA 30341: E-mail: pxb5@cdc.gov.

Some of this material has been previously published in: Shefer A, Briss P, Rodewald L, et al. Improving immunization coverage rates: An evidence-based review of the literature. Epidemiologic Reviews 1999;20:96–142.

more detail later in this paper and elsewhere in this supplement.

The Task Force decided on the following steps to obtain and evaluate evidence and translate that evidence into recommendations: (1) form multidisciplinary chapter development teams; (2) develop a conceptual approach to organizing, grouping, and selecting the interventions evaluated in each chapter; (3) select interventions to be evaluated; (4) search for and retrieve evidence; (5) assess the quality of and summarize the body of evidence of effectiveness; (6) translate evidence of effectiveness into recommendations; (7) consider evidence other than effectiveness; and (8) identify and summarize research gaps.

Chapter Development Teams

Because of the broad and multidisciplinary character of many public health problems, the Task Force employs chapter development teams representing diverse perspectives. Approximately 4–10 individuals with methodologic or subject matter expertise lead the development of a chapter. An additional 15–20 subject matter experts, including practitioners, advise on chapter development. The broad experience of the team members is vital to: (1) ensure the usefulness and comprehensiveness of the conceptual approach to the chapter; (2) ensure knowledge of, and experience with, numerous types of interventions to increase the usefulness of the reviews of interventions ultimately selected; (3) reduce the likelihood that important information will be missed; and (4) reduce the likelihood of errors or biases in the interpretation of identified information.

Develop a Conceptual Approach to Organizing, Grouping, Selecting and Evaluating the Interventions Evaluated in Each Chapter

The breadth of each *Guide* chapter requires the chapter development team to identify key areas on which to focus. A logic framework is a diagram mapping out a chain of hypothesized causal relationships among determinants, intermediate, and health outcomes. The logic framework is used to identify links between social, environmental, and biological determinants and pertinent outcomes; strategic points for action; and interventions that might act on those points. Perhaps most important, logic frameworks provide a structure for chapter development team to describe the interventions that are available to reach specified public health goals and allow the team to determine which of the available options will be reviewed in the chapter. Examples of logic frameworks are shown elsewhere.[4,5,13] One example is shown elsewhere in this issue. Once interventions are chosen, a detailed analytic framework is developed for each one that shows hypothesized links between the intervention and the health and other effects. Analytic frameworks are essentially detailed analysis plans representing portions of the larger logic framework. Analytic frameworks map the plan for evaluating each intervention, thus guiding the team's search for evidence. Similar frameworks have also been used to guide other systematic reviews.[14]

To show evidence of effectiveness for the purposes of the *Guide,* empiric evidence must demonstrate that an intervention will improve health outcomes. This demonstration can be direct, i.e., water fluoridation could reduce the occurrence of dental caries. More often, the demonstration is indirect—increased tobacco prices could reduce tobacco use which could reduce morbidity and mortality. Where links between intermediate and health outcomes have been well-shown elsewhere (e.g., links between smoking and adverse health outcomes or between vaccination and reduced disease) this evidence can be referenced and the *Guide*'s search for evidence will focus only on the relationship of the interventions to the intermediate outcomes. An analytic framework makes these choices explicit.

Select Interventions to Be Evaluated

An intervention is characterized by what was done, how it was delivered, who was targeted, and where it was delivered. Interventions can be either single-component—using only one activity—or multicomponent—using more than one related activity. Because population-based interventions usually are heterogeneous, the chapter development team must make explicit judgments about the extent to which interventions will be considered in the same body of evidence. These judgments are based on characteristics of the intervention, depth of available literature, theory, and other considerations.

In making selections of types of interventions to assess within chapters, the teams consider the: (1) potential for reducing the burden of disease and injury; (2) potential for increasing healthy behaviors and reducing unhealthy behaviors; (3) potential to increase the implementation of effective interventions that are not widely used; (4) potential to phase out widely used less-effective interventions in favor of more-effective or more-cost-effective options; and (5) current level of interest among providers and decision makers. The perceived volume of available literature is not a criterion for selecting interventions. Interventions that meet one or more of the above criteria but that have not been well-studied should be systematically evaluated in order to document important gaps in current research. The process for selecting interventions is systematic but dependent on judgment; a different group of partici-

pants might choose a somewhat different set of interventions.

Systematically Search for and Retrieve Evidence

Analytic frameworks provide some of the inclusion criteria for identifying evidence by specifying the intervention(s) considered and the pertinent outcome(s). Other inclusion criteria are also specified (e.g., countries and years in which the study was conducted and languages in which it was communicated). Searches are performed for literature published in books and journals meeting the inclusion criteria and include searches of multiple computerized databases, reviews of reference lists, and consultation with experts. The necessity and the methods for identifying information not published or in other sources are considered on an intervention-by-intervention basis by the chapter development teams. Comprehensive searches are performed to reduce the chance that information supporting a particular conclusion will be preferentially identified while other information is missed.

Assess the Quality of and Summarize a Body of Evidence of Effectiveness

After the individual studies making up the body of evidence of effectiveness for an intervention are identified, they are evaluated, their results are extracted, the overall body of evidence is summarized, and the strength of the body of evidence (i.e., the confidence that changes in outcomes are attributable to the interventions) is assessed.

Each study that meets the explicit inclusion criteria is read by two reviewers who use a standardized abstraction form[15] to record information about: (1) the intervention being studied; (2) the context in which the study was done (e.g., population, setting); (3) the evaluation design; (4) study quality; and (5) the results. Any disagreements between the two reviewers are reconciled by consensus among the chapter development team during the process of summarizing results into evidence tables.

Each study is characterized based on both the suitability of study design for assessing effectiveness and the quality of study execution. Study designs are classified using a standard algorithm (Figure 1). Suitability of study design (Table 1) is characterized based on several characteristics that help to protect against a variety of potential threats to validity.

The *Guide*'s process requires that a study design include a concurrent or before after comparison for the study to be used to assess effectiveness. Knowing the extent of effectiveness in the intervention group is impossible without an assessment of the extent to which

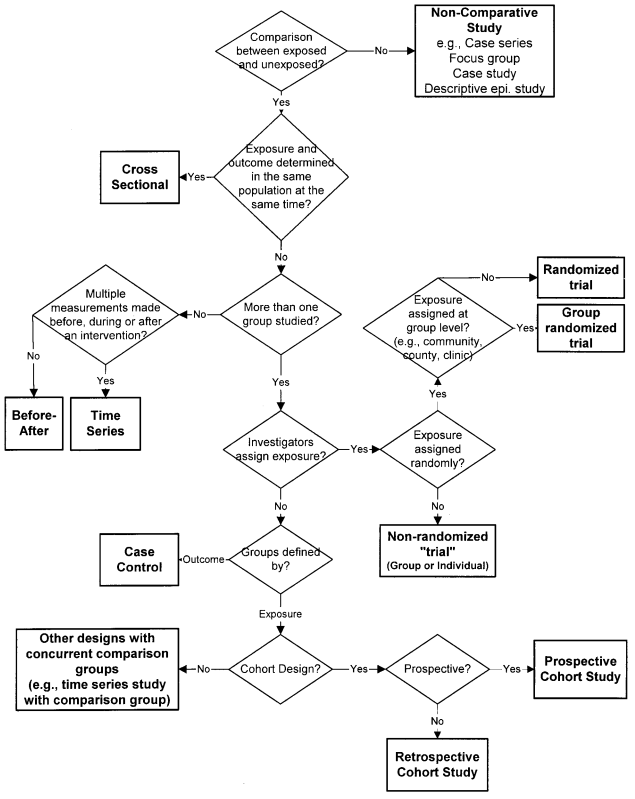

Figure 1. Classifying study design for the *Guide to Community Preventive Services*.

desired outcomes also occurred among persons unexposed to the intervention. Other characteristics of study design increase a study's suitability for assessing effectiveness. A study design with a concurrent comparison group protects against misinterpreting secular changes in outcomes that are not attributable to the intervention. To a lesser degree, studies with multiple outcome measurements made over time can also protect against such concurrent changes not attributable to the intervention. Study designs in which assessment of exposure precedes assessment of outcome protect against biased ascertainment of exposures.

Reviewers assess quality of study execution by considering six categories of threats to validity—study popu-

Table 1. Suitability of study design for assessing effectiveness in the *Guide to Community Preventive Services*

Suitability	Attributes
Greatest	Concurrent comparison groups *and* prospective measurement of exposure and outcome
Moderate	All retrospective designs *or* multiple pre or postmeasurements but no concurrent comparison group
Least	Single pre and postmeasurements and no concurrent comparison group *or* exposure and outcome measured in a single group at the same point in time

lation and intervention descriptions, sampling, exposure and outcome measurement, data analysis, interpretation of results (including follow-up, bias, and confounding), and other. Each category consists of several questions.[15] A total of 9 limitations are possible. Each study is categorized as having good, fair, or limited quality of execution based on the number of limitations noted, studies with 0–1, 2–4, and 5 or more limitations are categorized as having good, fair, and limited execution respectively. Studies with limited execution are not included in bodies of evidence to support recommendations. In general, information on quality of study execution is based only on information in published reports because bias could be introduced based on limited availability or variable quality of additional information from the authors and because collecting additional information from the authors may not be feasible.

Results across a group of related studies are summarized qualitatively and whenever possible are summarized using descriptive statistics such as the median and range or interquartile range of effect sizes. Depending on appropriateness and feasibility of a quantitative summary and the availability of statistical measures of variability or data from which to calculate them, formal procedures for statistical pooling might also be used to describe a summary measure of effect.

In addition, the body of evidence of effectiveness is characterized as strong, sufficient, or insufficient based on the number of available studies, the strength of their design and execution, and the size and consistency of reported effects (Table 2). Sufficient and strong evidence can be achieved in several ways that incorporate scientific rigor and the feasibility and appropriateness of evaluation for the wide range of interventions used in population-based approaches to improve health. The Task Force also retains the option of using expert opinion in rare circumstances when other evidence is not available and the intervention is deemed important enough and/or in widespread use that a recommendation must be made.

Several principles guided the designation of bodies of evidence of effectiveness as strong, sufficient, or insufficient evidence. Strong or sufficient evidence can be based either on a small number of studies with better execution and more suitable design or a larger number of studies with less suitable design or weaker execution (Table 2). For all designations of strong or sufficient evidence, study results must generally be consistent in direction and size. A single study could represent sufficient or strong evidence, but all other characteristics being comparable, a larger number of studies constitutes a stronger body of evidence because larger numbers of replications can help to reduce the likelihood that the results of individual studies are caused by chance, and may reduce the likelihood that the results of individual studies are due to bias if the potential flaws in the individual studies are not identical. In general, larger effect sizes (e.g., absolute or relative risks) are considered to represent stronger evidence of effectiveness than smaller effects. The Task Force makes judgments on the magnitude of effects on a case-by-case basis.

Translating Evidence of Effectiveness into Recommendations

Effectiveness

In general, strength of evidence of effectiveness (Table 2) links directly to strength of recommendation (Table 3). Evidence that is inconsistent in direction or size of effect based on definable characteristics of the population, setting, or the intervention should lead to separate recommendations for different situations. For example, some interventions could be recommended for urban populations but not for rural populations. Other interventions could be recommended in health department clinics but not in managed care organizations, and more-intensive interventions could be recommended over less-intensive ones. Insufficient or contradictory evidence, without evidence that the intervention is effective in one or more definable contexts, should lead to a determination that evidence is insufficient to assess effectiveness in any population. Documented ineffectiveness in some populations, without evidence that an intervention is effective in definable situations, would lead to recommendations *against* the intervention in all populations.

Evidence other than effectiveness is sometimes incorporated in Task Force recommendations and is routinely summarized for users.

Applicability

To help users determine the likelihood that available information will or will not apply to their local situations, chapter development teams: (1) define target populations and settings for which the intervention might be considered; (2) assess whether available studies have evaluated the intervention in those populations and settings; (3) assess the extent to which the populations or settings in those studies are likely to represent the target populations and settings of interest; and (4) make judgments about whether the intervention works better or worse in some populations and settings than in others. Based on that information, the Task Force will make a judgment about how widely the resulting recommendations should apply as well as identifying areas for further research.

Other Effects

Health interventions intended to influence one or more health outcomes can sometimes result in other effects (i.e., side effects, these include important outcomes of the intervention that are not already included in the assessment of effectiveness whether they are harms or benefits, intended or not intended, and relate to health or non-health outcomes). For example, an intervention that resulted in cleaner waterways might have recreational as well as health benefits. *Guide* reviews can define important other effects, and systematically search for and evaluate the strength of evidence supporting these using the same process as is used for effectiveness. If other effects are demonstrated, they will be discussed in the chapter. Documented harms that outweigh benefits will lead to recommendations that interventions not be used. Interventions for which important distributional considerations exist (i.e., effective in some populations but harmful to one or more "side populations") could lead to more narrowly targeted or less positive recommendations than would otherwise be made. Any available information on documented other effects will be shown; however, non-health other effects by themselves (without sufficient or strong evidence of improved health outcomes) will not generally be used to justify positive recommendations.

Economic Evaluations

Chapter development teams systematically search for available economic evaluations for recommended and strongly recommended interventions. When such studies are identified, they are assessed and abstracted using a standardized abstraction form[16] and, to the extent possible, standardized to the reference case recommended by the Panel on Cost-Effectiveness in Health and Medicine.[17,18] Standardization of results is useful because the design and reporting of economic evaluations varies greatly, which seriously limits the comparability of results across studies.

Resources for interventions to improve health are always constrained. Economic analyses can provide useful information to decision makers about the resources required for various interventions to improve health outcomes and could help those decision makers allocate their resources in a way that maximizes the health improvements achieved. Available economic information will be summarized in the *Guide* as a decision-making aid. However, the economic information will not routinely affect *Guide* recommendations because of limitations in the availability and quality of data and because different users will bring different values to bear regarding how and whether economic information should be incorporated into decision making.

Summarizing Barriers to Implementation of Interventions

Each intervention evaluated in a *Guide* review will include a discussion of applicable information on barriers that have been encountered when implementing interventions. This information is primarily included for decision-makers to consider when selecting inter-

Table 2. Assessing the strength of a body of evidence on effectiveness of population-based interventions in the *Guide to Community Preventive Services*

Evidence of effectiveness[a]	Execution—good or fair[b]	Design Suitability—Greatest, moderate, or least	Number of studies	Consistent[c]	Effect size[d]	Expert opinion[e]
Strong	Good	Greatest	At Least 2	Yes	Sufficient	Not Used
	Good	Greatest or Moderate	At Least 5	Yes	Sufficient	Not Used
	Good or Fair	Greatest	At Least 5	Yes	Sufficient	Not Used
	Meet Design, Execution, Number and Consistency Criteria for Sufficient But Not Strong Evidence				Large	Not Used
Sufficient	Good	Greatest	1	Not Applicable	Sufficient	Not Used
	Good or Fair	Greatest or Moderate	At Least 3	Yes	Sufficient	Not Used
	Good or Fair	Greatest, Moderate, or Least	At Least 5	Yes	Sufficient	Not Used
Expert Opinion	Varies	Varies	Varies	Varies	Sufficient	Supports a Recommendation
Insufficient[f]	A. Insufficient Designs or Execution		B. Too Few Studies	C. Inconsistent	D. Small	E. Not Used

[a]The categories are not mutually exclusive; a body of evidence meeting criteria for more than one of these should be categorized in the highest possible category.
[b]Studies with limited execution are not used to assess effectiveness.
[c]Generally consistent in direction and size.
[d]Sufficient and large effect sizes are defined on a case-by-case basis and are based on Task Force opinion.
[e]Expert opinion will not be routinely used in the *Guide* but can affect the classification of a body of evidence as shown.
[f]Reasons for determination that evidence is insufficient will be described as follows: A. Insufficient designs or executions, B. Too few studies, C. Inconsistent. D. Effect size too small, E. Expert opinion not used. These categories are not mutually exclusive and one or more of these will occur when a body of evidence fails to meet the criteria for strong or sufficient evidence.

Table 3. Relationship of strength of evidence of effectiveness and strength of recommendations

Strength of Evidence of Effectiveness	Recommendation
Strong	Strongly recommended
Sufficient	Recommended
Insufficient empirical information supplemented by expert opinion	Recommended based on expert opinion
Insufficient	Available studies do not provide sufficient evidence to assess
Sufficient or strong evidence of ineffectiveness or harm	Discouraged

ventions. Although the Guide is not intended to be a manual for intervention implementation, this section may also be useful to practitioners who must implement interventions. In general, information on barriers will not affect recommendations.

Summarizing Research Gaps

The systematic reviews in the *Guide* identify existing information on which to base public health recommendations. An important additional benefit of these reviews is the identification of areas where information is lacking or of poor quality.

Discussion

The *Guide* builds on considerable previous experience in systematic reviews[8–11] and somewhat less experience in linking evidence to recommendations.[2,19–21] The *Guide* shares with those processes a commitment to a systematic process and to explicitness.

Population-based interventions differ from individual interventions that have been the focus of many previous health-related systematic reviews because of the level of scale at which the interventions are implemented and at which outcomes are measured; in the greater number, variability, and complexity of available intervention options; and in the complexity and variability of the contexts in which the interventions are delivered. Because of these differences between individual and population-based interventions, the Task Force has adapted existing methods for assessing evidence and linking evidence to recommendations.

Randomization is an effective tool for controlling confounding in research, and thus, randomized clinical trials and group-randomized trials are classified as two of several designs with the greatest suitability for evaluating effectiveness in the *Guide*. However, randomization is sometimes not feasible or ethical in population-based research. Also, group-randomized trials often cannot feasibly randomize sufficient numbers of units to ensure even distributions of potential confounders among groups[22] and, for that reason, might lack some of the advantages over alternative strategies for controlling confounding that randomized clinical trials have in studies of clinical efficacy and effectiveness.

The *Guide*'s approach to suitability of study design allows studies to collect either individual or ecologic data. This choice was made because many community interventions are applied across populations and could be difficult, impossible, or inappropriate to study with individual-level data. In addition, individual- and population-level data have different strengths and limitations and might provide information that is complementary. Studies using ecological data are sometimes limited by attenuated effects or difficulties controlling confounding,[23] when applicable, these limitations are considered under quality of study execution.

The *Guide* assesses the quality of study execution in detail and considers it along with study design. This allows, e.g., a well-conducted case control or prospective cohort study to receive greater weight than a poorly-conducted randomized trial. Many different methods of measuring quality of study execution exist.[24–26] Methods for assessing quality differ in the characteristics considered (e.g., measures of internal, external, statistical, and construct validity) and in the way that these characteristics are measured. The impact of various approaches to quality on reported results of studies[25] and on the results of meta-analyses are areas of considerable current research interest; at present there is no gold standard. Methods for addressing quality of non-randomized studies (probably the bulk of studies that will be addressed in the *Guide*) are perhaps even less developed than those for randomized studies. For the *Guide*, we have developed an approach to study execution that reflects methodologies from other systematic reviews; reporting standards established by major health and social science journals; the evaluation, statistical and meta-analytic literature; and expert opinion and review. The face and content validity of the approach are strengthened by the method of its development. Nonetheless, this approach is one among many reasonable choices and we expect that approaches to assessing quality of study execution will continue to develop.

The Task Force has balanced several objectives in choosing its methods: (1) to obtain and use the best available empirical evidence to support decision making; (2) to set standards that will improve the availability and quality of evidence over time; (3) to make recommendations without requiring unobtainable data quality; (4) to balance the need for a consistent approach throughout the *Guide* with a need to have an evaluation approach that is appropriate and feasible across subjects; and (5) to cope with constraints on time and resources.

To develop a systematic, consistent process, a num-

ber of methodologic decisions were made. For example: (1) which issues to consider in the systematic reviews; (2) what evidence is appropriate to address those issues (e.g., what are better or worse study designs and executions for assessing effectiveness, for collecting economic data, or for describing barriers to intervention implementation); (3) how to judge study quality; (4) how to distinguish insufficient, sufficient, and strong evidence; and (5) how to relate evidence to recommendations (e.g., how to weigh effectiveness versus harms, costs, or barriers).

The Task Force believes that the rationale for its choices is sound and the use of the process results in conclusions that are reasonable and defensible. Nonetheless, implementation and evaluation of population-based interventions to improve health and prevent disease are dynamic areas undergoing rapid development. Methods used in the *Guide* to explicitly and reliably evaluate population-based interventions have evolved over the *Guide*'s history and will continue to evolve to keep up with developments in the field.

Performing comprehensive searches for information is time-consuming and difficult. Presently, no easily accessible, searchable, and well-indexed sources of information exist for finding information on population-based interventions. Existing reviews and recommendations are usually good sources of references but otherwise can be hard to incorporate in the *Guide*'s process because many review methods are not explicit and because of differences in inclusion criteria, outcomes of interest, or purposes. A complete body of evidence should ideally include information published in books and journals as well as other information (e.g., government reports, conference proceedings, and graduate school theses and dissertations). Because information published in journals is more likely than other information sources to show positive effects[27] limiting a review to published data could overestimate effects (i.e., publication bias). However, the ability to identify and include other sources of information must be weighed against practical constraints on finding and abstracting such information and the possibility that some types of information could be of lower quality than that published in journals. Finally, the *Guide*'s process primarily included English-language literature and could result in missing some applicable information.

Defining and categorizing interventions for the purpose of evaluation raises a number of difficult issues. Grouping and summarizing across heterogeneous interventions unavoidably results in the loss of some detail regarding the characteristics of the interventions and the context in which they are carried out. This challenge is multiplied for multicomponent interventions. Because of this, the *Guide* is not intended to be a manual for intervention implementation. However, judicious summaries of evidence regarding similar interventions will allow description of the types of activities that are being undertaken and an informed consideration of their advantages and disadvantages. Further, an enhanced understanding of the extent to which interventions actually achieve desired outcomes is possible. The ability to compare and contrast experiences should lead to improved decision making about which interventions to implement.

The *Guide* should not be viewed as the sole source for informed public health decision making because local contextual information is also important. Many issues not addressed in the *Guide* will affect which interventions are implemented (e.g., resource availability, social justice, community participation, cultural appropriateness, local burden of diseases and risk factors, and political considerations). However, the *Guide* provides systematically collected and detailed information on several issues of importance to public health practitioners and decision makers; information which is difficult or inefficient to develop locally. *Guide* reviews and recommendations will be most useful in conjunction with a participatory community planning process that clarifies needs and goals and that considers the *Guide*'s evidence reviews and recommendations in conjunction with additional applicable community-specific information.

Evidence-based recommendations are not without potential drawbacks. Evidence-based approaches could result in difficulties in making recommendations because of a lack of available information, problems with quality of available information, important outcomes that are difficult to measure or are far in the future, or emerging approaches that have not yet been adequately tested. Alternatively, evidence-based approaches could result in interventions being recommended too broadly or on the basis of too little information. In spite of the theoretical appeal of evidence-based methods, the question of whether and to what extent evidence-based methods result in changes in professional practice or improvement in health outcomes relative to other approaches to guideline formation is still open.

The Task Force concludes that the advantages of evidence-based methods exceed the potential drawbacks. Systematic and participatory processes will reduce errors in how information is collected and interpreted and reduce the likelihood that recommendations reflect only selected information or a limited point of view. Providing a clear analytic rationale for recommending certain interventions will enhance the ability of *Guide* users to assess whether recommendations are valid and prudent from their own perspectives, whether recommendations make sense in their local contexts, and whether the recommendations are likely to achieve goals of importance to them. Finally, evidence-based reviews explicitly show any limitations and uncertainties in available data, thereby creating opportunities to improve the quality of research and to

stimulate research that will close important gaps. Over time, all of these advantages could help to increase agreement regarding appropriate community health strategies and help to increase their implementation.

References

1. Pappaioanou M, Evans C. Development of the Guide to Community Preventive Services: a U.S. Public Health Service initiative. J Public Health Mgmt Practice 1998;4(suppl 2):48–54.
2. US Preventive Services Task Force. Guide to clinical preventive services. 2nd ed. Alexandria, VA: International Medical Publishing, 1996.
3. Zaza S, Lawrence RS, Mahan CS, et al. and the Task Force on Community Preventive Services. Scope and organization of the Guide to Community Preventive Services. Am J Prev Med 2000;18(suppl 1):27–34.
4. Shefer A, Briss P, Rodewald L, Bernier R, Strikas R, Yusuf H, et al. Improving immunization coverage rates: an evidence-based review of the literature. Epidemiol Rev 1999;20:96–142.
5. Briss PA, Rodewald L, Hinman A, et al. and the Task Force on Community Preventive Services. Reviews of evidence for interventions to improve vaccination coverage in children, adolescents and adults. Am J Prev Med 2000;18(suppl 1):97–140.
6. Task Force on Community Preventive Services. Vaccine-preventable diseases: improving vaccination coverage in children, adolescents and adults. A report on recommendations of the Task Force on Community Preventive Services. MMWR 1999;48(RR-8):1–16.
7. Task Force on Community Preventive Services. Recommendations on interventions to improve vaccination coverage in children, adolescents and adults. Am J Prev Med 2000;18(suppl 1):92–96.
8. Cooper H, Hedges LV, eds. The handbook of research synthesis. New York: Russell Sage Foundation, 1994.
9. Mulrow CD, Oxman AD, eds. Cochrane collaboration handbook. In: The Cochrane library [CDROM, updated September 1997]. The Cochrane Collaboration. Oxford: Update Software, 1997:Issue 4.
10. The Cochrane Collaboration. The Cochrane database of systematic reviews, 1999 vol 2 [online database]. Available at http://www.updateusa.com/clibip./clib.htm. Accessed July 22, 1999.
11. Agency for Health Care Policy and Research, Rockville MD. Evidence-based practice centers: overview, December 1998. Available at: http://www.ahcpr.gov/clinic/epc/. Accessed July 22, 1999.
12. CDC Guidelines: Improving the Quality. 1996. Centers for Disease Control and Prevention, Atlanta, GA.
13. The Task Force on Community Preventive Services. The guide to community preventive services. Available at: http://web.health.gov/communityguide. Accessed July 22, 1999.
14. Woolf SH. An organized analytic framework for practice guideline development: using the analytic logic as a guide for reviewing evidence, developing recommendations, and explaining the rationale. In: McCormick KA, Moore SR, Siegel RA, eds. Clinical practice guideline development: methodology perspectives. Washington, DC: US Dept. of Health and Human Services, Agency for Health Care Policy and Research, 1994:105–13.
15. Zaza S, Wright-De Agüero LK, Briss PA, et al. Data collection instrument and procedure for systematic reviews in the Guide to Community Preventive Services. Am J Prev Med 2000;(suppl 1):44–74.
16. Carande-Kulis VG, Maciosek MV, Briss PA, et al. and the Task Force on Community Preventive Services. Methods for systematic review of economic evaluations for the Guide to Community Preventive Services. Am J Prev Med 2000;18(suppl 1):75–91.
17. Gould MR, Siegel JE, Russell LB, Weinstein MC, eds. Cost-effectiveness in health and medicine. Oxford: Oxford University Press, 1996.
18. Haddix AC, Teutsch SM, Shaffer PA, Dunet DO. Decision analysis and economic evaluation. Oxford: Oxford University Press, 1996.
19. Novick LF. Public health practice guidelines: a case study. J Public Health Management and Practice 1997;3:59–64.
20. Gyorkos TW, Tannenbaum TN, Abrahamowicz M, Oxman MD, Scott EAF, Millson ME, et al. An approach to the development of practice guidelines for community health interventions. Can J Public Health 1994;85(suppl 1):S8–S13.
21. Canadian Task Force on the Periodic Health Examination. Canadian guide to clinical, preventive health care. Ottawa, Canada: Canada Communication Group, 1994.
22. Murray DM. Design and analysis of group-randomized trials: a review of recent developments. Ann Epidemiol 1999;7(suppl 7):S69–S77.
23. Rothman KJ. Modern epidemiology. Boston: Little, Brown, and Company, 1986.
24. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. Control Clin Trials 1995;16:621–13.
25. Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998;352:609–13.
26. Wortman PM. Judging research quality. In: Cooper H, Hedges LV, eds. The handbook of research synthesis. New York: Russell Sage Foundation, 1994:97–109.
27. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. JAMA 1992;267:374–8.