Methods for Conducting Systematic Reviews of the Evidence of Effectiveness and Economic Efficiency of Interventions to Promote Healthy Social Environments

Laurie M. Anderson, PhD, MPH, Jonathan E. Fielding, MD, MPH, MBA, Mindy T. Fullilove, MD, Susan C. Scrimshaw, PhD, Vilma G. Carande-Kulis, PhD, MS, and the Task Force on Community Preventive Services

Overview: The social and physical surroundings in which people live affect their health. Knowing what basic conditions and opportunities in communities advance or impede improvement of community health can inform public health practice and policy. This article describes the methods for conducting systematic literature reviews of three community interventions to promote healthy social environments: early childhood development programs, programs to promote affordable family housing in safe neighborhoods, and interventions to increase the cultural and linguistic competence of healthcare systems. Existing methods, established for conducting systematic reviews for the Guide to Community Preventive Services, were applied to these interventions to promote healthy social environments. (Am J Prev Med 2003;24(3S):25–31)

Introduction

The Task Force on Community Preventive Services (the Task Force) chose the topic “the social environment” for inclusion in the Guide to Community Preventive Services (the Community Guide) to address broad determinants of health from an ecologic perspective.1,2 This perspective recognizes the contribution of the community environments in which people live and interact, including the social, political, and economic circumstances that characterize these environments.3–6

The social environment reviews for the Community Guide evaluate cross-cutting and comprehensive approaches to solving multiple health problems by focusing on social resources across the life span that can improve an array of poor health outcomes.7,8 In these systematic reviews the Task Force assessed evidence of the effectiveness of social environmental interventions to promote healthy early childhood development and school readiness, to provide affordable housing in safe neighborhoods to low-income families, and to reduce the linguistic and cultural barriers to health care faced by multicultural communities.

The Community Guide’s methods for conducting systematic reviews and for linking evidence to recommendations have been described elsewhere.9 In brief, for each Community Guide topic, an interdisciplinary team, representing such fields as medicine, public health, economics, education, sociology, anthropology, and psychology, and practicing in diverse settings, conducts a review by

- developing a conceptual framework to organize, group, and select appropriate interventions for the health issues under consideration and to choose the outcomes used to define success for each intervention;
- systematically searching for and retrieving evidence;
- assessing the quality of and summarizing the strength of the body of evidence of effectiveness;
- summarizing information about other evidence (i.e., evidence of applicability, economic evidence, other effects, and barriers to implementation); and
- identifying and summarizing research gaps.

This report describes the application of these methods in the systematic literature reviews of center-based...
early childhood development programs, housing subsidy programs that avoid concentrating low-income families in impoverished neighborhoods, and interventions to increase the cultural and linguistic competency of healthcare systems.

**Systematic Review Development Team**

Three groups of individuals served on the systematic review development team:

- The coordination team consisted of three Task Force members, methodologic experts in systematic reviews and economics from the Community Guide Branch (Epidemiology Program Office, Centers for Disease Control and Prevention [CDC]), and experts in community-based research addressing broad determinants of health. This team drafted the conceptual framework for the reviews; oversaw the data collection and review process; and drafted the evidence tables, summaries of evidence, and reports of the findings.

- The consultation team reviewed and commented on materials developed by the coordination team and, in conjunction with the coordination team, identified and set priorities for evaluating the interventions that were the subject of these systematic reviews. The consultants are experts in social determinants of health, with backgrounds in epidemiology, public health, medicine, health promotion, community development, social services, and health policy and program evaluation. They practice in public health settings, academic organizations, federal agencies, and community-based organizations.

- The abstraction team collected and recorded data from studies for possible inclusion in the systematic reviews (see Evaluating and Summarizing the Studies). The abstraction team members hold masters or doctoral degrees in epidemiology, medicine, behavioral sciences, and health administration.

In this report and in the accompanying articles, the term “team” refers to the coordination and consultation teams only, as the abstraction team’s role was primarily that of abstracting data from candidate studies.

**Conceptual Approach**

Early in the systematic review process, a conceptual (or logic) framework was developed to illustrate the entire
Table 1. Sociocultural environment reviews: proposed set of interventions for review, based on priority-setting process

<table>
<thead>
<tr>
<th>Rank</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Child development programs (e.g., Head Start, Healthy Start)</td>
</tr>
<tr>
<td>2</td>
<td>Adequate public investment in education</td>
</tr>
<tr>
<td>3</td>
<td>Minimum wages that move working families above poverty</td>
</tr>
<tr>
<td>4</td>
<td>Access to quality health care for all</td>
</tr>
<tr>
<td>5</td>
<td>Mixed income housing to decrease residential segregation by social class</td>
</tr>
<tr>
<td>6</td>
<td>Programs to reintegrate stigmatized populations</td>
</tr>
<tr>
<td>7</td>
<td>Parenting classes in schools, churches, health agencies, etc.</td>
</tr>
<tr>
<td>8</td>
<td>Schools as site of numerous supportive programs</td>
</tr>
<tr>
<td>9</td>
<td>Family-focused, one-stop health and human service centers</td>
</tr>
<tr>
<td>10</td>
<td>Programs to increase collaboration between community groups and organizations</td>
</tr>
</tbody>
</table>

public health context in which interventions might take place (Figure 1). This framework identifies resources, both physical and social, as root determinants of the health of communities. In this model, indicators of access to resources are observed through six community characteristics, or intermediate outcomes. Community interventions related to any of these characteristics can produce intermediate outcomes leading to changes in community health status. The framework illustrates that access to resources is fundamental to sustaining a healthy community and provides a model for understanding how community interventions can promote healthy social environments. It also allows others to judge whether important determinants, outcomes, or areas for intervention have been included. The systematic review team used the logic framework to generate an inclusive list of candidate interventions to promote healthy social environments, presented in this issue (Table 1 in the accompanying article). Intervention to ameliorate underlying social causes of premature death and disability requires collaboration between public health and local communities, businesses, other government agencies, nonprofit community and national organizations, faith organizations, and philanthropies.

Candidate Interventions

The extensive list of interventions generated by the team (Table 1 in the accompanying article) aimed at each of the six intermediate conditions identified in the logic framework (Figure 1). The goal of these interventions is to

- support healthy community norms, customs, and processes;
- strengthen community social cohesion, civic engagement, and collective efficacy; and
- assure access to health promotion, disease and injury prevention, and healthcare opportunities.

Standard criteria, agreed upon by the Task Force, are used for selecting priority interventions for systematic review in the Community Guide and include

- burden of disease, injury, impairment, or exposure;
- preventable disease burden;
- related health initiatives (e.g., Healthy People 2010); and
- areas of interest to the Community Guide audience (e.g., increasing investment in research or interventions; activity in state or local public health departments and in the private sector).

An iterative ranking process, based on the above criteria, was used to select a proposed set of priority interventions (Table 1), which was approved by the Task Force for analysis in the social environment systematic reviews. The systematic reviews of three of these intervention topics have been completed—early childhood development programs, family housing programs, and access to culturally competent healthcare systems—and are described in the accompanying evidence review articles. Review of other priority interventions is pending.

Analytic Frameworks

Analytic frameworks designed for each of the three intervention topic areas chosen for review (early childhood development, family housing, and culturally competent healthcare systems) are presented in the accompanying review articles. Analytic frameworks are used to illustrate the key health and other outcomes that might result from the intervention (and on which the literature search was to concentrate), the potential effect measures for each of those outcomes, and the target population and settings for the intervention. Specific outcome and effect measures used for determining effectiveness are described in each of the accompanying evidence review articles.

Evaluating and Summarizing the Studies

Each study that met the inclusion criteria was evaluated using a standardized abstraction form (available at www.thecommunityguide.org/methods) and was assessed for suitability of the study design and threats to validity. On the basis of the number of threats to validity, studies were characterized as having good, fair, or limited execution. Studies with limited execution were not included in the summary of the effect of the
intervention. The remaining studies (i.e., those with good or fair execution) were considered “qualifying studies.” Estimates of effectiveness are based on those studies.

Where possible, for studies that reported multiple measures of a given outcome, the “best” measure with respect to validity and stability was chosen according to consistently applied rules.\(^9\) Measures that were adjusted for the effects of potential confounders were used in preference to crude effect measures. For studies in which adjusted results were not provided, net effects were derived when possible by calculating the difference between the changes observed in the intervention and comparison groups. Among similar effect measures, the median was calculated as a summary measure.

Bodies of evidence of effectiveness were characterized as strong, sufficient, or insufficient on the basis of the number of available studies, the suitability of study designs for evaluating effectiveness, the quality of execution of the studies, the consistency of the results, and the effect size.\(^9\)

### Other Effects

The Community Guide systematic reviews of intervention effectiveness routinely sought information on other effects (i.e., positive and negative health or nonhealth “side effects”). Evidence of potential harms was ascertained if they were mentioned in the effectiveness literature or if the team thought they were important considerations. For example, in the reviews of tenant-based rental vouchers, the team conducted additional literature searches to determine if the intervention had negative consequences for the neighborhoods of poverty from which families moved (i.e., disruption of social ties and networks, depleting neighborhoods of human capital, and furthering neighborhood decline). Unanticipated positive effects were also noted if mentioned in the effectiveness literature.

### Evaluating Economic Efficiency

When the Task Force recommends an intervention, the team conducts systematic reviews of the evidence of economic efficiency.\(^9\) These reviews are provided to help decision makers choose among recommended interventions.

Methods for conducting systematic reviews of economic efficiency have been previously reported\(^ {16,17}\) and are summarized here as they were adapted for the review of interventions to promote healthy social environments. The four basic steps are

- searching for and retrieving economic evidence,
- abstracting and adjusting the economic data,
- assessing the quality of the identified economic evidence, and
- summarizing and interpreting the evidence of economic efficiency.

### Searching for and Retrieving Economic Evidence

The databases Medline, TRIS, CHID, NTIS, Embase, EI Compendex, PsycINFO, Social Science Search, Sociological Abstracts, ECONLIT, and Dissertation Abstracts were searched for the period 1970–2000. In addition, the references listed in all retrieved articles were reviewed and experts were consulted. Most of the included studies were either government reports or published in journals. To be included in the review a study had to

- be a primary study rather than, for example, a guideline or review;
- take place in an Established Market Economy\(^a\);
- be written in English;
- meet the team’s definitions of the recommended intervention;
- use economic analytic methods such as cost analysis, cost-effectiveness analysis, cost-utility analysis, or cost-benefit analysis (see Appendix A); and
- itemize program costs and costs of illness or injury averted.

### Abstracting and Adjusting Economic Data

Two reviewers read each study that met the inclusion criteria. Any disagreements between the reviewers were reconciled by consensus of the team members. A standardized abstraction form (available at www.thecommunityguide.org/methods/econ-abs-form.pdf) was used for abstracting data. For those studies in which cost-effectiveness or cost-utility analyses were conducted, results were adjusted to approximate the analysis to the reference case suggested by the Panel on Cost-effectiveness in Health and Medicine.\(^ {18}\) Results from cost-benefit analyses were adjusted for currency and base-year only. When feasible, results were recalculated if the discount rate used in the study was other than 3%.

### Assessing the Quality of the Evidence

Quality of study design and execution was systematically assessed\(^ {17}\) across five categories: study design, cost data, outcome measure, effects, and analysis. By subtracting points for each limitation from a perfect score of 100, study quality was characterized as very good (90–100), good (80–89), satisfactory (60–79), or unsatisfactory

\(^a\)Established Market Economies as defined by the World Bank are Andorra, Australia, Austria, Austria, Belgium, Bermuda, Canada, Channel Islands, Denmark, Faeroe Islands, Finland, France, Germany, Gibraltar, Greece, Greenland, Holy See, Iceland, Ireland, Isle of Man, Italy, Japan, Liechtenstein, Luxembourg, Monaco, the Netherlands, New Zealand, Norway, Portugal, San Marino, Spain, St. Pierre and Miquelon, Sweden, Switzerland, the United Kingdom, and the United States.
Results from unsatisfactory studies are not presented.

Summarizing the Body of Evidence

The findings for the economic efficiency of interventions are presented in summary tables. The summary tables include information on seven aspects of each included study. Table 2 provides an example of a summary table.

Ratios or net present values (i.e., the discounted net benefit or net cost obtained from cost-benefit analysis) are pooled in ranges when the intervention definition, population at risk, and comparator match across studies.

Barriers

Information about barriers to implementation of the interventions was abstracted from reviewed studies, evaluated on the suggestion of the team, or both. Information on barriers did not affect the Task Force recommendations, but it is provided to assist readers contemplating implementation of the interventions.

Translating Strength of Evidence into Recommendations

The Task Force recommendations presented in the accompanying article are based on the evidence gleaned from the systematic reviews conducted in accordance with the methods presented here. Each recommendation includes the strength of the evidence of effectiveness. Other types of evidence can also affect a recommendation. For example, evidence of harms resulting from an intervention might lead the Task Force to recommend against use of the intervention if adverse effects outweigh improved outcomes. In general, the Task Force does not use economic information to modify recommendations.

A finding of insufficient evidence to determine effectiveness should not be seen as evidence of ineffectiveness, but rather it reflects the fact that our systematic

### Table 2. Example of an economic evaluation summary table

<table>
<thead>
<tr>
<th>Authors</th>
<th>Reported or calculated summary measure</th>
<th>Study location</th>
<th>Intervention studied</th>
<th>Base year</th>
<th>Costs included</th>
<th>Benefits included (for CBA) or reported effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barnett</td>
<td>Cost-benefit analysis</td>
<td>Low-income area in Ypsilanti, Michigan (follow-up wherever participants later lived)</td>
<td>Intervention group received 2.5 hours of classroom time with four teachers each weekday, with one 1.5-hour home visit each week from a teacher; program lasted 30 weeks</td>
<td>1992</td>
<td>Costs included teacher and support staff salaries, school district overhead, classroom supplies, and future educational expenses (college)</td>
<td>Benefits included lifetime salary differential, avoided welfare, and costs of criminal activity</td>
</tr>
<tr>
<td>1996</td>
<td>Net benefit (NB) per program participant</td>
<td>Preschool and homes throughout low-income community</td>
<td>No program</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Authors</th>
<th>Analytic method</th>
<th>Population description</th>
<th>Follow-up period</th>
<th>Adjusted base year</th>
<th>Adjusted value summary measure</th>
<th>Quality category</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barnett</td>
<td>Cost-benefit analysis</td>
<td>Low-income area in Ypsilanti, Michigan (follow-up wherever participants later lived)</td>
<td>Preschool and homes throughout low-income community</td>
<td>1997 U.S. dollars</td>
<td>--</td>
<td>Very good</td>
<td>100</td>
</tr>
<tr>
<td>1996</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
review did not identify enough information for the Task Force to determine whether the intervention was or was not effective. Findings of insufficient evidence are important for identifying areas of uncertainty that require additional research. In contrast, sufficient or strong evidence of ineffectiveness would lead to a recommendation against use of the intervention.

Summarizing Research Gaps

Systematic reviews in the Community Guide identify existing information on which to base public health decisions about implementing interventions. An important additional benefit of these reviews is the identification of areas in which information is lacking or of poor quality. To summarize these gaps, remaining major research questions were identified for each intervention evaluated. Where evidence of effectiveness of an intervention was sufficient or strong, remaining questions about effectiveness, applicability, other effects, economic consequences, and barriers were summarized. Where evidence was insufficient to determine the effectiveness of an intervention, remaining questions about only effectiveness and other effects were summarized. Applicability issues were summarized only if they affected the assessment of effectiveness. In general, the Task Force believes that it would be premature to identify research gaps in economic evaluations or barriers before effectiveness is demonstrated.

For each category of evidence, the team identified unresolved issues that had emerged from the review. When a conclusion was drawn about evidence, the team decided if additional issues remained.

• If effectiveness was demonstrated using some but not all outcomes, all other possible outcomes were not necessarily listed as research gaps.
• If the available evidence was thought to be generalizable, all subpopulations or settings where studies had not been done were not necessarily identified as research gaps.
• Within each body of evidence, the team considered whether there were general methods issues that would improve future studies in that area.

The Reviews of Evidence

This article describes the general methodologic approach used in the systematic reviews of interventions to promote healthy social environments. Community Guide methods for classifying studies, tabulating effects, and determining sufficiency of evidence have been detailed elsewhere and are not described here or in the accompanying review articles. The accompanying articles present the supporting evidence on which the Task Force based its recommendations. They describe the scope and extent of the problem studied, discuss the conceptual approach to the review of evidence for the interventions studied, and present additional information about methodology specific to the review of those interventions, in addition to giving a detailed report on the findings for each intervention.

We would like to thank our Consultation Team: Regina M. Benjamin, MD, MBA, Bayou La Batre Rural Health Clinic, Bayou La Batre, Alabama; David Chavis, PhD, Association for the Study and Development of Community, Gaithersburg, Maryland; Shelly Cooper-Ashford, Center for Multicultural Health, Seattle, Washington; Leonard J. Duhl, MD, School of Public Health, University of California, Berkeley, California; Ruth Enid-Zambrana, PhD, Department of Women’s Studies, University of Maryland, College Park, Maryland; Stephen B. Fawcette, PhD, Work Group on Health Promotion and Community Development, University of Kansas, Lawrence, Kansas; Nicholas Freudenberg, DrPH, Urban Public Health, Hunter College, City University of New York, New York, New York; Douglas Greenwell, PhD, The Atlanta Project, Atlanta, Georgia; Robert A. Hahn, PhD, MPH, Epidemiology Program Office, CDC, Atlanta, Georgia; Camara P. Jones, MD, PhD, MPH, National Center for Chronic Disease Prevention and Health Promotion, CDC, Atlanta, Georgia; Joan Kraft, PhD, National Center for Chronic Disease Prevention and Health Promotion, CDC, Atlanta, Georgia; Nancy Krieger, PhD, School of Public Health, Harvard University, Cambridge, Massachusetts; Robert S. Lawrence, MD, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland; David V. McQueen, National Center for Chronic Disease Prevention and Health Promotion, CDC, Atlanta, Georgia; Jesus Ramirez-Valles, PhD, MPH, School of Public Health, University of Illinois, Chicago, Illinois; Robert Sampson, PhD, Social Sciences Division, University of Chicago, Chicago, Illinois; Leonard S. Syme, PhD, School of Public Health, University of California, Berkeley, California; David R. Williams, PhD, Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

References

Appendix A

Making people safer and healthier is a goal that everyone shares. Fortunately, several choices are available to achieve that goal. Each choice involves using scarce resources (e.g., medical services, drugs, medical personnel, people’s time). Therefore, tough decisions about these choices must be made. The Community Guide provides information on choice and efficiency.

In economics, two concepts of efficiency are distinguished: technological and allocative. A choice is said to be technologically efficient if inputs are combined to achieve a certain health outcome in such a way that resources are not wasted. For example, if a wound infection can be treated using 4 bandages or 5 bandages with no difference in outcome, then, other things being equal, using 4 bandages is considered technologically more efficient than a treatment using 5 bandages. Allocative efficiency (or economic efficiency) refers to maximizing health outcomes given a fixed budget. Alternatively, allocative efficiency also refers to minimizing costs given a target health outcome. For example, an economically efficient strategy is the cheaper of two alternative treatments that are medically equivalent. If the medical benefits and costs are different, the choice of the economically efficient treatment will depend on the relative costs and medical benefits of one treatment compared with another. Identifying economically efficient choices is the crux of economic evaluation and, therefore, crucial for promoting health and quality of life. The more efficiently resources are allocated, the more resources will be available for other uses.

To help identify whether a particular program is economically efficient, the Community Guide provides information on efficiency from four types of economic evaluation studies: cost analysis (CA), cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), and cost-utility analysis (CUA) as a special variant of CEA.

Cost analysis involves the systematic collection, categorization, and analysis of all resources consumed by a health intervention or program. Summary measures of a CA include total costs, average costs, and marginal costs. Total costs include all the resource costs incurred in producing a given level of output or health outcome and are calculated by summing all of the costs identified as part of the production process. Estimates of total costs can be used to rank programs in terms of economic burden or to measure the true cost of a program relative to its budget. The average cost is the cost per unit of output (e.g., the cost per patient treated or the cost per child vaccinated), and can be calculated by dividing total costs by the outcome of interest. Marginal cost is the resource cost associated with producing one additional (or one less) unit within the same intervention or program (e.g., additional cost per additional child using a safety seat), and can be calculated by dividing the change in total costs by the change in the outcome of interest.

Cost-benefit analysis converts all costs and benefits of a program into dollars, which allows the decision maker to compare programs that have disparate outcomes (e.g., a smoking cessation program with a program that reduces lead levels in housing). The summary measures for a CBA included in the Community Guide are net benefits or net present value. Net benefits are calculated by subtracting net costs of a program (net present value of costs) from net benefits of a program (net present value of benefits). A program with positive net benefits is considered a good economic investment. When comparing two or more programs, the program with the higher net benefits is considered the better economic investment.

Cost-effectiveness analysis has been used to compare the costs of different types of health interventions that affect the same (typically, intermediate) outcome (e.g., cost per quitter or cost per case prevented). Cost-utility analysis is a special variant of CEA that allows the decision maker to compare programs that address different health problems (e.g., breast cancer screening and diabetes control), because the (typically, final) outcomes are converted into a common health metric, such as quality-adjusted life years. Summary measures for a CEA include average cost-effectiveness ratios, marginal cost-effectiveness ratios, and incremental cost-effectiveness ratios, where results are presented in the form of net costs per health outcome, such as “cost per case prevented” or “cost per quality-adjusted life year saved.” Net costs in the numerator of a CEA ratio include program or intervention costs (as described in the cost analysis section above) minus the cost savings from averted illness or injury. In average analyses, the comparator (i.e., the alternative that the intervention or program is compared with) is the baseline, status quo, or a “no program” option. In marginal analyses, the comparator is the same program or intervention, expanded or downsized by some marginal amount (e.g., from one more hour of operation or from one more day of screening). In incremental analyses, an intervention is compared to the next best viable alternative.