Chapter Three

IS THERE ENOUGH EVIDENCE TO SUPPORT AN EVIDENCE-BASED APPROACH?

An evidence-based approach requires evidence, and the primary source of scientific evidence is the published literature. We undertook a broad review of the evaluation and effectiveness research literature to determine whether it can support evidence-based decisionmaking for community-based health interventions. Because we were interested in the "state of the art" we focused the review on evaluation of programs in the United States in the past 15 years, and we attempted to identify studies of effectiveness and studies of costs. In conducting the review, we sought to understand better how community-based programs were evaluated and what criteria of effectiveness or success were used in the published literature. Furthermore, we wanted to approximate what would happen if an informed community organization were to attempt to apply an evidence-based approach to deciding which community-based health programs to initiate.

Before summarizing the procedures we used and our findings, we address a number of limitations to our review. First, our literature review includes only evaluations that have been published. It is thus probably biased toward evaluations of high-visibility programs or those conducted by evaluators who routinely publish in the academic literature. Furthermore, we may have excluded programs that failed to achieve their initial goals if unsuccessful projects are, as many believe, less likely to be published.

Second, we did not attempt a comprehensive literature review of community-based health interventions analogous to the systematic

clinical literature reviews and formal meta-analyses central to the evidence-based medicine approach (Mulrow et al., 1997; Hunt and McKibbon, 1997; Bero and Jadad, 1997; Cook et al., 1997; Sackett and Rosenberg, 1995; L'Abbe et al., 1987). Our goals were too broad to apply such an approach rigorously. Nevertheless, we have used published protocols for systematic reviews and meta-analyses as guides for our search.

Third, we excluded clinical interventions from consideration. Thus, most interventions requiring clinical staff such as immunization programs and prenatal interventions, which might be undertaken by a community group or health system, were not reviewed.

Finally, resource limitations undoubtedly resulted in our missing a number of important and relevant publications over the time period we targeted. Even narrowly defined MEDLINE literature searches focusing on clinical trials of specific interventions may miss up to half of published trials (Dickersin et al., 1994). By polling senior investigators in the field and applying other standard methods, a more intensive review might capture missed published papers, studies that were not published in venues included in computerized databases, or unpublished evaluations. Nevertheless, we believe the results of our broad review provide important insights into the obstacles that currently exist for the application of an evidence-based approach in this new arena.

METHOD

The literature review was divided into five tasks:

1. **Initial search for community-based health program articles.** We used the built-in search capabilities of two computerized commercial literature databases. We searched the HEALTH PLAN database (1980-present), which includes all health services research literature, and a majority of the relevant literature from MEDLINE. Keywords for the search were community, health intervention, evaluation, effectiveness, and cost (and combinations thereof). Articles matching those key words were downloaded into the bibliographic software package ProCite (R), which allowed us to manage and print full citations and (where available)

abstracts. The bulk of this task was conducted in December 1996. We updated it in February 1997 and added those parts of MEDLINE that are excluded from HEALTH PLAN.

2. **Keyword search of task 1 output to identify articles discussing cost, cost-effectiveness, effectiveness, or evaluation.** The task 1 search netted over 2,000 articles, so we screened on a smaller set of keywords. This cut the citation listing down to 1,067 citations.

3. **Review of citations and (where available) abstracts and rating of each article as to whether or not it should be included in an in-depth review.** For each article in the task 2 output list, we printed the full citation, which often included an abstract. All three authors reviewed the citations to determine whether or not the full paper should be obtained and a more detailed review completed. For this "gross" review, we used a series of simple inclusion and exclusion criteria (Table A.1, Appendix A). To assess the reliability of these criteria, we selected one-year of citations (1996) and compared our ratings ("yes," "no," or "maybe"). All three authors agreed on the ratings of 42 percent and two of three agreed on another 49 percent. We discussed each of the citations for which there was not complete agreement. The discussion produced consensus and on that basis we added items to the inclusion and exclusion criteria. In addition, we determined that there were certain review articles or editorials that should be obtained for background but not further reviewed for the literature search.

   The output of this task was a list of 119 intervention studies and review articles (see Appendix A). Sixteen of these articles could not be located, leaving a total of 103 that were copied for in-depth review. Of these, 19 were review articles, leaving a total of 84 articles thought to be studies of community health interventions that might be useful in applying an evidence-based approach.

4. **Development of criteria for the in-depth review.** The focus of our review was on scientific evidence, specifically evidence of effectiveness of community-based health interventions. Thus, for an article to be considered as "relevant," the study had to include at least a quasi-experimental design with either a simultaneous comparison group or before-after comparisons for an intervention group. It also had to have a description of the intervention and a quantitative outcome measure. We also included general

health screening studies aimed at identifying clinical conditions among members of the community.

5. **In-depth review of articles according to criteria from Task 4.** The articles reviewed in detail were classified as relevant in applying an evidence-based approach or not. Of the 84 intervention studies identified and obtained in task 3, a total of 30 articles met our relevance criteria.

FINDINGS

We summarized the final 30 articles along several dimensions, namely, target population, setting, area of health focus, study design, and outcomes measured (see Table A.2, Appendix A). In terms of target population, about two-thirds of the interventions targeted children or adolescents with the remaining spread among the elderly, the general population, young adults, and subpopulations with specific diseases. About half of the interventions were implemented in an educational setting—a primary or secondary school or college. The remainder were either broad community interventions (e.g., public health promotion media campaigns) or were implemented at community-based organizations (e.g., Boys and Girls Clubs).

The focus of most community health interventions was on behaviors that affect health. The specific behaviors of interest varied widely across the studies, with drug and alcohol abuse being most common; others included injuries and violence, HIV prevention, cardiovascular risks, nutrition and weight loss, and exercise. For example, one strong study assessed the effect of a community-wide intervention program using a range of strategies to reduce alcohol-impaired driving, related driving risks, and traffic deaths and injuries (Hingson et al., 1996). Two studies evaluated health screening and promotion projects whose primary goal was identification of disease rather than behavioral change (Brink and Nader, 1984; Rogers et al., 1992). Although these projects were implemented in community settings, they perhaps should have been eliminated from our review because of their more clinical focus.

The most common study design was a quasi-experimental, pre-/post-intervention comparison between one or more intervention groups and a control group. However, a substantial minority of

the studies involved randomization between experimental and control groups, usually at the level of the class, school, or school district or, less commonly, the community. A smaller number of studies made only a pre/post comparison for an intervention group.

A broad range of outcome measures was used to assess the effect of interventions. The most common outcomes were measures of knowledge and self-reported behavior. For example, another strong study assessed the effect of a school intervention with and without a connected media intervention designed to reduce smoking among students in grades four through six (Flynn et al., 1994). The outcome in this study was self reports of cigarettes smoked within a week of a survey. A substantial minority of studies included objective measures of health outcomes such as incidence of injuries or cholesterol and blood pressure measures. For example, an evaluation of the Stanford Adolescent Heart Health Program measured self-reported behavior such as exercise and diet as well as body mass index and resting heart rate (Killen et al., 1989). Aside from studies focusing on injuries, none measured markers of incidence of conditions such as HIV transmission rates.

Only one of the 30 articles included an analysis of costs (Rogers et al., 1992). The intervention in this study consisted of a clinical community-based program of health examinations and screenings combined with case management and education. The investigators calculated average direct monthly costs per participant over a two-year period. No other costs such as indirect costs of a participant's time were included in this analysis. Two other studies reported some measures of cost, although no detailed analyses (Hingson et al., 1996; Mayer et al., 1992). In describing the interventions evaluated, several studies reported details such as staffing requirements that could be of limited help in assessing resource requirements. In one study, a community-based weight reduction intervention was specifically described as a "low-cost" alternative to more expensive clinic-based interventions without providing any evidence of having actually assessed costs (Del Prete et al., 1993).

Typically, success was measured in terms of a statistically significant difference either between intervention and control groups or before and after the intervention. Relatively few studies discussed in detail the meaning of the magnitude of the difference in outcome measures

between the intervention and the control. Only two of the papers clearly stated that the evaluation had concluded that the intervention had not had any of the desired effects (Emery and Gatz, 1990; O'Brien and Anderson, 1987). One other study concluded that the intervention had not had an effect on one of its primary outcomes (McLoughlin et al., 1982).

The studies varied considerably in the level of detail provided on the content of the intervention. Many referred to other publications for details about the intervention, but most did not themselves provide detailed descriptions.

Of the 54 articles read for the in-depth review but considered not very relevant to applying an evidence-based approach, a substantial number were found to be review articles that were not identified as such in the task 3 screening. Some of these articles might be very useful for pointing to specific examples to review in more detail or for synthesizing a large body of research, but they typically did not provide detailed information on each intervention. Several were very general and often contained only vague discussions of public policy about specific types of interventions. Several papers addressed methodologic issues in the evaluation of community health interventions.

A small number of the papers in this category were labeled as formal process or qualitative evaluations. We found these very difficult to assess. Attributions about specific results of interventions were often not backed by sufficient evidence to warrant acceptance of the conclusion. For example, one study of a community-based health promotion program stated that "churches tended to be the most effective means of reaching the target audience," yet offered no evidence to support that claim (Doyle et al., 1989). Many of these studies focused on describing what was done rather than on providing evidence of effect; some provided no evidence, quantitative or qualitative, of effect. None of the evaluations in this category involved any comparison between a control group (either pre/post or simultaneous) and the intervention participants.

DISCUSSION

The literature search for effectiveness research on community-based health interventions yielded a wide range of papers, a minority of which would be immediately useful in applying an evidence-based approach to decisionmaking for community health interventions. Only one of the papers considered relevant included an analysis of costs, and that was relatively crude. The review raised a number of important issues regarding the feasibility of applying an evidence-based approach in this setting.

Scientific Rigor of Evaluation and Effectiveness Research

Any comprehensive review of the scientific literature must ultimately make judgments about the validity of scientific findings. If, as in clinical reviews, randomized trials are accepted as the gold standard because of their ability to reduce bias, the literature in this field for the most part does not meet this standard. Perhaps this should not be surprising, as randomization for community health interventions may be more difficult than randomization of patients in a clinical setting. The exception to this generalization may be schools, and these were indeed the most common setting where randomization at the population level did occur. Attention should be given to recent calls for broader use of randomization to increase the strength of scientific evidence in community health research (Green, 1997; Fortmann et al., 1995; Green et al., 1995; Koepsell et al., 1992). At the same time, the difficulty of randomization in community interventions should also lead to more careful consideration of ways in which strong evidence regarding effect might be drawn from quasi-experimental designs. While in the clinical research setting considerable effort has been devoted to developing a consensus about a paradigm for ranking of scientific evidence (e.g., Guyatt et al., 1995), we found little evidence that such effort has occurred in the area of community health research.[1]

---

[1]Indeed, there is considerable ambiguity even about the terms we have used in our review. The investigators began the review with the assumption that evaluation research included studies of effectiveness. However, surveys suggest that evaluation research may be considered to be distinct from research to determine effectiveness (Carpinello et al., 1992).

Even more difficult than deciding what to make of quasi-experimental designs is deciding what to make of qualitative or process evaluations of community health interventions. We recognize that some effort has been devoted to the development of criteria for assessing the scientific validity of qualitative research on interventions (e.g., Muir Gray, 1997a). Process evaluations play an important role in understanding how interventions work in real communities (Goodman et al., 1993). They may also provide evidence of the effect of interventions on outcomes that are difficult to measure (Light and Pillemer, 1984), such as cultural orientation of communities toward health promotion activities. Nevertheless, we believe that these evaluations are of limited use in an evidence-based model. Without clearer guides for assessing this type of evidence, it is extremely difficult to judge the scientific validity of claims made in such analyses.

Assessing prevention effectiveness requires quantitatively analyzing the effect of interventions. As one standard evaluation text (Rossi and Freeman, 1993) states:

> [A]ssessing impact in ways that are scientifically plausible and that yield relatively precise estimates of net effects requires data that are quantifiable and systematically and uniformly collected.

Outcome Measures and Comparisons Across Potential Interventions

The most common outcome measure among the studies we considered relevant was self-reported behavior. This may not provide an accurate picture of behavior because participants may be inclined to give what they consider to be desired answers. Gold-standard measures of outcomes such as biochemical markers of smoking or even more remote outcomes such as morbidity are more expensive and may be inappropriate or impractical in a community setting (Fishbein, 1996; Mittlemark et al., 1993). Therefore, proximal measures, such as measures of behavioral change, are reasonable outcomes for most community health interventions. Nevertheless, the strength of the evidence of community health interventions will almost certainly improve with the broadened use, when feasible, of more objective and valid measures of the desired outcomes. Our review suggests the need for readily available and comprehensive

guides as to the appropriateness of various outcome measures for different types of community health interventions. Such standards would facilitate assessing the evidence of effectiveness in these settings.

Even in the studies we considered relevant that used valid measures of effect, measures of similar outcomes varied with the study and typically, resource use was not reported. This impedes substantive comparisons across interventions or explicit comparisons of costs versus effectiveness or benefit. For example, if a community-based organization wanted to choose between a school-based intervention targeted at drug abuse and a community-wide intervention to target HIV prevention, the literature as it currently exists would not provide much guidance. The use of standard outcome measures and the provision of cost data would substantially improve the ability of a community-based organization to make such choices.

Access to Information on Implementation

Even if a community-based organization were to identify in the literature an intervention that it wished to implement, it might have difficulty locating details about it. Often, there appeared to be no easy way to obtain information on how to implement the programs described in the studies short of attempting to contact the authors or funding agencies directly. (It might be an interesting exercise to try to find detailed information about how to implement some of these interventions.)

The difficulty of ready access to information on implementation has been recognized as a major problem in promoting community health interventions. The Health Resources and Services Administration of the U.S. Department of Health and Human Services recently published a compendium of primary care health programs entitled *Models That Work* (Bureau of Primary Health Care, 1995). One aim of this publication (which was not identified in our literature searches) was to improve the dissemination of information about successful interventions to local organizations by providing general descriptions and addresses for program contacts. Of note, the interventions were chosen by an advisory panel of health professions and are described as "cost-effective" in the publication, but the publication unfortunately gives no rigorous evaluation of cost or effectiveness evidence. Never-

theless, publications such as this are an important step in the direction of implementing an evidence-based approach to decisions regarding community health interventions.