Evaluating the Public Health Impact of Health Promotion Interventions: The RE-AIM Framework

Russell E. Glasgow, PhD, Thomas M. Vogt, MD, MPH, and Shawn M. Boles, PhD

Although the field of health promotion has made substantial progress, the advancement of research methods used to evaluate the public health impact of our programs. However, with few exceptions, evaluations have restricted their focus to 1 or 2 of 5 dimensions of quality we believe to be important.

There is a great need for research methods that are designed to evaluate the public health significance of interventions. The efficacy-based research paradigm that dominates current notions of science is limiting and not always the most appropriate standard to apply. A reductionistic scientific paradigm oversimplifies reality, and in the quest to isolate efficacious treatments, most clinical trials focus on eliminating potential confounding variables and involve homogeneous, highly motivated individuals without any health conditions other than the one being studied.

This approach provides important information and strong internal validity; from an external validation perspective, however, it results in samples of nonrepresentative participants and settings.

Similarly, the emphasis on developing clinically significant outcomes often produces interventions that are intensive, expensive, and demanding of both patients and providers. These interventions tend to be studied in the rarefied, "controlled" atmosphere of specialty treatment centers using highly standardized protocols. This "efficacy paradigm" does not address how well a program works in the world of busy, understaffed public health clinics, large health systems, or community settings.

Our medical culture emphasizes pharmacological interventions that produce immediate results and whose dosage can be easily defined and controlled. There is little research on interventions that address whole populations, are long lasting, or become "institutionalized.

Indeed, many interventions that prove efficacious in randomized trials are much less effective in the general population.

In this commentary, we describe the RE-AIM evaluation model, which emphasizes the reach and representativeness of both participants and settings, and discuss the model's implications for public health research. The representativeness of participants is an important issue for outcome research. The representativeness of settings—clinics, worksites, or communities—for public health interventions is equally important. Many evaluations, such as the otherwise well-designed Community Intervention Trial for Smoking Cessation (CIT), explicitly restrict selection of participating communities (and research centers) to those most motivated, organized, and prepared for change. This results in experts, highly motivated research teams and settings, which are, by definition, unrepresentative of the settings to which their results are to be applied.

Recognizing some of the foregoing issues, both the National Cancer Institute and the National Heart, Lung, and Blood Institute have proposed sequential "stages" of research. These steps move from hypothesis generation to testing under controlled conditions, evaluations in "defined populations," and, finally, dissemination research. Interventions found to be efficacious then undergo "effectiveness" evaluations, and programs that prove to be effective—"especially cost-effective"—are selected for dissemination research.

There is often difficulty, however, in making the transition across phases. We think this may be due to a flaw in the basic model, in that many characteristics that make an intervention efficacious (e.g., level of intensity of the intervention and whether it is designed for motivated, homogeneous populations) work against its being effective in more complex, less advantageous settings with less motivated patients and overworked staff.

Low-intensity interventions that are less efficacious but that can be delivered to large numbers of people may have a more pervasive impact.
Abrams and colleagues defined the impact of an intervention as the product of a program's reach, or the percentage of population receiving the intervention, and its efficacy ($E$). We expand on this "RE" (Reach × Efficacy) concept by adding 3 dimensions that apply to the settings in which research is conducted (Adoption, Implementation, and Maintenance: "AIM") to more completely characterize the public health impact of an intervention.

**RE-AIM Model**

We conceptualize the public health impact of an intervention as a function of 5 factors: reach, efficacy, adoption, implementation, and maintenance. Each of the 5 RE-AIM dimensions is represented on a 0 to 1 (or 0% to 100%) scale. This framework is compatible with systems-based and social-ecological thinking as well as community-based and public health interventions. A central tenet is that the ultimate impact of an intervention is due to its combined effects on 5 evaluative dimensions. The RE-AIM model expands on earlier work and is summarized in Table 1.

**Reach**

Reach is an individual-level measure (e.g., patient or employee) of participation. Reach refers to the percentage and risk characteristics of persons who receive or are affected by a policy or program. It is measured by comparing records of program participants and complete sample or "census" information for a defined population, such as all members in a given clinic, health maintenance organization, or worksite. If accurate records are kept of both the numerator (participants) and the denominator (population), calculation of participation rates is straightforward.

Reach (as well as adoption) also concerns the characteristics of participants. Assessing representativeness is challenging. It requires demographic information and preferably psychosocial, medical history, or case mixture information on nonparticipating participants. Detailed information on nonparticipating is often challenging to collect and raises ethical issues in that nonparticipating are typically not consented to be studied. Cooperative arrangements that permit investigation of the extent to which participants are representative of the larger "denominator" population should be a priority for future research.

Unfortunately, participants in health promotion activities sometimes are those who need them least (e.g., the "worried well," those in the more affluent segments of the population, and nonsmokers). With the increasing gap between the "haves" and "have-nots" in our country, and the dramatic impact of socioeconomic status on health status, understanding the degree to which a program reaches those in need is vital. Because public health interventions are addressed to large numbers of people, even small differences in risk levels between participants and nonparticipants can have a significant impact on cost-effectiveness.

**Efficacy**

Entire textbooks have been devoted to evaluating the efficacy of interventions. We discuss 2 specific issues: the importance of assessing both positive and negative consequences of programs and the need to include behavioral, quality of life, and participant satisfaction outcomes as well as physiologic endpoints.

**Positive and negative outcomes.** Most population-based evaluations focus on improvement in some targeted health or risk indicator. Interventions delivered to large populations can also have unanticipated negative effects. Labeling someone with a potential illness may have profound social and psychological consequences. Many effective services remain underdelivered, while others are delivered that are not necessary or effective in the groups receiving them. Even services that cost only a few dollars can have substantial negative (as well as positive) societal effects, including misused resources and large opportunity costs, when delivered to millions of people. It is critical not only to determine benefits but also to be certain that harm does not outweigh benefits.

**Outcomes to be measured.** Clinical research emphasizes biologic outcomes—in particular, disease risk factors—and concerns about limited resources have led to an increasing emphasis on health care use. Such outcomes are important, but a public health evaluation should include more than simply biologic and use measures. Two other types of outcomes merit inclusion. First, behavioral outcomes should be assessed for participants (e.g., smoking cessation, eating patterns, physical activity, for staff who deliver an intervention (approaching patients, delivering prompts and counseling, making follow-up calls), and for the payers and purchasers who support the intervention (adopting an intervention, changing policies). Second, a participant-centered quality-of-life perspective should be included to allow evaluation of patient functioning, mental health, and consumer satisfaction, since these factors provide a critical check on the impact of delivery practices.

**Adoption**

Adoption refers to the proportion and representativeness of settings (such as worksites, health departments, or communities) that adopt a given policy or program. There are common temporal patterns in the type and percentage of settings that will adopt an innovative change. Adoption is usually assessed by direct observation or structured interviews or surveys. Barriers to adoption should also be examined when nonparticipating settings are assessed.

**Implementation**

The term effectiveness is used to describe evaluations conducted in real-world settings by individuals who are not part of a research staff. Implementation refers to the extent to which a program is delivered as intended. It can be thought of as interacting with efficacy to determine effectiveness (Efficacy × Implementation = Effectiveness). There are both individual-level and program-level measures of implementation. At the individual level, measures of participant follow-through or "adherence" to regimens are necessary for interpreting study outcomes. At the setting level, the extent to which staff members deliver the intervention as intended is important. Stevens et al. demonstrated that differential levels of protocol implementation were, in large part, the reason that a brief hospital-based smoking-cessation program was more successful when implemented by research staff than by hospital respiratory therapy staff. Implementation research is crucial in determining which of a set of interventions may be practical enough to be effective in representative settings.

**Maintenance**

A major challenge at both individual and organization-community levels is long-term maintenance of behavior change. At the individual level, relapse following initial behavior change is ubiquitous. Equally essential is the collection of program-level measures of institutionalization, that is, the extent to which a health promotion practice or policy becomes routine and part of the everyday culture and norms of an organization. Recently, there have been
advances in identifying factors related to the extent to which a change is institutionalized. At the community level, maintenance research is needed to document the extent to which policies are enforced over time (e.g., laws concerning alcohol sales, no-smoking policies). Maintenance measures the extent to which innovations become a relatively stable, enduring part of the behavioral repertoire of an individual (or organization or community).

Combining Dimensions

The public health impact score, represented as a multiplicative combination of the component dimensions (Table 1), is probably the best overall representation of quality. The RE-AIM model is silent on the choice of efficacy measure; any outcome that is quantifiable, reliable, valid, and important to scientific, citizen, and practitioner communities is admissible. Examples include hypertension, mammography screening, and smoking status.

Implicit in the constructs of implementation and maintenance is the length of the period during which data are collected: a minimum of 6 months to 1 year for implementation and 2 years or longer for maintenance. Frequency of assessment should be based on the particular issue, goals, and setting. If RE-AIM dimensions are assessed multiple times, then a RE-AIM profile can be plotted. Repeated measurements and visual displays can enhance understanding of intervention effects and be used to compare different interventions (Figure 1). (Additional tables and figures related to application of the RE-AIM framework are available at www.oroinfo.org/shawn/public/reainf/reaim.long.pdf.)

Discussion

The past several years have seen a variety of provocative articles on changing paradigms of health care. Unfortunately, there have been few discussions of evaluation models for these new population-based paradigms. Even economic analyses and outcomes research do not address several of the core evaluation issues and key dimensions of these evolving approaches. Evaluation methods must match the conceptual issues and interventions being studied.

With the shift to a multiple causation and holistic or systems approach to medical science, recognition of the complexity and various levels of disease determinants is required.

While classic randomized controlled trials have significantly advanced our knowledge of pharmacotherapy and medical interventions, they have limitations when applied to behavioral issues and, especially, community interventions. Randomized controlled trials emphasize efficacy to the de facto exclusion of factors such as adoption, reach, and institutionalization. RE-AIM provides a framework for determining what programs are worth sustained investment and for identifying those that work in real-world environments. RE-AIM can be used to evaluate randomized controlled studies as well as studies with other designs, and it is compatible with evidence-based medicine. RE-AIM asserts, however, that evidence should be broadened to include dimensions in addition to efficacy. The model can also be used to guide qualitative research efforts by focusing inquiry on each of these issues. To the extent that RE-AIM dimensions are incorporated into evaluations, decision makers will have more complete information on which to base adoption or discontinuance of programs.

Data collected via the RE-AIM model can serve several evaluative purposes: (1) assessing an intervention's overall public health impact, (2) comparing the public health impact of an intervention across organizational units or over time, (3) comparing 2 or more interventions across RE-AIM dimensions (Figure 1), and (4) making decisions about redistributing resources toward more effective programs.

Future Research and Policy Issues

There are several implications of the RE-AIM model. Empirical evaluations involving these implications would greatly help in assessing the utility of the model and informing policy and funding decisions.

1. From the RE-AIM perspective, we expect that programs that are very efficacious (under highly controlled, optimal conditions) may have poor implementation results. Such an inverse relationship between program efficacy and implementation, or between reach and efficacy, has significant implications for the types of interventions that should receive high priority.

2. The RE-AIM model does not explicitly include economic factors. However, cost issues are addressed in 3 ways. First, we think that cost is often a major factor determining whether a program will be adopted, implemented consistently, or maintained. This hypothesis should be tested and substantiated or refuted. Second, cost-effectiveness and cost-benefit are certainly appropriate outcomes. They determine how well resources are being used and whether or not more good could be accomplished through alternative uses (opportunity costs). Finally, a population-based cost-effectiveness index could be calculated by dividing the resulting public health impact by the total societal costs of a program. Dividing each RE-AIM component score by the costs relevant to that dimension could help identify areas of efficiency and waste. There is a need for further work on similar formulas and evaluation of the extent to which providing decision makers with information on RE-AIM dimensions influences decisions.

3. Systematic reviews that determine the extent to which different research fields have studied—or neglected—each of the RE-AIM dimensions are needed. We hypothesize that adoption and maintenance—institutionalization will be the most understudied dimensions, but this needs to be documented for different research topics.

Limitations of RE-AIM

The precise nature of the relationships among the 5 RE-AIM dimensions, and how
they combine to determine overall public health impact, is unknown. We have represented these factors as interacting multiplicatively because we believe that this is closer to reality than an additive model. A highly efficacious program that is adopted by few clinics or that reaches only a small proportion of eligible citizens will have little population-based impact. In future research, it will be necessary to determine the precise mathematical functions that best characterize the interplay of these dimensions.

In this initial model, we have implicitly assumed, in the absence of data to the contrary, that all 5 RE-AIM dimensions are equally important and therefore equally weighted. This may not always be the case. In situations in which 1 or more of the RE-AIM dimensions are considered most important, differential weights could be assigned. Similarly, it may not be necessary to assess all RE-AIM components in every study.

Finally, the time intervals we have suggested for assessing implementation (6 months–1 year) and maintenance (2+ years) are arbitrary. Future research is needed to determine whether there are necessary or optimal intervals for evaluating these dimensions.

Conclusion

Public health interventions should be evaluated more comprehensively than has traditionally been done.4,25,31 Dimensions such as reach, adoption, and implementation are crucial in evaluating programs intended for wide-scale dissemination. We hope that the RE-AIM framework, or similar models that focus on overall population-based impact, will be used to more fully evaluate public health innovations. Such an evaluation framework helps remind us of the key purposes of public health, organizational change, and community interventions.44,86,98 It is time to re-aim our evaluation efforts.

Contributors

R. E. Glasgow originated the idea for the RE-AIM framework and served as editor. T. M. Voight drafted sections on implications and coined the acronym RE-AIM. S. M. Boles drafted content on the various uses of the model and contributed the figure and ideas about displays. All authors contributed substantially to the writing of the paper.

Acknowledgments

Preparation of this manuscript was supported by grants RO1 DK51581, RO1 DK55524, and RO1 HL52538 from the National Institutes of Health.

Appreciation is expressed to Drs Gary Cutter, Mark Dugan, Ed Lieberson, and Deborah Toobert for their helpful comments on an earlier version of this article.

References


22. Flay BR. Efficacy and effectiveness trials and other phases of research in the development of
of the Society of Behavioral Medicine; March 1998, Rockville, Md.