The Future of Health Behavior Change Research: What Is Needed to Improve Translation of Research Into Health Promotion Practice?

Russell E. Glasgow, Ph.D.
Kaiser Permanente Colorado and
AMC Cancer Research Center
Denver, Colorado

Lisa M. Klesges, Ph.D.
University of Tennessee Health Science Center

David A. Dzewaltowski, Ph.D.
Kansas State University

Sheana S. Bull, Ph.D.
University of Colorado Health Sciences Center and
AMC Cancer Research Center
Denver, Colorado

Paul Estabrooks, Ph.D.
Kansas State University

ABSTRACT

Background: *It is well documented that the results of most behavioral and health promotion studies have not been translated into practice.* ***Purpose:*** *In this article, reasons for this gap, focusing on study design characteristics as a central contributing barrier, are discussed.* ***Methods:*** *Four reviews of recent controlled studies in work sites, health care, school, and community settings are briefly discussed and summarized. Their implications for future research and for closing the gap between research and practice are then discussed.* ***Results:*** *These reviews come to consistent conclusions regarding key internal and external validity factors that have and have not been reported. It is very clear that moderating variables and generalization issues have not been included or reported in the majority of investigations, and that as a consequence little is known about the representatives or the robustness of results from current studies.* ***Conclusions:*** *To significantly improve the current state of affairs, substantial changes will be required on the part of researchers, funding agencies, and review and editorial boards. In conclusion, recommendations for each of these entities are provided.*

(**Ann Behav Med** 2004, 27(1):3–12)

INTRODUCTION

Lifestyle practices of individuals, first recognized as predictors of cardiovascular disease in the early 1960s, brought the term *risk factor* into common use (1). Over the following 40 years, behavioral medicine and behavioral epidemiology have evolved to identify, explain, and address personal risk factors (2,3). Despite these 4 decades of research on "what works" to change risk factors, we as health care professionals are faced with a short supply of generalizable, effective, and sustainable interventions that have been translated into health promotion practice. In addition, there is inadequate documentation of the potential population or public health impact of behavioral change interventions. In this review we explore the extent of and potential reasons for this gap between research and practice and offer recommendations to bridge it.

A large body of evidence is available that identifies behavioral risk factors and demonstrates the efficacy of behavioral interventions. However, although establishing efficacy under carefully controlled analog conditions is an important step, it is not sufficient to judge the potential effectiveness of health promotion interventions (4–6). As typically conducted, efficacy studies involve narrowly drawn samples, lack sufficient description of how selected samples reflect the larger population under study, and involve tightly controlled intervention implementation. Because of these factors, overemphasis on such studies when conducting meta-analyses and making recommendations for practice can lead to biases in the conclusions. Specific examples of shortcomings of generalizations from these types of efficacy trials are documented elsewhere (7). In general, these biases limit the ability to establish the external validity of results, a key element for judging the potential effectiveness of interventions and for translating research to practice (see http://www.re-aim.org).

Efficacy and effectiveness studies can sometimes be difficult to differentiate, especially in community-based investigations. Most often they differ in the type, number, and representativeness of the participants, settings, or intervention agents involved. An example of the different conclusions that can be reached related to different intervention agents is provided by the work of Stevens and colleagues (8,9). They conducted an initial hospital-based smoking cessation efficacy study that demonstrated significant benefit of a brief intervention administered by experienced smoking cessation counselors (8). This same intervention however, when replicated in the same hospi-

Reprint Address: R. E. Glasgow, Ph.D., P.O. Box 349, Canon City, CO 81215. E-mail: russg@ris.net

tals and supervised by the same intervention team, was not significantly more effective than usual care when implemented by regular hospital respiratory therapists (9).

As we document in this article, the behavioral health promotion evidence base also generally lacks information on the feasibility and costs of interventions as well as assessments of what works across various targeted groups, under assorted conditions, and in diverse settings. Such investigation is crucial to improve translation of efficacious strategies to sustainable health promotion programs. For example, additional evidence to assess the replicability, robustness, and effectiveness of many health promotion interventions is needed to make clear recommendations for community preventive services (10,11). Deficits in the evidence base are receiving attention among researchers, community practitioners, funding organizations, editors, and other individuals who advocate health promotion. Incorporation of different design features; new funding opportunities; greater emphasis on external validity issues by review groups; and a larger evidence base that addresses feasibility, economic, and generalization issues will be necessary to bridge this gap between research and practice.

What Are the Implications of the Gap in Translation of Research to Practice?

A major limitation of the way much efficacy research is conducted is that it "decontextualizes" an intervention effect by studying narrowly selected participants and applying narrowly specified intervention strategies that are not applicable in settings that are representative or are not capable of being implemented by typical intervention agents (4). As a result, intervention effectiveness across different types of persons, settings, and times (12) is inadequately understood, and this narrows the range of application and potential translation of findings.

Strict inclusion and exclusion criteria give the sense of studying well-defined populations; however, these widely used procedures can create problems in interpretation. Recruitment methods often result in a sample of motivated volunteers or organizations that express interest and meet eligibility requirements. Cook and Campbell (12) forewarned that "being a member of a class does not necessarily imply being representative of that class" and stressed that the extent to which a clinical sample of participants reflects the entirety of the potentially eligible population should be addressed as a key aspect of generalization. Consideration of the reach or representativeness of the targeted population of persons or settings is critical to judge the potential public health impact of an intervention strategy and should be addressed in both efficacy and effectiveness studies.

For example, an intervention strategy that has been shown to improve dietary patterns, increase physical activity, or decrease tobacco use may be moderated by participant characteristics. The intervention may work only with individuals who respond to advertisements for participation in research studies and not be effective for individuals who screen their telephone calls. The intervention may be moderated by site characteristics as well; that is, there may be characteristics of the social and physical environments where intervention studies are delivered, such as the number of community physical activity opportunities or local health policies, that interact with the effectiveness of the intervention.

Health behavior interventions can also be judged in terms of the level of intervention implementation and fidelity. Process evaluation of intervention fidelity is especially important to differentiate true treatment failure from poor intervention implementation, or "Type III" error (13). Many interventions found successful in research studies prove impractical to implement in applied settings that have limited time, few resources, and many competing demands. Clinical trials often require intensive patient contact and monitoring protocols to ensure high levels of implementation, which result in large effects but entail sizable costs. Unfortunately, economic data are seldom reported, and without sufficient description of these intervention costs, the relative efficiency of an intervention cannot be determined. Furthermore, cost may interact with community conditions, such as level of resources, organization, or experience. Finally, an emphasis on short-term behavior change present in most efficacy studies provides little insight on what is required to maintain or institutionalize an intervention.

To help address these limitations and improve translation, an evaluative framework that draws attention to issues important for translation would be helpful. A good framework would promote a comprehensive approach and broader acceptance of critical elements of external validity, in addition to internal validity, that can assist in the translation of research to practice.

What Methods Are Available to Improve the Translation of Research to Practice?

Several guidelines are available to judge the clinical utility and efficacy of interventions directed at individuals. In fact, to improve evidence-based clinical practice, the Agency for Health Care Research and Quality has formed the National Guideline Clearinghouse (14), a "guideline on guidelines," to sort through the sometimes onerous number of published efficacy reviews. In addition, rules of evidence, such as the Consolidated Standards of Reporting Trials (CONSORT) statement (15), are available that focus on the conduct and reporting of randomized controlled trials.

Fewer resources are available to estimate the generalization or the public health impact of prevention interventions. Criteria have been developed and are being applied to reviews of evidence to make recommendations for community prevention interventions (10,11), and methods for calculating attributions of prevention interventions to estimate population health impact are emerging (16–18).

Less widely available are rules of evidence that attempt to evaluate dimensions of both internal and external validity and that describe elements of study design that are useful across various stages of research. One evaluative framework, the RE-AIM model (see http://www.re-aim.org), developed by Glasgow, Vogt, and Boles (19), has directed attention to the multiple criteria related to health behavior change research, including efficacy; reach and representativeness; im-

plementation feasibility; sustainability; and other factors important for public health decision making, such as quality of life and safety.

Specifically, the first two components of the RE-AIM model (see Table 1) are (a) *R*each, or the percentage and representativeness of patients who are willing to participate in a given program, and (b) *E*fficacy or *E*ffectiveness (depending on the study), or the impact of an intervention on important outcomes, including potential negative effects, quality of life, and economic outcomes. The three less often studied but equally important factors, which concern impact at the level of the intervention setting are the "AIM" dimensions: (c) *A*doption, or the percentage and representativeness of settings that are willing to adopt or try a health promotion program; (d) *I*mplementation, or how consistently various elements of a program are delivered as intended by different intervention agents, and the time and cost requirements of intervention; and (e) *M*aintenance, or the extent to which a program or policy becomes institutionalized or part of the routine practices and policies of an organization. Maintenance in the RE-AIM framework also has referents at the individual level. At the individual level, Maintenance refers to the long-term effects of a program on outcomes 6 months or more after the most recent intervention contact.

The RE-AIM framework can be applied in several capacities, including planning studies to maximize understanding of both internal and external validity characteristics, comparing the effectiveness of several interventions for policy decisions, and judging the level of transferability of findings to other settings and populations (20–22). It can also be used as an evaluative framework to review a body of literature (23,24).

In this article we apply the RE-AIM framework to summarize structured literature reviews of recently published health promotion research in four settings: (a) worksites, (b) schools, (c) health care, and (d) general community. We assess the extent to which health promotion studies in these various settings have described elements of internal and external validity along RE-AIM dimensions. On the basis of findings from this review, we present recommendations for researchers, funding agencies, review groups, and editorial boards that we hope will aid the translation of behavioral health promotion research to practice settings and help inform public health policy decisions.

METHOD

Our research group recently completed literature reviews using the RE-AIM evaluation framework to characterize the status of behavior change research. We analyzed the results of controlled intervention studies conducted during the years 1996–2000 and reported in one of several leading health behavior change journals. In these reviews, we discussed the status of outcome studies conducted in worksites (25), health care settings (26), schools (27), and community settings (28), respectively. Details of inclusion criteria, coding reliability, and other characteristics are reported in the individual studies, but in summary we coded the results of all studies reported in leading health behavior (nonspecialty) journals that evaluated a behavioral intervention compared with some type of control or comparison condition. Coding reliability on the various RE-AIM criteria was very high across the reviews. These reviews were conducted as part of the Behavior Change Consortium, so we reviewed all studies in the Behavior Change Consortium's focus areas of nutrition or dietary change (including weight loss), physical activity, and/or smoking cessation. The Behavior Change Consortium (http://www1.od.nih.gov/behaviorchange) is a collaboration among 15 intervention studies funded by the National Institutes of Health (NIH) that are theory based and address multiple behaviors (29).

Table 2 shows the number of articles that met inclusion criteria from each of 10 standard journals included in each of the

TABLE 1
RE-AIM Dimensions and Template Questions for Evaluating Health Education and Health Behavior Research

RE-AIM Dimension	*Questions*
Reach (individual level)	What percentage of potentially eligible participants: (a) were excluded and (b) took part, and (c) how representative were they?
Efficacy/effectiveness (individual level)	What impact did the intervention have on: (a) all participants who began the program; (b) on process intermediate and primary outcomes; and (c) on both positive and negative (unintended) outcomes, including quality of life?
Adoption (setting level)	What percentage of settings and intervention agents within these settings (e.g., schools/educators, medical offices/physicians): (a) were excluded and (b) participated, and (c) how representative were they?
Implementation (setting/agent level)	To what extent were the various intervention components delivered as intended (in the protocol), especially when conducted by different (nonresearch) staff members in applied settings?
Maintenance (both individual and setting level)	1. Individual level: (a) What were the long-term effects (minimum in 6–12 months following intervention)? (b) What was the attrition rate, were dropouts representative, and how did attrition impact conclusions about effectiveness? 2. Setting level: (a) To what extent were different intervention components continued or institutionalized? (b) How was the original program modified?

TABLE 2
RE-AIM Literature Review: Controlled Studies 1996–2000 Identified by Journal and Behavior

Journal	*Nutrition*	*Exercise*	*Smoking*	*Multiple*	*Total*
American Journal of Health Promotion	3	2	4	1	10
American Journal of Preventive Medicine	5	4	2	0	11
American Journal of Public Health	10	3	10	0	23
Annals of Behavioral Medicine	2	4	1	1	8
Australian and New Zealand Journal of Public Health	0	0	2	0	2
Canadian Journal of Public Health	1	0	0	0	1
Health Education & Behavior	11	1	4	3	19
Health Education Research	2	1	3	0	6
Journal of Behavioral Medicine	0	1	1	0	2
Preventive Medicine	8	5	8	5	26
Other[a]	5	2	3	1	11
Total	47	23	38	11	119

Note. RE-AIM = Reach, Efficacy, Adoption, Implementation, Maintenance.
[a]Included journals specific to the four settings reviewed, such as *Medical Care, Occupational Health and Medicine,* and *Journal of School Health.*

TABLE 3
Percentage of Studies Reporting on RE-AIM Dimensions Overall and by Setting of Research

RE-AIM Dimension/Measure	*Worksites*[a]	*Schools*[b]	*Health Care*[c]	*Community*[d]	*Overall Average*
Individual level					
Reach					
Participation rate[e]	88	59	69	88	76
Representativeness	9	7	28	11	14
Effectiveness					
Behavioral outcome measure	67	100	100	100	92
QOL or negative outcomes	0	NR	17	3	7
Maintenance					
≥ 6-month follow-up	4	26	86	30	36
Attrition at follow-up	54	74	87	100	79
Setting level					
Adoption					
Participation rate—site level	25	15	11	11	16
Representativeness of settings	0	0	0	7	2
Implementation					
Treatment delivery[f]	12	37	77	59	46
Time or cost	0	NR	31	63	31
Maintenance					
Setting continuation	4	0	6	0	2

Note. RE-AIM = Reach, Efficacy, Adoption, Implementation, Maintenance; QOL = quality of life; NR = not reported in review.
[a]$n = 24$. [b]$n = 32$. [c]$n = 36$. [d]$n = 27$. [e]Often from volunteer rather than population-based samples. [f]Often from efficacy studies where treatment delivered by research staff

reviews, and an "other" category of journals specific to each setting (e.g., *Journal of School Health* for school interventions; *Medical Care* for health care settings). As can be seen, a total of 119 studies were reviewed, with 23–47 in each of the three behavior change target areas but only 11 that reported on changes across multiple behaviors. Because results were very similar across the various target behaviors, the data summarized in this report are collapsed across target behaviors.

The top half of Table 3 summarizes results on the individual-level factors of the RE-AIM framework. As can be seen, some criteria were reported very consistently, including participation rate,[1] a behavioral outcome measure, and attrition rates. Studies inconsistently reported on follow-ups at least 6 months after last intervention contact, with worksite studies reporting the lowest percentage of such follow-up (4%) and health care settings most often reporting this length of follow-up (86%). Reports of the representativeness of participants versus nonpar-

[1]We note, however, that many of these rates were calculated from volunteer rather than population-based samples.

ticipants (a measure of Reach), and any measure of potential negative impact of intervention or a quality-of-life measure (a measure of effectiveness) were consistently low (less than 15% on average) across studies.

The lower half of Table 3 describes setting-level variables, and these results reveal quite a different pattern than the individual-level dimensions. As can be seen, few studies in any setting reported criteria important to external validity; adoption and setting-level maintenance were particularly infrequent. Only 16% of the studies reported any information on participation rate at the setting level (in contrast to 76% at the individual level), and only 2% reported any information on the representativeness of these settings (those including such data were all community-based studies). Similarly, very few studies reported any information on whether the setting in which the study was conducted continued the intervention, even in modified fashion, after conclusion of the study. Setting-level factors that were reported with moderate frequency were both measures of implementation, namely, extent of intervention delivery (or treatment fidelity) and the total time or costs associated with the intervention. Both of these factors were reported variably across the different types of settings, with worksites reporting least often on both criteria, health care settings reporting most often on intervention delivery, and community studies reporting most often on time or costs of the intervention.

Summary of Reviews

Across all four reviews, internal validity criteria, such as measures of effectiveness and intervention delivery, were reported much more often than measures of external validity. Information on representativeness of either individual participants or of settings studied was conspicuously absent. As a consequence, we can conclude very little about the robustness or impact of health promotion interventions across different settings or participant characteristics. Some have argued that internal validity should come first in the form of tightly controlled efficacy studies, to be later followed by effectiveness studies under more representative conditions (30–32). However, as we have demonstrated elsewhere (4), (a) this argument does not necessarily hold, because efficacy studies need to assess potential moderating factors that can limit or confound results (33), and (b) this transition from efficacy research to effectiveness studies that focus on generalizability, as demonstrated by the data in Table 3, is simply not taking place. This conclusion is consistent with results of other recent literature reviews using different databases and inclusion criteria (34–35).

The cumulative results for this review, along with other recent review articles (34–36), strongly support the conclusion that we need improvements in our attention to external validity and to setting factors in *both* efficacy and effectiveness studies. To remedy the present situation, researchers need to identify innovative ways to enhance Reach (especially representativeness and to underserved populations), Adoption, Implementation (adherence by staff), and Maintenance.

Another consistent finding was that individual level measures were reported substantially more often than were setting level variables. This condition also needs to be remedied. Information about the representativeness and actions of settings and intervention agents are just as important as data on the representativeness and behavior of individual participants.

In summary, these four reviews consistently concluded that the behavioral health promotion field needs much greater attention to external validity and to setting factors—in both efficacy and effectiveness studies. In a sentence, there is a need for studies that can determine the characteristics of interventions that can (a) reach large numbers of people, and especially those who can most benefit; (b) be widely adopted by different and representative settings; (c) be consistently implemented by staff members with moderate levels of training and expertise; and (d) can produce replicable and long-lasting effects (and minimal or no negative impacts) at reasonable cost.

RECOMMENDATIONS FOR FUTURE RESEARCH

The literature reviews summarized above make it clear that more effectiveness research is needed, and that research at all stages needs to pay increased attention to setting level and other moderating factors and to representativeness. Calls for increased attention alone will not likely change the present state of affairs, however. To produce meaningful change in the disturbing results summarized in Table 3, concerted and coordinated efforts will be needed not only by researchers, but also by funding organizations and by review panels and editorial boards. Based on our conclusions, this section makes recommendations for specific actions that each of these entities can take to help facilitate the translation of research to practice.

Recommendations for Researchers

To improve the information available relevant to translation issues, researchers can apply RE-AIM or similar frameworks in several capacities. The framework can be used to plan studies to maximize external validity; as a guide to report elements of both internal and external validity within all phases of research (4,6,21); review a body of evidence (23–24); and compare interventions to make policy decisions (22). For the purposes of this paper, RE-AIM framework and the results of our four recent reviews are used to offer specific recommendations to enhance the external validity of health promotion interventions (see Table 4).

It is recommend that researchers report RE-AIM dimensions in all studies, but in particular, in efficacy and effectiveness studies. This will improve extrapolation of results, aid qualitative and meta-analytic reviews, and assist in designing replication studies. It is especially important to implement study features that maximize external validity, such as those in Section 1 of Table 4, to improve the ability to translate study results to practice decisions. This process should begin in the planning phase of intervention studies (Section 1A of Table 4) and continue through to the reporting phase (Section 1B). In addition to these recommendations, we have recently developed an Internet resource to promote ease of use and understanding of the RE-AIM framework (http://www.re-aim.org). The site provides a

TABLE 4

Recommendations to Accelerate Transfer of Research to Practice for Researchers, Funding Organizations, and Editors/Reviewers Based on RE-AIM Criteria

1. Recommendations for Researchers

A. Designing and Conducting Research

Involve the target audience or "end consumers" and program providers in formative assessments and intervention design from the outset to enhance reach, adherence, adoption, implementation, and maintenance.

Design interventions so that they reach large numbers and representative portions of the intended target population.

Investigate recruitment methods and program features designed to enhance reach within populations of (a) participants and (b) settings.

Replicate intervention effects or include purposeful sampling across heterogeneous persons and settings to judge the robustness of the intervention.

Study the consistency of implementation and outcomes produced across a range of intervention modalities, settings, and delivery agents.

Validate interventions that are straightforward to implement and produce training materials so that a wide array of agents can successfully deliver the program.

Design interventions with theoretical constructs in mind and measure moderating and mediating variables at the individual and setting levels to validate hypothesized change processes.

Include criteria for success in addition to effect size, including measures of quality of life and potential negative outcomes of interventions.

Include a maintenance phase in research studies to improve understanding of long-term behavior change at the individual level and sustainability and institutionalization at the setting level.

B. Reporting Research Results

Describe results on reach, adoption, implementation, and maintenance—as well as effectiveness—in standardized ways to facilitate comparisons across studies.

Report the distribution of the targeted population(s), participation rates, and compare characteristics of participants and nonparticipants at both the individual and setting levels. If information on nonparticipants is not possible to collect, compare the sample to representative data from your area (see http://www.re-aim.org).

Specify recruitment methods and program features hypothesized to be key to enhancing or hampering program participation.

Carefully describe the context of the intervention in terms of persons, settings, and local history.

Report on characteristics of the intervention agents and the modalities of delivery so that these can be replicated and compared to other studies.

Document attrition from the study and describe characteristics of dropouts.

Report the maintenance of individual behavior change of at least 6 months duration, and use procedures to evaluate the impact of attrition on results.

Record and report costs of all aspects of the intervention, including intervention materials, training, and delivery.

Report on continuance or modification of the program after conclusion of study.

Report relationships among various RE-AIM components (e.g., Reach and Efficacy) and compare trade-offs between internal and external validity in various interventions.

2. Recommendations for Funding Organizations

Solicit proposals that investigate interventions in multiple settings and document that these settings are representative of those to which the program wishes to generalize.

Fund novel and innovative investigations of ways to enhance reach, adoption, implementation and maintenance of behavior change interventions.

Require standard and comprehensive reporting of exclusions, participation rates, and representativeness of both participants and settings.

Fund sequential program changes and time-series studies, replications, multiple baseline and cross-over designs, and other designs in addition to randomized controlled trials that can efficiently and practically address key issues in translation.

Fund programs that investigate and can demonstrate high levels of implementation and consistent outcomes across a wide range of intervention agents similar to those in applied settings.

(continued)

TABLE 4 (*Continued*)

Require a maintenance/sustainability phase in research projects and implementation of plans to enhance institutionalization once the original research has been completed. Alternatively, fund competing applications to test the long-term effects and generalization of initially successful programs.
Encourage innovation in intervention design and standardization in reporting on process and outcome measures at both the individual and setting/intervention agent levels.
3. Recommendations for Editors, Journal Reviewers, and Grant Reviewers
Provide editorial guidelines and templates (á la CONSORT criteria or National Institutes of Health reporting criteria on minority participation) to facilitate and reinforce reporting on reach, adoption, implementation, and maintenance and enhance critical appraisal of evidence.
Encourage and provide feedback on adequacy of reporting on representativeness of participants and settings as well as intervention agents.
Encourage reporting on implementation issues, including time, costs, and range of intervention agents who can successfully deliver programs.
Encourage reporting (and relax usual editorial criteria) for investigations of long-term results of programs at individual and setting levels, especially at both levels.
Publish editorials and calls for papers that address translation, dissemination, and external validity issues and publish review criteria that encourage reviewers to equally consider issues of internal and external validity.
Include generalizability and potential for translation to practice a review criterion for funding of behavior change research.

Note. RE-AIM = Reach, Efficacy, Adoption, Implementation, Maintenance; CONSORT = Consolidated Reporting of Clinical Trials.

detailed explanation of elements of RE-AIM and illustrates their application in study design, and it includes program planning and evaluation tools as well as an on-line community mechanism to exchange ideas among researchers.

Because clinical trials are generally designed to maximize internal validity, their limits in estimating external validity are quickly reached. Studies specifically designed to maximize external validity and that provide estimates of the robustness of an intervention across settings, target populations, delivery agents and modalities, and local history are needed to expand the health promotion evidence base (33). In addition, studies to compare the costs, cost effectiveness, and efficiency of programs are in short supply but are critical for making policy recommendations regarding preventive services. Conducting such investigations at all phases of the research continuum, especially the later phases, are imperative to address current deficiencies in the evidence base (4).

Recommendations for Funding Organizations

The middle section of Table 4 summarizes steps that private and governmental funding organizations can take to provide the necessary incentives and motivation for researchers to produce the types of research recommended in this review. First, funding organizations can require that proposals include implementation in multiple settings and that these settings be representative of those in which the program would be applied. Second, in much the same way that the NIH now requires grants to report on gender and race–ethnicity characteristics of participants, standard reporting criteria could be required to facilitate cross-study comparisons of exclusions, participation rates, and representativeness of both participants and of settings. Recent health promotion research has focused almost exclusively on producing large effect sizes; it is clear that other important dimensions of program outcomes have been neglected (34,35). Thus, Request For Applications on specific innovative methods of enhancing the Reach, Adoption, Implementation, and Maintenance of health promotion interventions are needed. Methodological innovativeness should also be encouraged, specifically through use of designs in addition to the usual randomized control trials, which are powerful but also constraining in terms of their ability to efficiently test new ideas or address setting-level factors. Few studies have evaluated the robustness of intervention programs across different intervention agents. A useful step before disseminating promising interventions might be to fund demonstrations of the impact of different intervention agents, and various training and supervision conditions, on both implementation and study outcomes.

It is doubtful that significant advances will be made in addressing the pressing problems of behavioral maintenance, at either the individual or the setting level, unless funding priorities change. Requirements that studies include at least a planning phase for long-term implementation and maintenance would substantially change current research practices in this area. An implication is that longer term grants—or sequenced grant phases, each contingent on satisfactory progress—would be needed. Funding organizations can do a great deal to change the present state of affairs. They should attempt to simultaneously provide encouragement of innovative interventions and study designs and at the same time standardized methods of reporting process and outcome results, to allow comparison across studies.

Recommendations for Reviewers

The last section of Table 4 provides recommendations for gatekeepers of the peer review process for research funding and publication. Specifically, these recommendations target journal editors and reviewers and members of grant review study sections who could facilitate the publication and funding of research that balances concerns of external and internal validity. If such recommendations were adopted, there would be increased reporting on factors that currently present barriers to the translation of research to practice.

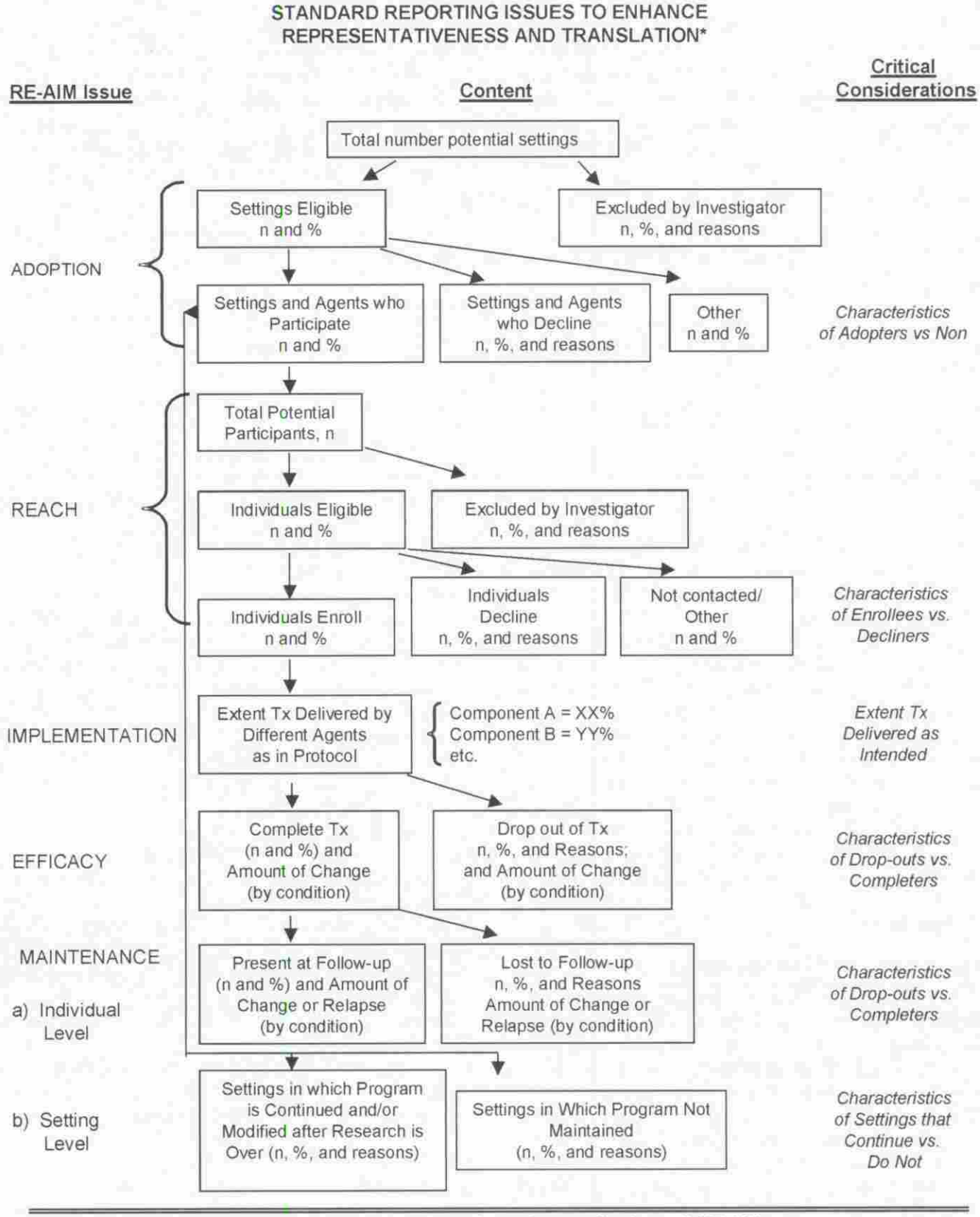

FIGURE 1 RE-AIM = *R*each, *E*fficacy, *A*doption, *I*mplementation, *M*aintenance; Tx = treatment. Reprinted from *American Journal of Preventive Medicine, 26,* Russell E. Glasgow, Sheana S. Bull, Cynthia Gillette, Lisa M. Klesges, and David A. Dzewaltowski, "Behavior Change Intervention Research in Healthcare Settings: A Review of Recent Reports With Emphasis on External Validity," 62–69, Copyright 2002, with permission from *American Journal of Preventive Medicine.*

Others share our concern for providing policy and decision makers with more high-quality scientific findings that can inform evidence-based prevention and intervention practice. To enhance the quality of behavioral science, the Evidence-Based Behavioral Medicine Committee of the Society of Behavioral Medicine researchers have recommended the adoption of a modified version of the CONSORT criteria for randomized controlled trials (37).

The CONSORT criteria consist of 22 standardized reporting items for articles illustrated by a flow diagram and a checklist. Although these reporting criteria emphasize internal validity, they generally fail to address external validity, generalization, or other factors that may moderate outcomes. Only 1 of the 22 current CONSORT items relates to external validity, and this item is less specific compared to the remaining items. We hope that editors and reviewers of health behavior journals will adopt the modified criteria of the Evidence-Based Behavioral Medicine Committee, which include seven elements from the RE-AIM framework (4). The adoption of the current CONSORT criteria without modification may contribute to a continued lack of translation of health behavior research to practice. We recommend that editorial and grant review guidelines and templates place balanced emphasis on internal and external validity factors (see Figure 1). Elbourne and Campbell (38) raised similar concerns in their extension of the CONSORT statement to cluster randomized trials. We suggest possible alternatives to the CONSORT checklist (4) and flow diagram as shown in Figure 1 (see also http://www.re-aim.org).

We also recommend that editors and reviewers encourage and provide feedback on the adequacy of the reporting on representativeness of participants, settings, and intervention agents. To increase the reporting of these issues, the proposed flow diagram includes representativeness of participants and settings/intervention agents. We also recommend the reporting of process and cost of implementation data by including time, costs, and the range of intervention agents who can successfully deliver programs. Editorial guidelines should facilitate and reinforce reporting the long-term results of programs at both participant and cluster level (agent, setting) and inclusion of possible moderating variables.

The recommendations in Table 4 are intended to direct attention to the balanced reporting of all of the RE-AIM elements to facilitate the integration of and conclusions that can be drawn from research findings. Generalization of an effect is probably best established by replication of studies in systematically sampled settings and among different types of research participants. Editors and reviewers should recommend including this information in research reports.

The final two recommendations focus on increasing the importance of generalization and translation of research to practice issues. We recommend editorials and calls for papers that address dissemination and external validity issues. We also recommend that generalizability and potential for translation to practice be included as a review criterion for funding of behavior change research (similar to the NIH's innovativeness criterion).[2] Because the current research culture underemphasizes issues of external validity (4), the gap between research findings and evidence-based practice will not be narrowed unless innovations are also made by journal reviewers, editors, and grant review panels.

[2]This criterion has recently been included in some NIH R18 proposal reviews.

CONCLUSIONS

If we are serious about evidence-based behavioral medicine (37,39) and about closing the gap between research findings and application of these findings in applied settings, we cannot continue "business as usual." Rather, as enumerated in Table 4, a different set of priorities will need to emerge. There have been some recent encouraging steps in this direction, including requests for proposals from the Robert Wood Johnson Foundation, the "dissemination" funding supplement pioneered by the National Cancer Institute, and the Translating Research into Practice RFAs from the Agency for Health Care Research and Quality, but much more comprehensive and systematic efforts are needed. We have provided some beginning steps in this article in the form of both resources to help researchers better address and report on key issues relevant for closing this gap and have recommended steps for reviewers and funding agencies to consider. As explicated elsewhere (4), we feel that these recommendations can be followed in all stages of health behavior research, from initial hypothesis generation and methods development continuing through efficacy, effectiveness, and dissemination studies. If even a few of these recommendations for each group were implemented, this would help to significantly improve the state of affairs summarized in Table 3. We hope that the issues raised in this article will stimulate debate and thinking about additional ways to achieve the goal of making behavioral medicine much more widely applied outside of research settings.

REFERENCES

(1) National Heart Lung and Blood Institute: *You Changed America's Heart: Framingham Heart Study.* Retrieved November 2, 2003, from http://rover.nhlbi.nih.gov/about/framingham/fhsbro.htm

(2) Glass T: Psychosocial interventions. In Berkman LF, Kawachi I (eds), *Social Epidemiology.* Oxford, England: Oxford University Press, 2000.

(3) Institute of Medicine: *Promoting Health: Intervention Strategies From Social and Behavioral Research.* Washington, DC: National Academy Press, 2000.

(4) Glasgow RE, Lichtenstein E, Marcus AC: Why don't we see more translation of health promotion research to practice? Rethinking the efficacy to effectiveness transition. *American Journal of Public Health.* 2003, *93:*1261–1267.

(5) Clarke GN: Improving the transition from basic efficacy research to effectiveness studies: Methodological issues and procedures. *Journal of Consulting and Clinical Psychology.* 1995, *63:*718–725.

(6) Estabrooks PA, Gyurcsik NC: Evaluating the public health impact of physical activity interventions. *Psychology of Sport and Exercise.* 2003, *73:*21–28.

(7) Collins R, MacMahon S: Reliable assessment of the effects of treatment on mortality and major morbidity: I. Clinical trials. *The Lancet.* 2001, *357:*373–380.

(8) Stevens VJ, Glasgow RE, Hollis JF, Lichtenstein E, Vogt TM: A smoking cessation intervention for hospital patients. *Medical Care.* 1993, *31:*65–72.

(9) Stevens VJ, Glasgow RE, Hollis JF, Mount K: Implementation and effectiveness of a brief smoking cessation intervention for hospital patients. *Medical Care.* 2000, *38:*451–459.

(10) Centers for Disease Control and Prevention: *The Guide to Community Preventive Services.* Retrieved November 2, 2003, from http://www.thecommunityguide.org

(11) Truman BI, Smith-Akin CK, Ilinman AR, et al.: Developing the *Guide to Community Preventive Services:* Overview and rationale. *American Journal of Preventive Medicine.* 2001, *18:*18–26.

(12) Cook TD, Campbell DT: *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally, 1979.

(13) Basch CE, Sliepcevich EM, Gold RS: Avoiding Type III errors in health education program evaluations. *Health Education Quarterly.* 1985, *12:*315–331.

(14) National Guideline Clearinghouse: Welcome home page. Retrieved November 2, 2003, from http://www.guideline.gov/

(15) Moher D, Schulz KF, Altman DG: The CONSORT statement: Revised recommendations for improving the quality of reports. *Journal of the American Medical Association.* 2001, *285:*1987–1991.

(16) Heller RF, Dobson AJ: Disease impact number and population impact number: Population perspectives to measures of risk and benefits. *British Medical Journal.* 2000, *321:*950–953.

(17) Heller RF, Page J: A population perspective to evidence based medicine: Evidence for population health. *Journal of Epidemiology and Community Health.* 2002, *56:*45–47.

(18) Rychetnik L, Frommer M, Howe P, Shiell A: Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health.* 2002, *56:*119–127.

(19) Glasgow RE, Vogt TM, Boles SM: Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health.* 1999, *89:*1322–1327.

(20) Glasgow RE, McKay HG, Piette JD, Reynolds KD: The RE-AIM framework for evaluating interventions: What can it tell us about approaches to chronic illness management? *Patient Education and Counseling.* 2001, *44:*119–127.

(21) Glasgow RE: Evaluation of theory-based interventions: The RE-AIM model. In Glanz K, Lewis FM, Rimer BK (eds), *Health Behavior and Health Education* (3rd Ed.). San Francisco: Wiley, 2002, 531–544.

(22) Lando HA, Valanis BG, Lichtenstein EL, et al.: Promoting smoking abstinence in pregnant and postpartum patients: A comparison of 2 approaches. *American Journal of Managed Care.* 2001, *7:*685–693.

(23) Eakin EG, Glasgow RE, Riley KM: Counseling patients to increase physical activity: A review of the effectiveness of current strategies. *Journal of Family Practice.* 2000, *49:*158–168.

(24) France EK, Glasgow RE, Marcus AC: Smoking cessation interventions among hospitalized patients: What have we learned? *Preventive Medicine.* 2001, *32:*376–388.

(25) Bull SS, Gillette C, Glasgow RE, Estabrooks P: Worksite health promotion research: To what extent can we generalize the results and what is needed to translate research to practice? *Health Education and Behavior.* 2003, *30:*537–549.

(26) Glasgow RE, Bull SS, Gillette C, Klesges LM, Dzewaltowski DA: Behavior change intervention research in health care settings: A review of recent reports with emphasis on external validity. *American Journal of Preventive Medicine.* 2002, *23:*62–69.

(27) Estabrooks PA, Dzewaltowski DA, Glasgow RE, Klesges LM: How well has recent literature reported on important issues related to translating school-based health promotion research into practice? *Journal of School Health.* 2003, *73:*21–28.

(28) Dzewaltowski DA, Estabrooks PA, Klesges LM, Bull SS, Glasgow RE: Behavior change intervention research in community settings: How generalizable are the results? *Health Promotion International* (in press).

(29) Ory MG, Jordan PJ, Bazzare T: The Behavior Change Consortium: Setting the stage for a new century of health behavior change research. *Health Education Research.* 2002, *17:*696–699.

(30) Flay BR: Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine.* 1986, *15:*451–474.

(31) Greenwald P, Cullen JW: The new emphasis in cancer control. *Journal of the National Cancer Institute.* 1985, *74:*543–551.

(32) Mook DG: In defense of external invalidity. *American Psychologist.* 1983, *38:*379–387.

(33) Brewer MB: Research design and issues of validity. In Reis HT, Judd CM (eds), *Handbook of Research Methods in Social and Personality Psychology.* New York: Cambridge University Press, 2000, 3–39.

(34) Oldenburg B, Ffrench BF, Sallis JF: Health behavior research: The quality of the evidence base. *American Journal of Health Promotion.* 2000, *14:*253–257.

(35) Oldenburg BF, Sallis JF, Ffrench ML, Owen N: Health promotion research and the diffusion and institutionalization of interventions. *Health Education Research.* 1999, *14:*121–130.

(36) Lomas J: Improving research dissemination and uptake in the health sector: Beyond the sound of one hand clapping. C97-1 (November, 1997). 1997. McMaster University Centre for Health Economics and Policy Analysis/Policy Commentary.

(37) Trudeau KJ, Davidson K, for the Evidence-based Behavioral Medicine Committee: A CONSORT Primer. *Outlook: A Quarterly Newsletter of the Society of Behavioral Medicine.* 2002, *Winter:*5–8.

(38) Elbourne DR, Campbell MK: Extending the CONSORT statement to cluster randomized trials: For discussion. *Statistics in Medicine.* 2001, *20:*489–496.

(39) Davidson K, Goldstein M, Kaplan R, et al.: Evidence-based behavioral medicine: What is it and how do we achieve it? *Annals of Behavioral Medicine.* 2003, *26:*161–171.