

# The Future of Physical Activity Behavior Change Research: What Is Needed to Improve Translation of Research into Health Promotion Practice?

David A. Dzewaltowski,<sup>1</sup> Paul A. Estabrooks,<sup>2</sup> and Russell E. Glasgow<sup>2</sup>

<sup>1</sup>Community Health Institute, Kansas State University, Manhattan, and <sup>2</sup>Clinical Research Unit, Kaiser Permanente–Colorado, Denver

DZEWALTOWSKI, D. A., P. A. ESTABROOKS, and R. E. GLASGOW. The future of physical activity behavior change research: What is needed to improve translation of research into health promotion practice? *Exerc. Sport Sci. Rev.*, Vol. 32, No. 2, pp. 57–63, 2004. *There is a gap between physical activity intervention research and the delivery of evidence-based programs in practice. We describe individual and setting level factors important for translation that moderate the impact of interventions and often are not reported in the literature. The Reach, Efficacy/Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework provides a useful way of organizing and reporting these factors.* **Keywords:** exercise, evaluation, generalization, recommendations, methods, intervention

## INTRODUCTION

Although several investigations have shown that physical activity reduces the risk of chronic disease and premature death and improves quality of life, many youth and adults are not meeting public health standards. Leaders in healthcare, worksites, schools, government, not-for-profit and other community settings are now investing time, effort, and monetary support to promote physical activity. However, the effectiveness and sustainability of these promotion efforts remain questionable.

There is a large, emerging body of research evidence on “what works” to improve physical activity. Recently, the Task Force on Community Preventive Services conducted a systematic review of physical activity interventions to provide information for decision-makers in translating evidence-based physical activity interventions into public health practice (11). Within health care, this challenge is labeled evidence-based medicine and evidence-based behavioral

medicine. Both evidence-based medicine and evidence-based behavioral medicine seek to improve the quality of practice through providing systematic information on proven intervention strategies.

As physical activity promotion practice moves into evidence-based behavioral medicine, lessons can be learned from other evidence-based efforts. A recent report by the Institute of Medicine titled “Crossing the Quality Chasm” noted the lack of translation of evidence-based medical and disease management intervention protocols into practice (2). Tunis *et al.* (15) recently called for more “practical clinical trials” that address generalization issues important for clinical and policy decisions. Similar problems have been documented in the delivery of prevention interventions to adults (12) and to youth through schools. Ennett *et al.* (4) compared current school practice against evidence-based standards for substance abuse intervention content and delivery methods. In a national sample of 1795 public and private middle schools, results indicated that only 14% of substance-abuse prevention providers used effective content and delivery methods. If researchers and practitioners do not change substantially the way we approach EEBM, then the physical activity promotion field is likely to repeat the same mistakes of these other fields.

We hypothesize and document below that, at both the individual and setting levels, variables that likely moderate

Address for correspondence: David A. Dzewaltowski, PhD, K-State Research and Extension, Community Health Institute, Umberger 101, Kansas State University, Manhattan, KS 66506 (E-mail: dadx@ksu.edu).

Accepted for publication: January 11, 2004.

0091-6331/3202/57–63

Exercise and Sport Sciences Reviews

Copyright © 2004 by the American College of Sports Medicine

the uptake, impact, and sustainability of interventions are seldom studied or reported in the physical activity literature. Consequently, we know little about the representativeness or robustness of the results from current physical activity promotion studies. Without this knowledge, the body of evidence that demonstrates the efficacy of physical activity interventions when delivered in controlled conditions may not generalize or be sustainable under the conditions in which these interventions must be implemented in practice. These real-world conditions include many factors that can interact with or moderate the reach, adoption, delivery, impact, or sustainability of an intervention.

In this paper, we describe the Reach, Efficacy/Effectiveness (depending on the research goal), Adoption, Implementation, and Maintenance (RE-AIM) framework as a tool to address and report internal and external validity information and estimate public health impact in the physical activity area (See Table 1); we summarize the data from four studies that support our hypothesis above; then we focus on two elements of RE-AIM (*i.e.*, Reach, Adoption) to illustrate the problem and to provide examples. To help bridge the gap between research and practice, the paper concludes with recommendations for future physical activity intervention research.

## THE REACH, EFFICACY/EFFECTIVENESS , ADOPTION, IMPLEMENTATION, AND MAINTENANCE FRAMEWORK

To enhance the potential for translating research to practice, one way to balance both internal and external validity in the planning, design, and evaluation of health behavior promotion interventions is to use the RE-AIM framework (8–10). Each element of the RE-AIM framework provides valuable information that may facilitate the translation of research to practice (see Table 1).

Individual participant level indicators within RE-AIM include reach and efficacy, whereas setting level (*e.g.*, schools, workplaces, medical offices) indicators include adoption and implementation. Maintenance is assessed at both an individual and setting level of impact. Each element of impact

provides valuable information that may moderate intervention effectiveness (see Table 1).

*Reach* is defined as the percent of potentially eligible individuals in the target population who participate in the intervention study, and how representative they are of the population from which they are drawn. Reach is calculated by determining the percent of participants in comparison with a target population, those exposed to recruitment, those responding to recruitment, and those eligible. (See [www.re-aim.org](http://www.re-aim.org) for a description of the calculation). The reach concept can help in setting realistic recruiting goals, tracking recruitment efforts, and increasing the external validity of program evaluation and research reports.

*Efficacy/Effectiveness* is the positive impact of the intervention and its possible unintended consequences on quality of life and related factors. Efficacy and effectiveness often are expressed as the effect size of an intervention on primary outcomes (*e.g.*, physical activity or fitness levels) in comparison with a control condition. *Adoption* is the percent of potentially eligible settings and staff participating in a study and how representative they are of targeted settings and staff. Thus, adoption is parallel to Reach, but at the setting level. Adoption of an intervention is characterized by the (a) absolute number of settings, (b) setting participation rate, and (c) representativeness of the sample of settings and intervention agents (see [www.re-aim.org](http://www.re-aim.org) for calculation formulas and detailed description).

*Implementation* refers to the quantity and quality of delivery of the intervention's various components. Implementation information commonly is included in intervention process evaluations and often is referred to as *treatment fidelity*.

The *Maintenance* dimension includes individual and setting level indices. At the individual level, Maintenance is defined as the longer-term efficacy and effectiveness of an intervention. Outcomes at 6 or more months after intervention contact reflect longer-term individual maintenance. The setting level definition of Maintenance refers to the institutionalization or sustainability of a program and is assessed by the percent of settings that continue the intervention program, in part or in whole, beyond the study duration or initial funding period.

**TABLE 1**  
Definition of RE-AIM dimensions

Dimension	Definition
Reach (individual level)	The absolute number, proportion, and representativeness of individuals who participate in a given initiative compared with those who decline (or to the general population in that area).
Efficacy/effectiveness (Individual level)	The impact of an intervention on physical activity and other health outcomes, including potential negative effects, quality of life, and economic outcomes.
Adoption (setting level)	The absolute number, proportion, and representativeness of settings and intervention providers who are willing to initiate a physical activity promotion program.
Implementation (setting level)	Implementation refers to the intervention provider's fidelity to the various elements of an intervention's protocol, including consistency of delivery as intended and the time and cost of the intervention.
Maintenance setting	The extent to which a physical activity promotion program or policy becomes institutionalized or part of routine organizational practices and policies.
Individual	The long-term effects of a physical activity promotion program on behavior after 6 or more months after the most recent intervention contact.

## CURRENT STATE OF THE LITERATURE

We conducted four reviews of the literature using the RE-AIM framework to critique controlled studies published from 1996 through 2000 in leading public health journals (8). In these reviews, we discussed the status of outcome studies conducted in worksite (1), health care (7), schools (5), and community (3) settings testing interventions promoting physical activity, nutrition, or tobacco cessation in comparison with some type of control or comparison condition.

In these reviews, across 119 studies summarized in Table 2 (23 targeting physical activity), we consistently concluded that health behavior intervention studies seldom have reported on external validity or generalizability information related to individual level impact (reach and representativeness) or setting level impact (adoption). In contrast, there is general consistency in the reporting of individual level impacts relating to internal validity (*i.e.*, efficacy and attrition) are regularly reported across behavioral physical activity, dietary, and tobacco intervention studies, whereas representativeness of the study sample and the intervention setting and delivery staff are reported rarely).

Of most concern is that some data suggest that physical activity studies may report even less often on these issues related to translation than studies in other health behavior areas (5). For example, only 14% of school-based physical activity intervention studies reported on *any* issue related to adoption of the physical activity program by targeted schools. This is in contrast to 40% of school-based smoking prevention studies that reported on adoption issues.

## THE PROBLEM

The lack of reported data on these issues raises questions about the generalizability and feasibility of the studied interventions in typical work-site, health-care, school, and community settings. To illustrate how RE-AIM type data can be collected in the course of most studies, below we provide a

hypothetical example of the problem and then discuss exemplary studies that have provided a report of two RE-AIM dimensions central to the generalizability of research results, Reach and Adoption.

The following hypothetical illustration describing how researchers often apply portions of the RE-AIM framework (Reach, Effectiveness, Adoption, Implementation, Maintenance) demonstrates challenges inherent in translating typical efficacy-based programs into practice (Table 3; Fig. 1). Take, as an example of the RE-AIM perspective, physical activity intervention research undertaken through primary care offices. If all of the primary care offices in the United States offered the program, it potentially could Impact 100% of the population (assuming, unrealistically, that everyone had access to health care). Let us assume that the intervention was moderately Effective in that 40% of participants achieved clinically significant improvements. Therefore, if the program were offered in 100% of primary care offices, approximately 40% of the population could benefit from the program. Next, assume that an unrealistically large 40% of all the health care settings in the United States Adopted this innovation. Because 40% of the patients who attend these offices will benefit (Effectiveness), now only 16% of the potential target population will be impacted (Fig. 1). However, clinicians do not universally adopt the program. Assume generously that 40% of all possible clinicians within these settings attempted the innovation (adoption by staff). Now, only 6.4% of the population potentially will be impacted (Fig. 1). Further, assume that a very encouraging 40% of all patients of these clinicians (Reach) took part in the relatively intensive program—down to 2.6% of the target population. Because of many competing demands, the average clinician is able to Implement only 40% of the rather complex program components, which proportionally reduces the effectiveness of the intervention—at this point, 1% of the population is impacted. Finally, assume that an encouraging 40% of the patients making successful initial changes are able to Maintain these improvements over time. The end result is that less than half of one percent of the target population will actually benefit in a meaningful way from this “evidence-based” intervention (Table 3; Fig. 1).

The point of this exercise is not to induce pessimism about translation, but rather (a) to illustrate the need to attend to all RE-AIM dimensions when selecting interventions for translation, not just to effectiveness of change or effect size, and (b) to demonstrate that if improvements were made along two or more of these other dimensions, the resultant public health benefits could be increased dramatically. As illustrated above and documented elsewhere (1,3,5,7,8), the vast majority of physical activity research has focused on efficacy, largely ignoring other RE-AIM dimensions.

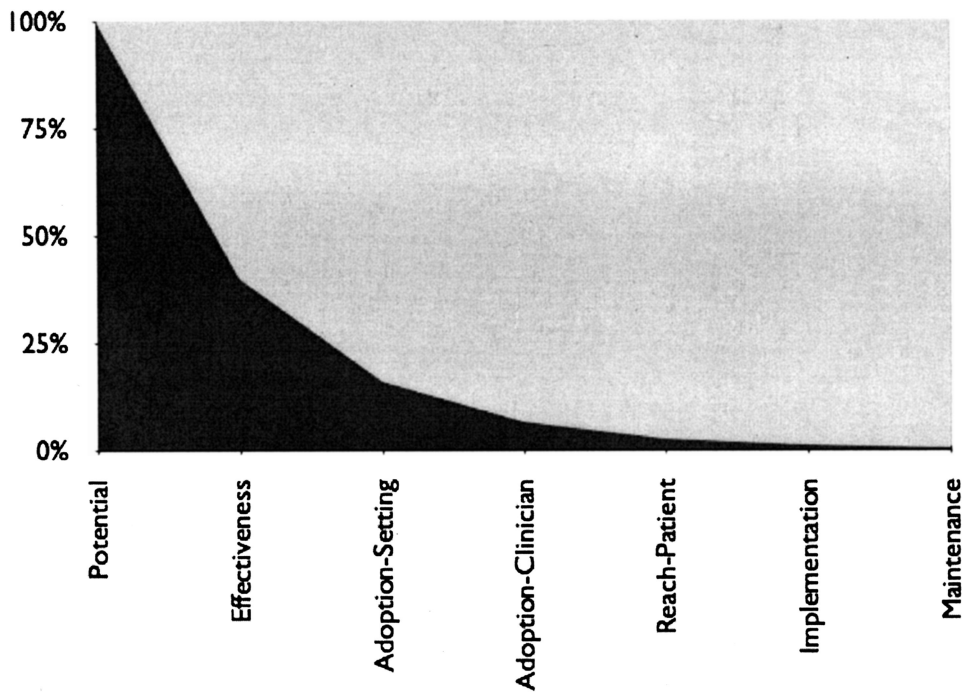
### Reach Example: Individual Level Participation

Physical activity intervention trial studies regularly report sample size and participation rate. For example, the Fitness Arthritis Seniors Trial (6) was designed to examine the impact of physical activity programs on disability in older adults with knee osteoarthritis. The trial involved three different treatment arms of aerobic training, resistance training, or health education for 18 months. This study reported

**TABLE 2**

Percent of studies reporting on RE-AIM dimensions overall and by setting of research (8)

RE-AIM Dimension/Measure	Overall Average
Individual level	
Reach	
Participation rate	76%
Representativeness	14%
Effectiveness	
Behavioral outcome measure	92%
Quality of life or negative outcomes	7%
Maintenance	
>6-month follow-up	36%
Setting level	
Adoption	
Participation rate—site level	16%
Representativeness of settings	2%
Implementation	
Treatment delivery <sup>2</sup>	46%
Time or cost	31%
Maintenance	
Setting continuation	2%



**Figure 1.** Public health impact of hypothetical evidence-based intervention after taking into account Reach, Efficacy/Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) dimensions. ■, Population-Impact; □, Population.

that a total of 841 people, of 4575 contacted directly via telephone, met the eligibility criteria for participation in the intervention. Of this total, 402 declined to participate. This left a total sample size of 439 participants and a 52% study participation rate.

In contrast to the regular reporting of sample size and study participation rate, physical activity intervention trials rarely report the degree to which study samples are representative of the target population with respect to basic demographic information. This is most likely because of the difficulty, which has recently increased with the enactment of the Health Insurance Portability and Accountability Act, of obtaining information on eligible individuals who decline to participate in studies. Often times, such information is simply not available, and when it is available, there are ethical issues to consider (*e.g.*, gaining consent for using the information). Nonetheless, the information is important to determine, or estimate at the very least, because those who choose to participate may be very different in socioeconomic status, age, race or ethnicity, and gender from those who do not. If not feasible in a given study to collect data on actual nonparticipants, it may be possible to use representative health surveys or Behavior Risk Factor Surveillance Survey data to compare the participants with comparable citizens in that geographic locale. If differences do exist, a given intervention may have a differential impact based on these variables that cannot be determined because of the lack of representativeness of the sample. If differences do not exist, then a stronger case for the generalizability of the intervention into real-world settings may be made.

Morey *et al.* (13) provided a good example of determining the representativeness of study participants. Their physical activity intervention was developed to target older adults who were at risk of losing functional independence. To examine representativeness, the 134 participants who agreed to randomization were compared with 100 participants who either declined participation after an introductory telephone

interview (those exposed to recruitment;  $n = 76$ ) or dropped out during subsequent screening ( $n = 24$ ). This comparison revealed that the intervention participants were significantly younger than those who declined. No other differences were identified among the variables that were measured. As such, the authors concluded that, with the exception of age, the study had good generalizability to the greater target population. Note that it is likely that studies that target populations with a large public health impact and with fewer exclusionary criteria will have more difficulty reaching those less motivated and most in need and have poorer rates of reaching a representative sample. Representativeness is important to evaluate because many critics of health promotion and physical activity programs have characterized most programs as reaching those who need it the least.

### Adoption Example: Setting Level Participation

The inclusion of adoption in the RE-AIM framework was influenced by the rationale that evaluating the impact of interventions solely at the individual level (*i.e.*, Reach and Efficacy) is not sufficient. Adoption can be described as the proportion and representativeness of intervention delivery channels (*i.e.*, settings or personnel) that participate in a study. Because different settings (*e.g.*, worksites, medical offices, schools, communities, governing agencies) and intervention staff (*e.g.*, teachers, physicians, health educators) can vary on the amount of resources, level of expertise, time available, competing demands, and commitment to intervention programs, understanding the adoption of interventions into various settings or by different types of staff is critical to the impact of an intervention. However, with the exception of the absolute number of settings involved, researchers seldom report on issues of setting-based adoption of the intervention. For example, most school-based health behavior intervention studies do not report the participation rate of the schools involved in the study (*i.e.*, the number of schools that chose to enroll in



**TABLE 3**  
Stages of translating an efficacious program into real-world settings—example

Issue	RE-AIM Dimension	Results—Multiplier	Population-wide Impact
Exciting evidence-based program			100%
Potential program results	Effectiveness (on main outcome)	0.4	40%
Clinic participation rate	Adoption	0.4	16%
Clinician participation rate (within clinics)	Adoption (part 2)	0.4	6.4%
Patient participation rate	Reach	0.4	2.6%
Intervention delivery fidelity	Implementation	0.4	1%
Longer-Term effects	Maintenance (individual level)	0.4	0.4%

the study divided by the number approached for participation), and fewer still report any comparisons between schools that participate and those that do not.

A notable exception was the Sports, Play & Active Recreation for Kids (SPARK) trial that targeted elementary students (14). In this study, the principals of 16 elementary schools were approached for participation in the trial. Twelve of the 16 schools were willing to participate; however, because economic restraints, only 7 of the 12 were selected for study participation. Although there were no tests of representativeness of school resources, location, staff-to-student ratio, or other school-level variables, the researchers did document that the seven smallest schools were selected for participation. Based on this information, one could conclude that the effects of the intervention could generalize to other small schools with similar resources, but not to larger schools.

## THE FUTURE

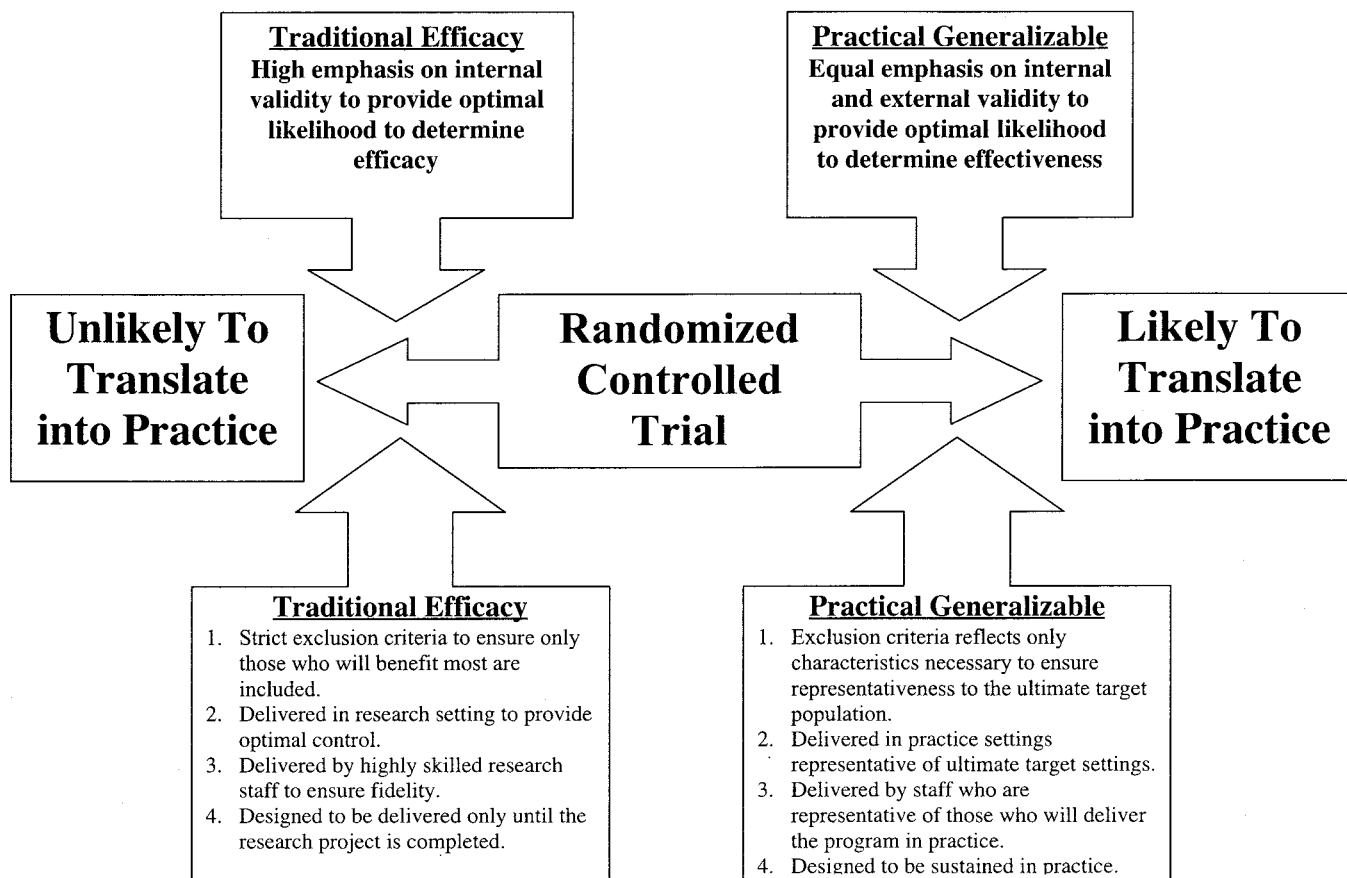
The middle column of Table 4 summarizes the type of evidence-based research on which guidelines and clinical recommendations are typically based, be they for physical activity promotion (11) or other preventive services. This table uses the RE-AIM framework to describe characteristics of such efficacy studies—usually randomized controlled trials—that most of us have been trained to conduct that maximize internal validity and rule out alternative explanations of intervention effects (9). The right-hand column illustrates a different type of study that is seldom conducted, but that would provide significantly more information for translation into practice and

that would answer many more questions that clinicians and decision makers have about programs—a more “practical clinical trial,” as leaders of the Centers for Medicare & Medicaid Services and the Agency for Health-care Research and Quality have recently called for (15).

The efficacy-type study usually has a very limited Reach—it relies on motivated volunteers to self-select and excludes patients having other medical conditions or complex cases (in many clinical trials, the exclusion rate can be as high as 95% to maximize internal validity). In contrast, a study that applies greater portions of the RE-AIM framework would explicitly keep exclusion criteria to a minimum and attempt to recruit as wide a range of participants as possible (10). Effectiveness (or more technically correct, efficacy) studies have strong internal validity—the intervention is usually implemented by expert research staff under controlled conditions. Such interventions often very are intensive and time consuming because a number of studies indicate that intensity is related to stronger effect sizes. In contrast, the translation study is likely to produce somewhat weaker effects because a variety of setting and implementation factors may influence the intervention, which is likely to be less intensive, and the patients are less selected and likely less motivated. Another important lesson related to efficacy or effectiveness to be learned from evidence-based behavioral medicine is also to assess potential unintended or negative consequences of a program (e.g., injuries, or decreased quality of life). As illustrated in numerous reports of usual practices (12), Adoption of most efficacy-based interventions is very low. There are numerous reasons for this (9), but the translation-type research programs that require fewer re-

**TABLE 4**  
Distinctive characteristics of efficacy and ‘translation’ intervention studies, using RE-AIM dimensions

RE-AIM Dimensions/Issue	Efficacy Study	Translation Study
Reach	Low: homogeneous, highly motivated sample; excludes those with complications, other comorbid problems	High: broad, heterogeneous, representative sample; often using a defined population
Efficacy or effectiveness	Intensive, specialized interventions that attempt to maximize effect size; very standardized; randomized designs	Brief, feasible interventions, not requiring great expertise; adaptable to setting; randomized, time series, or quasiexperimental designs
Adoption	Usually one setting to reduce variability; for settings with many resources and expert staff	Appeals to and works in multiple settings; adaptability to fit setting; tested in variety of settings
Implementation	High: by research staff closely following specific protocol	Lower: by variety of different staff with competing demands, using adapted protocol
Maintenance and cost	Often not an issue or concern; focus on individual level; both relapse rates and costs often are high	Major issues; focus on institutionalizing program; setting level maintenance equally important as individual



**Figure 2.** Designing randomized controlled trial studies to improve the translation of research into health promotion practice.

sources, demand less staff expertise, take less time, and allow greater flexibility and adaptation should produce significantly higher adoption rates and be more generally applicable. Recently, Tunis *et al.* (15) proposed that there should be a consistent effort to conduct clinical trials to meet the needs of decision makers. They defined practical clinical trials as studies for which the hypothesis and design are developed specifically to meet the needs of decision makers. It is likely that individuals making decisions regarding the adoption of evidence-based physical activity promotion programs will not have the option to choose interventions that are very intense and delivered by highly trained personnel in controlled settings.

Finally, Maintenance effects are more speculative; often substantial relapse is observed in efficacy type studies when the intensive intervention ends and external supports are withdrawn. To the extent that translation-type studies truly incorporate their interventions into usual care (*e.g.*, through family practice visits), one may expect better maintenance.

The discussion above is admittedly somewhat of a “straw program” debate because most programs fall somewhere between the two examples in Table 4 (see Fig. 2). It is intended, however, to focus attention on five of the key issues that need much greater attention if we are serious about moving the fruits of our research into real-world settings. It is not our intent to criticize efficacy-type studies or to say that they have no place, but rather to point out that the literature is predominantly of this type of study (1,3,5,7,8,11) and that reporting ONLY this

type of study may itself be a barrier to translation (9). Many have argued that there should be a progression of research from basic laboratory research on mechanisms to efficacy studies to effectiveness and finally to dissemination research (for a review, see (9)). Although such translation stage models are no longer used by National Institutes of Health institutes, they were popularized especially by the National Cancer Institute in earlier years. From a RE-AIM perspective, there IS value in basic and in tightly controlled (and less generalizable) efficacy studies. However, if the goal is to produce an intervention that is practical (15) and feasible to apply in a broad cross-section of settings and by a variety of staff, such factors need to be considered when planning an intervention and study design—in efficacy as well as effectiveness studies. Otherwise the “gap” between efficacy research and effectiveness studies—and the changes that need to be made from the original efficacy intervention—become so daunting that this “efficacy to effectiveness transition” is never made (9).

## CONCLUSIONS

We conclude that variables at the individual and setting level, likely to moderate the impact of interventions, are seldom studied or reported. Our reviews of the literature have supported this thesis. Would there be an increase in the translation of research to practice if researchers conducted more studies of “practical interventions” in representative

settings and reported both internal and external validity information? RE-AIM provides a framework to address this issue and should be taken into account by editors, reviewers, researchers, and practitioners (1, 3, 5, 7, 8).

## Acknowledgments

The authors have been supported for this work by a grant from the Robert Wood Johnson Foundation. From our work with the NIH Behavior Change Consortium and with funding from the Robert Wood Johnson Foundation, we have developed a website, WWW.RE-AIM.ORG, that targets both researchers and community leaders to help improve the translation of research into practice. The website includes a description of RE-AIM and resources to support the model.

## References

1. Bull, S.S., C. Gillette, R.E. Glasgow, and P. Estabrooks. Worksite health promotion research: To what extent can we generalize the results and what is needed to translate research to practice? *Health Educ. Behav.* 30:537–549, 2003.
2. Committee on Health Care in America, Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21<sup>st</sup> Century*. Washington, D.C.: National Academy Press, 2001.
3. Dziewaltowski, D.A., P.A. Estabrooks, L.M. Klesges, S.S. Bull, and R.E. Glasgow. Behavior change research in community settings: How generalizable are the results? *Health Promot. Internation.* 19:235–245, 2004.
4. Ennett, S.T., C.L. Ringwalt, J. Thorne, L.A. Rohrbach, A. Vincus, A. Simons-Rudolph, and S. Jones. A comparison of current practice in school-based substance use prevention programs with meta-analysis findings. *Prev. Sci.* 4:1–14, 2003.
5. Estabrooks, P., D.A. Dziewaltowski, R.E. Glasgow, and L.M. Klesges. Reporting validity from school health promotion studies published in 12 leading journals, 1996–2000. *J. Sch. Health.* 73:21–28, 2003.
6. Ettinger, W.H., R. Burns, S.P. Messier, W. Applegate, W.J. Rejeski, T. Morgan, S. Shumaker, M.J. Berry, M. O'Toole, J. Monu, and T. Craven. A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. The Fitness Arthritis Seniors Trial (FAST). *JAMA.* 277:25–31, 1997.
7. Glasgow, R.E., S.S. Bull, C. Gillette, L.M. Klesges, and D.A. Dziewaltowski. Behavior change intervention research in health care settings: A review of recent reports, with emphasis on external validity. *Am. J. Prev. Med.* 23:62–69, 2002.
8. Glasgow, R.E., L.M. Klesges, D.A. Dziewaltowski, S.S. Bull, and P. Estabrooks. The future of health behavior change research: What is needed to improve translation of research into health promotion practice? *Ann. Behav. Med.* 27:3–12, 2004.
9. Glasgow, R.E., E. Lichtenstein, and A. Marcus. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy to effectiveness transition. *Am. J. Public Health.* 93:1261–67, 2003.
10. Glasgow, R.E., T.M. Vogt, and S.M. Boles. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *Am. J. Public Health* 89:1322–1327, 1999.
11. Kahn E.B., L.T. Ramsey, R.C. Brownson, G.W. Heath, E.H. Howze, K.E. Powell, E.J. Stone, M.W. Rajab, and P. Corso. The effectiveness of interventions to increase physical activity. A systematic review. *Am. J. Prev. Med.* 22:73–107, 2002.
12. McGlynn, E.A., S.M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E.A. Kerr. The quality of health care delivered to adults in the United States. *N. Engl. J. Med.* 348:2635–2645, 2003.
13. Morey, M.C., M. Schenkman, S.A. Studenski, J.M. Chandler, G.M. Crowley, R.J. Sullivan, et al. Spinal-flexibility-plus-aerobic versus aerobic-only training: Effects of a randomized clinical trial on function in at-risk older adults. *J. Gerontol. A. Biol. Sci. Med. Sci.* 54:M335–M342, 1999.
14. Sallis, J.F., T.L. McKenzie, J.E. Alcaraz, B. Kolody, N. Faucette, and M.F. Hovell. The effects of a 2 year physical education program (SPARK) on physical activity and fitness in elementary school students. *Am. J. Public Health.* 87:1328–1334, 1997.
15. Tunis, S.R., D.B. Stryer, and C.M. Clancy. Practical clinical trails: Increasing the value of clinical research for decision making in clinical and health policy. *JAMA.* 290:1624–1632, 2003.