

# RE-AIMing Research for Application: Ways to Improve Evidence for Family Medicine

Russell E. Glasgow, PhD

**Objective:** To outline changes in clinical research design and measurement that should enhance the relevance of research to family medicine.

**Methods:** Review of the traditional efficacy research paradigm and discussion of why this needs to be expanded. Presentation of practical clinical and behavioral trials frameworks, and of the RE-AIM model for planning, evaluating, and reporting studies.

**Results:** Recommended changes to improve the external validity and relevance of research to family medicine include studying multiple clinical practices, realistic alternative program choices, heterogeneous and representative patients, and multiple outcomes including cost, behavior change of patients and staff, generalization, and quality of life.

**Conclusions:** The methods and procedures discussed can help program planners, evaluators and readers of research articles to evaluate the replicability, consistency of effects, and likelihood of widespread adoption of interventions. (J Am Board Fam Med 2006;19:11–9.)

Family medicine is by nature pragmatic and contextual. It deals with making practical decisions on complex and multiple issues in ways that are congruent with family values and situations.<sup>1,2</sup> Like other areas of medicine, it is also adopting evidence-based medicine as a key feature of its current and future direction.<sup>1</sup> Unfortunately, the evidence available for family medicine often does not address the above issues. Most available evidence comes from studies that attempt to rule out threats to validity by studying isolated issues and controlling or standardizing contextual factors.<sup>3,4</sup> This creates a gap between the available evidence and the situations and context in which the evidence needs to be applied.

This article has 2 primary purposes. First, it discusses how future primary care research might be “Re-Aimed” to be more relevant and practical. Second, it provides a series of questions that family physicians can ask to determine the applicability of

research reports to their setting and to help plan primary care programs that have broad impact.

## *Why Change and What Might Be Changed?*

The gap between research and practice has been extensively documented<sup>5,6</sup> and is increasingly the topic of meetings and initiatives.<sup>7–9</sup> However, few projects have addressed one of the fundamental causes of the gap between research and practice: Many family physicians and health system decision makers do not see much of the available research evidence as applicable to their setting. Specific issues concern the types of patients, settings, and resources available (including time), and outcomes assessed.

Table 1 summarizes key differences between the traditional “efficacy study” evidence most often available, and the evidence from practical effectiveness studies needed to help integrate research into practice.<sup>4</sup> There are both conceptual and philosophy of science differences, and methodological/design differences between available efficacy studies and those that are needed to inform family medicine. As shown, the traditional efficacy approach attempts to maximize internal validity by isolating causes so that treatment or theoretical mechanisms can be identified. In contrast, practical effectiveness studies aim to identify widely applicable, replicable programs that will work in a variety of different contexts. For heuristic purposes, Table 1 presents efficacy and effectiveness studies as a

Submitted 12 July 2005; revised 24 August 2005; accepted 29 August 2005.

From Kaiser Permanente Colorado, Denver, CO

This article is based on a presentation made at the 2005 Convocation of Practices, hosted by the American Academy of Family Physicians National Research Network and the Federation of Practice-based Research Networks, Colorado Springs, CO, March 2005.

*Conflict of interest:* none declared.

*Corresponding author:* Russell E. Glasgow, PhD, Kaiser Permanente Colorado, 335 Road Runner Lane, Penrose, CO 81240 (E-mail: russg@ris.net).

**Table 1. Purpose, Intent, and Elements of Traditional Efficacy Studies and Practical Effectiveness Trials**

Issue	Traditional Efficacy Study	Practical Effectiveness Trial
I. Purpose and intent		
Design goal	Isolate treatment mechanism; unique approach	Replicable treatment
Key criteria	Internal validity	External validity
Purpose/world view	Isolate cause	Understand program in context
Decisions regarding	Theory	Practice and policy
II. Study elements		
Participants	Homogeneous, motivated	Heterogeneous, representative
Settings	Homogeneous, single or few under control of investigator	Heterogeneous, multiple to evaluate generalization across settings
Intervention	Intensive, highly structured, complex	Less intensive, moderate structure, and flexible to permit options
Intervention staff	Research staff or highly trained experts	Regular staff in representative settings

dichotomy, whereas in reality, there is a continuum of research designs and many studies have elements of both efficacy and effectiveness research. Because of the preponderance of studies toward the efficacy end of the continuum in the literature and which form the basis for current practice recommendations, this paper focuses on how we might change such designs. It is recognized that not all studies need to be “complete” effectiveness studies, but movement in this direction would generally enhance the relevance of research data.

Design differences between efficacy and effectiveness studies impact the inferences that can be made from a given study. To maximize internal validity and chances of finding a treatment effect, efficacy studies tend to recruit homogeneous, highly motivated patients to participate in highly structured, intensive interventions that are conducted in one or a few settings. In contrast, effectiveness studies focus on heterogeneity and representativeness of both patients and settings, and emphasize interventions that are more flexible to address unique issues.

As Larry W. Green has said, “If we want more evidence-based practice, then we need more practice-based evidence.”<sup>10</sup>

**Practical Clinical Trials**

Tunis et al,<sup>11</sup> have proposed criteria for “practical clinical trials,” that should also be more relevant for family medicine. There are 4 key characteristics of practical trials. They study representative patients; are conducted in multiple settings; use reasonable alternative intervention choices as controls rather than placebos or “usual care;” and report on out-

comes relevant to patients, clinicians, potential adoptees, and policy makers.<sup>11</sup>

Tunis et al,<sup>11</sup> recommend having diverse samples of both patients and clinical settings. In particular, they recommend using few exclusion criteria so that the complex, comorbid cases seen in primary care are included. Their recommendations for inclusion of community practice settings are very compatible with practice-based research network research approaches within primary care.<sup>12</sup> Inclusion of a variety of different settings also permits investigation of variations in both processes and outcomes of care. In particular, it is recommended that studies include practices in small, rural, mixed payer and safety net settings as well as those that are part of larger health systems.

The third characteristic of practical clinical trials is that they use realistic alternative treatment comparisons—not just no treatment, placebo, or “usual care.”<sup>11,13</sup> The rationale for this is that clinicians and policy makers need to make decisions among alternative interventions, and including direct comparisons provides more valuable information on intervention strengths and weaknesses than does just knowing that a number of alternative treatments are each better than placebo or no treatment.

Tunis et al,<sup>11</sup> stress that it is important to collect multiple outcomes. Family medicine investigations could accelerate translation if more studies would collect the types of measures discussed below. My colleagues and I have proposed a comprehensive, yet feasible, set of measures summarized in Table 2.<sup>14,15</sup> The first 4 measures listed can be collected without adding any burden to patients. Contextual factors and moderating variables are important de-

**Table 2. Proposed Translational Research Measurement Package**

Contextual factors
Implementation consistency
Generalization (reach, adoption, maintenance)
Economic outcomes
Behavior change (including clinical staff behavior, such as quality of care delivery, if applicable)
Biologic change or clinical outcomes
Quality of life

terminants of intervention outcomes. Information on practice characteristics can add a great deal to understanding how interventions work—or do not work. The closely related issue of how interventions are implemented in practical trials is also critical to interpretation. Often not all components of a program are delivered equally, and some elements in complex programs may be delivered differently over time. In addition, it is important to understand the breadth of application or the facets of patients, practices, time and contexts across which study results generalize.<sup>16,17</sup> Program effectiveness often varies across settings and subgroups of users. We need to report on such contextual effects, especially those related to health disparities.

#### *Cost and Economic Measures*

One of the greatest evaluation needs is for more systematic collection of economic measures. Until more is known about the costs and cost-effectiveness of new interventions, it is unreasonable to expect decision or policy makers to adopt or reimburse such programs. Admittedly, comprehensive economic analyses that determine outcomes such as cost-benefit or cost offsets<sup>18,19</sup> require considerable time and expertise, and may be beyond the scope of many primary care studies. However, it should be feasible for almost all projects to collect measures of intervention costs, and to estimate what Meenan et al,<sup>20</sup> have called replication costs, which estimate what it would cost to deliver the intervention in other settings. One caveat regarding economic measures is that “costs are not costs are not costs.” Thus, potential adoptees may want to see a breakout of costs by category, because many settings have different budgets for upfront versus gradually accrued costs, and for equipment or software versus personnel costs.

The last 3 measures in Table 2 do require patients’ involvement, but are essential to evaluating outcomes

and for decision making. Measures of behavior change, biological changes or clinical outcomes, and quality of life (and/or potential negative outcomes) are recommended. Clinical outcomes are widely accepted and almost universally included in primary care studies and will not be discussed further.

#### *Behavior Change*

Because the intent of many primary care interventions is to assist patients in changing their health behaviors (eg, exercise more, stop smoking, take medication regularly) or to have staff change their care behaviors, it is important to directly assess behavior change. It is not sufficient to simply measure knowledge or biological outcomes, and assume that behavior change occurs.<sup>15,21</sup> Although there are usually linkages among these measures, knowing what happened on one outcome does not necessarily permit inference about results in other domains.

A major challenge to collecting behavioral outcomes has been the length of assessments required. Two relatively recent developments have combined to change this situation. First, brief forms of measures have been developed that perform almost as well as longer forms.<sup>22–24</sup> Second, when the intent of a measure is to assess intervention effects, the most relevant criterion for selecting a measure is its sensitivity to change: it does not necessarily need to have extraordinarily high levels of internal consistency (often obtained by having lengthy surveys). Glasgow, Ory, Klesges et al,<sup>23</sup> have recently recommended measures for dietary change, physical activity, risky drinking, and smoking that should be sufficiently sensitive, yet brief enough to be used in primary care interventions requiring brief measures. A rule of thumb recommendation is that an instrument should be capable of detecting an intervention effect, if present, with 100 patients per condition.

#### *Quality of Life and Potential Adverse Effects*

There are multiple reasons to recommend collection of quality-of-life measures. The first is that well-validated, quality-of-life measures provide a common metric on which to compare interventions for different problems and different behaviors. Several authors have argued that improving quality of life is the ultimate goal of health care.<sup>25,26</sup> Especially if quality of life can be converted to quality adjusted life years,<sup>19</sup> it provides a convenient and widely understood metric for comparing diverse programs. There are now several well-validated, brief quality-of-life measures,

such as the WHO-5<sup>27</sup> and the CDC Healthy Days measures<sup>28</sup> that are sensitive to change and appropriate for diverse cultural groups.

Quality-of-life measures can also evaluate whether a program inadvertently creates adverse outcomes or unintended consequences. It is now apparent that many health care interventions have created unintended adverse consequences.<sup>29</sup> We cannot assume that because programs were well intended, that they will not cause harm. Quality-of-life measures can assess whether an intervention does more harm than good. Given limited time and the competing demands faced by both patients and health care providers,<sup>30,31</sup> devoting greater attention to one health risk factor may mean doing less of some other valuable activity. Primary care programs, especially those not collecting quality-of-life measures, may want to collect measures of nontargeted health behaviors or of HEDIS items<sup>32</sup> to ensure that quality of care in nontargeted areas is not adversely affected.

Summarizing this section, practical studies should: study representative patients and settings (clinics); use comparison conditions that include alternative interventions (especially if one wants to claim that their program is superior to existing programs); collect a broad range of measures (Table 2); and present those results in a way that is understandable to decision makers and stakeholders.<sup>11,13,14</sup>

### **RE-AIM Evaluation Framework**

For innovators who wish to have their program widely adopted, it can be helpful to follow a translation framework throughout the planning, implementation, analysis, reporting, and refinement of their product. It is beyond the scope of this paper to discuss the relative advantages of different frameworks,<sup>33–36</sup> but almost all are influenced by the pioneering work of Rogers' Diffusion of Innovations model,<sup>35</sup> and of Green and Kreuter's PRECEDE-PROCEED model.<sup>33</sup>

This article discusses implications and recommendations from the RE-AIM framework.<sup>36,37</sup> RE-AIM is an acronym that stands for Reach (participation rate and representativeness of participants); Effectiveness (on both primary outcomes and quality-of-life/negative consequences); Adoption (participation rate and representativeness among settings and staff that begin or attempt a program); Implementation or program delivery, and Maintenance or sustainability at both patient and setting levels ([www.re-aim.org](http://www.re-aim.org)). Each dimension is impor-

tant for determining the eventual population-based impact of a program, and different interventions probably have different patterns of results across these 5 dimensions.<sup>37,38</sup> For example, a simple mail-based program encouraging patients to take a preventive action (eg, go for cancer screening) will probably have high reach and be widely adopted by many offices, but by itself have limited effectiveness. In contrast, a more intensive, multisession medical group visit intervention that requires patients to return repeated times would probably have lower reach, might be adopted by fewer practices (because of cost and complexity), but will probably be more effective for those who participate.

Different clinicians may wish to emphasize one RE-AIM dimension over others or to make adoption decisions based on dimension(s) most important to their practice. However, it would be helpful to have a composite index to summarize the public health impact of different programs. At the individual user level, overall program impact may best be conceptualized as a product of the Reach of a program multiplied by its Effectiveness.<sup>39,40</sup> Reach is a function of both the participation rate and the representativeness of those users.<sup>41</sup> Effectiveness is a function of multiple components, including: the median effect size on primary outcome(s) for a given program (effect size serves as a common metric across diverse content areas); adjusted for any adverse impacts on quality of life or other outcomes; and for differential impact across population subgroups,<sup>41</sup> with special reference to groups identified in health disparities research.<sup>42</sup> Most decisions are influenced not only by the overall impact of a treatment, but also by its cost. Therefore, based on reasoning by Green and Kreuter,<sup>33</sup> an "Efficiency Index" is calculated as the cost of an intervention divided by its composite Individual Impact score.<sup>41</sup>

The RE-AIM framework considers results not only at the individual level, but also at the setting (clinic) level. Setting level impact is determined by the number and types of practices that will attempt an innovation (adoption) and how well they deliver the innovation (implementation). A Summary Setting Level Impact score is calculated by multiplying Adoption times Implementation,<sup>41</sup> parallel to the Reach times Effectiveness score at the individual level. Adoption is a function of both the participation rate among settings as well as the representativeness of these settings (eg, do low resource practices and rural clinics participate in equal rates to other settings?). Implementation is a composite

**Table 3. Implications and Example of Application of RE-AIM Framework**

Goal: To identify interventions that can:	Example Smoking Study*	Example Diabetes Study†
Reach large numbers of people, especially those who can most benefit.	76% participation among low income, young female smokers. Participants appeared representative.	50% participation among primary care diabetes patients. Participants representative on key variables.
Be effective in producing targeted outcomes at reasonable cost and produce minimal negative impacts, relative to alternatives.	11% vs. 7% cessation at 6 weeks, $P < .05$ Quality of life and/or adverse consequences were not measured. Intervention time reported but costs not calculated.	Significant improvement on both preventive assessments and behavioral counseling aspects of care. Both conditions improved on quality of life. Intervention costs estimated at \$222 per patient.
Be widely adopted by many types of settings.	4 of 4 clinics with most diverse populations in the metro area.	6% of family medicine and internal medicine physicians throughout Colorado. Those participating were representative on variety of practice characteristics.
Be consistently implemented by staff members with moderate levels of training and expertise.	85%–100% implementation by usual clinical staff for all treatment components except calls (43%).	97% completion of key intervention components.
Produce replicable and long-lasting maintenance effects.	Cessation differences (11.6% vs. 8.5%) no longer significant.	Effectiveness measures were of equal magnitude and significance at 12-month follow-up as initial 6-month assessment.

\* Information from Glasgow et al.<sup>45</sup>

† Information from Glasgow et al.<sup>46</sup>

variable that reflects both the median level of implementation of different components of an intervention, and consistency of delivery across different settings and staff.<sup>41</sup>

### Implications

There are at least 3 implications from the RE-AIM framework for family medicine research. The first is that representativeness is important at multiple levels—patient, health care team, and organizational setting. Although representativeness has been largely ignored at the setting level,<sup>43</sup> it is just as important as patient level representativeness.

The second message from the RE-AIM framework, is to remember the “3 Rs” of translation and dissemination research: representativeness, robustness, and replicability. Representativeness has been covered above, but the other Rs deserve further comment. Robustness, or generalization of effects, is important from health disparities, methodological, and program understanding perspectives. For more detail, see Cronbach et al,<sup>16</sup> who refer to generalizability across persons, time, measures, situations, and program modifications, and Shadish, Cook, and Campbell.<sup>17</sup>

Replicability refers to whether the results of a program can be duplicated in settings in addition to

those in which they are originally produced. Replication is an important, but often under-emphasized criteria for strength of evidence.<sup>44</sup> It also helps to ensure that findings are not restricted to a unique practice, set of physicians, or context.

Finally, it is recommended that future family medicine research focus on identifying interventions that (Table 3): *reach* large and representative numbers of patients, especially those who are most in need or are underserved; are *effective* and produce minimal negative impacts at reasonable cost; are widely *adopted* across settings, especially those having fewer resources; are consistently *implemented* and do not require staff with high levels of expertise; and produce replicable, long-lasting effects for patients, and are sustainable at the practice level.

### Research Challenges and Examples

The reader may be thinking, “These issues are worth considering, but is it really feasible to integrate all them into a typical study, and without a huge budget?” The answer is “yes”: it is possible. Many of the recommendations above, such as specifying denominators of settings and patients approached, tracking costs, and analyzing representativeness and robustness require few financial resources and do not involve any patient burden. They can be addressed by

simply doing a systematic job of keeping project records. Other issues such as assessing patient quality of life and nontargeted behaviors do require additions to typical assessment batteries. The payoff from the ability to answer questions critical to decision makers should be well worth the added items required, especially now that brief, validated scales are available. The following section illustrates how many of the RE-AIM issues can be integrated into practical effectiveness studies. These examples are both drawn from the author's area of specialization, health behavior change; but the RE-AIM concepts and recommendations should apply across most types of intervention.

Glasgow et al,<sup>45</sup> conducted a randomized effectiveness study of a brief primary care-based smoking cessation intervention using the RE-AIM model. The settings were 4 Planned Parenthood clinics having the most diverse populations in the Portland, Oregon, metro area (all 4 clinics invited participated and were involved in program planning). Participants were low-income female smokers who were attending either general primary care or contraceptive visits. Table 3 summarizes the results along RE-AIM dimensions.

All current smokers were invited to participate, regardless of intent to quit and 76% participated. Participants were representative of smokers in these clinics and less than one-third intended to quit in the next month. The intervention consisted of a brief written assessment of readiness to quit and barriers to cessation, watching a 9-minute video developed for the project, clinician advice to quit, brief cessation counseling by regular clinical staff trained in motivational interviewing, and follow-up phone calls. All intervention components were implemented with 85% to 100% of participants, with the exception of phone calls (43%), which were challenging to complete.

A 6-week follow-up assessment by independent assessors found that more smokers had quit in the intervention than in the randomized advice-to-quit only control condition (11% vs 7% cessation,  $P < .05$ ). At 6-month follow-up, both intent-to-treat analyses of self-reported quits (11.6% vs 8.5%) and biochemically confirmed cessation favored intervention, but were no longer significant. This study did not assess cost (the intervention took approximately 15 minutes of total time from nurse practitioners, medical assistants, or physician assistants) or quality of life. No adverse outcomes were noted (HEDIS measures of other preventive services

were not collected, however), and participants in the intervention condition who did not quit smoking had greater reductions in smoking rate than those in the control condition,<sup>45</sup>  $P < .05$ .

The third column of Table 3 summarizes the results of a randomized, quality improvement study that used RE-AIM to evaluate outcomes among 886 type 2 diabetes patients of 52 primary care physicians throughout Colorado.<sup>46</sup>

All type 2 patients in these practices were sent a letter from their physician inviting them to participate in the computer-assisted diabetes care project during their next regular office visit. The intervention involved a touchscreen computer session to inform patients about diabetes care recommendations and identify areas they wished to discuss with their physician. The physician, as well as the patient, received a print-out highlighting overdue services and issues the patient wished to discuss. Finally, the physician encouraged the patient to meet with a care manager to go over plans for needed preventive services and to assist with the computer-generated behavior change plans as relevant for smoking, healthy eating, and exercise.

This study produced moderately high levels of Reach, with 50% of patients participating, including those who were older, had multiple comorbidities, and were Hispanic. Implementation was excellent with regular office staff completing on average 97% of intervention activities, and the intervention was effective in significantly improving both laboratory assessment and self-management counseling aspects of NCQA/ADA provider recognition care criteria relative to a computer-assisted assessment and health risk feedback condition.<sup>46</sup> Both conditions improved on quality of life, and there were no apparent negative effects of the program.

We estimated that the program cost \$222 per participant, and \$57 more per participant than the health risk feedback comparison condition. The downside was that only 6% of physicians invited to participate took part (Adoption), probably because of the need to incorporate the program into their office procedures. Finally, maintenance of the program was generally good with quality of care and implementation results being almost as good at a 12-month follow-up as at the initial 6-month assessment.

As the RE-AIM approach is a relatively new development, there has not been sufficient time for many publications to appear that use it. The national WISEWOMAN project, which is using clin-

**Table 4. RE-AIM Questions to Ask and Ways to Enhance Overall Impact**

RE-AIM Dimension	Questions to Ask	Ways to Enhance Impact
Reach (individual level)	What percentage of the target population comes into contact? Does program reach those most in need? Will participants be representative of the patients in your practice?	Formative evaluation with potential users with those declining Small-scale recruitment studies to test methods Identify and reduce barriers to participation Use multiple channels of recruitment
Effectiveness (individual level)	Does program achieve greater key targeted outcomes compared with other treatments? Does it produce unintended adverse consequences? How will or did it impact quality of life (QoL)?	Incorporate more tailoring to individual Reinforce via repetition, multiple modalities, social support, and systems change Use stepped-care approach; less intensive intervention first Evaluate adverse outcomes and QoL for program revision
Adoption (setting/organizational level)	Will organizations having underserved or high-risk populations use it? Does program help the organization address its primary mission?	Conduct formative evaluation with adoptees and settings that decline Recruit settings that have most contact with target audience Provide different cost options and customization of intervention Develop recruitment materials outlining program benefits and required resources
Implementation (setting/organizational level)	How many staff within a setting will try this? Can different levels of staff implement the program successfully? Are different components delivered as intended?	Provide delivery staff with training and technical assistance Provide clear intervention protocols Consider automating all or part of the program Monitor and provide staff feedback and recognition for implementation
Maintenance (individual and setting levels)	Does the program produce lasting effects at individual level? Can organizations sustain the program over time? Are those persons and settings that show maintenance those most in need?	Reduce level of resources required; make contacts extensive, not intensive Incorporate “natural environmental” and community supports Conduct follow-up assessments and interviews to characterize success at both levels Incorporate incentives and policy supports

ic-based approaches to improve the health of low-income women,<sup>47</sup> has used RE-AIM. They have successfully operationalized the RE-AIM constructs using both quantitative and qualitative methods. Partially in response to their concerns, the efficiency index has been added to incorporate costs into the model. The most common challenge that persons attempting to apply the RE-AIM model seem to have concerns what to do when “denominators” are not known for calculating Reach and Adoption. The [www.re-aim.org](http://www.re-aim.org) website provides detailed suggestions for such cases, but in general it is usually possible to estimate denominators and characteristics of intended target populations when not known using secondary data such as census information, large scale health surveys, state or/ local Behavioral Risk factor Surveillance Survey data or increasingly, GIS data bases.

**Questions to Ask and Conclusions**

The paper concludes with recommendations for a) questions for practitioners to ask when they are

reading research reports or planning their own programs, and b) strategies to enhance program impact pertinent to each of the RE-AIM dimensions.

Asking the questions in the middle column of Table 4 should help to determine the relevance of a research report to one’s setting, and to conduct a self-assessment of planned programs for one’s clinic. If answers to these questions are “no,” readers may want to consider one or more of the strategies in the right-hand column to enhance that RE-AIM dimension. These questions and recommendations are intended to summarize this article, and comments are made only on issues not previously discussed.

Given continuing health disparities,<sup>42</sup> programs should be evaluated for their reach among low-income, racial and ethnic minority, and low-health literacy patients. In terms of effectiveness, it is especially important to assess unintended consequences of practice changes.<sup>48–50</sup> In addition, stepped-care approaches that apply low-intensity interventions for all patients, and reserve more costly and intensive intervention for those who still need additional assistance

might be considered.<sup>40</sup> When evaluating program adoption, it is useful to consider the compatibility of the program to the mission and values of each clinic and its staff,<sup>35</sup> as well as logistic issues such as impact on patient flow.

Consistency of implementation in busy primary care clinics is a major challenge and practices should consider sharing the load across all clinic staff rather than the physician trying to do everything him or herself. Ways to provide automated prompts and reminders, or interactive programs for patients that inform the patient-clinician interaction but do not take additional time, should be considered.<sup>30,51</sup> Maintenance at both the practice and patient level can often be enhanced by establishing strong and reciprocal linkages to community resources relevant to patients' social environment.<sup>52,53</sup>

In conclusion, patient health behaviors, and family medicine, are complex, contextual, and multiply determined. Our programs, questions, and research designs should also incorporate these characteristics if they are to help integrate research and practice.

## References

1. Institute of Medicine. Primary care: America's health in a new era. Washington (DC): National Academy of Sciences; 1996.
2. Starfield B. Primary care: balancing health needs, services, and technology. Oxford University Press; 1998.
3. Flay BR. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med* 1986;15:451-74.
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. *Am J Public Health* 2003;93:1261-7.
5. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003;348:2635-45.
6. Institute of Medicine, Committee on Quality of Health Care in America. Crossing the quality chasm: a new health system for the 21st Century. Washington (DC): National Academy Press; 2001.
7. Zerhouni E. Medicine. The NIH Road Map. *Science* 2003;302:63-72.
8. Finney JW, Willenbring ML, Moos RH. Improving the quality of VA care for patients with substance-use disorders: the Quality Enhancement Research Initiative (QUERI). *Med Care* 2000;38(6 Suppl):1105-13.
9. Farquhar CM, Stryer D, Slutsky J. Translating research into practice: the future ahead. *Int J Qual Health Care* 2002;14:233-49.
10. Green LW, Ottosen JM. From efficacy to effectiveness to community and back: evidence-based practice vs. practice-based evidence. Proceedings from conference: From Clinical Trials to Community: The Science of Translating Diabetes and Obesity Research: National Institutes of Diabetes, Digestive and Kidney Diseases; 2004.
11. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials. Increasing the value of clinical research for decision-making in clinical and health policy. *JAMA* 2003;290:1624-32.
12. Nutting PA, Beasley JW, Werner JJ. Practice-based research networks answer primary care questions. *JAMA* 1999;281:686-9.
13. Glasgow RE, Davidson KW, Dobkin PL, Ockene J, Spring B. Practical behavioral trials to advance evidence-based behavioral medicine. *Ann Behav Med* In press 2006.
14. Glasgow RE, Magid DJ, Beck A, Ritzwoller D, Estabrooks PA. Practical clinical trials for translating research to practice: design and measurement recommendations. *Med Care* 2005;43:551-7.
15. Glasgow RE. Translating research to practice: lessons learned, areas for improvement, and future directions. *Diabetes Care* 2003;26:2451-6.
16. Cronbach LH, Glesser GC, Nanda H, Rajaratnam N. The dependability of behavioral measurements: theory of generalizability for scores and profiles. New York: John Wiley & Sons; 1972.
17. Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental design for generalized causal inference. Boston: Houghton Mifflin; 2002.
18. Kaplan RM. Need for continuing cost-effectiveness and cost utility studies in diabetes care. *Diabetes Spectrum* 1995;8:252-3.
19. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press; 2003.
20. Meenan RT, Stevens VJ, Hornbrook MC, La Chance PA, Glasgow RE, Hollis JF, et al. Cost-effectiveness of a hospital-based smoking cessation intervention. *Med Care* 1998;36:670-8.
21. Bennett-Johnson S. Behavioral aspects of diabetes. In: Byrne DG, Caddy GR, Editors. Behavioral medicine. Norwood: Ablex Publishing Corporation; 1992. p. 317-52.
22. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;35:1381-95.
23. Glasgow RE, Ory MG, Klesges LM, Cifuentes M, Fernald DH, Green LA. Practical and relevant measures of health behavior for primary care settings. *Ann Fam Med* 2005;3:73-81.
24. Lorig K, Stewart A, Ritter P, Gonzalez V, Laurent D, Lynch J. Outcome measures for health education

- and other health care interventions. Thousand Oaks (CA): Sage Publications; 1996.
25. Kaplan RM. The significance of quality of life in health care. *Qual Life Res* 2003;12(Suppl 1):3–16.
  26. Lorig KR, Sobel DS, Stewart AL, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Med Care* 1999;37:5–14.
  27. Bech P, Olsen LR, Kjoller M, Rasmussen NK. Measuring well-being rather than the absence of distress symptoms: a comparison of the SF-36 mental health subscale and the WHO-Five Well-Being Scale. *Int J Methods Psychiatr Res* 2003;12:85–91.
  28. Centers for Disease Control and Prevention. CDC Healthy Days Website. Available at: <http://www.cdc.gov/hrqol/monograph.htm> 2005.
  29. Leape LL, Berwick DM. Five years after To Err is Human: What have we learned? *JAMA* 2005;18:2384–90.
  30. Stange KC, Woolf SH, Gjeltema K. One minute for prevention: the power of leveraging to fulfill the promise of health behavior counseling. *Am J Prev Med* 2002;22:320–3.
  31. Goldstein MG, Whitlock EP, DePue J. Multiple health risk behavior interventions in primary care: summary of research evidence. *Am J Prev Med* 2004;27(2 Suppl):61–79.
  32. National Committee for Quality Assurance (NCQA). Health plan employer data and information set 3.0. Washington (DC): National Committee for Quality Assurance; 1996.
  33. Green LW, Kreuter MW. Health promotion planning: an educational and ecological approach. 4th Ed. Mountain View (CA): Mayfield Publishing Company; 2005.
  34. Rotheram-Borus MJ, Flannery ND. Interventions that are CURREN: Cost-effective, useful, realistic, robust, evolving, and sustainable. In: Rehmschmidt H, Belfer M, Goodyear I, Editors. Facilitating pathways: care, treatment, and prevention in child and adolescent health. New York: Springer; 2004. p. 235–44.
  35. Rogers EM. Diffusion of innovations. 5th Ed. New York: Free Press; 2003.
  36. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 1999;89:1322–7.
  37. Glasgow RE. Evaluation of theory-based interventions: the RE-AIM model. In: Glanz K, Lewis FM, Rimer BK, editors. Health behavior and health education. 3rd Ed. San Francisco: John Wiley & Sons; 2002. p. 531–44.
  38. 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 Educ Couns* 2001;44:119–27.
  39. Prochaska JO, Velicer WF, Fava JL, Rossi JS, Tsoh JY. Evaluating a population-based recruitment approach and a stage-based expert system intervention for smoking cessation. *Add Behav* 2001;26:583–602.
  40. Abrams DB, Orleans CT, Niaura RS, Goldstein MG, Prochaska JO, Velicer W. Integrating individual and public health perspectives for treatment of tobacco dependence under managed health care: a combined stepped care and matching model. *Ann Behav Med* 1996;18:290–304.
  41. Glasgow RE, Klesges LM, Dziewaltowski DA, Estabrooks PA, Vogt TM. Evaluating the overall impact of health promotion programs: using the RE-AIM framework for decision making and to consider complex issues. *Health Educ Res* In press 2006.
  42. Institute of Medicine. Unequal treatment: confronting racial and ethnic disparities in health care. Washington (DC): National Academies Press; 2003.
  43. Glasgow RE, Klesges LM, Dziewaltowski DA, Bull SS, Estabrooks P. The future of health behavior change research: what is needed to improve translation of research into health promotion practice? *Ann Behav Med* 2004;27:3–12.
  44. Rothman KJ, Greenland S, editors. Modern epidemiology. 2nd Ed. Philadelphia PA: Lippincott, Williams, and Wilkins; 1998.
  45. Glasgow RE, Whitlock EP, Eakin EG, Lichtenstein E. A brief smoking cessation intervention for women in low-income Planned Parenthood Clinics. *Am J Public Health* 2000;90:786–9.
  46. Glasgow RE, Nutting PA, King DK, et al. A practical randomized trial to improve diabetes care. *J Gen Intern Med* 2004;19:1167–74.
  47. Will JC, Farris RP, Sanders CG, Stockmyer CK, Finkelstein EA. Health promotion interventions for disadvantaged women: overview of the WISEWOMAN projects. *J Womens Health* 2004;13:484–502.
  48. Crabtree BF, Miller WL, Stange KC. Understanding practice from the ground up. *J Fam Pract* 2001;50:881–7.
  49. Stange KC, Goodwin MR, Zyzanski SJ, Dietrich AJ. Sustainability of a practice-individualized preventive service delivery intervention. *Am J Prev Med* 2003;25:296–300.
  50. Miller WL, McDaniel RRJ, Crabtree BF, Stange KC. Practice jazz: understanding variation in family practices using complexity science. *J Fam Pract* 2001;50:872–8.
  51. Glasgow RE, Bull SS, Piette JD, Steiner J. Interactive behavior change technology: a partial solution to the competing demands of primary care. *Am J Prev Med* 2004;27:80–7.
  52. Glasgow RE, Toobert DJ, Barrera M, Jr., Strycker LA. The Chronic Illness Resources Survey: cross-validation and sensitivity to intervention. *Health Educ Res* 2004;20:402–9.
  53. Riley KM, Glasgow RE, Eakin EG. Resources for health: a social-ecological intervention for supporting self-management of chronic conditions. *J Health Psychol* 2001;6:693–705.