

What Types of Evidence are Most Needed to Advance Behavioral Medicine?

Russell E. Glasgow, Ph.D.

Published online: 12 February 2008
© The Society of Behavioral Medicine 2008

Abstract

Background This editorial presents a perspective on the types of evidence most needed to advance behavioral medicine given the current status of the field.

Purpose The paper argues that the types of evidence most needed at present are evidence that is contextual, practical, and robust.

Methods Each of the above issues is discussed with attention to characteristics of interventions; representativeness at the multiple levels of setting, clinical staff, and participants; and research design and measures. Arguments are made from philosophy of science, status of the literature, and future directions perspectives.

Results The current dominant paradigm of reductionistic studies focused predominantly on internal validity using highly homogenous patients and academic settings is not and will not produce the desired translation to real-world practice and policy. Instead, broader “practical” clinical and behavioral trials are needed that address the influence of the context in which programs are conducted, that include outcomes important to decision makers and communities, and that focus on moderating, mediating, and economic issues.

Conclusions To create programs that will be disseminable, a greater focus is needed on external validity and transparency of reporting. We need to realize that the world is complex and

embrace and study this complexity to produce further progress. Such an approach can produce evidence that is both rigorous and relevant.

Keywords Research methods · Design · Dissemination · Translation · Effectiveness · RCT

Introduction

Given our current state of knowledge in behavioral medicine, my short answer to the question posed in the title of this article is, “evidence that is contextual, practical, and robust.” I provide the rationale for these recommendations throughout this editorial, but the bottom line is that if we want to influence health care, health policy, and real-world decisions to accomplish the mission of the Society of Behavioral Medicine (SBM), these are the types of evidence that are most needed.

To provide the background for this conclusion, I have organized my arguments under three considerations that led to my recommendations. These perspectives involve considerations of philosophy of science issues, the current state of scientific evidence in behavioral medicine, and key future directions for SBM.

Philosophy of Science Issues

The philosophy of science issues can be summarized as contextualism vs reductionism world views [1–3]. The key question is whether we are likely to make more progress by following a model of science that approaches statements of causality by isolating, simplifying, and holding constant key conditions (reductionism that attempts to understand effects by controlling or removing all potential confounds)

R. E. Glasgow (✉)
Center for Dissemination and Implementation Research,
Institute For Health Research, Kaiser Permanente Colorado,
P.O. Box 378066, Aurora, CO 80237-8066, USA
e-mail: russg@re-aim.net

R. E. Glasgow
335 Road Runner Lane,
Penrose, CO 81240, USA

or by studying programs in their context and investigating the impact of different contextual factors (contextualism)?

It is of historical and conceptual interest that those advocating a linear, reductionistic perspective on science typically do so to more closely resemble or emulate the “hard sciences” (e.g., physics, chemistry, biology) or medicine in order for behavioral sciences to be more widely accepted. The irony in this is that the “hard” and “leading sciences” have long ago moved beyond a reductionistic approach [2–4]. They indeed did make progress using a reductionistic approach. However, at least two to three decades ago, those fields realized that further progress was not going to be made using these models and have since evolved much more contextual approaches and research paradigms that employ models such as systems theory [4] and complexity/chaos theory [3, 5]. These approaches reject the possibility of understanding and advancing science—given our current knowledge—by studying things in isolation. Rather, their focus is on inter-relationships, the importance of starting points and contextual factors, etc. They also have much greater openness to different ways of knowing and types of evidence rather than positing that there is only one correct, right approach or “true answer.”

Only some branches of medicine, and by extension behavioral medicine advocates who feel progress will best be made by applying the methods of medicine—and especially the randomized drug efficacy trial (RCT)—cling today to a scientific approach that claims that the “best science” and the highest levels of knowledge are achieved by employing measures to control and precisely manipulate factors so as to isolate treatment effects and hold constant all “threats to internal validity” [6].

I have nothing against RCTs—or any other design. I do, however, think that it is unlikely that further progress will be made by relying predominantly on the type of reductionistic RCT that attempts to remove context, to study factors under non-representative and highly artificial conditions, and that only studies highly motivated individuals who do not have “confounding factors” (which often excludes the vast majority of persons to whom such drugs or procedures are eventually applied). Rather, I advocate “practical clinical” [7, 8] and practical behavioral trials [9]—and many other experimental designs [10] that help us to explore the relevance of different interventions for different populations, under different conditions [11, 12].

Glass and McAtee [13] conclude in their recent provocative review that “behavioral science...especially in the U.S., has focused primarily on individual health-related behaviors, without due consideration of the social context in which health behaviors occur (page 1664).” In his new book on an integrated approach to health and to primary care, Paul Thomas [3] concludes that “linear thinking

dominates...it isolates factors that are really complex...(in contrast)...systems thinking sees dynamic interactions between related things...knowledge generated in one context may not be relevant in others.”

Although space precludes further discussion, I note that a contextual approach to science is also much more congruent than a positivist reductionistic approach with following developments within behavioral medicine and health care: socio-ecologic models [14]; multi-level programs and multi-level analyses [15, 16]; systems thinking and dynamic modeling [13]; complex interventions [17] and complexity theory [3, 5]; and finally, transdisciplinary approaches.

Story

Another concern with linear reductionistic approaches to evidence is that they investigate only a small proportion of the issues central to program success. Realizing that humans learn best through stories [18], I illustrate this point through a story.

Imagine, as many scientists and citizens hope, that a specific genetic basis for obesity (or cancer or diabetes) was discovered and that a major pharmacogenetic company rapidly developed and proved the efficacy of a targeted pharmacogenetic intervention in record time. Imagine further that the FDA, after reviewing the key double-blind RCT efficacy study which demonstrated a large effect size—a 50% reduction in obesity compared to a double-blind placebo control condition—decided to rush this new drug to market because of the public health need.

This exciting breakthrough would then need to be translated into practice to actually impact public health. Here is where the story gets interesting and where the enormous impact of other behavioral, social, economic, and policy factors come into play. As summarized in Table 1, further assume that the government and the pharmaceutical company combine forces and resources in an unprecedented manner to rush the drug into widespread use. Table 1 describes what are likely realistic to optimistic estimates of the actual impact of a nationwide dissemination effort to promote use of this breakthrough drug. The right-hand column of Table 1 shows the bottom line public health impact or percent of all obese persons who would benefit from such an effort.

The left-hand column in Table 1 summarizes the series of steps involved in translating any basic science breakthrough into real-world practice. The second column labels the step according to its categorization in the RE-AIM framework. The third column displays the “success rate” for that step, and I have used estimates that vary from 40 to 60% for each stage to bracket the likely overall impact. For most steps, a 40–60% success rate would be considered

Table 1 The reality of translating an evidence-based (fill in blank) intervention

Issue	RE-AIM element ^a	Success rate (%)	Population-wide impact (%)
Persons having genetic risk factor	Population prevalence	40–60	40% – 60%
Healthcare settings that participate	ADOPTION— Setting level	40–60	16–36
Physicians who prescribe	ADOPTION— Clinician level	40–60	6–22
Patients who accept	REACH	40–60	2–13
Delivery/ medication adherence	IMPLEMENTATION (follow regimen correctly)	40–60	0.8–8
RCT efficacy results	EFFECTIVENESS (percent lost significant weight in RCT)	40–60	0.3–5
Continued longer term effects	MAINTENANCE (individual level)	40–60	0.1–3

^a www.re-aim.org

very good for results from a nationwide campaign over a 1- to 2-year period, and especially if the 40–60% impacted were representative and included those most at risk (which unfortunately is often not the case).

As can be seen, we begin with the assumption that 40–60% of the obese population has the genetic profile that puts them at risk. This would be at least ten times higher than the vast majority of genetic disorders to date, but let us be optimistic for purposes of illustration. If 40–60% of all healthcare clinics in the USA were to adopt this new drug approach to obesity, that would be a phenomenal success. To accomplish this, a terribly convincing case would need to be made to diverse organizations that would include the VA, managed care organizations, community health centers, the Indian Health Service, etc.—many which are under-resourced already.

The third row in Table 1 illustrates the impact of physician reaction to a newly approved medication, and again, optimistically assumes that 40–60% of physicians would test patients and prescribe this medication to *all* of their eligible patients. The reader can follow the remaining rows of Table 1 to see the impact of later steps in this sequential story of the national rollout of a new obesity wonder drug.

Three points should be made in summary: (1) The 40–60% estimates for the percent of patients who could accept/could pay for what would likely be an expensive medication; who would take the medication as prescribed over a sufficient period of time (and this assumes no side

effects or unanticipated negative consequences—such as ignoring healthy lifestyle behavior patterns); and who would continue to maintain benefits long-term are likely overestimates. (2) Only in the next to last row do the results of the groundbreaking RCT come in play—the issues in all the other rows are typically ignored in an efficacy-style RCT. (3) Finally, the “bottom line” impact is that after 1 to 2 years, approximately 0.1–3% of the obese population in the USA would substantially benefit in a lasting way from this revolutionary breakthrough in pharmacogenetics.

The purpose of this exercise is not to disparage pharmacogenetic approaches—the same issues apply to real-world application of behavioral interventions. The point is that our focus and evidence needs to expand beyond the narrow domain of studying only the impact on a single primary dependent variable. There is also a more subtle but optimistic message embedded in Table 1.

This message is that there are numerous and multiple opportunities—represented by each row of Table 1—to enhance the ultimate success rate in the bottom right of the table. Improving any of the steps of adoption, reach, implementation, or maintenance could also substantially increase the public health benefit. These various stages also make apparent the opportunities for transdisciplinary collaboration to address healthcare issues—the potential contributions of diverse fields such as social marketing, health communication, behavioral approaches to adherence and maintenance, patient-provider communication, risk and decision analysis, health economics, and health policy should be obvious.

State of Our Science

This section addresses the current imbalance between the attention to and quality of data reported on internal validity vs external validity in behavioral medicine. Virtually every evaluation of this question—including very recent reviews [19] have concluded that published articles report to a much greater extent on internal validity than on issues related to external validity [20, 21]. There are multiple reasons for this, including the much greater emphasis on basic than applied research by NIH, and the perspectives of study sections and manuscript reviewers [22, 23].

There is an underlying scientific perspective that operates at all levels of the grant announcement, proposal review, paper submission, manuscript review, and literature synthesis stages that is responsible for perpetuating this imbalance between internal and external validity reporting. This perspective is related to the worldview discussed above that always values randomized designs and tight control over internal validity as being superior to alternative or complementary approaches to science [13, 24]. This

perspective and value judgment is evident in guidelines such as the CONSORT criteria that are used by the vast majority of health journals [25]. The CONSORT criteria have been very helpful in increasing the quality of reporting on internal validity in health publications (www.consort-statement.org). Because only 1 of the 22 CONSORT checklist items say anything about external validity, however, it is understandable that authors continue to report on internal validity at the expense of external validity.

Let me be clear: I am not opposed to internal validity or to controlled studies. Quite the opposite; but there is also a science of external validity and of dissemination, and this science continues to be systematically ignored or minimized not only in the CONSORT criteria but also by such factors as “quality rating systems” and in many grant and manuscript reviews [26].

A group of 13 editors of leading health promotion and behavioral medicine journals, including *Annals of Behavioral Medicine*, recently came together to discuss what could be done about this situation. The journal editors agreed that steps need to be taken to encourage better reporting of external validity issues. Several of these journals are in the process of publishing editorials [27] or of taking other steps to evaluate the external validity of submissions. For more information on these developments, see www.re-aim.org.

Table 2 summarizes the key issues that need to be addressed regarding external validity [22, 23]. These issues include: reach and sample representativeness (at multiple

levels, including patient, clinician, and setting); implementation consistency and how programs are adapted to fit different settings and cultures [28] and over time [29]; reporting of generalizability across outcomes that are important to healthcare decision and policy makers (including costs); and finally, level of maintenance and organizational sustainability over time [30].

In summary, Rothwell [26] concluded a thorough analysis of this issue with the observation that “what little systematic evidence we now have confirms that RCTs often lack external validity...this issue is neglected by current researchers, medical journals, funding agencies, and governmental regulators alike (page 90).”

Practical Trials

It is easier to criticize the current state of affairs than to recommend realistic alternatives. Fortunately, there are feasible alternatives that can be implemented now and that can retain internal validity while substantially enhancing external validity. These strategies have been referred to as “practical trials” [7–9].

Such designs can be RCTs or they can be other experimental designs such as interrupted time series or multiple baseline across settings designs that control for threats to internal validity. The distinguishing characteristics of such designs, however, are that they address four key issues relevant to external validity. These issues include (1) *representative patients*—instead of selecting the most motivated, least complex patients that have the fewest “confounding factors” and that are the most homogeneous, samples are purposefully selected [6] to represent the range of patients encountered in the real-world settings to which one wants to generalize.

A second factor is that (2) the *interventions are conducted in multiple settings*. The emphasis is on including a range of settings that reflect those in typical practice—in contrast to only the settings that have the greatest expertise, the most resources and the highest chances of successfully delivering an intervention. The third factor is one of the most significant ways in which practical clinical [8] and behavioral studies [9] differ from “research as usual”. This criterion is that (3) *comparison conditions represent current standards of care*, or alternative treatments—rather than no treatment or placebo controls. The rationale for this criterion is that to justify changes in practice, the additional education and quality control modifications necessary and the frequently much higher costs of a new treatment, the innovation should be significantly better than current, familiar and less expensive interventions.

The final criterion reflects back to our story and is that (4) *multiple outcomes, and especially outcomes relevant to*

Table 2 Types of external validity evidence needed

-
1. Program or policy *reach* (number and percent of persons) and sample *representativeness* at multiple levels of
 - a. Patients or citizens
 - b. Clinicians or change agents
 - c. Settings
 2. Program or policy *implementation and adaptation*
 - a. Consistency of delivery of key intervention components
 - b. Consistency of delivery across different staff
 - c. Adaptation of the program to fit local settings, populations, and over time
 3. *Outcomes important to decision makers*
 - a. Effectiveness on multiple measures
 - b. Robustness across subgroups, especially those related to health disparities
 - c. Cost and cost-effectiveness
 - d. Impact on quality of life
 - e. Assessment of potential adverse outcomes
 4. *Maintenance and institutionalization*
 - a. Long-term effects at minimum 6–12 months post-program
 - b. Extent to which program or policy is sustained as originally designed, discontinued entirely, or adapted and modified after evaluation is completed.
-

clinicians and decision makers (and the community) should be included. These concerns address factors such as the feasibility, implementation requirements, costs, expected return on investment, range of applicability, and impact on quality of life [31] or benefit relative to alternative uses of scarce resources. In summary, practical trials provide important information on the influence of *contextual* factors and external validity that is often missing from traditional efficacy RCTs.

Research–Practice Integration

My final perspective on the types of evidence now needed to advance our field is based on my perception that the greatest current need in behavioral medicine is for enhanced integration of research findings into practice and policy [10, 32]. The gap, or as the IOM has concluded, “chasm” between what is known and what is routinely applied in healthcare is huge and does not appear to be narrowing [33–35]. Albert Bandura [36] has stated that we “need to examine the efficacy of alternative modes of diffusion with the same care and rigor as is devoted to the development of the models being diffused”, and Green and Ottosen [37] have advised that “If we want more evidence-based practice; we need more practice-based evidence”.

While most would likely agree with these statements in the abstract, my concern is that the research community needs to do more to actively help narrow this gap. There are some important exceptions such as the recent trans-NIH Program Announcement (PAR-07-0086) Dissemination and Implementation in Health and the AHRQ Translating Research into Practice program, but in general, our study sections, editorial boards, and training programs tend to reinforce the status quo [24, 26]. They tend to reward

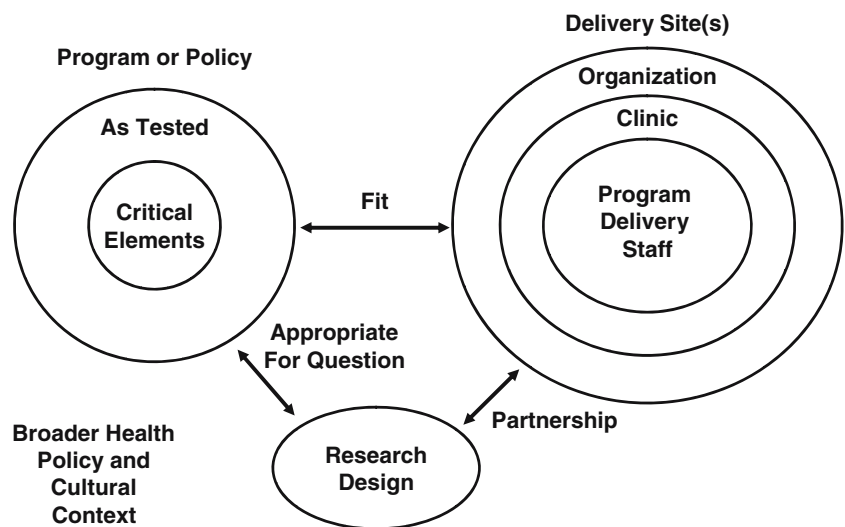
research and researchers that continue to produce results that are unlikely to translate into practice.

Even rigorous, thorough, and step-by-step approaches that follow all the recommendations of the linear phases of research model [38, 39] fail to advance programs successfully to the next “stage”. For example, Stevens et al. [40] conducted a well-controlled smoking cessation efficacy RCT for hospitalized smokers in a large hospital and found it to be efficacious when delivered by experienced smoking counselors [40]. However, translation of this exact intervention, in the very same hospital, when intervention was delivered in an effectiveness study by well-trained and supervised respiratory therapists failed to produce significant treatment effects [41]. Relatedly, Hallfors et al. [42] selected one of the SAMHSA-recommended model efficacious programs (based on an efficacy RCT), carefully implemented it in a real world drug treatment setting, and found it to be ineffective.

New and different models—and types of evidence—are needed to make successful translation to real world implementation and dissemination contexts. These approaches need to assess factors such as level of involvement of key stakeholders from the outset; feasibility, cost and practicality; the balance between fidelity and local adaptation [28]; and the “3 R’s” of translation research—representativeness (who participates, at all levels), robustness (especially impact on health disparities and in low resource settings), and replicability.

As summarized in Fig. 1, behavioral medicine programs are complex and are embedded in multiple layers of context. As shown in the figure, we need to focus on the best “fit” for a given question in a given setting among research design, the program being investigated, and the setting in which it is being implemented [43]. No single design is always the best answer, and we need more multi-

Fig. 1 Simplified systems model for translational research



method studies that combine the benefits of quantitative and qualitative studies [44, 45].

Conclusion

In summary, the world is complex and multiply determined; we ignore and attempt to oversimplify this complexity at our peril. We need to recognize that all models (and designs) are wrong [46]—and have greater appreciation for creative approaches to the new challenges we face. I believe that by focusing on research that is contextual, practical, and robust, we can advance our science with approaches that are both rigorous and relevant.

Acknowledgments A presentation version of this paper was presented as part of a debate entitled, “What Types of Evidence are Needed,” during the March, 2007, Annual Meeting of the Society of Behavioral Medicine, Washington, DC.

References

- Kuhn TS. *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press; 1962.
- Biglan A. *Changing Cultural Practices: A Contextualist Framework for Intervention Research*. Reno, NV: Context Press; 1995.
- Thomas P. *Integrating Primary Health Care*. Oxford, England: Radcliffe Publishing; 2006.
- Gharajedaghi J. *Systems Thinking: Managing Chaos and Complexity*. 2nd ed. Boston: Elsevier; 2006.
- McNulty T, Ferlie E. *Reengineering Health Care: The Complexities of Organizational Transformation*. Oxford, UK: Oxford University Press; 2002.
- Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-experimental Design for Generalized Causal Inference*. Boston: Houghton Mifflin; 2002.
- 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(6): 551–557.
- Tunis SR, Stryer DB, Clancey CM. Practical clinical trials. Increasing the value of clinical research for decision making in clinical and health policy. *J Am Med Assoc*. 2003; 290: 1624–1632.
- Glasgow RE, Davidson KW, Dobkin PL, Ockene J, Spring B. Practical behavioral trials to advance evidence-based behavioral medicine. *Ann Behav Med*. 2006; 31(1): 5–13.
- Glasgow RE, Emmons KM. How can we increase translation of research into practice? *Ann Rev Pub Health*. 2007; 28(1): 413–433.
- Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review: A new method of systematic review designed for complex policy interventions. *J Health Serv Res Pol*. 2005; 10(S1): S21–S39.
- Paul GL. Behavior modification research: Design and tactics. In Franks CM, ed. *Behavior Therapy: Appraisal and Status*. New York: McGraw-Hill; 1969: 29–62.
- Glass TA, McAtee MJ. Behavioral science at the crossroads in public health: Extending horizons, envisioning the future. *Soc Sci Med*. 2006; 62(7): 1650–1671.
- Stokols D. Social ecology and behavioral medicine: Implications for training, practice, and policy. *Behav Med*. 2000; 26: 129–138.
- Cohen J, Cohen P, West SG, Aiken LS. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. 3rd ed. London: Erlbaum; 2003.
- Murray DM. Statistical models appropriate for designs often used in group-randomized trials. *Stat Med*. 2001; 20(9–10): 1373–1385.
- Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. Framework for design and evaluation of complex interventions to improve health. *B Med J*. 2000; 321: 694–696.
- Goodman A. Storytelling as best practice: How stories strengthen your organization, engage your audience, and advance your mission. Andy Goodman; 2003.
- Klesges LM, Dzewaltowski DA, Glasgow RE. Childhood obesity prevention: Reviewing the translation potential of intervention evidence. *Am J Prev Med*. 2008; in press.
- Oldenburg B, Ffrench BF, Sallis JF. Health behavior research: The quality of the evidence base. *Am J Health Promot*. 2000; 14(4): 253–257.
- Glasgow RE, Klesges LM, Dzewaltowski 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(1): 3–12.
- Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: Issues in external validity and translation methodology. *Eval Health Prof*. 2006; 29(1): 126–153.
- Glasgow RE, Green LW, Klesges LM, et al. External validity: We need to do more. *Ann Behav Med*. 2006; 31(2): 105–108.
- Holmes D, Murray S, Perron A, Rail G. Deconstructing the evidence-based discourse in health sciences: Truth, power and fascism. *International Journal of Evidence-Based Healthcare*. 2006; 4: 180–186.
- Mohrer D, Schulz KF, Altman DG, Lepage L. The CONSORT statement: Revised recommendations for improving the quality of reports. *JAMA*. 2001; 285: 1987–1991.
- Rothwell PM. External validity of randomised controlled trials: To whom do the results of this trial apply? *Lancet*. 2005; 365: 82–93.
- Glasgow RE, Green LW, Ammerman A. A focus on external validity. *Eval Health Prof*. 2007; 30(2): 115–117.
- Castro FG, Barrera M, Jr., Martinez CR, Jr. The cultural adaptation of prevention interventions: Resolving tensions between fidelity and fit. *Prev Sci*. 2004; 5: 41–45.
- Rotheram-Borus MJ, Flannery D, Duan N. Interventions that are CURRES: Cost-effective, useful, realistic, robust, evolving, and sustainable. In Rehmschmidt H, et al, eds., *Facilitating Pathways: Care, Treatment, and Prevention in Child and Adolescent Health*. New York: Springer; 2004; 235–244.
- Goodman RM, McLeroy KR, Steckler A, Hoyle R. Development of level of institutionalization scales for health promotion programs. *Health Ed Q*. 1993; 20: 161–178.
- Kaplan RM. The significance of quality of life in health care. *Q Life Res*. 2003; 12(Suppl 1): 3–16.
- 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 Publ Health*. 2003; 93(8): 1261–1267.
- Berwick DM. A user's manual for the IOM's “Quality Chasm” report: Patients' experience should be the fundamental source of the definition of “quality”. *Health Aff*. 2002; 21: 80–90.
- Institute of Medicine, Committee on Quality Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press; 2003.
- McGlynn EA. An evidence-based national quality measurement and reporting system. *Med Care*. 2003; 41(1 Supp): I8–I15.
- Bandura A. *Self-efficacy: The Exercise of Control*. New York: W.H. Freeman; 1997.
- 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.
38. Greenwald P, Cullen JW. The new emphasis in cancer control. *J Natl Cancer Inst*. 1985; 74(3): 543–551.
 39. Flay BR. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med*. 1986; 15: 451–474.
 40. Stevens VJ, Glasgow RE, Hollis JF, Lichtenstein E, Vogt TM. A smoking cessation intervention for hospitalized patients. *Med Care*. 1993; 31: 65–72.
 41. Stevens VJ, Glasgow RE, Toobert DJ, Karanja N, Smith KS. One-year results from a brief, computer-assisted intervention to decrease consumption of fat and increase consumption of fruits and vegetables. *Prev Med*. 2003; 36(5): 594–600.
 42. Hallfors D, Cho H, Sanchez V, Khatapoush D, Kim HM, Bauer D. Efficacy vs. effectiveness trial results of an indicated “model” substance abuse program: Implications for public health. *Am J Publ Health*. 2006; 96(12): 2254–2259.
 43. Estabrooks PA, Glasgow RE. Dissemination, knowledge exchange, or knowledge integration: Explaining the gap between medical office-based physical activity intervention research and practice. *Am J Pub Health*. 2005; 31(4 Suppl): S45–S56.
 44. Stange KC. One size doesn’t fit all. Multimethod research yields new insights into interventions to increase prevention in family practice. *J Fam Pract*. 1996; 43(4): 358–360.
 45. Curry L, Shield R, Wetle T, eds. *Improving Aging and Public Health Research: Qualitative and Mixed Methods*. Washington, DC: Public Health Association; 2006.
 46. Sterman J. All models are wrong: Reflections on becoming a systems scientist. *Syst Dyn*. 2002; 18: 501–531.