

In Defense of the Randomized Controlled Trial for Health Promotion Research

Laura Rosen, PhD, Orly Manor, PhD, Dan Engelhard, MD, and David Zucker, PhD

The overwhelming evidence about the role lifestyle plays in mortality, morbidity, and quality of life has pushed the young field of modern health promotion to center stage. The field is beset with intense debate about appropriate evaluation methodologies. Increasingly, randomized designs are considered inappropriate for health promotion research.

We have reviewed criticisms against randomized trials that raise philosophical and practical issues, and we will show how most of these criticisms can be overcome with minor design modifications. By providing rebuttal to arguments against randomized trials, our work contributes to building a sound methodological base for health promotion research. (*Am J Public Health*. 2006;96:1181–1186. doi:10.2105/AJPH.2004.061713)

COMPELLING EVIDENCE

about the critical role lifestyle and environmental factors play in mortality, morbidity, and quality of life has contributed to the growing popularity of the field of health promotion, which has an overarching aim of “enabling people to increase control over, and improve, their health.”^{1(p1)} Because of the multidisciplinary nature of health promotion, research is influenced by many fields, including education, policy, social science, anthropology, and epidemiology.²

While the randomized controlled trial (RCT) is the gold standard in today’s medical world, the role randomization plays in health promotion research is a topic of hot debate, which in part reflects the ongoing debate in nonmedical fields about the relevance of randomized trials.^{3–7} Many health promotion researchers have asserted that RCTs are irrelevant to, or unworkable in, their field for a variety of philosophical and practical reasons.^{8–10} We refute the philosophical arguments and present approaches to handling the practical issues. Although we do not claim that all the practical problems can be entirely eliminated, we submit that a wisely designed RCT is a far superior evaluative mechanism for answering specific types of questions compared with the evaluation approaches that are currently popular in the health promotion field.

Our recent field study of a hygiene health promotion program in Jerusalem preschools is an illustrative example. This study

was a controlled trial with randomization at the preschool level: 40 preschools were randomized equally to either intervention or control.¹¹ The program addressed various hygiene issues, with a primary emphasis on hand washing. The program used a multi-pronged approach that included elements aimed at staff, children, parents, and school nurses, as well as hygienic changes to the classroom environment.

A total of 1029 children participated in the trial. The program was implemented in the intervention preschools during December 2000. Data on absenteeism due to illness was collected from children in both groups from January 2001 through April 2001. In May 2001, the program was implemented in the control preschools. Thus, all participating preschools eventually received the program in a “phased” manner. This design technique, which is better known for being used with individual randomization schemes, contributed considerably to the success of the trial. Details about this design technique have been published elsewhere.^{12,13}

CRITICISMS OF THE RANDOMIZED CONTROLLED TRIAL AND COUNTERARGUMENTS

Recent World Health Organization recommendations to policymakers state, “Use of randomized controlled trials to evaluate health promotion is, in most cases, inappropriate, misleading,

and unnecessarily expensive.”^{8(p2)} The International Union for Health Promotion has gone even further. Its message to the health promotion researchers in 1999 was, “Randomized controlled trials or corresponding experimental designs should not be used to measure the effectiveness of health promotion interventions.”^{14(p180)} We believe that health promotion as a field is harmed by such a view. We address both the philosophical and the practical arguments that underlie the rejection of RCTs.

Argument 1

Withholding a program from some individuals or groups on the basis of randomization is unfair.^{10,15} This argument is only germane if the intervention has a proven benefit. In medical research, discussion about the ethics of conducting a randomized trial is generally centered on the issue of clinical equipoise. Most agree that it is ethical to randomize when it is uncertain whether a new intervention is superior to an older one after benefits, risks, and costs have been taken into account.¹⁶ When assessing the degree of uncertainty, it is critical to distinguish between personal hunches and clear objective knowledge and to recognize that an initial hunch may be wrong. Indeed, several studies have shown that physicians’ initial hunches turn out to be incorrect more than 50% of the time.¹⁷

Moreover, there have been a number of instances where a widespread medical practice was

shown—after rigorous evaluation in a randomized trial—to be inappropriate. For example, observational studies suggested that hormone replacement therapy prevents cardiovascular events, but a more rigorous evaluation in the Women’s Health Initiative trial revealed that this treatment slightly increases the risk for a cardiovascular event.^{18,19} Similarly, after observational studies suggested a preventive effect of beta carotene and vitamin A on lung cancer and cardiovascular disease, a controlled trial was stopped early due to interim results that showed a possible adverse effect on those endpoints.²⁰

In the health promotion field, assessing the uncertainty about the effect of an intervention is often skewed. There is a tendency to assume that all health promotion programs are good and wholesome; many health promotion workers seem to discount the possibility of an equipoise situation. Opposition to randomized trials stems in part from this preconception that the utility of the program is a foregone conclusion.

To further complicate matters, the notion of what constitutes positive evidence in the health promotion field is fraught with confusion.⁶ Some researchers claim widespread support for the “adoption of methodological pluralism in the discipline as a whole.”^{5(p659)} The hierarchy of evidence from the US Task Force on Community Preventive Services ranks randomized and non-randomized trials with concurrent controls at the same level.²¹ Other researchers question the use of the term *evidence* in health promotion and suggest that the establishment of a linear hierarchy of evidence is premature.²²

When benefit is uncertain, conducting research—as opposed

to merely providing a program—is indeed ethical. In this case, a concurrent, comparable control group is essential. In any research setting, there is always some mechanism for deciding which groups or individuals receive the intervention; the mechanism differs across research efforts. Randomization of individuals or groups leads to more ethical allocation than allocation on the basis of politics, friendship, convenience, or geography.

In light of the foregoing considerations, the ethical concerns about not offering the control population a possibly beneficial intervention lose their force. Moreover, in many cases it is possible to design a trial that overcomes this particular problem. The Jerusalem project used a phased allocation scheme under which all participants received the program eventually. Other types of randomized trials that overcome the problem include (1) trials with 2 distinct interventions, each of which serves as the other’s control; (2) trials that allocate high- and low-intensity interventions; and (3) trials that randomize to wait lists.²³

Argument 2

Because health promotion programs are complex and multifaceted,²⁴ they cannot be evaluated with the randomized controlled trial’s single simple outcome. The multifaceted nature of health promotion programs does not preclude evaluating such programs on the basis of well-defined simple outcome measures. Nor is the RCT limited to simple interventions. In the Jerusalem study, for example, the intervention program was complex and included teacher training sessions, an educational program for the children, the provision of basic supplies

(liquid soap, paper towels, individual cups, and dispensers), and a home component. Yet, the primary outcomes were simple measures of hand washing behavior and illness-related absenteeism. More generally, simple measures such as changes in body mass index, cigarette consumption, or hours of vigorous weekly exercise are commonly used in health promotion studies.^{24–26} Defining and measuring simple but important outcome measures is an important scientific issue for the investigator.

Moreover, as can easily be seen from the literature, the RCT design does not preclude assessing the effect of an intervention on a range of outcome measures. At the same time, it is important to specify in advance which outcome measure or measures will be given primary emphasis to avoid the pitfall of data dredging.²⁷

Argument 3

Health promotion programs cannot be expected to produce changes in “hard” disease outcomes within a short time frame; therefore, randomized controlled trials are not practical.^{28,29} Finding an intervention effect on a long-term outcome is not a problem that is unique to randomized controlled trials. Whether or not a randomized design is used, studying long-term outcomes is inherently expensive and complicated.

Health promotion researchers have been frustrated by their failure to affect morbidity and mortality measures. One response to this frustration has been to reject standard epidemiological indicators as outcomes altogether. Tones said, “Epidemiological indicators (e.g., mortality and morbidity) should never be used to assess health promotion programs.”^{9(p93)}

Others disagree with this approach, claiming that “to deny the centrality of examining the effect of health promotion on health related outcomes . . . is to raise serious questions about the legitimacy of some health promotion activity.”^{30(p703)} Some health promoters prefer to concentrate on behavioral changes. Others—who may have been encouraged by the popularity of Prochaska’s Stages of Change Model, which postulates a staged process of behavioral change—have focused attention on the “softer” earlier endpoints, such as intention to change.³¹ However, this approach has come under fierce attack in recent years, because the evidence for correlation of the earlier endpoints with behavioral change is limited.³²

Randomized designs can be used to study outcomes of interventions on a range of effects, including attitudes, health behavior, morbidity, and mortality. Some epidemiological measures can indeed be changed within a relatively short time frame, such as those that involve communicable diseases. Far from precluding the study of either intermediate or long-term endpoints, the randomized trial can study all of these, or some of them, with a single design. The Jerusalem design examined a range of endpoints, with attitudes and beliefs as early endpoints, hand washing compliance as an intermediate behavioral endpoint, and illness-related absenteeism—a consequence of communicable disease—as a “hard” endpoint. The choice of realistic outcomes, which is a key research challenge, is unrelated to the use of randomization.

Argument 4

Randomized controlled trials are inappropriate for the types of

questions typically addressed in health promotion research.^{8,28} We agree that for certain questions that arise in the health promotion field, research methodologies other than RCT are indeed more appropriate. Programs that seek to change legislation, organizational practice, or public policy are best evaluated with observation rather than with experimentation. Questions about the determinants of health or changing population norms may be best answered with prospective cohort studies or cross-sectional studies that include high-quality surveys. Process evaluation and identification of factors that influence the degree of success in implementing a program can be conducted with various qualitative methods.

However, we submit that the RCT is in fact the most appropriate research tool for many of the questions that arise in health promotion. There is an abundance of health promotion activity on the basis of programs aimed at changing knowledge, attitudes, behavior, risk factors, morbidity, and mortality, and the RCT—with appropriate process evaluation—is the best mechanism for generating valid scientific evidence.^{33,34}

Argument 5

Randomized controlled trials are of use only when studying narrowly defined groups of people under artificial laboratory conditions, and external validity is questionable.^{10,35} Some of the practical problems that are encountered when conducting RCTs in real-world conditions pertain to potential trial participants and trial settings. Clinical trials are often conducted with specific types of patients under highly controlled conditions, and inferences to the wider population involve a certain degree of

difficulty. Criticism of the RCT in the health promotion world reflects concern about this issue, because health promotion is often directed at entire populations. However, the RCT design is not restricted to artificial laboratory conditions.

The Jerusalem study was conducted with an ethnically and socioeconomically diverse group of participants while they were engaged in their normal daily activities. Other randomized trials in the health promotion field, such as the Community Intervention Trial for Smoking Cessation (COMMIT; community-based), the Child and Adolescent Trial for Cardiovascular Health (CATCH; school-based), the Eating Patterns Study (clinic-based), and the Mwanza HIV trial (community-based), were aimed at the general population while the population functioned in its normative environments.^{36–39} The polio vaccine trial, which is still the largest clinical trial on record, provided the vaccine to some (but not all) school classes in a natural school setting.⁴⁰ Several of the major breast cancer screening trials were conducted in field settings and included the randomization of geographic regions, workplaces, or clinics.⁴¹

An innovative approach to testing improvements in medical care settings allows patients in regular practice settings to choose between a standard treatment, a new treatment, or randomization to one of the treatments.⁴² This approach—which seeks a realistic clinical population rather than a “perfect patient” population—is likely to be particularly useful for health promotion trials that do not involve pressing medical conditions.

In health promotion, as in medicine, due consideration

must be given to both internal and external validity. Of the 2, internal validity calls for greater weight, because a research finding that is broad in scope is of no value if the finding itself is unreliable.⁴³ Internal validity is severely compromised by the quasi-experimental approach, where the investigator controls the treatment allocation. This approach is directly open to bias, albeit usually unintentional. External validity depends on the degree to which the communities chosen for randomization or that agreed to be randomized differ in some systematic way from the communities that were not chosen for the trial. Both investigator bias and volunteer bias threaten external validity. We submit, however, that this issue is far less serious than the severe internal validity problems of the quasi-experimental approach.

External validity may actually be higher in community trials than in traditional individually randomized trials because of reduced volunteer bias. Typically, community-level implementation automatically covers the entire community, with no need to ask individual members to sign up for the trial. This is especially true when data are collected at the aggregate level rather than at the individual level. External validity is further enhanced when the trial data are analyzed with the intent-to-treat approach, which is standard in clinical trials and is natural with aggregate data. Zucker discussed the intent-to-treat approach in the context of community-based trials,^{44(p86,87)} where participants who deviate from the intervention protocol for any reason (e.g., treatment noncompliance by participants or physicians) are nonetheless included in the analysis within their initially

assigned treatment group. This analysis approach, while it has its limitations, offers distinct advantages over other approaches, particularly because it shows how the intervention works under real-world conditions.

The RCT in the real-world community setting has the potential for achieving high levels of internal and external validity.

Argument 6

Randomized controlled trials focus on the individual, and health promotion is concerned primarily with the community.¹⁰ Cluster randomization overcomes the problem of focusing on the individual.^{3,45} This method, which was used in the Jerusalem study, uses the group (cluster) as the unit of randomization and analysis. Members of a cluster (e.g., workplace, classroom, clinic, or community), who might naturally influence one another or be affected as a group by prevailing conditions, are randomized together to a given treatment arm. In addition to overcoming the philosophical objections, this technique also decreases the possibility of contamination, particularly if the study is planned so that clusters are geographically separated.⁴⁵ COMMIT, CATCH, and many other trials have used cluster randomization.^{36,37}

Practical difficulties with randomizing communities are sometimes cited as further arguments against randomization. Problems may occur because of difficulties at either the community level or the individual level within the community. At the community level, randomization may be seen as interference to building partnerships and coalitions; however, such interference can be avoided if the justification for randomization is presented clearly to all

relevant parties. There was no resistance to randomization in the Jerusalem study, which involved partnerships with Hebrew University, Hadassah Hospital, the Jerusalem Municipality, the Ministry of Health, the Ministry of Education, and the local Parents' Committee. The fact that all participants eventually received the intervention probably contributed to the acceptability of randomized allocation.

At the individual level, a randomized controlled design may be unappealing to individual participants, but steps can be taken to make the research design palatable. In the Jerusalem study, teachers were informed that classrooms would be randomized to either earlier or later implementation: at least 88% of the teachers and 95% of the parents agreed to participate.

Argument 7

Randomized controlled trials preclude or discourage tailoring the intervention to local needs.^{4,8} The claim here is that the RCT is too rigid to handle the flexible intervention programs that are commonly used in health promotion. Many health promotion programs entail active participation of subjects and adaptation of the actual intervention package to the needs of the individual community. Opponents of the RCT assert that the RCT is applicable only to highly standardized interventions and is therefore useless for assessing flexible interventions.

This assertion is unfounded. The RCT in no way precludes testing a flexible intervention. In the Jerusalem study, the active participation of subjects, particularly staff, was an integral part of the intervention. Modifications to the educational program were

made in classrooms at the discretion of the teachers. The COM-MIT program likewise was built to allow program development in accordance with community needs at different sites.³⁶ If the same flexibility is allowed in a trial as part of the program dissemination, the trial results will accurately reflect the program's effectiveness under real-world conditions of implementation.

Argument 8

Randomized controlled trials are too expensive.⁸ The expense of the RCT and the limited funds available for health promotion activities are sometimes cited as reasons why RCTs should be avoided,⁸ but the cost of randomization itself is inconsequential. The real costs of the RCT are a function of the large sample sizes often necessary for detecting the modest intervention effects commonly associated with health promotion programs and the resources required for gathering valid information.⁴⁶ In fact, the real size of other types of trials or designs is nearly always much bigger if the various sources of bias are to be appropriately accounted for and if a difference of the same magnitude is to be detected. Peto et al. examined the cost of evaluating interventions associated with heart disease and cancer.⁴⁷ They concluded that the most efficient design for detecting moderate effects that have the potential for a large impact on public health is a large, simple randomized trial, where a small amount of essential data is collected from a large number of individuals. Randomization is necessary for avoiding even moderate biases that are the result of systematic differences between intervention groups, and large sample sizes are necessary for making the

random error small enough to detect moderate effects. An editorial in the *American Journal of Public Health* echoed this sentiment for community trials and advocated simple, large randomized trials with broad eligibility criteria for effective international health care planning.⁴⁸

One way to cut costs when a promising but unevaluated program is being disseminated to clusters (e.g., clinics or schools) is to randomize to phased implementation. This approach significantly decreases the cost of running a trial, particularly if a simple, easily assessed outcome measure is chosen, because the cost of the intervention itself is already covered.

ALTERNATIVE METHODOLOGIES USED IN HEALTH PROMOTION RESEARCH

There are alternatives to RCTs that are commonly advocated by health promotion researchers. One method—the simple before-and-after approach—has achieved such popularity that a periodic systematic review added this category of study design to the inclusion criteria for its 1998–2000 review.⁴⁹ This design provides only weak evidence for effectiveness, because underlying societal trends or changes that arise from natural cycles (e.g., seasonal allergies or influenza rates) can easily be mistaken for program effects. In the Jerusalem trial, this approach was not feasible, because illness absenteeism was closely associated with an unpredictable and dynamic underlying rate of communicable illness, which made comparisons with a concurrent control group imperative.

Quasi-experimentation is another alternative approach often

used when doing community intervention trials.⁶ One type of quasi-experimentation is a study where there is a concurrent control group and where the allocation could have been random but the investigators allocated interventions through a nonrandom mechanism. The Minnesota Heart Trial, the Stanford Five Cites Trial, and the North Karelia project are studies in which a health promotion program was implemented in several communities, with treatment allocation defined on a “convenience” basis.^{26,50,51} The Minnesota and Stanford studies were primarily demonstration projects, and the allocation scheme was not a major concern, but when scientific evaluation is the goal, the allocation scheme is critical. Quasi-randomization is analogous to physician-determined treatment allocation in clinical trials; this approach has been discredited for many years in medical research. When the investigator determines allocation, the study groups differ with respect to pre-intervention prognosis, which makes it difficult or impossible to ascertain the intervention effect. It is unclear why many health promotion researchers claim that allocation by investigator decision is more ethical than randomization. The potential for yielding valid and useful scientific information from a nonrandomized design is substantially compromised; thus, it is actually less ethical than a truly randomized trial.

Triangulation, which brings multiple types of observational evidence to bear on a given question—similar to the judicial process—has recently been suggested as the gold standard to replace the RCT.⁹ Triangulation can be convincing in certain cases, such as the association

between smoking and lung cancer or cholesterol and heart disease. However, much health promotion research does not rise to this level. Furthermore, it does not make sense to make do with observational evidence alone unless experimental studies cannot be conducted. When experimentation is possible, it should be done, and the experimental design should be as rigorous as possible to maximize the reliability of the results.

Use of randomized designs prevents or reduces biases of many kinds, particularly investigator bias. It ensures balance on average between groups, including balance for unmeasured and unmeasurable variables, such as differences over time and differences between geographic areas. It also provides a valid basis for statistical tests of significance without reliance on statistical modeling assumptions.⁵² Investigators should be reluctant to forego the RCT in favor of less scientifically consistent approaches for assessing program effectiveness.^{33,34,53,54}

Nonrigorous evaluation can lead to misleading results. This fact has been expressed in the *stainless steel rule of evaluation*, which states that the better the evaluation, the lower the chances of positive results.⁵⁵ Many members of the medical research community, who realize the high stakes of the interventions, have adopted the RCT to ensure that their recommendations are made on the basis of reliable data. Health promotion researchers are well advised to do the same.

THE PRICE OF AVOIDING RANDOMIZED CONTROLLED TRIALS

It is well known that health promotion research is often

conducted at a reduced level of scientific rigor and that community health interventions are often poorly evaluated.^{4,56} Ethical and practical arguments that supposedly override the need for rigorous evaluation can backfire in several ways, partly because the original assumption of intervention benefit was incorrect. The program may produce effects that are opposite to the intended effects. An Australian educational effort that attempted to delay smoking, alcohol use, and analgesic use instead increased these behaviors among the 1700 schoolchildren who were exposed to the intervention.⁵⁷ The HIV epidemic continues in Vancouver, British Columbia, despite the city's needle exchange program, which is the largest such program in North America.⁵⁸ A program to decrease unsafe sexual behavior among adolescent boys was found to actually increase the practice of unsafe sexual behaviors.⁵⁹ Other undesirable effects also may occur. In the Jerusalem study, for example, ritual hand washing before meals decreased among the religiously observant population, even though hand washing with soap increased. This was unintended, unforeseen, and unacceptable to the target population.

Furthermore, all programs cost something in terms of both money and time. Regarding time, it must be recognized that people can adopt only a limited number of health promotion activities. If people are flooded with flimsy health recommendations, they may well come to view them as "junk health" and thus treat them like junk mail. Proper evaluation to weed out ineffective programs will prevent waste of money and effort and will maintain public respect for health promotion. Finally, a program may be beneficial but

not cost-effective, and this knowledge can lead to the development of an improved or alternate approach.

CONCLUSIONS

The key objections to the use of RCTs in health promotion research stem mainly from a limited understanding of the RCT design. Many of these objections can be eliminated through a better grasp of the basics of RCTs and their proper implementation and with a better understanding of research ethics in general. The RCT framework is not as narrow as many health promotion researchers imagine, and it offers advantages that should not be cavalierly forfeited. Other objections to the RCT often can be overcome with minor modifications to the RCT design. A design that includes the combination of cluster randomization with phased intervention delivery, which was successfully used in the Jerusalem study and several other trials, is one example of a modified RCT design. Avoiding the RCT will ultimately weaken health promotion and diminish its potential benefits.

The ability to reverse a current public health catastrophe—obesity, which is a health promotion challenge of the first order⁶⁰—will be much enhanced by continuing to develop randomized designs that are appropriate for community-based health promotion research. The worldwide casualty rate from tobacco continues to rise as successful tools for tobacco control remain elusive.⁶¹ The failure of the European medical community to promote lifestyle changes among patients who had heart disease is another pressing reason for the development of effective preventive techniques.⁶² Rather than staking a position against

randomization and expending research efforts on alternative methods, health promotion researchers should attempt to develop randomized designs that are both appropriate and feasible. This approach will better serve the goals of health promotion and disease prevention. ■

About the Authors

Laura Rosen and Orly Manor are with the Hebrew University School of Public Health, Jerusalem, Israel. Dan Engelhard is with the Hebrew University-Hadassah Medical Center, Jerusalem. David Zucker is with Hebrew University, Jerusalem.

Requests for reprints should be sent to Laura Rosen, PhD, Hebrew University School of Public Health, PO Box 12272, Jerusalem, Israel 91120 (e-mail: lrosen@vms.huji.ac.il).

This article was accepted April 26, 2005.

Contributors

L. Rosen originated the article and was the lead author. D. Zucker helped with the writing. O. Manor and D. Engelhard edited the article.

Acknowledgments

The authors wish to thank Rebecca Gelman of Harvard University for her helpful comments.

Human Participant Protection

No protocol approval was needed for this study.

References

- Ottawa Charter for Health Promotion, First International Conference on Health Promotion, Ottawa, 1986. Available at: http://www.who.int/hpr/nph/docs/ottawa_charter_hp.pdf. Accessed April 10, 2006.
- Egger G, Spark R, Lawson J. *Health Promotion Strategies & Methods*. Hong Kong, China: McGraw-Hill Book Co; 1990.
- Murray D. *Design and Analysis of Group Randomized Trials*. New York, NY: Oxford University Press; 1998.
- Nutbeam D. Evaluating health promotion—progress, problem and solutions. *Health Promotion Int*. 1998;13:27–44.
- Tilford S. Evidence-based health promotion. *Health Educ Res*. 2000;15:659–663.
- Speller V, Learmonth A, Harrison D. The search for evidence of effective health promotion. *BMJ*. 1997;315:361–363.
- Shadish W, Cook T, Leviton L.

- Foundations of Program Evaluation*. Newbury Park, Calif: Sage Publications; 1991.
8. World Health Organization (WHO). *Health Promotion Evaluation: Recommendations to Policymakers. Report of the WHO European Working Group on Health Promotion Evaluation*. Copenhagen, Denmark: WHO; 1998.
 9. Tones K. Evaluating health promotion—beyond the RCT. In: Norheim L, Waller M, eds. *Best Practices, Quality and Effectiveness of Health Promotion*. Helsinki, Finland: Finnish Centre for Health Promotion; 2000: 86–101.
 10. Fuller S. Research methods. In: *Oral Health Promotion: A guide to Effective Working in Pre-School Settings*. London, UK; Health Education Authority; 1999: 96–97.
 11. Rosen L, Manor O, Engelhard D, et al. Can a handwashing intervention make a difference? Results from a randomized controlled trial in Jerusalem preschools. *Prev Med*. 2006;42:27–32.
 12. Gortmaker SL, Peterson K, Wiecha J, et al. Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. *Arch Pediatr Adolesc Med*. 1999;153:409–418.
 13. Roberts L, Smith W, Jorm L, Patel M, Douglas R, McGilchrist C. Effect of infection control measures on the frequency of upper respiratory infection in child care: a randomized, controlled trial. *Pediatrics*. 2000;105:738–742.
 14. Rimpela A. Challenging current evaluation approaches: Lessons from the conference for the research community. In: Norheim L, Waller M, eds. *Best Practices, Quality and Effectiveness of Health Promotion*. Helsinki, Finland: Finnish Centre for Health Promotion; 2000:180.
 15. Learmonth A. Utilizing research in practice and generating evidence from practice. *Health Educ Res*. 2000; 15:743–745.
 16. Last JM, *A Dictionary of Epidemiology*, 4th ed., New York, NY: Oxford University Press; 2001.
 17. Gilbert JP, McPeck B, Mosteller F. Statistics and ethics in surgery and anesthesia. *Science*. 1977;198:684–689.
 18. Stampfer M, Colditz G. Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. *Prev Med*. 1991;20:47–63.
 19. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. *JAMA*. 2002;288:321–333.
 20. Omenn G, Goodman G, Thornquist D, et al. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N Engl J Med*. 1996;334:1150–1155.
 21. Briss P, Aza S, Pappaioanou M, et al. Developing an evidence-based guide to community preventive services—methods. *Am J Prev Med*. 2000;18(1S): 35–43.
 22. McQueen D, Anderson L. What counts as evidence: issues and debates. In: Rootman I, Goodstadt M, Hyndman B, et al., eds. *Evaluation in Health Promotion: Principles and Perspectives*. WHO Regional Publications, European Series, No. 92. Copenhagen, Denmark: WHO; 2001:77.
 23. Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ*. 2000;321:694–696.
 24. International Union for Health Promotion and Education. *The Evidence of Health Promotion Effectiveness: Shaping Public Health in a New Europe, Part Two*. Brussels, Luxembourg: ECSC-EC-EAEC; 2000.
 25. Holtzman J, Schmitz K, Babes G, et al. *Effectiveness of Behavioral Interventions to Modify Physical Activity Behavior in General Populations and Cancer Patients and Survivors*. Rockville, Md: Agency for Healthcare Research and Quality; 2004. Summary, Evidence Report/Technology Assessment No. 102. AHRQ Publication No. 04-E027-1.
 26. Farquhar J, Fortmann S, Flora J, et al. Effects of communitywide education of cardiovascular disease risk factors: the Stanford five-city project. *JAMA*. 1990;264:359–365.
 27. Friedman L, Furberg C, DeMets D. *Fundamentals of Clinical Trials, 3rd ed*. New York, NY: Springer-Berlag Inc; 1998.
 28. Perkins ER. Evidence. In: Perkins E, Simnett I, Wright L, eds. *Evidence-Based Health Promotion*. West Sussex, England: John Wiley and Sons; 1999.
 29. MacDonald G. A new approach for evaluating effectiveness in health promotion interventions. In: Norheim L, Waller M, eds. *Best Practices, Quality and Effectiveness of Health Promotion*. Helsinki, Finland: Finnish Centre for Health Promotion; 2000:160.
 30. Sheldon T, Sowden A, Lister-Sharp D. Systematic reviews include studies other than randomised controlled trials. *BMJ*. 1998;316:704.
 31. Prochaska J, Reding C, Evers K. The transtheoretical model and stages of change. In: Glanz K, Lewis FM, Rimer BK, eds. *Health Behavior and Health Education: Theory, Research, and Practice, 2nd edition*. San Francisco, Calif: Jossey-Bass Inc; 1997:101.
 32. Whitelaw S, Baldwin S, Bunton R, Flynn D. The status of evidence and outcomes in stages of change research. *Health Educ Res*. 2000;15:707–718.
 33. Chalmers I. Unbiased, relevant, and reliable assessments in health care. *BMJ*. 1998;317:1167–1168.
 34. Stephenson J, Imrie J. Why do we need randomized controlled trials to assess behavioral interventions? *BMJ*. 1998;316:611–613.
 35. Britton A, Thorogood M, Coombes Y, Lewando-Hundt G. Search for evidence of effective health promotion. *BMJ*. 1998;316:703.
 36. COMMIT Research Group. Community Intervention Trial for Smoking Cessation (COMMIT): I. Cohort results from a four-year community intervention. *Am J Public Health*. 1995;85:183–192.
 37. Zucker DM, Lakatos E, Webber LS, et al, for the CATCH Study Group. Statistical design of the Child and Adolescent Trial for Cardiovascular Health (CATCH): implications of cluster randomization. *Control Clin Trials*. 1995;16:96–118.
 38. Beresford SA, Curry SJ, Kristal AR, Lazovich D, Feng Z, Wagner EH. A dietary intervention in primary care practice: the Eating Patterns Study. *Am J Public Health*. 1997;87:610–616.
 39. Hayes R, Mosha F, Nicoll A, et al. A community trial of the impact of improved sexually transmitted disease treatment on the HIV epidemic in rural Tanzania: 1. Design. *AIDS*. 1995;9:919–926.
 40. Francis T, Napier R, Voight R, et al. Evaluation of 1954 field trials of poliomyelitis vaccine. In: Buck C, Llopis A, Najera E, Terris M, eds. *The Challenge of Epidemiology*. Washington, DC: Pan American Health Organization; 1995: 838–854. Scientific publication no. 505.
 41. Semiglazov VF, Manikhas AG, Misenko VM, et al. Results of a prospective randomized investigation to evaluate the significance of self-examination for the early detection of breast cancer. *Vopr Onkol*. 2003;49:434–441.
 42. Hillis A, Rajab H, Baisden C, Villamaria F, Ashley P, Cummings C. Three years of experience with prospective randomized effectiveness studies. *Control Clin Trials*. 1998;19:419–426.
 43. Elwood JM. *Causal Relationships in Medicine*. New York, NY: Oxford Medical Publications; 1988.
 44. Zucker DM. Cluster randomization. In: Geller N, ed. *Contemporary Biostatistical Methods in Clinical Trials*. New York, NY: Marcel Dekker; 2000.
 45. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research*. New York, NY: Oxford University Press; 2000.
 46. Merzel C, D'Afflitti J. Reconsidering community-based health promotion: promise, performance, and potential. *Am J Public Health*. 2003;93:557–574.
 47. Peto R, Collins R, Gray R. Large-scale randomized evidence: large, simple trials and overviews of trials. *J Clin Epidemiol*. 1995;48:23–40.
 48. Green S. The Eating Patterns Study—the importance of practical randomized trials in communities. *Am J Public Health*. 1997;87:541–544.
 49. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs the worksite: 1998–2000 update. *Am J Health Promotion*. 2001;16:107–116.
 50. Mittlemark M, Luepker R, Jacobs D, et al. Community-wide prevention of cardiovascular disease: education strategies of the Minnesota Heart Health Program. *Prev Med*. 1986;15:1–17.
 51. Puska P, Salonen J, Tuomilehto J. Changes in coronary risk factors during comprehensive five-year community programme to control cardiovascular disease (North Karelia project). *BMJ*. 1979;2:1173–1178.
 52. Byar D, Simon R, Friedewald W, et al. Randomized clinical trials: perspectives on some recent ideas. *New Engl J Med*. 1976;295:74–80.
 53. Green S. The advantages of community-randomized trials for evaluating lifestyle modification. *Control Clin Trials*. 1997;18:506–513.
 54. Neuhauser D, Green S. Efficient clinical research. *Control Clin Trials*. 1998;19:427–429.
 55. Petticrew M. Why certain systematic reviews reach uncertain conclusions. *BMJ*. 2003;326:756–758.
 56. Smith P, Moffatt M, Gelskey S, Hudson S, Kaita K. Are community health interventions evaluated appropriately? A review of six journals. *J Clin Epidemiol*. 1997;50:137–146.
 57. Hawthorne G, Garrard J, Dunt D. Does Life Education's drug education programme have a public health benefit? *Addiction*. 1995;90:205–15.
 58. Strathdee S, Patrick D, Currie S, et al. Needle exchange is not enough: lessons from the Vancouver injecting use study. *AIDS*. 1997;11:F59–F65.
 59. Christopher FS, Roosa MW. An evaluation of an adolescent pregnancy prevention program: is "just say no" enough? *Fam Relations*. 1991;39:68–72.
 60. The catastrophic failures of public health [unsigned editorial]. *Lancet*. 2004;363:745.
 61. Secker-Walker RH, Gnich W, Platt S, Lancaster T. Community interventions for reducing smoking among adults. In: *The Cochrane Library*, Issue 3, 2004. Art. No.: CD001745. DOI: 10.1002/14651858. CD001745.
 62. EUROSPIRE I and II Group. Clinical reality of coronary prevention guidelines: a comparison of EUROSPIRE I and II in nine countries. *Lancet*. 2001; 357:995–1001.