Commentary

The Challenges of Interpreting Research for Use by Practitioners

Comments on the Latest Products from the Task Force on Community Preventive Services

Mark W. Lipsey, PhD

The announced purpose of the *Guide to Community Preventive Services* (the *Community Guide*) is to "promote evidence-based public health practice" by providing "public health practitioners and decision makers with recommendations regarding population-based interventions to promote health and to prevent disease, injury, disability, and premature death in communities."[1] In this regard, it constitutes another initiative in what has been one of the most striking developments in program evaluation over the last decade or so—a proliferation of research review and synthesis endeavors aimed at advising practitioners and policymakers about "what works." In health care, the Cochrane Collaboration is the oldest and most highly developed of these efforts, with overlap into many of the public health domains encompassed in the *Community Guide*. Similar undertakings in other intervention domains include the Campbell Collaboration, the U.S. Department of Education's new What Works Clearinghouse, and various sponsored lists of evidence-based program recommendations, such as Blueprints for Violence Prevention; the National Registry of Effective Programs; Exemplary and Promising Safe, Disciplined and Drug-Free Schools Programs (U.S. Department of Education); and the Compendium of HIV Prevention Interventions with Evidence of Effectiveness (Centers for Disease Control and Prevention).

These various attempts to package the results of intervention research in ways that influence program practice generally fall into one of two categories. The model program approach identifies programs very specifically, virtually by "brand name," and reviews studies of their effectiveness. A program that shows positive effects on the outcome(s) of interest in studies that meet stipulated methodologic standards, and perhaps other criteria, is dubbed "effective" or "promising" or a "model," and given a place on the sponsor's list of recommended programs. Olds's Nurse-Family Partnership program[8] and Chamberlain's Multidimensional Treatment Foster Care program,[9] both included in the reviews in this collection, for example, appear on most such lists.

The model program approach has the advantage that the programs recommended are specific and individuated. Practitioners and policymakers wishing to adopt one of these programs thus have a well-defined model to follow and, usually, an easily identified program developer or sponsor to contact for operational details. Moreover, the research supporting the effectiveness of the program was conducted on that same program model and not on a variant that is similar but could also differ in some locally important manner.

There are also several drawbacks to the model program approach. The body of research demonstrating the effects of any one model program is usually rather thin, often consisting of only a few studies at most. In addition, those studies are frequently conducted by the program developer and may investigate implementations of the program configured mainly for research purposes and not for continuing practice. These circumstances raise questions about the generalizability of the results to other situations and applications, especially when the program is implemented under circumstances of routine practice without the close involvement of the developer.

Model programs also present rather specific program recipes to be followed. If those protocols are not followed with high fidelity, it greatly diminishes the basis for expecting that the results found in the supporting research will follow. In practice, however, program models are frequently modified and adapted to at least some extent when implemented locally.[2,3] Model programs thus embody sufficient specificity to permit implementation with fidelity but may require extraordinary organizational effort to ensure that they are implemented that way.[4]

The other major approach to the translation of research evidence about effective programs into advice for practitioners and policymakers is systematic synthesis or meta-analysis. In this approach, what constitutes a program is typically defined more broadly than in the

From the Center for Evaluation Research and Methodology, Vanderbilt Institute for Public Policy Studies, Vanderbilt University, Nashville, Tennessee

Correspondence and reprint requests: Mark W. Lipsey, Director, Center for Evaluation Research and Methodology, Vanderbilt Institute for Public Policy Studies, 1207 18th Avenue South, Nashville TN 37212. E-mail: mark.lipsey@vanderbilt.edu.

0749-3797/05/$–see front matter
doi:10.1016/j.amepre.2004.09.026

model program approach so that it encompasses a set of acceptable variations around a group of specified common characteristics. In this context, school-based peer-mediation training for conflict resolution would be a program as would bans on specified firearms or ammunition, as described in the paper on firearms laws in this collection. By using these more generic definitions of programs, the research synthesis approach provides practitioners with correspondingly less distinct blueprints to follow if they wish to implement one that has been found effective. How to identify authoritative sources about how the program is supposed to be conducted may also be ambiguous.

A significant advantage of the synthesis approach over the model program approach, however, is that it generally draws on a larger body of research to investigate program effectiveness. In addition, the broader definition of what constitutes a program means that relevant studies will encompass some diversity of program variations, participants, settings, and researchers. If the synthesis finds reasonable consistency in the effects across this diversity, practitioners have some assurance that program effects may be robust enough to appear despite inexact replication of a given program model. If the synthesis finds differential effects with program variation, it may be possible to determine which characteristics are associated with the strongest effects, and to inform practitioners of those relationships. The meta-analysis techniques now available for such applications provide sophisticated tools for investigating both the consistency of effects and the characteristics associated with their differences.[5,6]

Reviews for the *Guide to Community Preventive Services*

In this context, the approach used to summarize research evidence for the reviews in the *Community Guide* is an interesting hybrid. The Task Force presents the methodology it has developed as a form of systematic synthesis.[7] As the papers in the collection presented here illustrate, however, those methods may be applied to sets of studies grouped according to such tight program definitions that they virtually constitute model programs. The review of "program-intensive therapeutic foster care for chronically delinquent juveniles" in this volume, for instance, includes three studies, all conducted by Chamberlain on her Multidimensional Treatment Foster Care program. Moreover, the review procedures for a few studies of this sort operate to first screen for methodologic quality, and then draw conclusions largely from the findings of the individual studies found acceptable, as is typical of a model program approach.

The Task Force methodology also includes structured cross-study synthesis techniques that can be applied to larger groups of studies. A specialized percentage change effect size is computed for each study, median values are reported for the group of studies and selected subgroups, and the interquartile range is examined to assess the consistency of study effects. In the reviews in this volume, for example, these techniques come into play for the studies on the effect of home visitation on child maltreatment.

The advantage of this scheme is its flexibility—it can function as a systematic synthesis approach when a relatively large number of studies is available under the specified program definition, or as a model program approach when there are very few. The downside, as with any hybrid system, is that it is not optimized for either application. As a model program approach, the procedures for assessing the evidence of effectiveness emphasize consistency across multiple studies that are not constrained to deal with replicates of the same program model, and discount evidence based on a single large, high-quality study. Furthermore, when all the studies reviewed are of a single program, as in the case of Chamberlain's Multidimensional Treatment Foster Care mentioned above, the program definition and associated recommendation to practitioners do not highlight the named model program represented but, rather, maintain the more generic description typical of the systematic synthesis approach.

As a systematic synthesis approach, the Task Force methodology uses thoughtfully selected but nonstandard methods for analyzing and summarizing the effects observed across studies. The limitations of these techniques are evident in the review of the 21 studies of the effect of home visitation on child maltreatment. Most notably, no systematic account is taken of the varying sample sizes of the studies represented as would be done in conventional meta-analysis. The median percentage change effect sizes reported give the same weight to the study with 20 subjects in each experimental group as the one with over 1300 per group. They are not supported with confidence intervals or formal tests of the homogeneity of the values contributing to those medians. Instead, the interquartile range of unweighted effect values is interpreted as an index of both consistency and the plausibility of a null effect, supplemented by vote counting of the number of effects in positive and negative directions. Moderator variables, such as professional versus paraprofessional visitors and duration of service, are examined with simple descriptive comparisons or rank order correlations, and ignore the large sample size differences among the studies and potential confounding among the moderators. By contemporary meta-analysis standards, these are crude and potentially misleading analyses. The Task Force's rationale for preferring these techniques to the more sophisticated ones readily available for analyzing empirical findings from moderate to large sets of studies is not clear.

These comments are not meant to argue that the analytic approach used by the *Community Guide* is unjustified, nor that the recommendations for practitioners that follow are necessarily flawed. The simple truth is that we do not yet know much about how best to translate research findings into advice for practitioners that they can and will actually use effectively. The gap filled by focusing on population-based prevention programs, the careful and thoughtful conceptual and methodological framework, and the distinctive approach of the Task Force makes the *Community Guide* an important experiment in its own right. The ultimate test of any such endeavor is whether it strengthens prevention practice and policy in ways that yield tangible social benefits.

On this latter point, some apprehension may be warranted. All attempts to package research findings for the purpose of advising practitioners share the assumption that practice would be improved if such advice were heeded. Specifically, these initiatives assume that practice will have positive effects like those found in the respective research studies if it emulates the programs investigated in those studies. However, much of the research we have available on prevention programs and related practices has been conducted as demonstration projects, frequently by the program developer, and under circumstances amenable to research with whatever special characteristics that entails. The various schemes for providing advice to practitioners and policymakers based on such evidence have not yet been validated with their own outcome studies. Validation will require evidence showing that attainable implementations of programs shown effective in research will produce the expected effects in routine practice. At present, we have little evidence about the effects of taking research-based programs to scale in public health and related areas of mental health, education, welfare, and criminal justice.

References

1. Truman BI, Smith-Akin CK, Hinman AR, et al. Developing the Guide to Community Preventive Services—overview and rationale. Am J Prev Med 2000;18(suppl 1):18–26.
2. Gottfredson DC, Fink CM, Skroban S, Gottfredson GD. Making prevention work. In: Weissberg RP, Gullotta TP, Hampton RL, Adams GR, eds. Healthy children 2010: establishing preventive services. Thousand Oaks CA: Sage, 1997:219–52.
3. Ringwalt CL, Ennett S, Johnson R, et al. Factors associated with fidelity to substance use prevention curriculum guides in the nation's middle schools. Health Educ Behav 2003;30:375–91.
4. Elliott DS, Mihalic S. Issues in disseminating and replicating effective prevention programs. Prev Sci 2004;5:47–52.
5. Cooper H, Hedges LV, eds. The handbook of research synthesis. New York: Russell Sage, 1994.
6. Lipsey MW, Wilson DB. Practical meta-analysis. Thousand Oaks CA: Sage, 2001.
7. Briss PA, Zaza S, Pappaioanou M, et al. Developing an evidence-based Guide to Community Preventive Services—methods. Am J Prev Med 2000;18(suppl 1):35–43.
8. Olds DL, Prenatal and infancy home visiting by nurses: from randomized trials to community replication. Prev Sci 2002;3:153–72.
9. Chamberlain, P. Treating chronic juvenile offenders: advances made through the Oregon Multidimensional Treatment Foster Care Model. Washington, DC: American Psychological Association, 2003.