Introduction: How Is the Growing Concern for Relevance and Implementation of Evidence-Based Interventions Shaping the Public Health Research Agenda?

The first set of reviews in this volume, under the Epidemiology and Biostatistics section, signals a growing recognition of the methodological challenges and underdeveloped opportunities for public health research. The thread weaves from an overview of research and evaluation designs for dissemination and implementation studies, through subsequent methodological articles on research designs and sources of evidence such as uncontrolled confounding in health sciences, to natural experiments, and through to surveillance systems. From there, the thread is picked up on the social and behavioral uses of network theory and the analysis and uses of surveillance systems in public health's attempts to track and evaluate obesity prevention efforts. Further traces of the thread are found in the Environmental and Occupational Health section in two reviews on assessing the exposome and in subsequent sections where articles address similar methodological issues in applying research to policy and practice.

The first review, by C. Hendricks Brown et al., poses the issues raised by the growing recognition of the gap between evidence and practice, culminating in the title of an Institute of Medicine report referring to the gap as a "quality chasm." Dissemination and implementation became compelling issues for the National Institutes of Health (NIH) when Congress challenged the agency in the early 1970s to account for the degree to which medical and public health practitioners were receiving and applying the apparent growth of knowledge. Congress had generously funded the research that provided the new knowledge, particularly in chronic diseases and more particularly in heart disease prevention and control, which would understandably concern the age group making up a majority of the US Congress. The response provided to Congress by the National Heart, Lung and Blood Institute's Director, Robert Levy, included a graphic illustration of the broad investment in basic research required to produce a narrower subset of promising interventions. These, in turn, needed applied research on applicability to medical and public health practice and, from there, to a narrower subset ready for dissemination and application. That graphic illustration of the flow of knowledge from basic to applied research through successfully narrowed scope and budgeting to the point of the arrow representing dissemination and application became known as the Levy Arrow. It was sometimes invoked by other Institutes of the NIH when they needed to justify their larger budgets for basic and clinical research relative to their small investments in dissemination and implementation. However, dissemination

doi: 10.1146/annurev-pu-38-030717-100001

and implementation science at the NIH has continued as a focus, as evidenced by an annual meeting on the topic (beginning in 2007), targeted funding announcements, and a dedicated study section.

The dissemination and implementation thread weaved through the sections of this *ARPH* volume culminate, by one reckoning, with the review by Laura Leviton on external validity. The external validity formulation of the problem of dissemination and implementation is one of relevance, applicability, and generalizability of evidence. How do we proceed from research to policy and practice in settings, populations, and circumstances other than those in which the research was conducted.

The uncomfortable irony seen through the external validity lens is that the more highly controlled the studies producing evidence with strong internal validity (controls on confounding variables and selection bias) are, the more they might have squeezed out some of the external validity of that evidence. This trade-off might have been tilted in favor of internal validity by three realities. One reality is the inescapable truth that external validity cannot exist without internal validity. Another is a more historical reality of evidence-based medicine having served medical practice very well in clearing the clinical and surgical repertoires of many unsubstantiated practices. The rules of systematic reviews for deriving and publishing "evidence-based practices" from clinical research were then applied to public health research and reviews. This strengthened public health research-to-practice and policy. But "translation" struggled with the greater heterogeneity of public health settings, populations, and circumstances as compared with those of medicine. The external validity or "applicability" question caused many policy makers and practitioners to question the applicability of the "best practice" recommendations to their settings, populations, and circumstances.

A third challenge relates to where and how evidence is generated. Often, our repositories of evidence-based programs and policies (best practices) are developed through university-based research projects. These projects must survive the gauntlet of peer review that favors internal validity over external validity. Most often, the approved and funded studies are conducted in the context of high resources (highly trained staff, elegant evaluations). The resulting evidence-based programs and policies are often difficult to implement by a public health agency or nongovernmental organization with limited resources or real-world political constraints.

Another closely related issue is the need to rely more heavily on natural experiments to inform practices and policies in public health. These are naturally occurring circumstances where different populations are exposed or not exposed to a potentially causal "exposure" (e.g., a stringent new school food policy, new tobacco policies) such that it resembles a true experiment in which study participants are assigned to exposed and unexposed groups. Although they are more subject to selection bias, natural experiments often provide higher relevance and greater timeliness for public health practice because they reflect what is happening in the real world rather than what a researcher decides to study. The review by Craig et al. highlights some challenges of relying on natural experiments and suggests how causal inference can be strengthened.

These issues routinely and increasingly arise in systematic reviews or meta-analyses, such as those performed by the US Preventive Services Task Force and the US Community Preventive Services Task Force. One way that they address the problem is by specifying the populations and circumstances with compelling evidence of intervention effectiveness. However, in some cases, these contexts have narrow boundaries

compared with the overall potential impact of the studied interventions in all relevant circumstances.

The issues we have outlined, which are well articulated in this volume, hold much weight for the future articulation of public health research, policy, and practice. Many of these articles grapple with these external validity issues in various ways. Some researchers attempt to strengthen the interpretation of the original evidence by recognizing its biases. Others suggest applying innovative or alternative research methods to the original research. Others argue the need to adapt or tailor and evaluate the adaptations of the evidence in real time, in real practice settings, and in real uncontrolled populations.

Lawrence W. Green
Ross C. Brownson
Jonathan E. Fielding