AcademyHealth

Evaluating Complex Health Services Interventions:

Challenges, Goals, and Proposals for Progress

June 2016

Report from an AcademyHealth Meeting
December 16, 2015
Washington, DC

Introduction/Problem Statement

Recently, the rate of change in health policy at all levels has increased dramatically. From programs to implement health system reform to the responses to emerging threats such as Ebola and Zika, policymakers, health professionals, and community and patient advocates are being pushed to respond to public health and clinical issues with ever more complex interventions adopted and implemented on ever shorter timelines. All this must happen in an atmosphere of polarized public debate, underscoring more than ever the importance of an evidence base to justify decisions.

In response, AcademyHealth convened a meeting, *Evaluating Complex Public Health Interventions*, to learn from policymakers, advocacy groups, research funders, and researchers how to strengthen the evidence base on which policy can rest. The goals of the meeting were to understand current obstacles to evaluating policies and complex interventions and to chart a path to achieving both rigor and relevance in future research.

In preparation for the meeting, AcademyHealth invited nine presentations to learn about different options for generating policy-relevant evidence (see box: Presentations Delivered at the *Evaluating Complex Public Health Interventions* Meeting; see appendices for the agenda and list of participants). AcademyHealth then convened 54 experts on December 16, 2015, to identify strategies to make research more timely and impactful. More than half the participants could be thought of as decision makers—people whose jobs are to set policies or implement health care and public health programs and interventions, design advocacy strategies, or fund research. The remaining participants were researchers from a wide variety of disciplines, representing experience with a broad range of research methods. There was a focus on inviting researchers with particular experience in ensuring that scientific results reach decision makers and are included in policy discussions.

The ensuing discussion generated myriad observations and lessons for the producers, funders, translators, and users of health services research. This report begins with a description of current challenges to achieving an evidence-based approach to identifying complex health services interventions that will improve outcomes. It then lists several general goals recommended by the meeting participants to increase the rigor of research. The meeting conversation then turned to strategies for enhancement, which are grouped into five categories:

(1) Establishing relationships between decision makers and researchers to increase the usefulness and impact of research
(2) Planning research collaboratively
(3) Optimizing research design to answer questions faced by decision makers
(4) Choosing data analytic methods that make research more useful
(5) Getting research into the hands of decision makers

The meeting concluded with discussion of proposals for change. They are described in the penultimate section of this report, followed by proposed next steps.

Background

AcademyHealth has been working on a number of initiatives to strengthen rigor in evaluating complex health interventions. In June 2015, AcademyHealth collaborated with several partners in the United Kingdom—The Health Foundation, Medical Research Council, National Institute for Health Research, and Universities UK—to conduct a symposium in London on evaluation. The symposium titled *Evaluation London 2015: Evaluating Services and System Innovations in Health Care and Public Health* focused on new and emerging approaches to evaluation, including strengths, applications, and limitations as well as key challenges and future research directions. Proceedings from the symposium have just been published in an e-book to inform future approaches to health evaluation, *Challenges, Solutions and Future Directions in the Evaluation of Service Innovations in Health Care and Public Health* (http://dx.doi.org/10.3310/hsdr04160).[1] AcademyHealth continues to push the frontier of evaluation methods in conjunction with UK partners through joint summits at our 2016 Annual Research Meeting and the 2016 Health Services Research UK Symposium.

In addition, AcademyHealth has engaged in advancing research and evaluation designs implemented in patient-centered outcomes research in conjunction with the Patient-Centered Outcomes Research Institute (PCORI). More specifically, by evaluating and providing thorough assessments of PCORI's Draft Methodology Report (2012) developed by PCORI's Methodology Committee, AcademyHealth has shaped and promoted the importance of the contexts of interventions along with a diversity of data sources, rigorous study designs, and analytic methods in evaluations to determine program effects. AcademyHealth continues to engage with PCORI to augment its New and Revised Methodology Standards and works with PCORI through additional communication channels to foster rigor in evaluation.

Presentations Delivered at the *Evaluating Complex Public Health Interventions* Meeting

The Rationale and Potential for Pursuing Evidence-Based Health Policy

- Role of Evidence in Federal Policymaking. Ron Haskins, Brookings Institution
- Bridging the Gap between Research, Policy, and Practice. Vivian Tseng, William T. Grant Foundation

Overview of Innovative Methods

- Key Innovations in Design and Methods—Examples from UK. Rosalind Raine, University College London
- Evaluation Context and Mechanisms. Brian Mittman, U.S. Department of Veterans Affairs

Panel of Possibilities

- The Promise of Randomized Evaluations in Public Health. Mary Ann Bates, J-PAL North America
- Bayesian Analysis as an Innovative Approach to Evaluation of Complex Interventions. Nancy McCall, Mathematica Policy Research
- Analytic Approaches for Complex Multicomponent Interventions. Glen Mays, University of Kentucky

Reactor Panel: Hearing from Users at the Community Level

- Government Public Health Professional. Bowen Chung, Los Angeles County Department of Mental Health
- Nonprofit Community Service Agency. Loretta Jones, Healthy African American Families

I. Evaluating Complex Health Services Interventions: Where We Are Now

Challenges to Using Research as Currently Performed

In 2015, a bipartisan group in Congress introduced the Evidence-Based Policymaking Commission Act, calling for initiatives to ensure that government policies are based on research.[2] The act should be beneficial to both researchers and decision makers in the future. However, currently, it is difficult for decision makers to identify policies that can be adopted based on evidence.

A common refrain from government officials and consumer advocates was that the available research often does not answer the questions faced by decision makers because the research gives average effect sizes without explaining how effects vary across groups. Others noted that research reports often do not include enough information about a given study's context, making it difficult for decision makers to ascertain whether the findings are applicable to their own communities. In addition, a public health professional said that new research programs often progress too slowly and that their answers arrive too late. Finally, a federal official expressed concern about the frequent absence of information about long-term outcomes.

Challenges to Doing the Research Decision Makers Want

It is not easy for researchers to know that they are engaged in projects that will affect real-world decisions. Few researchers reported long-term relationships with decision makers that help them choose research questions and study designs. Those that did have such connections noted that achieving the optimal level of comfort and communication required substantial time and that funding to support that time was rarely available.

In the absence of such relationships, researchers may find it difficult to generate evidence at the pace preferred by and on the topics of most interest to decision makers. Common issues include implementing the intervention before the evaluation is planned or even funded. This can cause or be exacerbated by difficulty in obtaining important data about the intervention, the population served, or the context in which the intervention is being implemented.

There are also critical issues of study design. Randomized evaluations (also known as RCTs or randomized controlled trials) are widely viewed as the most efficient and reliable way to determine clearly the outcomes of particular interventions. The benefit of using random assignment to determine who is offered a program and who is assigned to a control group is that it allows researchers to attribute any differences observed between the groups to the program itself or to chance rather than to the characteristics of the people who chose to enroll in the program. (Research has shown that the effect of who selects or is selected to be in a program can have more impact than the intervention itself and that selection bias can cause researchers to reach erroneous conclusions based on observational data.[3])

While randomized evaluations are a powerful tool for research, they are not suited to every problem and can be challenging to implement in some cases. In general, randomized evaluations are best suited to cases where a program has limited resources and must limit participation or to programs that are rolling out new and unproven variations on existing programs. In some cases, it can be difficult to convince policymakers or organizations to assign program participants to a program randomly—because randomization would either (1) represent a change in the status quo and would give those running the program less ability to choose who is selected to participate or (2) likely delay full implementation of a program that organizational leaders are confident they should pursue.

In the cases where randomized evaluations are strong, they can prove to be simple and inexpensive methods of building research into existing programs. For example, if a program has only limited resources to enroll participants, a random lottery is often the fairest and most transparent way to choose participants. (The alternative is often giving the program to the first people who apply, who are often wealthier and more advantaged than people who might apply later.[4]) Given that lotteries select random treatment and control groups, adding an evaluation component can be relatively simple. In other cases, programs are interested in rolling out a change in their operations and constrained in how quickly they can do so—

once again, they can randomly choose who gets the new program. Another situation where randomized evaluations are often particularly strong is when programs are looking to test new, unproven variations on existing programs and want to compare the new variations to the status quo.

As with any analytic tool, there are also situations where randomized evaluations are not well suited. Given the clear advantages of randomized evaluations, there have been some groups that have overemphasized the use of this approach rather than treating it as a tool with particular applications. For example, if those running the program have strong preferences about who participates (for instance, they want to target carefully selected people), it can often be difficult to assemble a large enough group of potential participants for a randomized evaluation. Narrow enrollment criteria also limit generalizability. In addition, if the optimal approach to randomization is at the organizational level (for instance, randomizing medical groups to receive a new payment method), it can be very difficult to convince organizations to accept randomization. In general, conducting randomized evaluations or other studies with sample sizes that are too small is not advisable, as there is often little potential for producing useful results and a large potential for lacking the statistical power to detect the program's impact.

Another priority for researchers using randomized evaluations is to recognize that, despite the advantages of the method, randomized evaluations can fall prey to problems that can affect any scientific study. In all scientific studies, researchers need to be careful about the potential for the study to create unintended effects, the difficulty of generalizing research findings, and the need to consider both short-term and long-term effects (which may have contradictory implications for policymakers). Several examples of these general scientific problems follow:

- In 2004, most hospitals in Ontario agreed to be randomized to release (or not) a public report card about cardiovascular care. The hospitals knew that public reporting for all hospitals was coming within two years. In this context, control hospitals improved their performance substantially (administration of beta-blockers from 32 to 71 percent and of statins from 58 to 86 percent). Intervention institutions also improved dramatically, but their response was swamped by the secular trend of improving the measures in anticipation of universal public reporting. In the end, the intervention produced statistically significant improvements among the treatment versus control group on only one of 13 measures.[5]

- Another challenge for scientific studies is determining whether or not they are generalizable beyond the original setting. If evaluations—whether they are randomized or draw on quasi-experimental methods—are conducted in an idiosyncratic situation (and many settings are in some way idiosyncratic), people who learn of the results may doubt that their setting is similar enough for the trial results to be applicable. For instance, although the Ontario trial produced improvement in one measure, a thoughtful decision maker might ask whether public reporting would have had any effect at all if the implementing hospitals were not already poised to change. If other hospitals were not similarly motivated, would committing resources to public reporting have any effect? These types of questions arise with every scientific study, and the lack of easy answers can be unsatisfying to decision makers. Therefore, it is important to have a framework for understanding what types of findings and lessons are most likely to translate across contexts.

- All prospective studies must also contend with the challenge of obtaining funding to follow participants for more than a year or two. In some cases, short-term gains themselves may be enough to justify programs, but that is not always the case. In other cases, programs might cause disruption effects (and apparent negative outcomes) when first introduced but also produce strong and enduring positive effects. Tracking a program over time is a challenge for all prospective evaluations, and an important one. Even those programs that do engage in long-term tracking may face the criticism that circumstances have changed since the intervention began. For instance, in a recent trial, a 25-year follow-up of a program of breast cancer screening found no difference in mortality for women age 40 to 59 who had annual mammography versus those who did not.[6] Critics argued that the findings were not relevant because mammography and other screening tests have improved since the trial started.[7]

- Moreover, the desire for understanding long-term outcomes often conflicts directly with the other priority of decision makers to get research results fast. How does one generate evidence quickly about long-term outcomes?

These examples demonstrate that building evidence through science and evaluation is a difficult proposition. Researchers are well served in being transparent about the limitations of their work in general. Researchers and decision makers must also be realistic about the trade-offs in time and money that different types of studies suggest. In some cases, randomized evaluations may be cheaper and faster than other study designs. In other cases, they may be more expensive and time-consuming. The timeliness, cost, and generalizability of a study will be determined by the methods, measures used, and duration of the data collection efforts attached to the study design. A randomized evaluation or other study design is generally conducted with far less cost when using data that are already being collected (administrative data) than if the researchers need to collect original data specifically for their research project.

Treating methods of evaluation as a set of tools rather than as a ranked set of "better" or "worse" methods can help researchers and implementers remain realistic about the limits and trade-offs inherent in different forms of research. Randomized evaluations are the tool best suited to show that a program caused a particular outcome, but other methods can have different advantages. For any study of rare events or long-term outcomes, creating appropriately randomized, prospectively followed populations of adequate size and without crossover is very difficult and expensive and hence is almost never achieved.[8] In these settings, observational studies can be superior. For instance, post-marketing surveillance registries may be required to recognize uncommon but serious medical device side effects that are not frequent enough to be noted in randomized trials required for device approval. Similarly, observational retrospective trials may be preferable for understanding the association of a given intervention with very long-term outcomes (such as whether early childhood health or educational programs are associated with health outcomes during adulthood). In addition, observational studies may be the only option for studying a large number of organizations and may be important for understanding the impact of organizational characteristics or contextual factors on the outcomes of interest.

In addition to quantitative analyses, qualitative studies may generate stakeholders' hypotheses about the relationships between and among population, organization, and contextual variables and the impact of an intervention. However, because these approaches do not involve random assignment, they are not as well suited to proving causality.

Finally, several researchers at the meeting stated that the struggles to produce relevant research are exacerbated by lack of funding, noting that government invests significantly less in assessing its services than occurs in other sectors of the economy. A recent National Science Foundation (NSF) report supports this observation. The NSF found that private industry commits, on average, 3.3 percent of its expenditures to research and development[9] in order to figure out what the next product or service should be or how current products and services could be made better. Many service industries spend even higher percentages. The federal government, in contrast, commits just over 1 percent of expenditures to research (little of which goes to assessments of the government's "product"—the services it provides), while state and local governments spend almost nothing (even though they account for about one-third of government expenditures nationwide).

Moving toward Evidence-Based Policies

Meeting attendees noted several opportunities to make research more helpful to decision makers. Advancing the field, however, necessarily depends on a clear statement of what the end product should be. Participants identified five general goals for improving the rigor and relevance of health services research:

(1) Ask the right question (the question a decision maker needs answered)
(2) Complete research quickly
(3) Be able to assess applicability of the research to a specific local context
(4) Get evidence to decision makers at the time of decision-making
(5) Reduce the cost of evaluations

Many participants felt that strategies to achieve those goals were available and could find support among researchers and from funders.

Participants offered many specific approaches, which may be broadly grouped into the following strategies: (1) establishing relationships that foster better research, (2) collaboration between researchers and decision makers on research planning, (3) optimizing research design, (4) adopting state-of-the-art data analysis, and (5) ensuring that decision makers can access research. The next sections of this report review these strategies and describe their links to the goals identified to improve research (Table 1).

Table 1: How Strategies Outlined during the Meeting Relate to Identified Goals for Advancing Health Services Research

	Identified Goals for Improving Health Services Research				
Strategy from Meeting	**Ask the Right Question**	**Complete Research Quickly**	**Be Able to Assess Applicability to Local Context**	**Get Evidence to the Research User at the Right Time**	**Reduce the Cost of Evaluations**
Establishing relationships that foster better research	√		√	√	
Collaboration on research planning	√	√	√	√	√
Optimizing research design		√	√		√
Data analysis innovations		√	√		√
Ensuring that decision makers can access research				√	

II. Establishing Relationships between Decision Makers and Researchers that Increase the Usefulness and Impact of Research

As with many situations in which the two sides of a relationship are not getting all that they want, several participants felt that communication could be improved. One researcher described a long-term relationship with local decision makers that had generated a series of research projects. She felt that, over time, the researchers had developed a progressively deeper understanding of decision makers' needs, while the decision makers were more aware of how research could inform their activities. However, such relationships are rare, in part because they take time to nurture.

A government official proposed to address the relationship issue by creating "idea incubators"—forums established by government or research funders in which decision makers and researchers can interact. Another idea was the "embedded researcher" who works within a service-providing organization to learn when and how research is used and what questions are asked. There are already some programs that create idea-incubating environments, although such programs are rare and mostly not sustained. The Health Services Research (HSR) Learning Consortium, supported by AcademyHealth, is a forum for educators, employers, and trainees to address training needs for the field of health services research. A key goal of the HSR Learning Consortium is to improve communication between stakeholders in the field and encourage partnerships to ensure that HSR training meets the needs of employers in the public and private sectors as well as in academia. The HSR Learning Consortium meets every year at the AcademyHealth Annual Research Meeting. An example of another approach was the Public Health Practice-Based Research Network (PBRN) program funded by the Robert Wood Johnson Foundation. The PBRNs formalized and funded long-term relationships between public health practitioners and researchers (see box: *Incubating Public Health Ideas in Practice-Based Research Networks*). Other programs also embed individual researchers, such as the AcademyHealth Delivery System Science Fellowship, which is designed to recruit and place well-trained researchers within health care delivery organizations to conduct research on topics driven by the information needs of clinical and system leaders.

In the United Kingdom, similar approaches have been embodied in the National Institute for Health Research's Collaborations for Leadership in Applied Health Research and Care (CLAHRC) program. Each CLAHRC is a partnership of the National Health Service, researchers, and local community organizations (including, but not limited to, providers of care). CLAHRCs identify research questions of interest to local stakeholders but then perform the research so that the findings will be relevant to the entire country.

Whatever the model, the goals of relationship building are fairly straightforward. A government official stated that stronger relationships should lead researchers to ask questions decision makers need answered and to be present to share evidence when decision makers face choices. In addition, a foundation executive commented that researchers could help shape research questions by bringing to the conversation theories of change or conceptual models about how policies affect health. A researcher pointed out that decision makers can increase the chances that research will be useful by making it clear to researchers what criteria will be applied during decision-making.

Incubating Public Health Ideas in Practice-Based Research Networks

In 2007, a team at the Robert Wood Johnson Foundation decided to address a critical paradox in public health in the United States. While most health care expenditures went to conditions that were preventable in theory, there had not been much research conducted about how public health professionals could prevent those conditions in practice.

The solution sought by the team was the establishment of long-term links between academic institutions and local public health practitioners. The overall structure of the program created by the team, Public Health Practice-Based Research Networks (PBRN), forged the needed links. Rather than relying on a program office at an academic institution, the program direction-setting task was the responsibility of three public health practice associations—the National Association of County and City Health Officials (NACCHO), the Association of State and Territorial Health Officials (ASTHO), and the National Association of Local Boards of Health (NALBOH)—and an academic partner.

Within this structure, the plan called for 12 PBRNs. However, the call for proposals generated so much interest that, while 12 PBRNs received technical assistance and funding, a further 16 PBRNs also received technical assistance only.

The focus on creating an infrastructure supporting long-term relationships was extremely successful. Over 900 local health departments, 20 state agencies, and 35 academic units joined a PBRN. The 12 PBRNs eligible for funding completed over 50 jointly designed research projects. Public health practitioners gained access to a database of articles addressing questions that PBRN members felt were important; the articles were housed in the Health Services and Sciences Research Resources Website at the National Library of Medicine (https://www.nlm.nih.gov/hsrph.html). The three public health practice associations developed a harmonized member survey that they have used for several years and from which they can now identify trends in public health and public health practice.

Crucially, the PBRN approach was also beneficial for the researchers. Some received funding, and others took on doctoral students, which increased research output. Some grantees advanced on university faculties. And, given that the research performed was developed with a research user, researchers knew that the work was used.

III. Planning Research Collaboratively

The next step to generating relevant research findings with a greater likelihood of being used by decision makers is to answer important questions. This involves first understanding decision makers' information needs and then identifying an approach to generate the evidence that would help decision makers act. This is often best accomplished collaboratively.

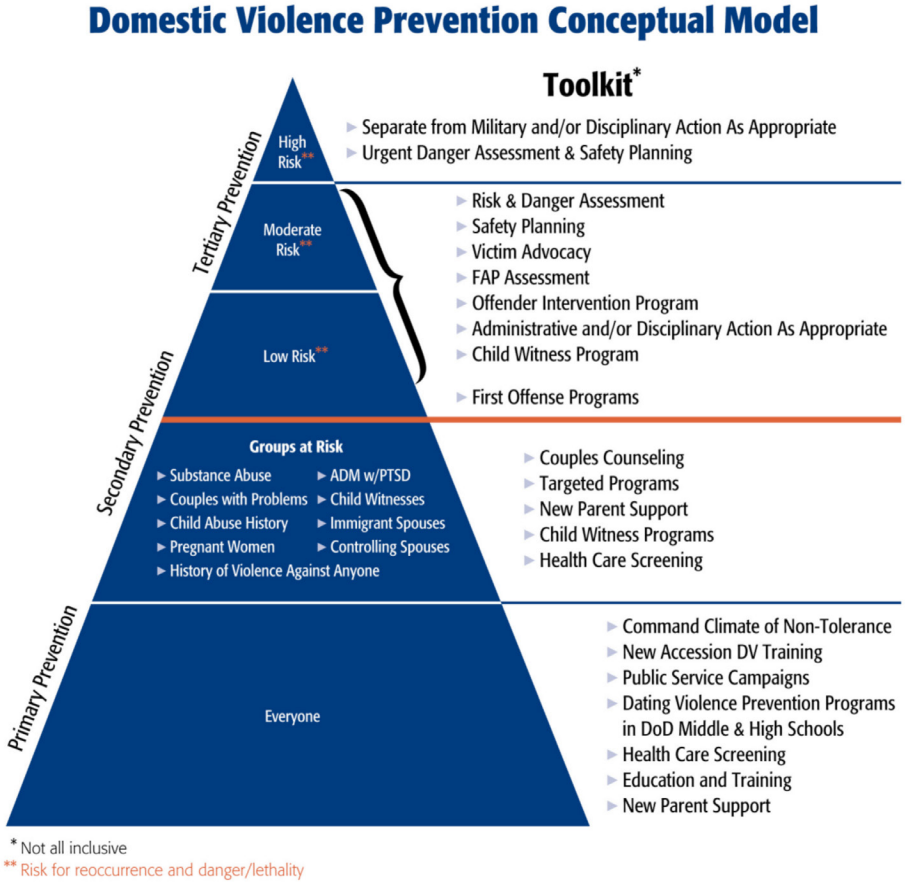
How Conceptual Models and Theories of Change Can Facilitate Collaboration

As one public health leader said, "I don't want to hear whether it did or did not work somewhere else in the past. I need to know whether it will work here in the future and under what circumstances and for whom." The implication for researchers is that users of evidence want researchers to understand the context in which their research will be performed and findings used. This means recognizing the local factors that will enable programs to be successful or that may pose barriers to success.

The task of jointly identifying effect mediators, however, is easier when situations are analyzed in the context of a conceptual model of the causes of the health problem (referred to in some disciplines as a causal pathway), a theory of how change can occur, and an understanding of the factors that modify the effect of an intervention.

An example comes from U.S. Department of Defense (DoD) efforts to address domestic violence. The Defense Authorization Act of 1999 mandated that the DoD produce a plan to address high rates of domestic violence in military families. DoD convened a Defense Task Force on Domestic Violence that included not only military leaders and victim advocacy groups but also researchers. Over three years, the group developed a strategic plan. As part of the plan development process, researchers guided the group through development of the conceptual model (see figure: *Domestic Violence Prevention Conceptual Model*).[10]

We highlight the conceptual model not to endorse it as the best lens through which to view the complex phenomenon of domestic violence. Rather, we include the model to illustrate that a researcher could use it as the springboard for conversations with military leaders about how to evaluate an intervention. For instance, if military leaders were interested in knowing whether additional public service announcements (PSA) might reduce rates of violence, researchers and military officials could use the model to decide what other data to collect besides attack rates. The model suggests that a "culture of non-tolerance" among local commanders might influence violence rates; therefore, a study randomizing bases to receive or not receive additional PSAs would need to account for local culture. In addition, the model suggests that some groups are at higher risk, so differences in the percentage of the population at each base



at each risk level should be monitored, with an a priori decision possibly made to calculate the impact of the PSAs in different risk groups. These steps would go a long way to helping commanders answer, "Will this work for DoD, under what circumstances, and for whom?"

Collaborative Research Planning as a Strategy to Reduce the Cost of Research

Collaboration may also reduce the cost of research, maximizing both the chances that the research is feasible (a researcher objective) and the resources available to provide services (a decision maker's objective). For example, the DoD model suggests that rates of domestic violence are higher for couples who are new parents. By including in the research design of a new study a plan to collect parental status from the military's electronic health record, it may be possible to collect a key covariate at low cost. Similarly, researchers studying Oregon's Medicaid expansion used Medicaid claims data, available at low cost, to assess the impact of the expansion on emergency room use.[11]

Other forms of collaboration also could reduce research costs. For example, researchers could collaborate with medical specialty societies to use their registries in randomized registry trials. An example is the American College of Cardiology's PINNACLE registry, which captures information about cardiology practices nationwide.[12] Given that most cardiologists are not researchers, few participating practices are academic. However, academicians could partner with these practices through PINNACLE, greatly reducing data collection costs.

IV. Optimizing Research Design to Answer Questions Faced by Decision Makers

The next step is to select a research design that is responsive to decision makers' needs. The meeting dialogue on design started with the observation that no one study design would be sufficient to generate evidence applicable to every situation. In addition, researchers and decision makers alike reported an increasing awareness that strong determinants of health outcomes lie outside the health system. Moreover, random assignment designs often cannot control for these determinants, and investigators pursuing randomized designs often have limited ability to measure them.

There was strong sentiment that randomized trials should continue to be part of the researchers' toolbox and that more evidence could be generated by random assignment with the use of methods such as factorial designs.[13] It was also clear that many other research designs are relevant as well.

Given that a fundamental benefit of randomized trials is that randomization increases the confidence that any effect is causal, it is helpful to think about non-randomized designs in terms of how much they can be used to address causality. For instance, natural experiments,[14] instrumental variables analyses,[15] and econometric models such as sample selection models[16] and treatment effect models[17]—none of which is randomized—may offer evidence about causality that is nearly as strong as the evidence from a randomized trial. In a different vein, evidence from qualitative studies, such as focus groups and semistructured interviews with key stakeholders, may generate insight into what patients or professionals see as causal patterns or mediating factors. While qualitative data do not offer proof of causation, such data may offer the best available basis for a decision maker facing a choice.

At the very least, qualitative studies can generate hypotheses to be tested and can complement experimental studies by helping to explain why the results occurred. Many observational studies can do the same. Because of the nature of qualitative studies and the conduct of such studies in complex systems, many researchers cautioned that it can be hard to understand the cause and even the strength of associations found in observational studies. A foundation officer noted that researchers have improved their ability to represent some of the complexity of the system by using multilevel modeling.[18]

Most meeting participants agreed that, ideally, important questions could be approached by using mixed methods as different approaches can lead to different types of insights. For instance, a randomized trial might generate data about the average effectiveness of an intervention, while interviews with local leaders might highlight local assets that increase the likelihood of success in implementation of the intervention, or focus groups with patient advocates might identify local subpopulations particularly interested in trying the intervention.

Several researchers and foundation leaders noted that a key first step in planning a research program is to base the choice of study designs on an understanding of both local context and general conceptual models. In particular, participants noted that research often is reported without sufficient information about the context in which the research was performed, making the results hard to interpret. When selecting contextual factors to measure and report, researchers could refer to state-of-the-art conceptual models during research design in order to reduce the likelihood of oversights in the interpretation of results.

V. Choosing Data Analytic Methods that Make Research Findings More Useful

Researchers and patient advocates noted that decisions about analyzing data could make research results more useful. For instance, qualitative studies could involve clinicians or public health providers in coding a sample of transcripts or discussing themes, as they might understand better than researchers what local stakeholders' statements mean. Clinicians and decision makers could then be included in the team that interprets the significance of the identified themes.

Analytic alternatives are also available for quantitative methods. In an observational study design, regression discontinuity designs[19] focus the analyses on individuals or organizations that fall just above or below a threshold for eligibility for an intervention. For example, this approach could be used to study outcomes from a disease management program among diabetics whose blood sugar is just above a cutoff that makes them eligible for the program versus diabetics whose blood sugar is just below the cutoff and therefore makes them ineligible for the program but who are otherwise similar to the intervention group in terms of disease severity.[20]

Interrupted time-series analyses, in contrast, use all study subjects but compare the rate of change in the outcome variable before and after an intervention in one study arm versus another study arm with no intervention followed for the same time period. This approach has been used recently, for instance, to assess the impact of a program to reduce prescription errors introduced at one hospital, with another hospital with no intervention as a control.[21]

Bayesian models,[22] which assume when sample sizes are small that effects are near what would be expected given prior knowledge, have much to offer in evaluating health interventions. These methods tend to shrink estimated effects toward the mean for specific clusters. For example, in a study in which an intervention was tested at several public health clinics, if a particular clinic saw a large impact from the intervention but among only a small number of recipients of that intervention, a Bayesian model would shrink that clinic's estimated effect toward the mean for all clinics. This approach has the advantage of not assuming that each individual research project occurs in a vacuum of knowledge but rather recognizes that new projects are usually conducted in a rich background of existing information. Bayesian analysis allows policymakers to make intuitive probability statements about the magnitude of program effects.

Another researcher noted that it was crucial to incorporate context into quantitative analysis in other ways. One example is a context in which a particular clinical quality issue (e.g., smoking cessation) had become such a common topic of conversation that providers were willing to be randomized to receive (or not) assistance with referring patients to smoking cessation clinics. A difference in difference methods should be used in comparing the two arms of the study to account for improvement in the control group that might reflect the context in which smoking cessation is broadly discussed. For example, other strategies (e.g., physicians and nurses devoting more of their own hours to call in referrals to the smoking cessation clinic or spending more time with the patient on other cessation methods, such as use of nicotine gum) could produce improvements among providers randomized not to receive the intervention.[23]

VI. Getting Research into the Hands of Decision Makers

A federal official noted that even the most relevant research has no impact if it does not reach decision makers while they are making choices. Establishing relationships between researchers and research users will help decision makers know where to go for help. In addition, some meeting participants called for repositories of research and conceptual models, that is, widely recognized and used places where decision makers could go when a new question arises that needs an answer urgently.

Foundation leaders affirmed that getting research into the hands of decision makers is critical to their mission. In particular, participants stressed that the ideal would be a rapid cycle of research use in policy decisions, followed by implementation of a policy that would then stimulate more research to assess the effect of the intervention. However, one researcher pointed out the contradiction between rapid-cycle research and learning about long-term outcomes.

Organizations like AcademyHealth that play a role in knowledge translation are working to bridge the gaps among research, policy, and practice.[24] The Translation and Dissemination Institute at AcademyHealth helps the field of health services research move its findings more effectively into policy and practice. The institute undertakes activities that help research producers better understand the needs of research users and serves as an incubator for innovative approaches to moving knowledge into action. In addition to listening to policymakers, delivery system leaders, and other health care stakeholders in order to identify pressing research needs, the institute focuses on learning from a variety of disciplines on how to advance the art and science of translating and disseminating research findings for policy and practice as well as on innovating by testing new approaches and tools. A variety of communication tools can be used in disseminating research results and translating research findings into practical applications for users' needs.[25] They range from traditional research articles in journals such as *Health*

Services Research to policy-oriented research articles in journals such as *Health Affairs* to policy briefs and practice change guides that synthesize research findings to address user needs and communicate information to the appropriate audience. For example, the Urban Institute's Quick Strike series has been producing policy briefs that provide timely analysis of health policy issues related to implementation of health care reform; the briefs allow researchers to respond in real time to policy questions and take advantage of existing data sets and research knowledge.

VII. Advancing Evaluations of Complex Interventions

As the meeting turned to identifying proposals for change, participants revisited previously identified themes, with the intention of identifying more specific suggestions. Table 2 presents the suggestions made and their relationship to the five goals identified to improve health services research.

Create Idea Incubators and Embed More Researchers

Participants noted the centrality of relationships, and the potential of idea incubators or embedding researchers seemed universally accepted. A federal official suggested that idea incubators would be more effective if they were augmented by specific, strategic investments in human capital. Such investments would focus on methods training or workshops to make decision makers aware of how few of their activities are evidence-based and to help them understand how research could inform their choices of interventions in the future. The same official envisioned idea incubators incorporating sessions for researchers to help them ask the right research questions. Such workshops might include interactions with decision makers to understand how policy is made and reports from other researchers about how and when they successfully communicated research findings to the policymaking process.

A repeated premise was that more exposure to community advocates and practicing health professionals would lead researchers to recognize the importance of addressing specific community needs and equity issues. By their structure, idea incubators and embedded researchers would increase opportunities to ask the right questions and get information to decision makers at the time they face choices.

Identify Custodians (not Curators) for Language, Theories of Change, and Conceptual Models

The conversation repeatedly highlighted the importance of conceptual frameworks that facilitate collaboration between researchers and decision makers. These frameworks are intended to reflect the

Table 2: How Proposals for Change Address Identified Goals for Improving Public Health Research

	Identified Goals for Improving Health Services Research				
Proposals for Change	**Ask the Right Question**	**Complete Research Quickly**	**Be Able to Assess Applicability to Local Context**	**Get Evidence to the Research User at the Right Time**	**Reduce the Cost of Evaluations**
Idea incubators and embedded researchers	√		√	√	
Methods workshops for public health practitioners and decision makers				√	
Methods workshops for researchers at idea incubators	√				
Wikis for conceptual models, theories of change, and glossaries	√		√		
Rapid, low-cost data access and prelinked data sets		√	√		√
Shorter time to approve access to data sets		√			
Shorter time to respond to funding requests		√			
Broader selection of study designs		√	√		√
Research funders adopting a conceptual model and a portfolio perspective	√		√		
Change incentives for academicians	√	√	√	√	

state-of-the-art understanding of why the world is how it is and theories of how it can be changed. In doing so, they use scientific language that captures key concepts clearly.

However, one federal official noted that researchers might understand scientific terms one way and policymakers another. In addition, a foundation leader pointed to the difficulty in finding a good theoretical model on which to base evaluations. The group then concluded that one of the most important next steps would be to develop shared language and models through which to describe the world in which health interventions are tested and used.

Unfortunately, language and theory are not static phenomena, so the solution cannot be a one-off project to build a glossary. However, participants did not accept that the only alternative is an academic free-for-all. Just as the Cochrane database addresses calls for increasing access to systematic reviews, similar steps could be taken for glossaries of terms and conceptual models in health services research. Foundations or government agencies could accept similar roles. That is, just as the Cochrane Collaboration is a nonprofit organization that runs a systematic reviews database, so could the Robert Wood Johnson Foundation maintain a database of terms and models relating to the creation of a Culture of Health. The Laura and John Arnold Foundation could do similar work focused on the impact of criminal justice reform on long-term health outcomes.

Critical to the utility of this conceptual framework approach is that the organizations that take on this function, view themselves not as curators (making decisions about the content of a collection) but rather as custodians (maintaining something that they do not own). A researcher proposed that custodians create a series of open-source wikis that allow registered participants to update term definitions, theories, and conceptual models. Anyone would be allowed to register, but all would have to share their credentials. Given that the updating of wikis could occur only after sign-in, users of the site could see who made which changes and decide whether they agree.

Especially for theories of change and conceptual models, it would be important to allow competing views to be expressed (there might be three conceptual models of the causes of domestic violence at the same time, with ongoing discussion). Although such an approach means that absolute consensus would never be achieved, creating a central location for a discussion to take place would make it easier for non-researchers to find the state-of-the-art in thinking about complex interventions.

Strengthen the Research Infrastructure

When the conversation turned to next steps, participants made several proposals to address the pace and cost of research by improving the local and national infrastructure. One university-based researcher suggested that creating a repository of prelinked anonymized data sets for researchers' access could reduce research cost. One example called for creating de-identified data sets from electronic health records across the country and linking them to claims data, behavioral data, and other sources of information about risk factors and exposures.

Another researcher responded that getting approvals to access data—even if previously collected—was a major source of expense and slow response. As routine a task as getting Medicare claims can rarely be accomplished in under four months, with six months more typical. A third researcher asserted that the greatest source of delay in research was waiting to find out if a proposal was funded, which requires a minimum of 9 months from the National Institutes of Health or the Agency for Healthcare Research and Quality but can take 18 months or more. In summary, a series of investments in research infrastructure, especially focused on increasing access to data and timely funding, could increase the speed with which research could be accomplished.

Use a Broad Range of Study Designs

Meeting participants were in agreement that there is no one optimal approach to study design. While some strongly believed that randomization was essential to evaluating interventions—even complex ones—a large group of researchers and decision makers felt that the limitations of feasibility and generalizability made an exclusive focus on randomized trials too narrow and that sometimes randomized trials were not necessary or appropriate. The same group encouraged funders to consider using a mix of methods across funded research, though always requiring the design of non-randomized methods to be based on causal pathways, theories of change, and/or conceptual models. They argued that such an approach would allow research to proceed more quickly and address contextual factors and local assets or barriers to implementation that might not have been considered in or reported by the authors of randomized trials. Most important, matching the most robust method to the nature of the research question and identifying comparison groups should be achieved.

Research Funders Should Adopt a Conceptual Model and a Portfolio Perspective

For research funders, generating a body of evidence that clearly explains to decision makers how complex interventions interact within a complex system is always going to be challenging. However, it is made more difficult if research funders have no overarching strategy for how research is used to answer policy questions.

For instance, there was much discussion that two studies of the same intervention might vary in key features of the intervention and that understanding the differences was critical. For instance, PSA campaigns about the importance of prenatal care for reducing preterm birth might differ in the total number of advertisements, distribution channels (magazines in general versus magazines targeting populations with higher rates of preterm birth), or how graphic individual advertisements are about the consequences of preterm birth. The ultimate impact of these strategies can be understood only by capturing all the characteristics of the set of interventions.

In addition, meeting attendees stressed the importance of understanding local contextual factors—enablers and barriers or assets and weaknesses—that mediate the effect for the same intervention applied in two different settings or with two subgroups of the population. For the preterm birth example, contextual factors might include the number of providers of prenatal services available in a community or the existence of electronic connections from primary care, urgent care, or emergency rooms that facilitate referrals to prenatal care.

Reflecting on these concerns about intervention and contextual factors, a small group of participants proposed that research funders should think about these issues not only when they decide about funding individual projects but also across their entire portfolio. That is, funders should have in mind a conceptual model and theory of change—perhaps drawn from the wikis described above—as they ask how a new proposal fits into their overall portfolio.

Reliance on a conceptual model and theory of change would allow funders to adopt strategies that would accelerate the pace at which the understanding of complex interventions grows. One strategy is sequential hypothesis testing[26] of characteristics of the intervention. Continuing with the preterm birth PSAs example, especially if a funder must start with a small campaign, it is important to choose a context in which the campaign has its best chance of success; that is, a community in which key enablers are present and barriers to prenatal care are few. Of note, success in that community is not generalizable; rather, the community was chosen for its high pretest probability of success. However, a failure in that same community is important because it strongly suggests that the intervention as designed for that first trial will not be effective elsewhere. If the first trial is successful, funders could then move to more difficult contexts to retest the intervention. If the first trial fails, funders could increase the "dose" of the intervention (for example by developing more advertisements or reframing the message). Choosing the initial and subsequent interventions and communities wisely, however, requires an understanding of the key characteristics of the intervention and contextual enablers and barriers (that is, having a conceptual model or theory of how change might occur).

For interventions that show potential, it then becomes valuable to understand the degree of importance of enablers and barriers and how much they mediate the effect of the primary intervention. Studies addressing this have been referred to as parallel hypothesis testing[27] because they take place in parallel with the sequential testing of different strengths of the primary intervention. Again, however, knowing what needs to be studied in parallel requires a conceptual model and, preferably through a highly transparent process such as open-access wikis, assurance that the model evolves over time. Taken together, these strategies can help research funders understand their portfolios overall and how that body of research is valuable to decision makers.

While not discussed at the meeting, the concept of funder portfolio management is also important across funders to better understand what research, using which methods, is being conducted and to increase the efficiency of and, ultimately, knowledge gained from research investments. For example, the National Library of Medicine supports HSRProj, a database of over 30,000 HSR studies funded in North America by over 250 public and private research funders. Funders rarely use this resource to understand who else is funding work in any particular domain before launching new initiatives or to assess their funding in the context of other funded work. One recent exception is the Patient-Centered Outcomes Research Institute (PCORI), which supports the Strategic Portfolio Initiative at AcademyHealth to analyze PCORI funding in specific areas.

Change Incentives for Researchers

A frequently stated obstacle to changing researcher behavior is the misalignment between the incentives offered to researchers and the five goals identified for improving research. Instead of being incentivized to build trusting relationships with decision makers and devote time and resources to getting their results into decision-making processes, researchers currently are rewarded for publishing in respected journals, being cited by peers, and receiving funding.

Several meeting participants suggested that research funders should work with academic institutions to recast researcher incentives. Criteria for funding future research could be adjusted to account for evidence of previous impact on policy. For instance, in 2014, the United Kingdom adopted a Research Excellence Framework[28] to be used to allocate funding to universities. The framework explicitly

allocates 20 percent of a university's score to impact on "economy, society, public policy, culture and the quality of life." However, the specific approach adopted for measuring impact has been controversial.[29] Universities submit impact case studies to the Higher Education Funding Council of England, with emphasis on proof of use of the research (such as incorporation into guidelines). The council in part uses automated text mining, including topic modeling, keyword searching, and information extraction (techniques to pull structured data from unstructured text of the case studies to assess use of research).[30] Some universities are unsure that such an approach accurately captures their impact. In addition, to the extent that impact is measured by media mentions, evidence suggests that media coverage is stronger for research whose design is easier to understand but lower quality (e.g., observational studies often receive more coverage than randomized trials).[31] Nonetheless, although the optimal methods for impact assessment for UK universities is unclear, the framework represents a commitment at the policy level to continue developing ways to evaluate and base research funding on impact.

Next Steps

Meeting participants identified a variety of challenges to formulating evidence-based policy, along with several actionable ideas to move the field forward. Natural next steps include pursuing implementation of the proposals for change outlined above.

Among the proposals, some may make more sense as first steps. In particular, the meeting on this topic itself could be viewed as a one-day idea incubator. Prioritizing proposals for change may be best done after gathering further input from more experts. Thus, AcademyHealth could consider extending the conversations initiated at the meeting into subsequent forums addressing each proposal for change and how it could be implemented. The conversations could take place online (through webinars and wikis), in person, or both. Establishing these forums is a role AcademyHealth has played previously. The process of creating the forums could start with the identification of interested partners who could engage in developing plans to help AcademyHealth and the field of health services research achieve the goals outlined in this first meeting.

About the Authors

R. Adams Dudley, M.D., M.B.A., is professor of medicine and health policy and director of the UCSF Center for Healthcare Value at the University of California, San Francisco. Ben Struhl, M.P.P., is policy manager at J-PAL North America, MIT Department of Economics.

Acknowledgments

AcademyHealth acknowledges Rosalind Raine, Ph.D., FFPH, M.Sc., M.B.B.S., B.Sc., professor of health care evaluation, head of department of applied health research, UCL, and director of NIHR CLAHRC North Thames, and Mary Ann Bates, M.P.P., deputy director of J-PAL North America, MIT Department of Economics, for their review of this report.

Suggested Citation

Dudley, R.A., & Struhl, B. "Evaluating Complex Health Services Interventions: Challenges, Goals, and Proposals for Progress," AcademyHealth, June 2016.

References

1. Raine, R, Fitzpatrick, R, Barratt, H, Bevan, G, Black, N, Boaden, R et al. Challenges, Solutions and Future Directions in the Evaluation of Service Innovations in Health Care and Public Health. Health Serv Deliv Res, 2016;4(16).
2. Clark, CS. Bipartisan Evidence-Based Policy Bill Passes House. July 28, 2015. Available at http://www.govexec.com/management/2015/07/bipartisan-evidence-based-policy-bill-passes-house/118661/.
3. Grodstein, F, Clarkson, TB, & Manson, JE. Understanding the Divergent Data on Postmenopausal Hormone Therapy. N Engl J Med., 2003 Feb 13:348(7):645-50.
4. Buchanan, Allen. Health Care Delivery and Resource Allocation. In *Medical Ethics*, 2d ed., ed. Robert M. Veatch. Boston, MA: Bartlett and Jones, 1996.
5. Tu, JV, Donovan, LR, Lee, DS, Wang, JT, Austin, PC, Alter, DA, & Ko, DT. Effectiveness of Public Report Cards for Improving the Quality of Cardiac Care. The EFFECT Study: A Randomized Trial. JAMA, 2009;302(21):2330-7.
6. Miller, AB, Wall, C, Baines, CJ, Sun, P, To, T, & Narod, SA. Twenty-Five Year Follow-Up for Breast Cancer Incidence and Mortality of the Canadian National Breast Screening Study: Randomised Screening Trial. BMJ, 2014;348:g366.
7. Javitt, MC. The Canadian National Breast Cancer Screening Study: Caveat Emptor. American Journal of Roentgenology, 2014;202:W505-W505.
8. Black, N. Why We Need Observational Studies to Evaluate the Effectiveness of Health Care. BMJ, 1996;312(7040):1215-8.
9. http://www.nsf.gov/statistics/2016/nsf16301/#chp2.
10. Defense Task Force on Domestic Violence. Year 3 Report. 2003. Available at http://www.ncdsv.org/images/Year3Report2003.pdf.
11. Taubman, SL, Allen, HL, Wright, BJ, Baicker, K, & Finkelstein, AN. Medicaid Increases Emergency-Department Use: Evidence from Oregon's Health Insurance Experiment. Science, 2014 Jan 17;343(6168):263-8.
12. http://cvquality.acc.org/en/NCDR-Home/Registries/Outpatient-Registries.aspx.
13. Allore, HG, & Murphy, TE. An Examination of Effect Estimation in Factorial and Standardly-Tailored Designs. Clin Trials, 2008;5(2):121-30.
14. Nikpay, SS. Federal Support for Family Planning Clinics Associated with Dramatic Gains in Cervical Cancer Screening. Womens Health Issues, 2016 Jan 9. pii: S1049-3867.
15. Kwok, MK, Yeung, SL, Leung, GM, & Schooling, CM. Birth Weight and Adult Cardiovascular Risk Factors Using Multiple Birth Status as an Instrumental Variable in the 1958 British Birth Cohort. Prev Med., 2015 Dec 31;84:69-75.
16. Heckman, JJ. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. Annals of Economic and Social Measurement, 1976;5(4):475-92.
17. Maddala, GS. A Survey of the Literature on Selectivity Bias as It Pertains to Health Care Markets. Advances in Health Economics and Health Services Research, 1985. 6:3-26.

18. Diez-Roux, AV. Multilevel Analysis in Public Health Research. Annual Review of Public Health, 2000;21:171-92.
19. Sood, N, & Wagner, Z. Impact of Health Insurance for Tertiary Care on Postoperative Outcomes and Seeking Care for Symptoms: Quasi-Experimental Evidence from Karnataka, India. BMJ Open, 2016 Jan 6;6(1):e010512.
20. Linden, A, Adams, JL, & Roberts, N. Evaluating Disease Management Programme Effectiveness: An Introduction to the Regression Discontinuity Design. J Eval Clin Pract, 2006 Apr;12(2):140-7.
21. Reynolds, M, Jheeta, S, Benn, J, Sanghera, I, Jacklin, A, Ingle, D, & Franklin, BD. Improving Feedback on Junior Doctors' Prescribing Errors: Mixed-Methods Evaluation of a Quality Improvement Project. BMJ Qual Saf., 2016 Apr 4. pii: bmjqs-2015-004717.
22. Pirikahu, S, Jones, G, Hazelton, ML, Heuer, C. Bayesian Methods of Confidence Interval Construction for the Population Attributable Risk from Cross-Sectional Studies. Stat Med., 2016 Jan 21. doi: 10.1002/sim.6870 [Epub ahead of print].
23. Hamman, MK, & Kapinos, KA. Mandated Coverage of Preventive Care and Reduction in Disparities: Evidence from Colorectal Cancer Screening. Am J Public Health, 2015 Jul;105 Suppl 3:S508-16.
24. Gold, M. Pathways to the Use of Health Services Research in Policy. Health Services Research, 2009; 44:4:1111-36.
25. Gold, M. Translation and Dissemination of Health Services Re¬search for Health Policy: A Review of Available Infrastructure and Evolving Tools. AcademyHealth, April 2015.
26. Dudley, RA. Pay-for-Performance Research: How to Learn What Clinicians and Policy Makers Need to Know. JAMA, 2005 Oct 12;294(14):1821-3.
27. Dudley, RA. Pay-for-Performance Research: How to Learn What Clinicians and Policy Makers Need to Know. JAMA, 2005 Oct 12;294(14):1821-3.
28. http://results.ref.ac.uk/.
29. Sheppard, J. Humanities Research Threatened by Demands for 'Economic Impact.' *The Guardian*, 2009 October 12. Available at http://www.theguardian.com/education/2009/oct/13/research-funding-economic-impact-humanities.
30. http://www.hefce.ac.uk/pubs/rereports/Year/2015/analysisREFimpact/.
31. Selvaraj, S, Borkar, DS, & Prasad, V. Media Coverage of Medical Journals: Do the Best Articles Make the News? PLoS One, 2014 Jan 17:DOI: 10.1371/journal.pone.0085355.

AcademyHealth is a leading national organization serving the fields of health services and policy research and the professionals who produce and use this important work. Together with our members, we offer programs and services that support the development and use of rigorous, relevant, and timely evidence to increase the quality, accessibility, and value of health care, to reduce disparities, and to improve health. A trusted broker of information, AcademyHealth brings stakeholders together to address the current and future needs of an evolving health system, inform health policy, and translate evidence into action. This meeting and report were made possible by generous support from the Robert Wood Johnson Foundation.

Appendix A: Meeting Agenda

Evaluating Complex Public Health Interventions: Invitational Meeting

December 16, 2015
8:00 a.m.–4:00 p.m. ET

Marriott Marquis Washington, Shaw/LeDroit Room (Meeting Level 3)
901 Massachusetts Avenue, NW, Washington, DC 20001

AGENDA

8:00–8:30 **Continental Breakfast**

8:30–10:00 **Welcome and Setting the Stage**

8:30–9:00 Welcome: Lisa Simpson, AcademyHealth

9:00–9:40 Introductory Panel Presentation
- Role of evidence in federal policymaking: Ron Haskins, Brookings Institution
- Bridging the gap between research, policy, and practice: Vivian Tseng, William T. Grant Foundation

9:40–10:00 Q&A/Discussion

10:00–11:00 **Introductions and Reflections**

10:00–10:40 Participant Introductions

10:40–11:00 Participant Reflections on Panel: Small-Group Discussions
- What am I seeing in my work; what do I think is important?
- What has changed: Demand for evidence, use of evidence, methods?
- How is this affecting the world of public health evaluation and the demands on it for better evidence of what works?

11:00–11:15 **Break**

11:15–12:15 **Overview of Innovative Methods**

11:15–11:45 Panel Presentation
- Key innovations in design and methods—Examples from UK: Rosalind Raine, University College London
- Evaluation context and mechanisms: Brian Mittman, U.S. Department of Veterans Affairs

11:45–12:05 Small-Group Discussions
Participants discuss their shared vision for the future related to evaluating complex public health interventions

12:05–12:15 Report Back from Small Groups

12:15–1:00 **Lunch**

Participants discuss critical components of evaluation at tables on the following topics: study design, measurement, analysis, external/internal validity, cost, and implementation

1:00–2:45 **Panel of Possibilities**

1:00–2:00 Examples of Rigorous Evaluation
Three robust evaluations with methodological diversity summarized as springboards for reflecting on key lessons learned, innovations, challenges, and solutions regarding design, execution, and communication of findings

- Mary Ann Bates, J-PAL North America
- Nancy McCall, Mathematica Policy Research
- Glen Mays, University of Kentucky

2:00–2:20 Small-Group Discussions
Participants discuss the most important next steps to chart the path for strengthening rigor and relevance of evaluation of policies and interventions to achieve population-level health effects

2:20–2:45 Report Back from Small Groups and Discussion

2:45–3:00 **Break**

3:00–3:30 **Reactor Panel**

3:00–3:20 Using Evidence for Policy and Practice: Hearing from Users at the Community Level

- Bowen Chung, University of California, Los Angeles; Los Angeles County Department of Mental Health
- Loretta Jones, Healthy African American Families

3:20–3:30 Q&A

3:30–4:00 **Now What? How to Move the Field Forward?**

Where do we go from here? How do we promote integration of rigorous methods into public health evaluation, policy, and decision-making?

- Karen Minyard, Georgia Health Policy Center
- Lisa Simpson, AcademyHealth

Wrap-up and next steps

Appendix B: Meeting Participants

Evaluating Complex Public Health Interventions: Invitational Meeting

December 16, 2015
8:00 a.m.–4:00 p.m. ET

Marriot Marquis Washington, Shaw/LeDroit Room (Meeting Level 3)
901 Massachusetts Avenue, NW, Washington, DC 20001

PARTICIPANTS

Parag Aggarwal, Ph.D.
Patient-Centered Outcomes Research Institute

Aaron Alford, Ph.D., M.P.H., P.M.P.
National Network of Public Health Institutes

Dawn Alley, Ph.D.
Centers for Medicare & Medicaid Services Center for Medicare and Medicaid Innovation

Jon Baron, J.D., M.P.A.*
Laura and John Arnold Foundation

Anirban Basu, Ph.D., M.S.
University of Washington

Mary Ann Bates, M.P.P.*
J-PAL North America

Jason Bauman, J.D.
J-PAL North America

Linda Bilheimer, Ph.D., M.A.*
Congressional Budget Office

Anupa Bir, Sc.D., M.P.H.
RTI International

Amanda Cash, Dr.P.H.
Office of the Assistant Secretary for Planning and Evaluation

Marshall Chin, M.D., M.P.H.
University of Chicago

Bowen Chung, M.D., M.S.H.S.
University of California, Los Angeles
Los Angeles County Department of Mental Health

Elizabeth Cole
AcademyHealth

Maureen Dobbins, Ph.D.
McMaster University

R. Adams Dudley, M.D., M.B.A.
University of California, San Francisco

Margo Edmunds, Ph.D.
AcademyHealth

Thomas Grannemann, Ph.D.
Mathematica Policy Research

Lawrence Green, Dr.P.H., M.P.H.
University of California, San Francisco

Stuart Guterman, M.A.
AcademyHealth

Ron Haskins, Ph.D.
Brookings Institution

Rachel Hogg, Dr.P.H., M.A.
University of Kentucky

Erin Holve, Ph.D., M.P.H., M.P.P.
AcademyHealth

Edward Hunter, M.A.
de Beaumont Foundation

Loretta Jones, M.A., Th.D.
Healthy African American Families

Marya Khan, M.P.H.
AcademyHealth

M. Rashad Massoud, M.D., M.P.H., FACP
University Research Co., LLC

Glen Mays, Ph.D., M.P.H.*
University of Kentucky

Nancy McCall, Sc.D., S.M.
Mathematica Policy Research

Edward Miech, Ed.D.
U.S. Department of Veterans Affairs

Carolyn Miller, M.A., M.S.
Robert Wood Johnson Foundation

Karen Minyard, Ph.D., M.S.N.**
Georgia Health Policy Center

Brian Mittman, Ph.D.*
U.S. Department of Veterans Affairs

Penny Mohr, M.A.
Patient-Centered Outcomes Research Institute

Andrew Murphy, M.H.S.A.
University Research Co., LLC

Robin Newhouse, Ph.D., M.S., M.G.A., RN
Indiana University

Kate Papa, M.P.H.
AcademyHealth

Gareth Parry, Ph.D., M.Sc.
Institute for Healthcare Improvement

Babak Pourbohloul, Ph.D.
University of British Columbia

Rosina Pradhananga, M.P.H.
AcademyHealth

Ron Prinz, Ph.D.
University of South Carolina

Richard Puddy, Ph.D., M.P.H.
Centers for Disease Control and Prevention

Rosalind Raine, Ph.D., FFPH, M.Sc., M.B.B.S., B.Sc.*
University College London

Danielle Robbio
AcademyHealth

Ana Diez Roux, M.D., Ph.D., M.P.H.
Drexel University

Elana Safran
White House Social and Behavioral Sciences Team

Katie Sellers, Dr.P.H., CPH
Association of State and Territorial Health Officials

James Shelton, M.D., M.P.H.
Global Health: Science and Practice
Johns Hopkins University

Lisa Simpson, M.B., B.Ch., M.P.H., FAAP*
AcademyHealth

Angela Snyder, Ph.D., M.P.H.
Georgia Health Policy Center

Kathy Stack
Laura and John Arnold Foundation

Michael Stoto, Ph.D.
Georgetown University

Vivian Tseng, Ph.D.
William T. Grant Foundation

Abraham Wandersman, Ph.D.
University of South Carolina

Jacqueline Zinn, Ph.D., M.B.A.
Health Services Research
Temple University

*Steering Committee Member, ** Steering Committee Chair