Guiding Principles & Practices
for Improving Evaluation
of Health Improvement Interventions

July 2018

In December 2017, health services researchers, program implementers (including physicians), funders, and policymakers met to consider the merits and limitations of current approaches to evaluating health improvement interventions. The conference, *Learning from Each Other: Getting Evaluations of Complex Health Interventions Right,* was organized by the Jewish Healthcare Foundation and AcademyHealth. Using case studies of what went right and wrong in the evaluations of four high profile health improvement initiatives in the U.S. and England, the day-long session gave participants an opportunity to engage in open and frank discussion.[i]

As thought leaders who care about the quality and impact of healthcare services, participants noted multiple issues, not just in the evaluation of improvement initiatives, but also in their design and implementation, that limit the optimal value and knowledge gained from the evaluation investment. These issues have meant that even after investing considerable resources in, and significant time awaiting evaluation results, stakeholders may be left knowing how an intervention occurred, but not whether it worked and why or why not.

As a result, efforts to get cost-effective, evidence-based interventions to patients and communities can be hindered. And, perhaps more importantly, opportunities for learning may be missed that could inform the adaptation, replication, and dissemination of effective interventions, as well as policy and practice reforms aimed at improving health and health care.

The issues surfaced during the meeting's deep dive into the art and science of evaluation and the overall evaluation ecosystem – including funders, program leaders, and evaluators – are worthy of further discussion. They can contribute to creative solutions, improved and/or more appropriate evaluation designs, effective implementation and evaluation of interventions, and – ultimately – to new and improved practices and policies. Towards that end, this discussion document outlines three major areas where there are opportunities for productive debate. Outlined below, each issue is followed by a proposed set of guiding principles and practices that, if widely adopted, could address these issues in the design, implementation and evaluation of health improvement initiatives in the future. While this document is based upon discussions held at the December 2017 meeting, it was developed by JHF and AcademyHealth staff to spur further discussion amongst the larger community of evaluation stakeholders so that our collective efforts have their optimal result: improved interventions resulting in real and sustainable improvement at reasonable costs.[ii] We welcome comments, suggestions and debate. Together, we make evaluation stronger!

I. Idea Generation & Strategic Thinking Issues: Current processes and incentives for generating health improvement research ideas tend to favor projects that add often relatively small, incremental details to a body of studies, rather than ideas that aim for meaningful transformation of current practice, address pressing evidence gaps, or inform dissemination.

a. Too few federal funding mechanisms employ Requests for Applications that are bold calls-to-action emerging from pressing health policy issues and evidence gaps or that drive practice improvement. Researcher responses to these may thus propose designs that favor the incremental accumulation of knowledge.

b. Evaluations of health improvement interventions are funded by multiple agencies and foundations, each hewing to its own organizational priorities. Too often, this results in limited funds available for evaluation research of discrete programs and initiatives, rather than coordinated opportunities to prioritize worthy ideas, to combine interests, and/or to leverage significant funds to test them.

c. Too often public and private funders require that a proposed intervention be evidence-based, yet up to half of existing practices are not evidence based and no accounting is made for generating that evidence base as part of the intervention.

Proposed Guiding Principles & Practices for Addressing Idea Generation & Strategic Thinking Issues

1. An annual or bi-annual conference serves as a platform for identifying pressing evidence gaps and research needs, sharing qualitative learning on improvement initiatives, and encouraging stakeholder collaboration around needed research.

2. Funders reduce the risk and expense of testing innovative ideas by considering phased, or step-implementation as a fast pre-test before providing longer-term funding.

3. Alternative funding mechanisms, like research networks with a coordinating center, support joint exploration of a problem for which funding is limited.

4. Funders and other stakeholders think creatively about adapting the principles embedded in CMS's "Coverage with Evidence Development" by providing support for collecting the data needed as part of the intervention.

II. Project Design & Planning Issues: Too often it is unclear why and for whom a health improvement initiative will be conducted. This lack of clarity may emerge from failures to surface and align the goals, interests and incentives of all stakeholders (e.g., funders, policymakers, implementers, patients, physicians). The resulting problems are numerous:

a. Conflicting and/or unrealistic expectations among stakeholders – including funders – are not surfaced early, when they could be resolved or used to better inform the design of the intervention, its outcomes, the selection of implementation sites, the appropriate evaluation methodologies, and ultimately, the interpretation of evaluation findings. When this happens, consternation, confusion and even anger on the part of one or more of the participant sites, communities, stakeholders or funders can result when the evaluation results are released, heightening the probability that findings may be contested once the evaluation is released. After all, the old adage is true: 'when they don't like your findings, they will attack you on your methods'.

b. Project design and evaluation approaches aren't developed to explicitly address the questions, concerns and interests of all stakeholders – including those of frontline providers, patients and the public at large. While engaging *all* potential stakeholders may be unrealistic, clarity on the most important stakeholder audience needs is critical.

c. Representatives from the implementing organizations or communities (either actual ones or exemplar types) are often not included among stakeholders at the early stage of intervention and evaluation approaches, contributing to the selection of project design and evaluation methods that may be at odds with implementers' capacity, organizational culture and/or goals.

d. Similarly, patients, consumers, community representatives, and/or the general public are often not included among stakeholders when they might help ensure that intervention and evaluation designs reflect real community health concerns and address barriers, as reflected in their lived experiences.

Proposed Guiding Principles & Practices for Addressing Project Design & Planning Issues

1. Program designers engage stakeholders early in collaborative meetings and/or key informant interviews to learn about their goals and incentives, working to align expectations around the "why" of doing a project.

2. End users provide input about what outcomes would be most useful to them, and project designers work to satisfy the multiple purposes of stakeholders with, for example, appropriate evaluation approaches and diverse measures of success.

3. Patients serve on grant review panels – for both private and public funders.

4. Evaluators are involved as stakeholders in project design and planning.

5. When involving potential implementers in project design and planning isn't possible, designers should take the time to find out about them (their context, capacity, interests, and perspectives) as critical input into model development.

III. Evaluation Selection Issues. Selection of evaluation methodologies may be guided by biases in favor traditional scientific methods and summative evaluation methodologies that prioritize internal validity and causation (answering the question of "does it work?") over external validity, contextual constraints, and accounting for intervention adaptations in situ. This risks ignoring the dynamic, novel and constantly-shifting environments in which all health improvement interventions occur, minimizing the observed impact by over-controlling for factors that are a necessary part of the causal chain, and decreasing the likelihood of successful, post-evaluation dissemination and uptake of the findings.

a. When funders prioritize some evaluation outcomes over others, or insist on adherence to strict interpretation of implementation guidelines, there may be limited resources and time available for learning about why an intervention did or did not work, limiting adaptation to local contexts, and curtailing full exploration of potential impact.

b. Failing to appropriately match the selected evaluation design and methods to the specific goals of the intended stakeholders and program, including a strong preference amongst leaders in academic and scientific communities for experimental and quasi-experimental quantitative methods, can lead to:

    i. Reliance on a narrow set of evaluation methods, including a preference for traditional controlled designs over mixed methodological (quantitative and qualitative) approaches;

ii. Unwillingness to alter project design and make mid-course corrections, even when it is clear that the initiative isn't working at some/all sites;

iii. Reluctance to think about evaluation as an integral part of a health improvement project; and

iv. Preference for deploying evaluators as outside, 'unbiased,' scientific observers whose professional standards would be violated if they worked with implementers to adapt the evaluation design.

c. To the extent that those who train the next generation of evaluators hew to the traditional biomedical model dominant in academic institutions and are slow to embrace emerging methods and innovations in study design, graduates will be ill-equipped to work in the very dynamic world of health improvement programs and initiatives.

d. The imperative to publish in the highest impact factor, peer-reviewed publications, and the publication bias toward positive findings, may force improvement scientists to trade off relevance, timeliness and learning for traditional norms in academia.

Proposed Guiding Principles & Practices for Addressing Evaluation Issues

1. Design and methodological advances (for example the field of dissemination and implementation research or the PCORI standards for studies of complex interventions) take into account the complexity of implementing interventions in community and clinical settings. The expanding toolbox of evaluation methodologies and designs answers questions beyond "does it work?" to address also "how did it work?" and "can it work here?" and thus yields far more learning to inform policy and practice change.

2. Learning, including failures, are valued, solicited and periodically shared among stakeholders during implementation, and are shared widely and quickly, in interim and final reports, using non-academic publications, platforms and venues.

3. Rigor in evaluation is important, but not in isolation from the goals of end users. Evaluation methods are selected to address those goals.

4. At a minimum, stakeholders – including peer reviewed journals and their reviewers, promotion and tenure committees, etc. – request and recognize the value of research that is relevant, uses appropriate methods (including qualitative rather than solely randomized studies or advanced quantitative analyses), and provides information on program context, as well as feedback from implementers and other stakeholders during program roll-out.

5. Evaluators take time to educate stakeholders, including funders, about the relative merits of alternative and complementary evaluation approaches and their respective strengths and limitations. Methods and findings are communicated effectively, as they don't speak for themselves.

Learning from Each Other:

Getting Evaluations of Complex Health Interventions Right

An Invitational Meeting by the Jewish Healthcare Foundation and AcademyHealth

December 7, 2017

PARTICIPANTS

Sharon Arnold, PhD, MSPH
Agency for Healthcare Research and Quality

David Atkins, MD, MPH
Department of Veterans Affairs

Andrew Bindman, MD
University of California, San Francisco

Anupa Bir, ScD, MPH
RTI International

Niall Brennan, MPP
Health Care Cost Institute

Randall Brown, PhD
Mathematica Policy Research

Amanda Cash, DrPH, MPH
Office of the Assistant Secretary for Planning and Evaluation

Jim Chase, MHA
Network for Regional Healthcare Improvement

Valerie Cheh, PhD
Mathematica Policy Research

Marshall Chin, MD, MPH
The University of Chicago Medicine

Susan Dentzer
Network for Excellence in Health Innovation

Kelly Devers, PhD, MA
NORC at the University of Chicago

Rachel Dungan, MSSP
AcademyHealth

Susan Elster, PhD
Jewish Healthcare Foundation

Karen Feinstein, PhD
Jewish Healthcare Foundation

Robert Ferguson, MPH
Jewish Healthcare Foundation

Sheila Fifer, PhD
Avalere Health

Robert Fornango, PhD
Health Services Advisory Group

Lauren Gerlach, MPP
AcademyHealth

Brian Gill, PhD, JD
Mathematica Policy Research

Michael Gluck, PhD, MPP
AcademyHealth

Don Goldmann, MD
Institute for Healthcare Improvement

Marc Gottlieb, MPA
Anthem, Inc.

Marya Khan, MPH
AcademyHealth

Vivian Lee, MD, PhD, MBA
University of Utah

Rashad Massoud, MD, MPH, FACP
USAID/University Research Co., LLC

Paul McGann, MD
Centers for Medicare & Medicaid Services

Renee Mentnech, MS, BSN
Centers for Medicare & Medicaid Services

Shawna Mercer, MSc, PhD
Centers for Disease Control and Prevention

Karen Minyard, PhD, MN
Georgia Health Policy Center

Barbra Rabson, MPH
Massachusetts Health Quality Partners

Rosalind Raine, PhD, FFPH, MSc, MBBS
University College London

Rohit Ramaswamy, PhD, MPH
University of North Carolina at Chapel Hill
Robert Saunders, PhD
Duke Margolis Center for Health Policy

Dennis Scanlon, PhD, MA
Penn State

William Shrank, MD, MSHS
UPMC Insurance Services Division

Lisa Simpson, MB, BCh, MPH, FAAP
AcademyHealth

Nancy Sonnenfeld, PhD
Centers for Medicare & Medicaid Services

Bradley Stein, MD, PhD, MPH
RAND Corporation

Michael Stoto, PhD
Georgetown University

Paul Wallace, MD
AcademyHealth

Nancy Zionts, MBA
Jewish Healthcare Foundation

FOR FURTHER INFORMATION:

Susan Elster, PhD, Jewish Healthcare Foundation; elster@jhf.org

Marya Khan, MPH, AcademyHealth; Marya.Khan@AcademyHealth.org

[i] A Summary of the discussion is available on request.

[ii] The American Evaluation Association has a set of guiding principles for evaluators, which can be accessed at: https://www.eval.org/p/cm/ld/fid=51.