Toward Evaluation in Public Services: Getting Past the Barriers

Gordon Harper, M.D.

Evaluation of public programs in mental health and in other fields is often blocked when "reasons not to" are cited. These include "HIPAA," "IRB," "not my job," "it's already evidence based," "we know what's right," "we don't know enough," "we don't have baseline data," and "there's too much to do."

Examining these reasons, the values thought to justify them, and possible ways to respond will facilitate evaluation research.

In public human services, it is easier to start programs than to evaluate them. Programs are often started by legislators and administrators who know little about what previous initiatives had achieved. Evaluation science, an invention of the 20th century, has a prominent and evolving role in health care of individuals (1,2), but its application to public human services, even before the current change in federal policy (3), has lagged.

What accounts for the omission of such a crucial component of public service? Such evaluation is difficult; recall the decades-long controversy over the benefits of Head Start. The Community Mental Health Center Act of 1963 specifies evaluation as an essential service, but most programs do not follow through. With regard to health services in general, the Center for Medicare and Medicaid Innovation (CMMI), part of the Centers for Medicare and Medicaid Services (CMS; https://innovation.cms.gov), has a charge to do robust evaluation of all payment and service delivery models. In particular, evidence about health outcomes, as opposed to participation in the health care system, has started to emerge' on several CMMI models, including Medicare Pioneer Accountable Care Organizations (4).

With regard to another significant change in public health care, the privatization of Medicaid, more is known about access and cost than about health outcomes. This gap may begin to be filled as CMS supports measurement science and data collection strategies and as states test their own payment and service delivery models under Medicaid, moving from fee-for-service reimbursement to accountable care or bundled payment arrangements (5).

In public mental health services, process often gets more attention than outcomes. For instance, a Canadian group applied a Cochrane analysis to the development of practice guidelines in child and youth mental health. They found that most guidelines used to guide care do not meet internationally established criteria for guideline development (6). But the role of practice guidelines in measuring outcomes played only a small part in their review. In another review, the U.S.-based advocacy organization Building Bridges (www.building-bridges4youth.org), which looked at residential treatment, generated "Recommendations for Outcome and Performance Measures" (7). The organization found no consensus in measuring outcomes; the search for outcomes easily got lost, the investigators found, amid many "performance indicators" at payer and provider levels.

Politics plays a role. Voters want to see something get done. The politician who raises questions about effectiveness may lose popular support. Elected officials want to show results to voters while still in office, not afterward, and they do not want to see unfavorable data emerge.

Such political realities are unavoidable. Translating public concern into political will and legislation is not easy. But the relevant barriers can easily take the form of "reasons not to." These reasons may or may not be stated explicitly.

Some reasons not to evaluate programs can be given names: "HIPAA," "IRB," "not my job," "it's already evidence based," "we know what's right," "we don't know enough," "we don't have baseline data," and "there's too much to do." Naming and describing the reasons for not evaluating programs allow us to recognize them, identify the core values on which they are felt to be based, and identify a response that can promote evaluation. [The eight "reasons not to" are summarized in a table available as an online supplement to this Open Forum.]

Deconstructing "Reasons Not to"

In "HIPAA," the value of patient privacy is invoked, with citations to the 1996 federal Health Insurance Portability and Accountability Act (HIPAA). But there are HIPAA-compliant

ways to evaluate outcomes: obtaining permission from individual patients or proxies, using anonymized data, and conducting evaluation as part of quality improvement.

In "IRB," the need for approval by an institutional review board (IRB), which protects the rights of human research subjects (again under federal law), is cited. But studies for quality improvement are exempt from IRB review. When uncertain, investigators can request exemption from an IRB.

"Not my job" cites fidelity to mission, defined in terms of a specified goal, as in "We were hired to provide a service, not to evaluate it. That would require another contract and more money." Although systematic evaluation requires substantial resources, evaluation can begin on a smaller scale with available resources. To begin on this level, the necessary ingredients are motivation and commitment, not money.

"It's already evidence based" cites a popular value, the use of evidence-based treatment. As the program being considered uses interventions previously tested with individuals, usually in randomized clinical trials, the results of those individual trials are seen to make program evaluation unnecessary. But the clinical trials may have been done in research settings with selected populations—applying the results to general populations is different (8). "Practical clinical trials" may meet this need (9). Translational research aims to take findings from laboratory into practice, with evaluation occurring in clinical settings (http://ncats.nih.gov/clinical). Most of what gets cited as evidence based was not done in such settings.

"We know what's right" invokes a popular ideology as justification to keep doing something consistent with that ideology, even if data supporting its effectiveness are lacking. For instance, the idea that community-based care is always good, and hospital-based care always bad, has been seen to justify continued closing of psychiatric hospital beds, in the absence of supportive data, or even in the presence of evidence of adverse effects (10).

"We don't know enough" is often invoked with regard to mental health services, particularly services for children. The challenges of tracking longer-term, not just short-term outcomes, are cited, as are the challenges of facilitating and measuring change among parents as well as children and of coordinating and evaluating interventions in the separate silos of health, mental health, education, and social services. Relevant, too, are the challenges of using first-person reports from children and parents, along with "objective" data from professionals, and of accounting for cultural differences of expectation and assessment. Substantial as these challenges are, however, they have been addressed (11,12), even in the global context (13).

"We don't know the baseline" cites lack of data from the years preceding the new program. Obviously, preprogram data would help. But there may be data from a previous era from some of the population under review. It may also be possible to implement the new program in one population, letting another untreated population ("treatment as usual") serve for comparison. A commitment to evaluating outcomes can take the form of starting with a part of the whole population or of evaluating the population for a limited time, with the goal of measuring at least some results.

"There's too much to do" cites the myriad parts of any public project, from concept to design to legislative approval and funding to implementation. Getting the program going, especially as policy or resources change, can be felt to be the first priority, precluding evaluation. The response is to acknowledge the urgent operational needs, to foster an ethic of evaluation among the parties, and to work incrementally, starting with some part of the project that is feasible to evaluate.

Those invoking reasons not to care about public service and cite important values. But leaving these reasons unchallenged robs us of something equally valuable, namely knowledge of the effectiveness of what we do. We need such knowledge in order to appreciate interventions that work and to learn what is not working so that we can better serve people who depend on public service. Formative evaluation may even improve the intervention while it is being implemented.

Future Directions

The need in the United States to base innovation on outcomes data, not on good intentions, principles, or ideology, has grown stronger as the coalition of stakeholders committed to expanding services and access to care has had to reckon with a new White House administration less clearly committed to public action on behalf of those in need (3,14).

AUTHOR AND ARTICLE INFORMATION

Dr. Harper is with the Department of Psychiatry, Harvard Medical School, Boston. Send correspondence to Dr. Harper (e-mail: gordon_harper@hms.harvard.edu).

REFERENCES

1. Djulbegovic B, Guyatt GH: Progress in evidence-based medicine: a quarter century on. Lancet 390:415–423, 2017
2. Frieden TR: Evidence for health decision making: beyond randomized, controlled trials. New England Journal of Medicine 377:465–475, 2017
3. Wilensky GR: The first hundred days for health care. New England Journal of Medicine 376:2407–2409, 2017
4. Sommers BD, Blendon RJ, Orav EJ, et al: Changes in utilization and health among low-income adults after Medicaid expansion or expanded private insurance. JAMA Internal Medicine 176:1501–1509, 2016
5. McCluskey PD: Here's what we know about the state's plan to revamp MassHealth. Boston Globe, Nov 8, 2016. https://www.bostonglobe.com/business/2016/11/07/here-what-know-about-state-plan-revamp-masshealth/5e3BSxYdu1UvxlY3dSPFsK/story.html
6. Bennett K, Gorman DA, Duda S, et al: On the trustworthiness of clinical practice guidelines: a systematic review of the quality of methods used to develop guidelines in child and youth mental health. Journal of Child Psychology and Psychiatry, and Allied Disciplines 57:662–673, 2016
7. Dougherty RH, Strod D: Building Consensus on Residential Measures: Recommendations for Outcome and Performance Measures.

Rockville, MD, Substance Abuse and Mental Health Services Administration, March 2014. http://www.dmahealth.com/pdf/BBI%20Building%20Consensus%20on%20Residential%20Measures%20-%20March%202014.pdf

8. Schorr LB: Broader evidence for bigger impact. Stanford Social Innovation Review, 2012. https://ssir.org/articles/entry/broader_evidence_for_bigger_impact

9. March JS, Shapiro M: Practical clinical trials: from medicine to psychiatry. American Academy of Child and Adolescent Psychiatry, 2016. https://www.aacap.org/AACAP/Medical_Students_and_Residents/Mentorship_Matters/DevelopMentor/Practical_Clinical_Trials_from_Medicine_to_Psychiatry.aspx?WebsiteKey=a2785385-0ccf-4047-b76a-64b4094ae07f

10. Tyrer P, Sharfstein S, O'Reilly R, et al: Psychiatric hospital beds: an Orwellian crisis. Lancet 389:363, 2017

11. McMorrow S, Howell E: State Mental Health Systems for Children: A Review of the Literature and Available Data Sources. Washington DC, Urban Institute, 2010

12. Fixing Mental Health Care in America: A National Call for Measurement-Based Care in Behavioral Health and Primary Care. Draft Issue Brief. Chicago, Kennedy Forum. https://thekennedyforum.org/wp-content/uploads/2017/06/Issue-Brief-A-National-Call-for-Measurement-Based-Care-in-Behavioral-Health-and-Primary-Care.pdf

13. Kieling C, Baker-Henningham H, Belfer M, et al: Child and adolescent mental health worldwide: evidence for action. Lancet 378: 1515–1525, 2011

14. Williams DR, Medlock MM: Health effects of dramatic societal events: ramifications of the recent presidential election. New England Journal of Medicine 376:2295–2299, 2017

Psychiatric Services Seeks Additional Reviewers

Psychiatric Services welcomes new reviewers.

Professionals in the mental health disciplines who are qualified to review are invited to contact the Editor, Lisa B. Dixon, M.D., M.P.H. (psjournal@psych.org). Qualifications include prior publication of articles in your specialty area in peer-reviewed journals; current knowledge of recently published research in that area; familiarity with *Psychiatric Services'* audience and recent publications in your field; and willingness to invest the time needed to thoroughly evaluate the manuscripts you agree to review. In your message, please include a list of representative publications and/or professional activities.

Further information for reviewers can be found at ps.psychiatryonline.org/ps_authors_reviewers#02.