How do we learn about improvement?

Danika Barry [1]
Healthcare Improvement Fellow, USAID ASSIST Project/URC

Commentary on the Quality & Performance Institute's Technical Meeting held on December 17, 2014. A full transcript is available here [3].

For our December Quality & Performance Institute Technical Meeting, we invited Dr. Frank Davidoff [4] and other thought leaders in the field of improvement science to comment on the issues raised in Davidoff’s recent article, “Improvement interventions are social treatments, not pills.” [5]

Davidoff comments on the limitations of the “gold standard” randomized controlled trial study design, meticulously followed by Goldman et al. [6] in an intervention to reduce emergency department visits and readmissions among elderly patients in an ethnically and linguistically diverse setting in Northern California. Davidoff writes, “The study method is a thing of beauty; its beauty, unfortunately, is also its curse.” As Davidoff commented in our session, “the beautifully designed, protocol-driven study tells us that the intervention didn't work, but it didn't tell us what improvement can possibly do to change things.” The researchers’ rigid adherence to the study protocol, while required for so-called “P value-based statistical inference,” also prevented the researchers from taking actions to increase the success of the intervention. This is the realm of improvement science [7], which promotes continuous learning and adaptation of the intervention to address social contexts, and new knowledge gained during implementation. As the moderator, Dr. Rashad Massoud [2] pointed out, this is the type of work we do with the USAID ASSIST Project. However, use of multiple, iterative interventions poses a challenge to convincingly demonstrate how we know whether the improvement is truly due to the intervention.

Davidoff identified two possible solutions, first to conduct mixed-methods analyses, which combine protocol-driven studies with qualitative research that can help us understand why an intervention works, and if not, how it might be improved. The second approach proposed was to switch to different kinds of statistical analyses, namely, those that recognize that improvement and healthcare delivery are time-dependent processes. Time-sensitive statistical analyses [8] like statistical process control [9] can help detect meaningful changes in outcomes as a result of an intervention. However, both of these solutions suffer from limitations to drawing causal inference, and are not particularly suited to control for confounders or bias in the way that protocol-driven study designs can.

Dr. Tom Bossert [10] noted these limitations, and emphasized the need to integrate elements like control groups or randomization to answer fundamental questions about causality.
Difference-in-Differences [11] analyses between a control and intervention group, as well as innovative randomization, like that used by Dr. Gary King [12] who randomized matched pairs of districts in an evaluation of a major health reform in Mexico [13], can help strengthen causal inference. Bossert also advocated for the importance of real-time data so that implementers can make improvements. While time-series run charts [14] are one way of using these data, a common pitfall is the mistake of not collecting enough baseline data, which can artificially inflate the perceived intervention effect, as the initial outcome is recorded as near zero.

Dr. John Øvretveit [15] emphasized the need to match methods [16] to the study aims, and to the time and money available, “the controlled trial is the best of designs and the worst of designs, depending on who it is for, and which decisions it is meant to inform.” He further commented, “the more we move towards complex social interventions, the more we have got to reinvent them.” We must question, as Cindy Brach did in her paper, “Will it Work Here?” [17] Especially for context-sensitive interventions like those covered in the Goldman article, Øvretveit commented that we are implementing principles, concepts or ideas that must be adapted to the local context. Similarly, as argued in an article [18] shared by Davidoff after the meeting, complex interventions require that intervention integrity be defined by the intervention function rather than their particular form, in order to address variation in local contexts, while simultaneously preserving an element of standardization required by randomized controlled trials, or other rigorous designs. Process or pathway analysis [19] or PDSA cycles [14], which Davidoff commented are essentially “social tools,” can be useful to inform the adaptation process. Øvretveit further indicated there is a lot to be learned from hybrid studies in other fields [20], including the logframe model, action evaluations or case study approaches used in program evaluation, implementation science, and welfare service evaluations in public health.

Dr. Edward Broughton [21] pointed out a similarity between controlled trials and improvement interventions, for example, in RCTs we do not control for every aspect of someone’s life, just one particular element. Similarly, in interventions for improvement, we do not control every part of the intervention, just the principles that need to be applied. In a similar way, the “seething dichotomy” between P value-based statistics and run chart analyses [22] belies the fact that elements of both can actually be combined. This point was expanded upon in Broughton and Bossert’s response to a question from Esther Karamagi [23], Senior Quality Improvement Advisor, in the USAID ASSIST Project office in Uganda. Karamagi asked about the role of controls in day-to-day decisions on how one change leads to improvement. Broughton acknowledged the Project’s constraints in collecting control data, as we cannot collect data from facilities that we are not working in. Therefore, we must determine a somewhat lower or more practical level of evidence to go forward with an intervention. However, Bossert added that there may be ways to strengthen the evidence by identifying controls among other project sites, as long as the patients have similar characteristics.

In conclusion, the panelists commented that hybrid approaches can be a powerful way to triangulate evidence for causal inference. Additionally, the panelists emphasized the need to advocate to donors to provide the resources and incentives to perform this sort of work in the field contexts that we operate.