What Makes a Realistic Evaluation?

Richard Byng

(Fam Med 2011;43(2):112-3.)

Realistic Evaluation (RE) is a framework for program or project evaluation and can utilize qualitative and quantitative data derived from experimental or descriptive designs. RE can provide guidance on developing “middle-range theories” that encompass individuals’ thoughts and behaviors, as well as organizational mechanisms, while taking into account a range of contextual factors from local geography to individual illnesses and beliefs. In contrast to the randomized control trial, which eliminates measurement of contextual effect, RE explicitly aims to understand the impact of context on the effectiveness of interventions and seeks to improve knowledge of what works for whom and in which situations.

RE is one of a group of theory-based evaluation methods that are currently having something of a resurgence. Another such evaluation method, “Theories of Change,” developed over many years in the United States, starting during the time of the New Deal, also aims to improve the theoretical understanding of how social programs work. Theories of Change starts by bringing together practitioners and recipients of services to map out and make explicit the theoretical model of how a program works. By collecting qualitative and quantitative data about the process of care, and its immediate impact, refinement of the model can be made. The RE approach can complement this by emphasizing the importance of analysing the impact of context on these intermediate outcomes.

So in what way is the paper, by Yoffe et al., published in this issue of Family Medicine, a RE?

I applaud the careful use of local data to demonstrate, quite convincingly for a non-randomised design, that information to parents can have an impact on emergency room use. However, I am not convinced that the design can really be termed an RE. Essentially the paper is based on before and after ratios of emergency to routine primary care use by children. The additional information on other services, demonstrating limited or no reduction in service use during the same time period adds weight to the observed drop in emergency use in the intervention period. I therefore disagree with the authors that the data on the other services is purely descriptive and contextual. I would argue that it does provide useful quantitative data, which allows us to make judgements about how likely it is that the intervention is effective. Other quantitative analyses could be made using more of the data.

The evidence collected about context is not used within the paper to help understand what works for whom and when. Instead, Cooper’s study uses a pretest, posttest design. The originators of RE, Pawson and Tilley, criticize the pretest, posttest design even more roundly than the randomized control trial, although they recognize that in some situations the randomized control trials are impossible, and pretest, posttest designs can provide important data. This kind of evaluation, however useful, should not be confused with a RE just because contextual data is collected. The additional quantitative data presented adds weight to the theory that providing information to parents...
can reduce emergency room use, and tells us about the location, but does not tell us anything about the context in which this intervention is most likely to work.

The authors could have carried out a true RE of their intervention in several ways. First, at a regional and local level the findings could have been taken up by other services, and the intervention could have been implemented fully in a number of other diverse clinics and settings in the region. It is possible that in all of these clinics similar results would have been seen. However, it is more likely that the intervention will work in some but not others, even if replicated carefully. A study of the contexts in which it worked could then be carried out. Contextual issues, such as the socio-demographics of the population, the population density, the makeup and the type of services, and other factors could have an impact on whether the intervention worked. These could be measured and impact trends mapped against such contextual factors. Regression analyses would require data from large numbers of clinics, but the impact of contextual factors could be assessed by examining context-outcome predictor matrices.4

Second, in addition to this study of organizational and geographical context it might be feasible to examine which families are likely to gain most from the intervention (some will already have achieved high levels of self care, and others due to illiteracy or other factors may not be helped). It would have been possible to do an in-depth qualitative RE by studying the successes and failures to see how the intervention had an impact on the thinking and behavior of parents of children attending the clinics. Do some families provide a more fertile context for this type of social intervention? Parents of children who have not re-attended and those who had re-attended could have been sampled purposefully and interviewed in order to ask the reasons for re-attendance and the impact if any of the information provided on their likelihood to re-attend and any improved home-based self care. This type of study could be done by a researcher or as part of a quality improvement program. Such a study could be done inexpensively and could provide good information to improve both the written material itself and the way in which it was provided to patients. This could be followed by a more quantitative questionnaire study aiming to prove or disprove a variety of potential middle-range theories as to how context has an impact on the intervention.

Even though this study was not a true RE, it seems clear that the authors were quite aware of the importance of context. Perhaps through informal conversations with practitioners and patients, the investigators noted that having clinicians give out the leaflets had far more impact than other delivery methods. In the world of clinical medicine, it is often impossible to control for all contextual factors. Randomized clinical trials are often impossible or impractical. There is an important role for methods such as RE in such settings, and we all have much to learn about how to conduct such research in a rigorous and systematic manner.

CORRESPONDENCE: Address correspondence to Dr Byng, University of Plymouth, Peninsula College of Medicine and Dentistry, N2 ITTC Building, Tamar Science Park, Plymouth, United Kingdom. richard.byng@pms.ac.uk.

References