

# Proposed standards for quality improvement research and publication: one step forward and two steps back

P Pronovost, R Wachter

“Dumbing down” the scientific expectation of journals and readers regarding QI research may not be the answer

The recent article by Davidoff and Batalden and the accompanying commentaries left us with mixed emotions.<sup>1-4</sup> On the one hand we applaud the authors’ efforts to improve the evidentiary base of quality improvement (QI) through publication guidelines. These guidelines will bring structure and rigor to the field of QI, which is a necessity.

However, the author’s plea and support for more rigor was placed in the midst of an argument to “dumb down” the scientific expectation of journals and readers regarding QI research. While we are all frustrated by the slow pace of improvement in quality of care and appreciate the authors’ and commentators’ arguments to relax traditional evidence based medicine standards for QI research and publication, we fear this approach would be detrimental to the field, waste essential and scarce resources, and lead providers and organizations down too many blind alleys.

The authors frame the challenges well. In short, they argue that many reviewers of high impact journals do not fully understand the challenges QI researchers face in conducting quality improvement programs with limited or no funding. Budgetary constraints, coupled with the difficulties of adhering to scientifically acceptable designs in the real world of healthcare delivery, make it impractical to expect that QI studies will traverse the threshold of scientific rigor that is generally expected of more traditional biomedical studies. The authors feel that failing to relax these standards will rob readers of the rich fruits gleaned from those toiling in the trenches.

To be certain, we too have struggled in publishing QI studies. While we believe the authors have correctly identified many of the signs and symptoms, they have misdiagnosed the disease and recommended the wrong therapy. True, QI work is often difficult, underappreciated, and undersupported. It is frequently difficult to determine if a

change resulted in an improvement and, if so, whether the intervention was the reason. The ability to generalize from successful experiences at a single site to other settings is tenuous. Because of these and other issues, the publication standards of traditional medical journals frequently result in rejection of well executed QI “studies”. Undoubtedly, some of these studies contain important insights, and their absence from high impact medical journals can be tragic for everyone.

Nevertheless, many of these findings, if pulled from poorly designed or conducted studies, would be wrong, and the choice not to publish them in traditional journals probably improves and protects public health. Reviewers’ concerns about the risk for bias and lack of validity for many QI results are often valid and prevent inappropriate uptake of questionable QI interventions.

Contrary to the view that most QI efforts are rigorous, our experience has revealed that rigorous QI efforts are rare. QI studies commonly lack clarity regarding the study population, interventions and co-interventions, outcome measurements and definitions, and quality assurance procedures for data collection. The amount of missing data is often significant (although generally not reported) and may exceed the data available, and what data are available may be poor in quality. As such, results from one organization that performs well are often cherry picked among others not performing well and presented as the generalizable results. Consequently, many QI studies have significant bias and have a high risk of misinforming readers.

In short, many studies are appropriately rejected for publication by journals despite the hard work performed by staff truly dedicated to improving patient care. We are not saying that disseminating the results of local QI efforts with uncertain methodological rigor is inherently bad. Framed correctly, such anecdotes can generate

new hypotheses, spur organizations to consider new approaches (“*We’re a pretty similar hospital to that one ... perhaps we should try that intervention here ...*”), and provide needed support and credibility to others in the field.

However, vehicles for disseminating study findings should not confuse readers regarding the scientific rigor of QI projects by bestowing the imprimatur of evidence-based medicine on study findings. Suitable vehicles to disseminate QI experiences and findings are readily available: conferences, newsletters, and websites contain a myriad of interesting anecdotes of improvement projects that can help guide and inspire others working in the field. While some of the findings from these efforts represent the “truth” and some will be generalizable, the absence of rigorous research methods makes it exceedingly difficult—if not impossible—to differentiate the wheat from the chaff.

Indeed, there are many examples illustrating the plausibility of conducting QI studies which are both scientifically sound and feasible. For example, over 120 ICUs in Michigan are participating in a QI collaborative, have collected data using rigorous methods, and more than 10 manuscripts have been submitted or are being prepared to share these findings. There are other examples of well done QI studies that have been published in high impact journals.<sup>5-7</sup> Studies such as these can stand toe to toe with the best in biomedicine on a methodological playing field.

Some might wonder whether it would be harmful if scientific standards for reporting QI studies were relaxed. After all, these QI interventions are not potentially toxic drugs or procedures that could cause harm if applied incorrectly. In fact, we believe that publishing unsound findings can and does cause harm. For example, recent evidence suggests that early observations regarding the value of computerized provider order entry (CPOE) systems may have been less generalizable than previously appreciated and may even, at least in the short term, be harmful.<sup>8</sup> Even when new practices seem unlikely to cause harm, there are significant opportunity costs involved in adopting them—money and attention could be applied elsewhere.<sup>9</sup> An example of this issue may be the widespread adoption of rapid response teams (RRTs) which is based on relatively scanty and mostly anecdotal evidence of benefit (the most rigorous clinical trial to date demonstrated no benefit).<sup>10</sup> Without data regarding the benefits and costs, hospitals may inappropriately allocate scarce resources to RRTs at the expense of

other interventions such as intensivist or increased nurse staffing, which may have stronger evidence, a larger impact on mortality, and more clear data on estimated costs.<sup>11,12</sup>

Moreover, relaxing scientific standards opens the door not only for unintended biases—that is, “researchers” failing to appreciate that their honestly observed result is not correct or generalizable—but other kinds of biases as well. QI efforts can be plagued by conflicts of interest.<sup>13</sup> Some organizations and individuals who set standards or publish QI results may benefit financially from selling services to healthcare organizations. As such, it is imperative that such recommendations have a strong evidentiary base and potential conflicts of interest remain transparent. In addition to protecting patients, institutional review board processes can help ensure scientific integrity and evaluate potential financial conflicts of interest with minimal burden on QI practitioners.

We do respect the view that, if the bar is set too high, some potentially important findings may not be disseminated. Conversely, lowering the bar will allow dissemination of too much biased information and promulgation of too much unexamined conflicts of interest. All in all, we believe the overall effect will be negative. A better course of action than to “dumb down” the field would be to push for increased support for rigorous QI studies (not only from external funders but from institutions themselves), better training in research methods for individuals working in the field, and more sensitivity on the part of

journals and their reviewers regarding the science of QI. If approached in this manner, stakeholders deciding to implement a QI intervention that is informed by the published literature can be confident that the conclusions from the literature they are emulating represent the truth. This will only occur by maintaining scientific integrity. To that end, creating QI specific standards and guidelines for publication (such as those offered by Davidoff and Batalden) represents one giant step forward. Unfortunately, this progress will be lost if we take two steps backward and relax the standards of scientific rigor that have served us so well in clinical medicine and science.

*Qual Saf Health Care* 2006;15:152–153.  
doi: 10.1136/qshc.2006.018432

#### Authors' affiliations

**P Pronovost**, The Johns Hopkins University, Department of Anesthesiology and Critical Care Medicine, Surgery, and Health Policy and Management, Baltimore, MD, USA  
**R Wachter**, University of California San Francisco, Department of Medicine; UCSF Medical Center, San Francisco, CA, USA

Correspondence to: Associate Professor P Pronovost, The Johns Hopkins University, Department of Anesthesiology and Critical Care Medicine, Surgery, and Health Policy and Management, 1909 Thames Street, 2nd Floor, Baltimore, MD 21231, USA; ppronovo@jhmi.edu

Funding: none.

Competing interests: Dr Pronovost has support from research grants and honoraria for

consulting and lecturing on quality improvement efforts.

#### REFERENCES

- 1 **Davidoff F**, Batalden P. Toward stronger evidence on quality improvement. Draft publication guidelines: the beginning of a consensus project. *Qual Saf Health Care* 2005;14:319–25.
- 2 **Berwick D**. Broadening the view of evidence-based medicine. *Qual Saf Health Care* 2005;14:315–6.
- 3 **Stevens D**. Why new guidelines for reporting improvement research? And why not? *Qual Saf Health Care* 2005;14:314.
- 4 **Thomson R**. Consensus publication guidelines: the next step in the science of quality improvement? *Qual Saf Health Care* 2005;14:317–8.
- 5 **Landrigan C**, Rothschild J, Cronin J, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med* 2004;351:1838–48.
- 6 **Kucher N**, Koo S, Quiroz R, et al. Electronic alerts to prevent venous thromboembolism among hospitalized patients. *N Engl J Med* 2005;352:969–77.
- 7 **Landon BE**, Wilson IB, McInnes K, et al. Effects of a quality improvement collaborative on the outcome of care of patients with HIV infection: the EQHIV study. *Ann Intern Med* 2004;140:887–96.
- 8 **Han Y**, Carcillo J, Venkataraman S, et al. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. *Pediatrics* 2005;116:1506–12.
- 9 **Werner RM**, Asch DA. The unintended consequences of publicly reporting quality information. *JAMA* 2005;293:1239–44.
- 10 **MERIT Study Investigators**. Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. *Lancet* 2005;365:2091–7.
- 11 **Needleman J**, Buerhaus PI, Stewart M, et al. Nurse staffing in hospitals: is there a business case for quality? *Health Aff (Millwood)* 2006;25:204–11.
- 12 **Pronovost P**, Angus D, Dorman T, et al. Physician staffing patterns and clinical outcomes in critically ill patients: a systematic review. *JAMA* 2002;288:2151–62.
- 13 **Brennan T**, Rothman D, Blank L, et al. Health industry practices that create conflicts of interest: a policy proposal for academic medical centers. *JAMA* 2006;295:429–33.