SOUNDING BOARD

The Tension between Needing to Improve Care and Knowing How to Do It

Andrew D. Auerbach, M.D., M.P.H., C. Seth Landefeld, M.D., and Kaveh G. Shojania, M.D.

The past 7 years have seen unprecedented interest in patient safety and the quality of health care.[1-3] As physicians whose careers are focused on improving quality and safety, we have welcomed this change. However, we have also witnessed recent initiatives that emphasize dissemination of innovative but unproven strategies, an approach that runs counter to the principle of following the evidence[4] in selecting interventions that meet quality and safety goals, as well as the idea that interventions should be tailored to local needs and resources.[5] These principles have been used as safeguards in helping us pursue practices that have clear benefits for patients and that can be implemented with local resources. This approach also reflects the recognition of how little we know about ways to improve care in a large number of settings.[4,6,7]

Our consideration of the rationale for rapid dissemination of novel quality and safety strategies has led us to identify a number of weaknesses inherent in approaches that consistently favor action over evidence. In this article, we outline the arguments in favor of rapid dissemination and the counterpoints to each of the arguments (Table 1). We conclude by proposing a framework for evaluating interventions to improve the safety and effectiveness of health care.

ARGUMENT 1: WE CANNOT WAIT

The most common argument in favor of prioritizing action over evidence is that the need to address quality and safety problems is urgent. Often, this need is summed up by the question, "How many times does outcome X need to occur before we implement intervention Y?"

This question seems particularly compelling because hundreds of thousands of patients (possibly millions) experience harm as a result of underuse, overuse, or misuse of medical therapies.[8] However, similar claims about the scale of morbidity and mortality could be made for heart disease, cancer, AIDS, depression, and many other disorders. Medical error may be the eighth leading cause of death in the United States,[2] but by proceeding largely on the basis of urgency rather than evidence, we exempt the eighth cause of death from standards applied to the top seven.

In addition, the question of how many instances of X outcome need to occur before we implement Y intervention assumes that we can define Y and X accurately, as well as connect Y to a decreased risk of X. Donabedian pointed out that Y can be either a structural element of health care (e.g., staffing ratios) or a process (e.g., administration of a drug) and emphasized the importance of establishing a connection between Y and the outcome of interest, X.[9] Unfortunately, connections between structural or process-based interventions and outcomes are usually presumptive,[7,10] and defining problems and solutions with respect to patient safety is generally difficult.[10,11]

For example, mandates to reduce residents' work hours reflect the view that tired residents cause errors that harm patients. However, evidence linking patient harm directly to care provided by a fatigued resident is indirect,[12,13] and although reductions in work hours do not appear to have harmed patients, evidence that reforms have met their goal of improving safety is tentative at best.[14-18] Furthermore, to be cost-effective, a reduction in work hours would have to result in greater improvement in safety than that reported for any other intervention.[19] Regardless of whether an 80-hour workweek ultimately improves patient safety, an intervention with a number of potential effects was introduced without a full understanding of its risks and benefits and without a plan to evaluate its effectiveness after implementation.

Promising initiatives and bold efforts at improvement can consume tremendous resources yet confer only a small benefit[15,20,21] or a benefit that

Table 1. Arguments for and against Rapid Dissemination of Quality-Improvement Interventions.

Argument	Why Proceeding Quickly Is Critical	Why Evaluation Is Critical
We cannot wait — the need to improve the quality of care is urgent.	Thousands of patients are injured or killed each year by medical errors.	The need to improve the treatment of many diseases is equally urgent, yet we demand rigorous evidence that a therapy works before recommending it widely.
Any effort to improve quality is better than the current state of affairs.	On balance, the harms of quality improvement are likely to be far less than those of the status quo.	Knowledge of the harms and opportunity costs of quality improvement is important for an understanding of the net benefit to patients and health care systems, which is often small.
Emulating successful organizations can speed effective improvement.	Emulation and collaboration provide an efficient means of disseminating potentially effective solutions.	Emulation and collaboration can incorrectly promote or even overlook interventions that have not worked.
The effectiveness of some quality-improvement strategies is obvious.	Insistence on evidence may lead us to underuse interventions that are obviously effective.	Even though many quality-improvement practices have a simple rationale, they may be less effective than expected and can be difficult to implement fully.
Innovation can be catalyzed by dissemination of strategies that have promise but are unproven.	Preliminary data provide an important opportunity to speed innovation and improve care rapidly.	Flawed, biased, or incomplete data may lead to adoption of interventions that are ineffective or harmful.
The framework of evidence-based medicine does not apply to quality improvement.	The nature of quality improvement exempts it from the usual strategies of assessment.	Given the complexity of quality and safety problems, the complexity of their causes, and how little we understand them, we should use rigorous study designs to evaluate them.
Developing evidence in quality improvement is too costly.	The resources and expertise required to evaluate quality and safety interventions rigorously make trials impractical, particularly when the field is moving so quickly.	As compared with the large opportunity costs incurred by wide implementation of ineffective quality and safety strategies, investments in better evaluation would be small.

is at best unclear.[22] How many such examples must we have before we decide to choose our efforts more wisely?

ARGUMENT 2: ANY EFFORT TO IMPROVE IS BETTER THAN THE CURRENT STATE OF AFFAIRS

Multiple problems in our flawed health care system lead to the view that any attempt at improvement is better than the status quo. Although understandable, this view ignores the possibility that quality-improvement efforts can cause harm.[23] Unfortunately, few studies have assessed this possibility. For example, only 12 of 66 reports on trials of strategies to improve care for patients with diabetes included rates of hypoglycemia.[24] However, in 7 of those 12 studies, hypoglycemia was more frequent in the intervention group than in the control group. Although hypoglycemia is an easily anticipated consequence of efforts to intensify the treatment of diabetes, adverse consequences of many other efforts at improvement of care have been less predictable, including errors introduced by computerized entry of physicians' orders,[25,26] bar coding,[27] and infection-control isolation protocols.[28] Side effects may seem inherently less likely with quality-improvement interventions than with drugs and devices. However, most quality-improvement interventions involve changes in the organization of complex systems, and the law of unintended consequences — long recognized as a side effect of complex change — tends to apply to such interventions.[29-31]

ARGUMENT 3: EMULATING SUCCESSFUL ORGANIZATIONS CAN SPEED IMPROVEMENT

A recommendation to emulate successful organizations reflects the reasoning that adopting fea-

tures of these organizations — the institutional culture, leadership styles, or specific improvement practices — will result in similar successes. Unfortunately, this reasoning ignores the possibility that many unsuccessful organizations also share these features, so that the truly critical determinants of success are not being targeted.

For instance, continuous quality improvement and quality-improvement collaboratives are often recommended on the basis of their adoption by successful organizations. However, systematic evaluations of these approaches have shown that they result in only modest improvements at best.[20,21,23,32] These disappointing findings probably reflect the overemphasis on success that is inherent in benchmarking and the collaborative approach, which tend to neglect an examination of unsuccessful organizations that share features of successful ones.[33]

Successful organizations may also have a vested interest in promoting their services or preferred quality-improvement strategies, further distorting the usefulness of emulating such organizations. Even when direct financial conflicts of interest do not exist, any organization that has undertaken a major campaign to improve the quality of care has little incentive to invest resources in a rigorous evaluation of the effects of its efforts. If anecdotal reports or superficial analyses are positive, the organization will understandably focus on advertising these measures of success rather than pursuing more rigorous evaluation.

Argument 4: The Effectiveness of Some Quality-Improvement Strategies Is Obvious

Some solutions appear to be so obviously beneficial that requiring evidence seems like asking for randomized trials of parachutes.[34] However, anyone who has undertaken a quality-improvement project understands that identifying an apparent solution to a problem is only a first step. Even with pilot testing and evaluative steps, implementing solutions in practice can present numerous challenges.

Hand washing is an example of a well-defined, effective solution to a problem (nosocomial infections), but strategies that consistently result in increased hand washing remain unestablished.[35] Unfortunately, many initiatives fall into the hand-washing category — that is, the case for improvement is obvious, but effective strategies for translating solutions into practice remain elusive.[8]

Changes in complex systems can have unanticipated consequences (as we note with respect to Argument 2), such as new problems[25,26,30,31,36] or simply the failure to achieve the desired goal. Until we advance the basic sciences in quality improvement (e.g., organizational theory and ergonomics),[4,7,23,32] we cannot assume that even the most apparently straightforward solutions can be seamlessly implemented. Without an understanding of not only what to do but also how to help people actually do it, many apparently obvious quality-improvement interventions have more in common with calls for world peace than with parachutes — the goal is not in question, but the path for achieving it is.

Argument 5: Promising but Unproven Strategies Can Catalyze Innovation

Many quality-improvement interventions have such strong face validity that their dissemination seems to be justified on the basis of early or preliminary evidence. This strategy will certainly speed dissemination, but it also carries substantial risks.

Early trials of medical emergency teams suggested a large potential benefit[37-40] — to the point that some observers regarded further study as unethical.[41] However, a large, randomized trial subsequently showed that medical emergency teams had no effect on patient outcomes.[42] The validity of the earlier positive studies has also been questioned,[43] but only after many hospitals introduced medical emergency teams (and have had no reason to switch from advertising the adoption of an innovation to questioning its usefulness in the first place — Argument 3).

There are many examples of drugs or devices that showed substantial promise on the basis of early findings, which were then modified or refuted by later-phase research. These often represent therapies for disorders that affect millions of people (as we note with respect to Argument 1). Yet we rarely sanction the widespread distribution of new drugs on the basis of preliminary data alone. It is therefore not clear why we favor approaches to quality improvement that foster change over appropriate evaluation.

It is worth emphasizing that when studies show no benefit of an intervention with strong face validity, as has occurred with rapid-response teams[42] and more recently with teamwork training,[44] one should not necessarily conclude that the intervention has no value. The finding may simply mean that the intervention had no effect in the form and setting that were studied. The crucial point is that without the randomized trial, we would have no way of knowing that implementation of the intervention in its current form confers no advantage over usual care (or confers a much smaller advantage than that suggested by preliminary studies) and that refinement is necessary.

ARGUMENT 6: THE FRAMEWORK OF EVIDENCE-BASED MEDICINE DOES NOT APPLY TO QUALITY IMPROVEMENT

A recent commentary argued that we would not require randomized trials to determine whether we have solved problems or learned skills in our daily lives.[45] By extension, according to this argument, evidence-based medicine may not apply to the processes that underlie many quality-improvement initiatives. Although it is true that we often do not need trials to test our acquisition of knowledge or skills, we do need them when choosing between alternative methods of acquisition — particularly when training is costly or the skill is of high value.

Rigorous evaluation does not always require randomized trials. Alternative designs (e.g., before-and-after studies that include concurrent control groups and time-series designs involving multiple preintervention and postintervention measurements) can sometimes provide robust results,[32,46] as can research that combines quantitative and qualitative approaches.[47] But anecdotal reports and simple before-and-after studies, although sometimes adequate to justify local quality-improvement efforts, are probably never sufficient to support widespread initiatives because of the risks of expending tremendous resources without obtaining a true benefit and possibly introducing new problems.

Randomized, controlled trials, although not always necessary,[46] remain highly relevant to quality improvement. The value of such trials lies in the random assignment of subjects with unknown characteristics that affect outcomes to intervention and control groups. In clinical medicine, important confounders are often well known and easily planned for, so that observational studies can adjust for these factors, thereby producing results that often agree with the results of randomized trials.[48,49] However, outcomes of quality-improvement interventions depend on many factors, related to patients, providers, and organizations, that remain poorly understood. Thus, the complexity of health care and the dearth of evidence with respect to how components of the system interact to influence outcomes provide a strong rationale for conducting randomized trials to evaluate quality and safety interventions whenever feasible.

ARGUMENT 7: DEVELOPING EVIDENCE IN QUALITY IMPROVEMENT IS TOO COSTLY

Many people have argued that with limited resources available for quality-improvement efforts, the costs of evaluation are untenable. However, one could also argue that we should not spend scarce resources on quality improvement unless we know it is effective. More important, there are tremendous opportunity costs. An institution that invests millions of dollars or expends hundreds of personnel hours in implementing an ineffective system almost certainly could have made other investments that would have benefited its patients. Moreover, if the investment at one hospital is multiplied by thousands of hospitals across the country, then surely spending several million dollars for evaluation is cost-effective, given the billions of dollars at stake with widespread implementation. In this sense, it is the absence of evidence — with respect to efficacy, possible harms, and strategies for implementation — that is too costly, not the efforts to generate such evidence.

CONCLUSIONS

The urge to favor action over evidence in efforts to improve the quality and safety of health care is understandable. However, we have seen in recent years that progress in quality improvement occurs just as it does in the rest of biomedicine: interventions that appear to be promising on the basis of preliminary studies often prove to have no benefit, and those that are beneficial typi-

cally result in modest improvements, not monumental breakthroughs. And quality-improvement interventions, like clinical therapies, can have untoward effects and both direct and indirect costs. These commonalities compel us to argue that interventions to improve the quality and safety of health care should meet the same standards that are applied to the adoption of all medical technologies.

In the rest of biomedicine, innovation begins with basic-science experimentation and proceeds through evaluative trials in successive phases. The basic sciences in quality improvement differ from those in the rest of biomedicine, but the framework for evaluating candidate interventions is largely the same. Clinicians often make decisions about treatment in individual patients on the basis of limited evidence or even just intuition. Similarly, individual hospitals may pursue promising quality-improvement strategies on the basis of scant evidence, including anecdotal reports or face validity. However, clinical practices based on such limited evidence would never become broad standards of care, much less requirements for accreditation or reimbursement. Similarly, recommending or mandating the widespread adoption of interventions to improve quality or safety requires rigorous testing to determine whether, how, and where the intervention is effective — just as in the rest of medicine. Clarification of this picture is critical because a number of widely promulgated interventions are likely to be wholly ineffective, even if they do not harm patients. Even worse, in the current environment, we will not know what these interventions are.

The movement to improve quality and safety has achieved substantial momentum in recent years and has begun to address the many errors of omission and commission that harm patients each day. Moreover, the visible moral leadership associated with these efforts[3] has played a crucial role in maintaining public trust. Although the scope of the problems may seem to favor action over knowledge, quality improvement is on common ground with the rest of biomedicine. The temptation is to circumvent traditional models of evidence when it comes to quality improvement, but this temptation has always existed in medicine for those seeking cures to conditions with high morbidity. Just as in the rest of medicine, we must pursue the solutions to quality and safety problems in a way that does not blind us to harms, squander scarce resources, or delude us about the effectiveness of our efforts.

Supported by a Patient Safety Research and Training Grant from the Agency for Healthcare Research and Quality (K08 HS11416-02, to Dr. Auerbach), an Academic Leadership Award from the National Institute on Aging (K07 AG00912, to Dr. Landefeld), and funding from the Government of Canada Research Program (to Dr. Shojania).

No potential conflict of interest relevant to this article was reported.

We thank Erin Hartman, M.S., Ralph Gonzales, M.D., M.S.P.H., and Chaim Bell, M.D., M.P.H., for their helpful comments on an early version of the manuscript.

From the University of California, San Francisco, Department of Medicine, San Francisco (A.D.A., C.S.L.); San Francisco Veterans Affairs Medical Center, San Francisco (C.S.L.); and Ottawa Health Research Institute, Ottawa (K.G.S.).

1. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academy Press, 2001.
2. Kohn LT, Corrigan JM, Donaldson MS, eds. To err is human: building a safer health system. Washington, DC: National Academy Press, 2000.
3. Berwick DM, Calkins DR, McCannon CJ, Hackbarth AD. The 100,000 Lives Campaign: setting a goal and a deadline for improving health care quality. JAMA 2006;295:324-7.
4. Brennan TA, Gawande A, Thomas E, Studdert D. Accidental deaths, saved lives, and improved quality. N Engl J Med 2005; 353:1405-9.
5. Nelson EC, Batalden PB, Ryer JC, eds. The clinical improvement action guide. Oakbrook Terrace, IL: Joint Commission on Accreditation of Healthcare, 1998.
6. Forster AJ, Shojania KG, van Walraven C. Improving patient safety: moving beyond the "hype" of medical errors. CMAJ 2005;173:893-4.
7. Shojania KG, Duncan BW, McDonald KM, Wachter RM. Safe but sound: patient safety meets evidence-based medicine. JAMA 2002;288:508-13.
8. Chassin MR, Galvin RW. The urgent need to improve health care quality: Institute of Medicine National Roundtable on Health Care Quality. JAMA 1998;280:1000-5.
9. Donabedian A. Evaluating the quality of medical care. Milbank Mem Fund Q 1966;44:Suppl:166-206.
10. Pronovost PJ, Miller MR, Wachter RM. Tracking progress in patient safety: an elusive target. JAMA 2006;296:696-9.
11. Hofer TP, Kerr EA, Hayward RA. What is an error? Eff Clin Pract 2000;3:261-9.
12. Gaba DM, Howard SK. Patient safety: fatigue among clinicians and the safety of patients. N Engl J Med 2002;347:1249-55.
13. Weinger MB, Ancoli-Israel S. Sleep deprivation and clinical performance. JAMA 2002;287:955-7.
14. Fletcher KE, Davis SQ, Underwood W, Mangrulkar RS, McMahon LF Jr, Saint S. Systematic review: effects of resident work hours on patient safety. Ann Intern Med 2004;141:851-7.
15. Laine C, Goldman L, Soukup JR, Hayes JG. The impact of a regulation restricting medical house staff working hours on the quality of patient care. JAMA 1993;269:374-8.
16. Shetty KD, Bhattacharya J. Changes in hospital mortality associated with residency work-hour regulations. Ann Intern Med 2007;147:73-80.
17. Horwitz LI, Kosiborod M, Lin Z, Krumholz HM. Changes in outcomes for internal medicine inpatients after work-hour regulations. Ann Intern Med 2007;147:97-103.
18. Goldman L, Fiebach N. Hippocrates affirmed? Limiting residents' work hours does no harm to patients. Ann Intern Med 2007;147:143-4.

19. Nuckols TK, Escarce JJ. Residency work-hours reform: a cost analysis including preventable adverse events. J Gen Intern Med 2005;20:873-8.
20. Landon BE, Hicks LS, O'Malley AJ, et al. Improving the management of chronic disease at community health centers. N Engl J Med 2007;356:921-34.
21. 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.
22. Wachter RM, Pronovost PJ. The 100,000 Lives Campaign: a scientific and policy review. Jt Comm J Qual Patient Saf 2006; 32:621-7.
23. Mittman BS. Creating the evidence base for quality improvement collaboratives. Ann Intern Med 2004;140:897-901.
24. Shojania KG, Ranji SR, McDonald KM, et al. Effects of quality improvement strategies for type 2 diabetes on glycemic control: a meta-regression analysis. JAMA 2006;296:427-40.
25. Koppel R, Metlay JP, Cohen A, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005;293:1197-203.
26. Han YY, Carcillo JA, Venkataraman ST, et al. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. Pediatrics 2005;116: 1506-12. [Erratum, Pediatrics 2006;117:594.]
27. Patterson ES, Cook RI, Render ML. Improving patient safety by identifying side effects from introducing bar coding in medication administration. J Am Med Inform Assoc 2002;9:540-53.
28. Stelfox HT, Bates DW, Redelmeier DA. Safety of patients isolated for infection control. JAMA 2003;290:1899-905.
29. Merton RK. The unanticipated consequences of purposive social action. Am Sociol Rev 1936;1:894-904.
30. Perrow C. Normal accidents: living with high-risk technologies. New York: Basic Books, 1984.
31. Reason J. Managing the risks of organizational accidents. Burlington, VT: Ashgate Publishing, 1997.
32. Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. Health Aff (Millwood) 2005; 24:138-50.
33. Denrell J. Selection bias and the perils of benchmarking. Harvard Business Review. April 2005:114-9, 134.
34. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 2003;327:1459-61.
35. Gawande A. On washing hands. N Engl J Med 2004;350: 1283-6.
36. Schnipper JL, Kirwin JL, Cotugno MC, et al. Role of pharmacist counseling in preventing adverse drug events after hospitalization. Arch Intern Med 2006;166:565-71.
37. Bristow PJ, Hillman KM, Chey T, et al. Rates of in-hospital arrests, deaths and intensive care admissions: the effect of a medical emergency team. Med J Aust 2000;173:236-40.
38. Buist MD, Moore GE, Bernard SA, Waxman BP, Anderson JN, Nguyen TV. Effects of a medical emergency team on reduction of incidence of and mortality from unexpected cardiac arrests in hospital: preliminary study. BMJ 2002;324:387-90.
39. Bellomo R, Goldsmith D, Uchino S, et al. Prospective controlled trial of effect of medical emergency team on postoperative morbidity and mortality rates. Crit Care Med 2004;32:916-21.
40. DeVita MA, Braithwaite RS, Mahidhara R, Stuart S, Foraida M, Simmons RL. Use of medical emergency team responses to reduce hospital cardiopulmonary arrests. Qual Saf Health Care 2004;13:251-4.
41. Kerridge RK, Saul WP. The medical emergency team, evidence-based medicine and ethics. Med J Aust 2003;179:313-5.
42. Hillman K, Chen J, Cretikos M, et al. Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. Lancet 2005;365:2091-7. [Erratum, Lancet 2005; 366:1164.]
43. Winters BD, Pham J, Pronovost PJ. Rapid response teams — walk, don't run. JAMA 2006;296:1645-7.
44. Nielsen PE, Goldman MB, Mann S, et al. Effects of teamwork training on adverse outcomes and process of care in labor and delivery: a randomized controlled trial. Obstet Gynecol 2007;109:48-55.
45. Berwick DM. Broadening the view of evidence-based medicine. Qual Saf Health Care 2005;14:315-6.
46. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials unnecessary? Picking signal from noise. BMJ 2007;334:349-51.
47. Bradley EH, Herrin J, Wang Y, et al. Strategies for reducing the door-to-balloon time in acute myocardial infarction. N Engl J Med 2006;355:2308-20.
48. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N Engl J Med 2000;342:1878-86.
49. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med 2000;342:1887-92.