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. 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 in selecting interventions that meet quality and safety goals, as well as the idea that interventions should be tailored to local needs and resources. 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.

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. 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, 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. Unfortunately, connections between structural or process-based interventions and outcomes are usually presumptive, and defining problems and solutions with respect to patient safety is generally difficult.

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, 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. 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. 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 or a benefit that...
is at best unclear.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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,\textsuperscript{25,26} bar coding,\textsuperscript{27} and infection-control isolation protocols.\textsuperscript{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.\textsuperscript{29-31}

**Argument 3: Emulating Successful Organizations Can Speed Improvement**

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

---

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>We cannot wait — the need to improve the quality of care is urgent.</td>
<td>Thousands of patients are injured or killed each year by medical errors.</td>
<td>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.</td>
</tr>
<tr>
<td>Any effort to improve quality is better than the current state of affairs.</td>
<td>On balance, the harms of quality improvement are likely to be far less than those of the status quo.</td>
<td>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.</td>
</tr>
<tr>
<td>Emulating successful organizations can speed effective improvement.</td>
<td>Emulation and collaboration provide an efficient means of disseminating potentially effective solutions.</td>
<td>Emulation and collaboration can incorrectly promote or even overlook interventions that have not worked.</td>
</tr>
<tr>
<td>The effectiveness of some quality-improvement strategies is obvious.</td>
<td>Insistence on evidence may lead us to underuse interventions that are obviously effective.</td>
<td>Even though many quality-improvement practices have a simple rationale, they may be less effective than expected and can be difficult to implement fully.</td>
</tr>
<tr>
<td>Innovation can be catalyzed by dissemination of strategies that have promise but are unproven.</td>
<td>Preliminary data provide an important opportunity to speed innovation and improve care rapidly.</td>
<td>Flawed, biased, or incomplete data may lead to adoption of interventions that are ineffective or harmful.</td>
</tr>
<tr>
<td>The framework of evidence-based medicine does not apply to quality improvement.</td>
<td>The nature of quality improvement exempts it from the usual strategies of assessment.</td>
<td>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.</td>
</tr>
<tr>
<td>Developing evidence in quality improvement is too costly.</td>
<td>The resources and expertise required to evaluate quality and safety interventions rigorously make trials impractical, particularly when the field is moving so quickly.</td>
<td>As compared with the large opportunity costs incurred by wide implementation of ineffective quality and safety strategies, investments in better evaluation would be small.</td>
</tr>
</tbody>
</table>
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. 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.

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. 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. 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.

Changes in complex systems can have unanticipated consequences (as we note with respect to Argument 2), such as new problems or simply the failure to achieve the desired goal. Until we advance the basic sciences in quality improvement (e.g., organizational theory and ergonomics), 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 — to the point that some observers regarded further study as unethical. However, a large, randomized trial subsequently showed that medical emergency teams had no effect on patient outcomes. The validity of the earlier positive studies has also been questioned, 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 and more recently with teamwork training, 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.

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. 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 as can research that combines quantitative and qualitative approaches. 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, 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. 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 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. 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 as can research that combines quantitative and qualitative approaches. 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, 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. 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 un-
toward effects and both direct and indirect costs. These commonalities compel us to argue that in-
terventions to improve the quality and safety of health care should meet the same standards that
are applied to the adoption of all medical tech-
nologies.

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 deci-
sions about treatment in individual patients on
the basis of limited evidence or even just intu-
ition. Similarly, individual hospitals may pursue
promising quality-improvement strategies on the
basis of scant evidence, including anecdotal re-
ports or face validity. However, clinical practices
based on such limited evidence would never be-
come broad standards of care, much less require-
ments for accreditation or reimbursement. Simi-
larly, 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 ef-
fective — just as in the rest of medicine. Clarifi-
cation of this picture is critical because a num-
er of widely promulgated interventions are likely
to be wholly ineffectve, 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

5. Chassin MR, Galvin RW. The urgent need to improve health
care quality: Institute of Medicine National Roundtable on

6. Donabedian A. Evaluating the quality of medical care. Mil-


10. Fletcher KE, Davis SQ, Underwood W, Mangrulkar RS, Mc-


12. Shetty KD, Bhattacharya J. Changes in hospital mortality as-

13. Horwitz LI, Kosiborod M, Lin Z, Krumholz HM. Changes in outcomes for internal medicine inpatients after work-hour regula-

Copyright © 2007 Massachusetts Medical Society.