One hundred fifty years ago, the British Parliament implemented a system of payment by results, championed by Robert Lowe, First Viscount Sherbrooke, to base school funding on student test scores. Even then, critics argued that basic tests could not capture students’ desire and ability to read intelligently. Despite similar objections along the way, the use of standardized metrics to improve organization performance has been widely adopted, spreading first to manufacturing in the 1970s and 1980s and then to many sectors of the economy. Nowhere has this approach been more avidly embraced than in health care. Many factors account for the enthusiastic adoption of performance measurement, including the recognition of wide variations in processes and outcomes of care, rapidly escalating costs, difficulty assessing value, and general public skepticism of government and large organizations. In addition, early experiences with using measurement to improve metrics, such as the Healthcare Effectiveness Data and Information Set, resulted in impressive improvements over short time frames, appearing to validate this approach.1,2

Over time, the number of performance metrics has exploded from a few dozen to thousands.3 In contrast to early measures of comparatively simple processes, such as ordering a test or administering a vaccine, measures now address complex outcomes such as enhancing patients’ satisfaction or reducing mortality. Moreover, the goals of performance measurement have expanded beyond public accountability and process improvement to supporting payment systems by insurers and incentive systems for executive and clinician salaries. Performance metrics are now insinuated into every aspect of health care delivery, especially the electronic health record. Clinicians and health systems are continuously rewarded or penalized based on grades.

Despite early successes, the proliferation of performance measures has occurred with limited evidence that this approach has improved the overall quality of health care or enhanced the value of health care. Wide variations in care persist and costs have continued to escalate. Nonetheless, many health care delivery systems furnish clinicians with elaborate reports and dashboards that display their performance relative to national or local benchmarks along with data about peers. A sizable literature shows that this form of feedback is modestly effective particularly when delivered frequently by a supervisor or respected colleague, includes both specific goals and action plans, and seeks to decrease a specific behavior.4 Moreover, it is more effective when baseline performance is low and recipients are not clinicians. However, overall effect sizes are generally small on the order of 10% improvement.5 Many health care systems augment this form of feedback with monetary incentives, so-called pay for performance, but studies indicate that these arrangements also produce very limited improvements that are mainly related to a limited number of process measures.6,7

Thus, other strategies need to be explored. Patel and colleagues8 studied whether application of behavioral economic theory could improve prescribing of statins by clinicians. The technique evaluated was active choice framing. Instead of merely presenting data to clinicians, they also provided a set of electronic orders that displayed mutually exclusive options for ordering statins and were easily implemented. Clinicians could also opt to not prescribe a statin but were then required to provide a rationale. If a prescription for a statin was selected, assistance was provided to ensure proper execution of the order and to furnish information to patients. This strategy precluded the opportunity for clinicians to defer a decision passively, which is often responsible for so-called clinical inertia. By minimizing the volume of feedback to the clinician, facilitating responses to the feedback,
and providing assistance to patients, the investigators sought to avoid the information overload and multiple mouse clicks that plague clinicians and contribute to distraction and burnout.

This active choice intervention was tested in a large, well-designed, 3-arm, cluster randomized clinical trial. In 1 arm, clinicians were exposed to active choice and in a second arm, this intervention was supplemented with peer comparisons that included kudos for top performers. In the third control arm, no communication was provided. At baseline, among patients for whom a statin was indicated, about 64% had received a prescription. Following the intervention, using an intention-to-treat analysis, this rose by 6.7% among the patients of clinicians exposed to the dashboard compared with 2.6% in the control arm, a difference that was not statistically significant after adjustment. Among patients assigned to clinicians exposed to both the dashboard and peer comparisons, statin prescribing rose by 8.0%, significantly different from the control group.

These findings augment the existing literature on using audit and feedback to improve clinical practice. The 6.4% absolute increase in statin prescriptions observed by Patel and colleagues is roughly consistent with the 10% mean observed effect size from meta-analyses, after taking into account the relatively high baseline rate of statin prescribing and targeting the intervention at clinicians. Patel and colleagues cogently argue that their approach is attractive because it is simple, short, and minimally intrusive on harried primary care providers. On the other hand, if this technique was used simultaneously to address numerous quality issues, it might come to be perceived as onerous in the same way that clinicians currently view clinical reminders in the electronic health record.

There is also the problem of effect dilution that plagues most trials with an encouragement design, in which clinicians are cued in some manner to take a recommended action. In this case, Patel and colleagues refer to it as a nudge. In the first step, the clinician must recognize the prompt. Only one-third to one-half of the clinicians accessed the dashboard in this trial. The clinician must then respond to the cue, and only about half of those who accessed the dashboard wrote and signed prescriptions. In other settings, more than 90% of clinicians who were presented with electronic orders that they trusted signed them immediately. Even after clinicians signed the prescriptions, patients had to fill them and take the medication. The drug must have the desired effect; 40 to 100 patients must take a statin for a year to prevent 1 myocardial infarction. This dilution blunts any opportunity to observe a positive clinical effect and is magnified more when the number needed to treat (or screen) is even higher, eg, for 13-valent pneumococcal vaccine, mammography, or colonoscopy, which are all an order of magnitude higher.

Does this mean we should be nihilistic about efforts to improve practice through audit and feedback? To the contrary, many small incremental improvements are necessary and as Patel and colleagues have done, efforts to refine those methods are worthwhile. However, to be fully effective these techniques must be integrated intelligently into clinical workflow via the electronic health record so that cues can be handled efficiently and in context. In addition, we should recognize situations in which clinicians require a shove rather than a nudge.
REFERENCES


