Clinical quality measurement remains an elusive goal, and it has the potential to result in adverse outcomes and unexpected consequences. Practicing physicians are wary of current efforts but should remain professionally committed to the development of effective, evidence-based quality measures.

In 34 years of practicing medicine, I have yet to be paid based on any measured quality outcome. Instead, my sense of professionalism, my interest in evidence-based decision making and clinical outcomes, and my respect for my patients have driven my interest in quality improvement activities. I have no reason to think that the experiences of my physician colleagues have been any different.

The role of quality measurement may change in the near future, however, when programs such as the Physician Quality Reporting System (PQRS) of the Centers for Medicare & Medicaid Services (CMS) move from merely requiring the reporting of processes and numbers to providing rewards when targets are met or imposing financial punishments when targets are missed [1]. Other payers and insurers in North Carolina and nationwide are gradually moving from “reward for reporting” programs to true pay-for-performance activities. The development of accountable care organizations in North Carolina, which are being created by commercial insurance companies, Medicare, and Medicaid, will certainly engage many more practices in quality improvement efforts.

From the perspective of a practicing physician, the rewards have been relatively small so far compared to the effort involved in process reporting. Although I care for many Medicare patients, my “reward” for PQRS reporting in 2011 was less than $1,200, and it was half that amount in 2012, when the incentive rate fell from 1% of claims to 0.5%.

Of course, reporting also has a cost. Using data from 8 diverse primary care practices in North Carolina, Halladay and colleagues [2] examined the costs of gathering and reporting data on quality care indicators to 4 major quality programs. Costs of reporting data included implementation expenses, ranging from less than $1,000 to $11,100 per clinician, and annual maintenance costs, ranging from less than $100 to $4,300 per clinician. The article states that practices with fewer care providers “appeared to be especially hard hit” by reporting requirements [2].

Many physicians are not convinced that reporting of quality measures has a major impact on quality. A survey of 4,934 physicians in 2009 revealed that only 38% were participating in the Physicians Quality Reporting Initiative (PQRI), which has since been renamed the PQRS. More than half of medical specialists and surgeons participating in the PQRI believed that it had no impact on quality; this opinion was shared by 40% of primary care physicians [3]. Overall participation rates in PQRS may soon increase, as failure to participate will result in a 1.5% reduction in allowable Medicare fees beginning in 2015.

Payers in North Carolina have offered some practices financial incentives for participation in quality improvement work through recognition programs such as Bridges to Excellence (BTE), Improving Performance In Practice, and the National Committee for Quality Assurance (NCQA). Blue Cross and Blue Shield of North Carolina has offered financial incentives for quality improvement to selected practices since 2006. In 2009, Blue Cross and Blue Shield of North Carolina expanded these incentives with its Blue Quality Physician Program (BQPP). Focusing on independent practices and offering improvements in fee-for-service schedules for completion of the program, BQPP provides a significant financial incentive for practices to participate in selected quality improvement programs. The measured elements include components of care and quality, which use entities such as BTE and NCQA to assess disease management; education of administrative and clinical staff members on issues such as cultural competence and end-of-life decision making; and practice organization elements, including continuous quality improvement, maintenance of certification, cost-saving practice efforts and policies, office hours and access, motivational interviewing, and practice efficiency. The flexibility of different pathways for recognition may well improve the overall value, participation, and applicability of the process to different practices.

Electronically published April 16, 2013.
Address correspondence to Dr. W. James Stackhouse, Goldsboro Medical Specialists, 1209 E Ash St, Goldsboro, NC 27530 (jim@stackhouse.com).
N C Med J. 2013;74(2):155-158. ©2013 by the North Carolina Institute of Medicine and The Duke Endowment. All rights reserved. 0029-2559/2013/74214
Efforts to measure clinical quality and outcomes accelerated after a 2001 report from the Institute of Medicine of the National Academies (IOM) highlighted quality gaps in medical care [4]. This finding in turn accelerated the proliferation of clinical guidelines. Although guidelines are not quality measures, they are often used as a quality yardstick.

More than 2,500 disease-specific guidelines are now available through the Agency for Healthcare Research and Quality’s National Guideline Clearinghouse [5]. Many of these guidelines are consensus statements from panels of recognized experts, drawn from various levels of evidence combined with the experts’ experiences and current practice. Best practices are evidence-based, but sometimes limited evidence is available. Further research to validate both guidelines and best practices occasionally results in significant revisions due to evidence of unintended harm. Ignoring the complex needs of patients with multiple medical issues, programs have often used guidelines for individual disease management as measures of physician performance, despite disclaimers from the developing organizations that the guidelines are not to be considered standards of care.

Although such guidelines are often used for payer authorization decisions, the evidence base supporting these guidelines is often limited, with many guidelines relying on expert consensus alone. Recently, 6 professional organizations, including the American College of Physicians and the American College of Cardiology Foundation, outlined 48 specific practice guidelines for management of ischemic heart disease, 18 of which were, in the authors’ own judgment, based on low-quality evidence [6]. In addition, guidelines may conflict when applied to patients with multiple significant disease processes, and authoritative panels often disagree about the efficacy of screening measures such as mammograms, prostate-specific antigen testing, and chest radiographs.

In 2012, Kung and colleagues screened 130 randomly selected guidelines from the National Guideline Clearinghouse Web site and concluded that fewer than half met at least 50% of the guideline standards set by the IOM in 2011. The authors used what they termed a “liberal” interpretation of the guidelines, meaning that they did not include 7 of the 18 IOM guidelines, which they deemed “too vague and subjective to be analyzed.” Information on conflicts of interest (COI) was provided in fewer than half of subspecialty society-developed guidelines. In guidelines that contained this information, 71.4% of committee chairpersons and 90.5% of cochairpersons had disclosed COIs [7]. A 2011 IOM report recommended that persons with COIs not fill those roles, whether or not the COIs were disclosed [8].

Connecting compensation to guidelines and/or quality measurements elicits new behaviors from organizations and individuals. Data on the effects of measuring and rewarding performance are available from the British National Health Service (NHS). In 2004, the NHS used tracking information from the NHS-funded electronic health record (EHR) to assess performance on 80 clinical indicators, 43 organizational indicators, 4 patient-experience indicators, and 8 other service indicators. General practices could earn up to 1,050 points by meeting these criteria, and financial awards were tied to performance. The average practice achieved 91% of the available 1,050 points, representing a bonus of $75,000 USD per physician. “Cherry picking,” or excluding certain patients from the scoring, did not occur to any significant extent. The NHS had only budgeted for about two-thirds of this performance level [9].

As far back as 2003, studies had begun to uncover the unintended consequences of pay-for-performance initiatives. Shen demonstrated that incentives offered to a non-profit substance abuse treatment center affected which patients were selected for treatment. Among those patients whose care was included in the incentive program, the implementation of incentives resulted in a significant decrease in the percentage of patients selected for treatment who were designated “most severe.” During the same period, the percentage of patients selected for treatment who were categorized as “most severe” increased among Medicaid patients, whose treatment was not included in the incentive program [10]. In an excellent recent commentary, Woolhandler and colleagues discussed the role of behavioral economics and warned that pay-for-performance may be incompatible with quality improvement efforts. They point to several potential shortcomings including issues with outcome measurement, risk adjustment, claims in excess of the level of needed or provided services, inappropriate consequences of quality measurement, and patient socioeconomic factors [11].

In 2011, a Cochrane review of 2,933 potentially relevant studies of pay-for-performance programs found only 7 studies that met the criteria for outcomes measurement [12]. An analysis of those 7 studies found that 6 of them “showed positive but modest effects on a minority of the measures of quality of care included in the study.” However, the studies suffered from a risk of selection bias. The authors concluded that “there is insufficient evidence to support or not support the use of financial incentives to improve the quality of primary health care,” and they urged caution in using incentive schemes [12].

My personal experience with recognition programs has been mixed. I conservatively estimate that completing the steps in the BQPP program has required more than 80 hours of my time, outside of patient care, over a period of 6 months. My office staff has spent just as much time participating in education activities and performing data extraction related to the program. While time consuming, I believe our participation in the BQPP program improved the quality of our patient care.

I am also well aware of the statistical shortcomings of reporting data from a small sample size. Since 1990, my practice has used Microsoft Excel software to track clinical details of the care of more than 225 patients with diabetes, aged 32-86 years. Reporting of data is limited to a smaller...
subset. PQRS requires me to report data on only my 30 most recently seen Medicare patients who are under the age of 75 years. Similarly, BTE and NCQA recognition programs for quality diabetes care require data from 36 sequential patients who are selected retroactively from a certain date. Although the average glycosylated hemoglobin level of my patients with diabetes is 7.1%, the frequency of visits for these patients varies. Patients whose blood glucose levels are well controlled may be seen only every 6 months, whereas patients whose levels are poorly controlled are seen much more frequently. The need for short-term follow-up visits after medication changes or acute illnesses means that those patients with the worst glycosylated hemoglobin levels may be included much more frequently in my list of recently seen patients.

Practices with more economically disadvantaged or uninsured patients, or simply older individuals, are also similarly penalized by such selection criteria. Some practices may attract or specialize in the care of patients with more difficult or complex conditions. For example, it is discouraging when a patient is taking 4 drugs for hypertension, has experienced a reduction in systolic blood pressure from 240 mm Hg to 140 mm Hg, and has a stable blood pressure maintained by medications that he can afford, yet he fails to meet the target criterion of “systolic pressure less than 140 mm Hg.”

Variations in performance standards are also important. Analysis of the same 36 patients might allow a physician to meet the recognition requirements for one agency but not the requirements of another agency. Differences in the date chosen to start measurement reporting can affect scoring not only due to sample size and more frequent visits by poorly controlled patients but also because there are significant seasonal variations in measurements such as glycosylated hemoglobin [13]. I schedule my poorly controlled patients for more frequent visits. Seasonal diet and activity changes also affect my patients’ level of control.

Self-insured health programs follow quality measures in practices, and scores from these reports and NCQA’s Healthcare Effectiveness Data and Information Set are returned to practices as a measure of “quality.” Because they are usually based on data about charges—not on medical record reviews—these reports are often incomplete. They frequently fail to capture specific blood tests performed in panels, for instance. Even more vexing, some of the measured parameters may extend over a period of 2-10 years and may include years during which data on charges are not available to the current payer. Submission of accurate chart data raised our practice’s rate for repeated colon cancer screening within 10 years from 37% to 65%, our rate of annual low-density lipoprotein (LDL) cholesterol screening in patients with ischemic heart disease from 30% to 89%, and our rate of annual LDL screening in diabetic patients from 63% to 93%. Given current limitations, data based on charges alone does not seem well suited for quality assessment.

For many payers, an individual practice does not have enough covered lives to make any data statistically sound, even if the data are properly collected. One of the goals of the North Carolina Health Quality Alliance is to interest major payers in sharing such data in a protected environment that is trusted by physicians, with the goal of gathering more accurate overall practice-specific information.

Almost inevitably, practices in the future will receive some payment or bonus that is based on performance measures, if not on quality outcomes. Many, but not all, practices are already engaged in activities such as installation of EHR systems to meet CMS’s standards of meaningful use. Adoption of EHR systems is a necessary step for these practices. It remains to be seen whether the cost of installation, subsequent inefficiencies, and painful changes to work flow will be balanced by the financial rewards of using the EHR, much less improved quality outcomes. Whether electronic practice data can be extracted and reported in a format that meets different agencies’ requirements, and at what cost, also remains to be seen.

Like many primary care physicians in North Carolina who are nearing retirement, I expect to retire not too long after Medicare begins levying financial penalties on practices that are not using an EHR system. I also do not expect to find that anyone interested in taking over my practice would want to use an EHR system that I chose now. Thus, I do not plan to implement an EHR system, as I personally have little incentive to invest in another source of office overhead and continuing costs, and I already have sufficient internal and external reporting measures to verify my belief that I am providing quality care.

For me, and for most physicians I know, quality measurement and performance improvement are not new concepts. Still, I am doubtful that most current activities can substantially change outcomes of care. However, we remain professionally committed to working toward a future in which quality of care can be accurately measured and then improved. NCMJ

W. James Stackhouse, MD, MACP Goldsboro Medical Specialists, Goldsboro, North Carolina.

Acknowledgment
Potential conflicts of interest. W.J.S. has no relevant conflicts of interest.

References
5. National Guideline Clearinghouse. Guidelines by Topic. Agency for...


STROKE TARGETS BY COLOR.

Know where you stand. The odds are African Americans are twice as likely to suffer a stroke as white Americans.

Beating the odds isn’t about winning. It’s about living.

You have the power to end stroke. 1-888-4-STROKE / StrokeAssociation.org