

Performance Measurement at a “Tipping Point”

Laurent G. Glance, MD,* Mark Neuman, MD, MSc,† Elizabeth A. Martinez, MD,‡
Kenneth Y. Pauker, MD,§ and Richard P. Dutton, MD||

After the release of the Institute of Medicine (IOM) report *To Err Is Human*,¹ improving the quality of health care in the United States (US) has become a national priority.² There is also a growing realization that the dramatic growth in health care costs is not economically sustainable. Total spending on health care is projected to increase to nearly 40% of the gross domestic product by 2050.³ The imperative to improve quality and cut costs has led the federal government to initiate public reporting with Hospital Compare,⁴ and to incentivize physician and hospital performance using pay-for-performance.⁵ The Centers for Medicare and Medicaid Services (CMS), along with key stakeholders such as the Joint Commission, the American Medical Association, the Physician Consortium for Performance Improvement (PCPI), and many medical societies such as the American Society of Anesthesiologists (ASA), are developing performance measures for hospitals and physicians.⁶

With the creation of the Anesthesia Quality Institute (AQI) and the National Anesthesia Clinical Outcomes Registry (NACOR), the ASA is creating the data infrastructure for quality benchmarking and outcomes improvement in anesthesiology.⁷ Performance measurement is potentially a “transformative tool”⁸ and is widely recognized as the cornerstone in the drive to improve health care quality.⁹ Performance measurement can be used to transform hospitals into “learning laboratories” to discover and implement best practices.^{9,10} In seeking to integrate AQI with the CMS Physician Quality Reporting Initiative, the ASA is also creating a mechanism to facilitate anesthesiologist participation in the physician quality reporting and payment incentive system legislated by Congress in the 2006 Tax Relief and Health Care Act. Performance assessment is now a core component of anesthesiologist credentialing through the American Board of Anesthesiology Maintenance of

Certification. Although to date only 3 of the 266 performance measures developed by the AMA Physician Consortium for Performance Improvement¹¹ are anesthesia-specific,^a there can be no doubt that performance measurement will become an integral part of the practice of anesthesia in the future.⁷

The implementation of public reporting and policies linking payment to health care outcomes has profound implications for health care delivery in the US.¹² In theory, publicly reporting health care outcomes and linking payment to quality will harness market and regulatory forces to promote greater quality and efficiency in the health care system. In practice, the success of using performance measures to fundamentally reshape health care is critically dependent on the accuracy and reliability of quality measurement. Although anesthesiologists are one of the key stakeholders in current efforts to reengineer health care delivery, most of us have, at best, only very limited knowledge in this area. Our goal in writing this white paper is to provide anesthesiologists with the background they need to actively participate in the development of performance measures for anesthesiologists. We will use the Donabedian structure-process-outcome conceptual model¹³ as a framework to describe the key concepts, strengths, and limitations of clinical performance measurement. Although performance measurement is a critical driver for improving health care, it is not a panacea and can lead to important unintended consequences if it is not developed and implemented using a careful and evidence-based approach. We intend for this white paper to serve as a useful primer for anesthesiologists on performance measurement.

OUTCOME MEASURES

The IOM defines quality as “the degree to which health services for individuals and populations increases the likelihood of desired health outcomes and are consistent with current professional knowledge.”¹⁴ This widely accepted definition equates the quality of health care with the effectiveness of health care in producing desired patient outcomes. At the core of the IOM definition of quality is the word “outcome.” The most important reason to measure outcomes is the simple fact that “patients care about results.”¹⁵ Patient outcomes represent the final common product of all clinical activity and are the cornerstone of performance measurement.

Unfortunately, an individual patient’s outcome is not simply the result of the effectiveness of medical care, but is also a complex function of a patient’s risk factors (how sick is the patient before receiving medical care) and chance (random) events.¹⁶ Thus, it is not possible to use crude mortality rates after coronary artery bypass graft (CABG) surgery to measure hospital quality because differences

*The 3 anesthesiology and critical care measures consist of (1) prevention of ventilator-associated pneumonia, (2) prevention of catheter-related bloodstream infections with a central venous catheter insertion protocol, and (3) perioperative temperature management.

Authors’ affiliations are listed at the end of the article.

Accepted for publication December 20, 2010.

Supported in part by grants from the Agency for Healthcare and Quality Research (ROI HS 16737 to Dr. Glance, and K08 HS 013904 to Dr. Martinez), and by the Department of Anesthesiology at the University of Rochester School of Medicine and Dentistry.

The authors declare no conflicts of interest.

Reprints will not be available from the authors.

Address correspondence to Laurent G. Glance, MD, Department of Anesthesiology, University of Rochester Medical Center, 601 Elmwood Ave., Box 604, Rochester, NY 14642. Address e-mail to Laurent_Glance@urmc.rochester.edu.

Copyright © 2011 International Anesthesia Research Society

DOI: 10.1213/ANE.0b013e31820e778d

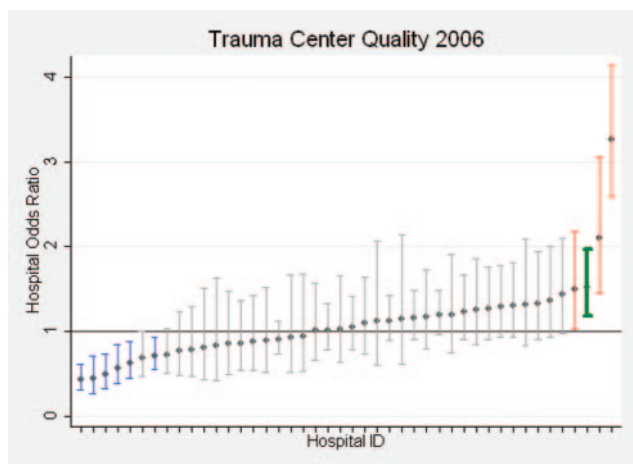


Figure 1. Sample hospital outcomes report for hospital cohort. Vertical bars represent 95% confidence intervals. Hospitals whose quality is below average are in red and hospitals with above-average quality are in blue. “Your” facility is shown in green. (Reproduced with permission from the *Journal of Trauma*.²⁰)

in hospital mortality may be attributable to differences in patient case mix as well as differences in hospital quality. In theory, risk adjustment allows us to account for patient risk factors and equate risk-adjusted outcomes with the effectiveness of care. In practice, however, “The Risks of Risk Adjustment”¹⁷ is that different risk adjustment models can yield different quality assessments for the same group of hospitals: a hospital can receive a different quality rating depending on which risk adjustment model is used.

Risk adjustment is necessary to avoid penalizing hospitals and physicians who treat high-risk patients.¹⁷ Using unadjusted mortality rates to evaluate hospital quality lacks face validity. Moreover, risk adjustment may make it less likely that physicians will avoid high-risk patients to boost their quality rating.¹⁷ For many physicians and policy makers, the methodology of risk adjustment is a “black box.”¹⁸ Understanding what is inside the black box is critically important to be able to meaningfully assess the validity of the risk adjustment methodologies used to construct quality measures.

Risk adjustment relies on a regression model that uses a patient’s clinical risk factors to predict the probability of death for a patient treated at an “average” hospital.¹⁹ This information is then used to calculate a hospital’s expected mortality rate (EMR) by averaging the predicted probability of death for each of the hospital’s patients. Hospital performance is then assessed by comparing a hospital’s observed mortality rate to its EMR, usually using the ratio of the observed mortality rate to the EMR, the O-to-E (OE) ratio. Hospitals with an OE ratio significantly >1 are labeled as high-mortality outliers, whereas hospitals with an OE ratio significantly <1 are labeled as low-mortality outliers. The comparative performance of hospitals can be displayed using a caterpillar graph such as the sample report card used in the Survival Measurement and Reporting Trial for Trauma²⁰ (Fig. 1). (The odds ratio and the OE ratio have similar interpretations in this application.) It is common to multiply the OE ratio by the unadjusted

population rate for the entire patient cohort to obtain the risk-adjusted mortality rate. For example, if the OE ratio for hospital A in the New York State (NYS) CABG report card is 0.5 and the overall CABG mortality rate in NYS is 2%, then hospital A would have a risk-adjusted mortality rate of 1%.

In order for a risk adjustment model to be useful for benchmarking performance, the data elements used for risk adjustment must be clearly defined and accurate. Poor-quality data cannot be used to generate unbiased quality reports, regardless of the level of statistical expertise.²¹ Report cards can be based either on administrative data or clinical data. Administrative data are extracted from discharge billing forms, and consist of demographic information, primary and secondary diagnoses coded using the International Classification of Diseases (ICD-9-CM) system, and procedure codes. Administrative data are readily available because they are routinely collected for billing purposes; these data can be used for benchmarking purposes without additional expense.²² However, administrative data frequently do not contain sufficient clinical content to perform risk adjustment,²³ might not be coded accurately, and are incomplete.²² The ICD9 coding system does not provide comprehensive clinical descriptions for ICD9 codes.²² For example, there are 40 diagnostic codes for anemia, but none specifies the hematocrit. Compared with clinical data, administrative data have low sensitivity for important comorbidities.²⁴ In one study based on heart failure patients, the sensitivity of administrative data for myocardial infarction (MI) was only 35.7%, chronic obstructive pulmonary disease 55.5%, cerebrovascular disease 24.1%, peripheral vascular disease 19.5%, and renal insufficiency 44.9%.²⁵ Chronic and asymptomatic conditions in sicker patients are undercoded in administrative data, resulting in some studies demonstrating a paradoxically protective effect for some diagnoses. In some cases there is a negative association between conditions such as hypertension or diabetes and mortality²¹; patients with hypertension or diabetes appear to have a lower mortality than patients without these conditions. Some key clinical variables, such as left ventricular ejection fraction, that are powerful predictors of mortality, are missing in administrative data.

One of the key limitations of administrative data is that they often fail to distinguish between preexisting conditions and complications that occur after hospital admission. For example, a patient who undergoes CABG surgery and has a postoperative anterior wall MI will receive the same ICD9 code as a patient who had his MI preoperatively. As a result, hospitals may receive “credit” for their complications causing their overall EMR to be higher than it should be.²⁶ This will falsely decrease their OE ratio and can lead high-mortality hospitals to be misclassified as average-mortality hospitals.²⁷ Under the Deficit Reduction Act of 2005, CMS is mandated to require hospitals to report a present-on-admission indicator, which identifies whether a secondary diagnosis represents a preexisting condition or a complication.²⁸ However, because the present-on-admission indicator will be used by CMS to eliminate payments for hospital-acquired conditions, it may be subject to data manipulation by hospitals attempting to optimize their billing.

Although clinical data are more accurate than administrative data, they are also much more expensive and less widely available.²³ One of the earliest and most highly regarded outcomes report is produced by the NYS Department of Health for CABG surgery and is based on audited clinical data.²⁹ Many of the most credible outcome reporting systems are also based on clinical data: the Society of Thoracic Surgeons (STS),³⁰ the Veterans Administration (VA) National Surgical Quality Improvement Program (NSQIP),³¹ the American College of Surgeons (ACS) NSQIP,³² and the Northern New England Cardiovascular Study Group.³³ However, some well-known statewide³⁴ and national reporting programs, such as the CMS Hospital Compare,³⁵ are based in part on administrative data.

Data quality issues notwithstanding, there remains considerable controversy surrounding the use of administrative versus clinical data for risk adjustment. No one disagrees that clinical data are preferable to administrative data if cost was not a concern. However, some (especially in the business community) view clinical data collection on a national scale as prohibitively expensive and not feasible. They argue that "perfection" (report cards based on clinical data) "should not be the enemy of good" (report cards based on administrative data).³⁶ The IOM states in *Envisioning the National Health Care Quality Report* that "administrative data, such as Medicare claims, represent one of the most practical and cost-effective data sources on selected components of health care quality available today."³⁷ Others, physicians in particular, argue that the poor quality of administrative data make it unusable for risk adjustment purposes.³⁶ Some medical societies, such as the American College of Cardiology³⁸ and the STS,³⁹ have adopted the position that report cards should be based on clinical data. This issue is far from being resolved.

Another potential source of bias in quality reporting is hospital differences in diagnostic practices. There are significant differences in the intensity of diagnostic practices across the US.⁴⁰ Hospitals that engage in more intensive diagnostic practices will diagnose more comorbidities per patient. These hospitals will appear to have a sicker patient case mix, and thus lower adjusted mortality rates. This bias will occur with both administrative and clinical data.

Data quality alone is not sufficient to ensure that risk adjustment will "level the playing field." The "right" set of risk factors must be selected for inclusion in the risk adjustment model. The challenge lies in determining which risk factors to include. Variable selection is guided by both clinical judgment and statistical criteria. The model must then be validated using measures of discrimination⁴¹ and calibration.⁴² In the seminal paper "The Risks of Risk Adjustment," Iezzoni¹⁷ revealed that different risk adjustment models (each with a different set of risk factors) frequently disagreed on hospital quality for patients hospitalized with acute MI (AMI), pneumonia, stroke, and CABG surgery. The finding that quality varies depending on the choice of risk adjustment model is a significant limitation of outcome measurement.

Assuming that data quality is perfect and that risk adjustment is perfect, can we then be confident that risk-adjusted outcome measures will reliably discriminate between high- and low-quality hospitals? Sample size (i.e.,

case volume) is also a potential problem. In an elegant article published in the *Journal of the American Medical Association*, Dimick et al.⁴³ demonstrated that, with the exception of CABG surgery, none of the surgical procedures recommended by the Agency for Healthcare Research and Quality (AHRQ) for mortality measurement (hip replacement, abdominal aortic aneurysm [AAA] repair, pediatric heart surgery, pancreatectomy, esophagectomy, and craniotomy) is performed frequently enough to be used as the basis for quality measurement. In this simulation, it was assumed that a hospital's observed mortality rate was its true mortality rate, meaning that all hospitals had the same case mix. Low-quality hospitals were defined as hospitals whose mortality rate was twice the average mortality rate for the overall cohort. With the exception of CABG surgery, the majority of hospitals did not have a sufficient case volume of surgery to permit hospitals with twice the average mortality to be identified as low-quality hospitals. Sample size is an even bigger problem when measuring physician performance.

One possible solution to the sample size problem is to aggregate similar procedures together to achieve adequate case volumes. This approach is used in the VA NSQIP and ACS NSQIP, which group together all major general and vascular surgical procedures. This means that the same risk adjustment model is used to estimate the probability of death for patients undergoing AAA repair, neurosurgery, and breast surgery, in addition to many other surgical procedures. The risk adjustment model accounts for surgical complexity by including work relative value units as one of the explanatory variables. However, the same set of risk factors and the same "weights" for those risk factors are used for all of the procedures included in the benchmarking report. A priori, it would seem unlikely that the same risk adjustment model would be the best model for procedures as different as AAA repair and breast surgery.

Another possible solution to the sample size problem is to use outcomes that occur more frequently (e.g., complications) as the basis for quality measurement. Using complications as outcome measures would increase the statistical power of performance measures to discriminate between low-quality and high-quality hospitals (and practitioners). NSQIP in fact does this: reporting both risk-adjusted mortality and morbidity for its member hospitals.³¹ However, using complications as the basis for quality measurement is not without limitations, especially if they are based on administrative data. Unlike mortality, which is relatively easy to define and capture irrespective of whether one is working with clinical or administrative data, defining what is a complication is much more subjective (and possibly more prone to data manipulation). In particular, administrative data are relatively insensitive for identifying complications. AHRQ developed a set of complication measures based on administrative data, the Patient Safety Indicators (PSIs),⁴⁴ which are increasingly being used for public reporting and pay-for-performance.⁴⁵ Using the NSQIP clinical data as the "gold standard," a validation study of select AHRQ PSIs demonstrated low sensitivities for detecting some complications, such as postoperative respiratory failure (19%), postoperative iatrogenic cardiac complications (17%), postoperative sepsis

(32%), and moderate to excellent sensitivity for other complications, such as postoperative physiologic/metabolic derangement (44%), postoperative pulmonary embolism/deep vein thrombosis (56%), and postoperative MI (81%).⁴⁵ One of the major barriers to using complications outcomes for performance measurement, especially when using administrative data, is that variability in hospital performance may reflect differences in coding or surveillance practices among hospitals, as well as "true" differences in outcomes.

In summary, the "quality" of quality measurement (based on outcomes) is a function of data quality, risk adjustment, sample size, and the accuracy of the outcomes themselves. However, the goal of performance measurement is not to make measures, but rather to improve outcomes. Having described the limitations and barriers to outcomes measurement, what is the actual evidence that measuring outcomes improves patient outcomes?

There is a substantial body of evidence that performance measurement is associated with improved outcomes. Beginning in 1989, the NYS Department of Health began to collect data on CABG surgery and provide feedback to hospitals and surgeons on their risk-adjusted outcomes. In addition to providing outcome reports, the department used the data to help guide quality improvement interventions in high-mortality hospitals. Between 1989 and 1992, the observed mortality rate for CABG surgery in NYS decreased by 21% and the risk-adjusted mortality rate decreased by 41%.⁴⁶ In a landmark study published in the *Journal of the American Medical Association*, the Northern New England Cardiovascular Disease Study Group showed that an organized intervention, consisting of outcomes feedback coupled with a structured quality improvement initiative, resulted in a 24% reduction in hospital mortality rate for CABG patients.⁴⁷ The VA conducted a parallel study for cardiac and noncardiac surgery: the VA NSQIP combined feedback with structured site visits and dissemination of best practices³¹ resulting in a 27% reduction in mortality and a 45% decrease in morbidity. This program was extended to the private sector under the sponsorship of the ACS. Hospitals participating in the ACS NSQIP also experienced significant reductions in mortality and morbidity.⁴⁸

Although at first glance the evidence connecting feedback and outcomes appears impressive, these studies are either observational in nature,^{46,49} or are prospective cohort studies that lack a control group.^{31,47,48} Nevertheless, there does seem to be a strong association between outcomes feedback and improved population outcomes. Despite the fact that these studies do not exclude the Hawthorne effect (i.e., outcomes tend to improve by virtue of being studied), the exact mechanism for the observed improvement may not matter nearly as much as the fact that outcomes did in fact improve when hospitals and physicians received feedback on their outcomes.

PROCESS MEASURES

Process measures describe what we do to our patients. The proportion of surgical patients who receive appropriate antibiotic prophylaxis within 1 hour before surgical incision is an example of a process measure. The use of a

process indicator to measure quality is predicated on the assumptions that the intervention captured by the process measure is associated with improved outcomes, and that it represents a "best practice." Physicians are frequently slow to adopt new evidence and incorporate it into their practice.^{10,50} It took 25 years after the publication of the Beta-Blocker Heart Attack Trial for β -blocker treatment post-MI to achieve nearly universal adoption.⁵¹ Performance measurement and pay-for-performance is a new paradigm for redesigning health care: public reporting is designed to accelerate the dissemination and adoption of best practices.⁵⁰ The lack of standardization and the widespread variability in the practice of medicine are believed to be factors responsible for the "quality gap."⁵² Promoting standardization through process measures is considered to be a "first step in ensuring quality improvement."⁵²

Unlike outcome measures, process measures typically do not require risk measurement as long as the appropriate target population is correctly specified. Risk-adjusted outcome measures can be biased because of improper risk adjustment, whereas measurement bias is rarely a problem for process measures.⁵³ In the case of the SCIP (Surgical Care Improvement Project) measures, whether a surgical patient receives appropriate antibiotic prophylaxis before skin incision does not depend on a patient's severity of disease. Because risk adjustment is not required, the data collection burden (and therefore the cost) for process measures is less than for outcome measures. Unlike outcome measures, process measures are directly actionable. If a hospital receives feedback that its performance on a SCIP measure for antibiotic prophylaxis is low, it can directly target this particular aspect of care for quality improvement. In contrast, if a hospital receives feedback from the ACS NSQIP that its mortality rate for noncardiac surgery is high compared with other hospitals, it will be much more difficult for a particular hospital to determine which care practices to target for improvement to reduce mortality.¹³

Process measures can improve quality by increasing adherence to measured processes of care that lead to improved outcomes. According to the Joint Commission, improvement in the performance of US hospitals on publicly reported process measures has been nothing short of dramatic. For example, in 2009, nearly 100% of hospitals administered β -blockers at discharge to MI patients, compared with fewer than 50% in 2002.⁵⁴ The primary limitation of process measures is that the evidence base linking specific medical practice and outcomes is very limited^{53,55}; there are simply too few agreed-upon "best practices" to choose from. An intervention is useful only if it results in better outcomes. Much of what we do when we care for patients is not truly evidence based, but rather grounded in expert opinion.⁵⁵ Even in areas of medicine such as cardiology, which has been the most studied, the evidence base is not particularly strong. A recent review of the scientific evidence underlying cardiovascular practice guidelines of the American College of Cardiology and the American Heart Association found that nearly 50% of the recommendations were based on expert opinion, case studies, or standards of care.⁵⁶ It can be argued that expert opinion and standards of care do in fact constitute "evidence to guide patient care."⁵⁷ However, the use of guidelines or

recommendations, based solely on expert opinion or standards of care, as performance measurement lacks face validity, especially when such measures are to be used as the basis for public reporting or pay-for-performance.⁵⁸ Turning guidelines into performance measures requires guidelines based on high-quality evidence that can be implemented using minimal clinical judgment.⁵⁸ However, many of our patients have multiple complex comorbidities that cannot always be accommodated by simple decision rules. Furthermore, guideline creation is consensus based and incorporates the implicit biases and values of expert panels. Using treatment guidelines as performance measures requires meticulous attention to minimizing these potential biases.⁵⁸

Even for process measures for which clinical trials suggest a strong link between a therapeutic intervention and mortality, the evidence that hospital compliance with best practices leads to better outcomes is somewhat mixed. CMS evaluates and publicly reports hospital performance using process measures for patients with AMI, heart failure, and pneumonia on the CMS Hospital Compare web site. These process measures include aspirin and β -blocker use in patients with AMI, angiotensin-converting enzyme inhibitor for left ventricular dysfunction, and timing of initial antibiotic use in patients with pneumonia.⁴ Individually, these processes of care have been shown to be strongly associated with outcomes in clinical trials.⁴ Despite this, hospital performance on these process measures was only “modestly correlated” with risk-adjusted mortality in one study, leading the authors to conclude that these process measures had questionable ability to discriminate between high- and low-quality hospitals.⁴ Another group of investigators using the same dataset, and a slightly different methodology, found that high performance on process measures was associated with a 7% to 15% reduction in the odds of death: hospitals in the top quartile of performance on the process measures had mortality rates that were 1% lower for patients with AMI, 0.4% lower for congestive heart failure, and 0.8% lower for pneumonia.⁵⁹

In the case of surgical care, CMS partnered with the American College of Surgeons, the American Society of Anesthesiologists, and other national groups to create SCIP. This program was established with the goal of reducing surgical complications by 25% by 2010. The SCIP measures are publicly reported on the CMS Hospital Compare web site, along with the measures for AMI, heart failure, and pneumonia. A recent study, based on 398 hospitals in the US, demonstrated that adherence with a composite all-or-none index of SCIP measures was associated with 14% lower odds of postoperative infections after cardiac and noncardiac surgery.⁶⁰ This is the first large-scale study to investigate whether adherence to the publicly reported SCIP measures is associated with improved outcomes. Taken together, these studies demonstrate a modest correlation between compliance with process measures and outcomes in both medical and surgical patients.

However, process measures also have potentially important unintended consequences. First, the emphasis on assuring adherence to measured processes (“teaching to the test”) may divert resources from other key clinical tasks, and may lead to worse outcomes in unmeasured areas.

Second, process measures may not directly measure the effectiveness or appropriateness of actual care. For example, a hospital can receive “credit” for a measured process even if the wrong drug dose is administered, or if the drug is used in a patient at risk for an undesirable drug interaction.¹⁵ Third, what constitutes best practices is a “moving target,” and the ability to innovate may be dampened unless there is a mechanism to refine and improve process measures over time.¹⁵ For example, the use of a 60-minute time window for antibiotic dosing, as specified in the SCIP prophylactic antibiotic measures, is not evidence based and may need to be modified as new research becomes available.⁶¹ Fourth, clinical care is frequently complex and nuanced, and cannot always be captured using simple process measures, making it difficult to determine whether care is consistent with practice guidelines.⁶² Fifth, process measures are easily gamed by excluding patients “who have questionable contraindications to standard therapies.”⁶³ Finally, process measures may cause physicians to focus too narrowly on improving adherence to a specific set of reported measures, encouraging tunnel vision, as opposed to striving to achieving better global outcomes.⁵³ Although process measurement may be effective for improving adherence to practices that are supported by high-level evidence, process measures do not measure clinical quality for most clinical activities, and are therefore not a substitute for outcome measures.¹⁵

STRUCTURAL MEASURES

According to Donabedian,¹³ structural measures refer to the properties of the hospital setting, including the “adequacy of facilities and equipment, the qualifications of medical staff and their organization, the administrative structure and operations of programs.” Use of the structure of the health care environment as a quality measure assumes that “given the proper settings ... good medical care will follow.”¹³ Typically, structural measures include organizational characteristics (hospital size, procedure volume, ownership, and accreditation status), human resources (practitioner credentialing and certification, and nurse staffing ratios), and technology (computerized physician order entry systems, and information infrastructure).³⁷ Structural measures are relatively inexpensive and easy to collect.³⁷ In some cases, such as procedure volume, there is a strong association between structure and outcome.^{64–66} In most cases, however, the evidence supporting a robust linkage between structure and outcome is not particularly strong.^{67,68} In addition, many structural measures such as procedure volume and ownership/teaching status are not actionable; individual hospitals, for example, cannot easily change their case volumes to “improve” quality.^{37,67,69} Finally, structural measures do not measure the performance of individual hospitals or physicians. For example, although some hospitals with high case volumes for CABG have better outcomes than lower-volume centers, some high-volume centers may have poor outcomes,⁷⁰ despite the fact that high volumes are generally associated with better outcomes. As a result, hospitals and surgeons view structural measures as “unfair”⁶⁹ to quantify performance. However, as in the case of process measures, using structural measures as performance measures provides

regulatory agencies and third-party payers the option to incentivize the adoption of innovative organizational processes, such as health information technology, computerized physician order systems, and intensivist-run intensive care units, that are viewed as best practices.

WHERE DOES ANESTHESIOLOGY FIT IN?

Anesthesiology is considered by many to exemplify the safe practice of medicine by its achievement of a reported 10-fold decrease in anesthesia-related mortality in the last several decades.⁷¹ If in fact the death rate from anesthesia is <1 in 200,000, further improvements in anesthesia mortality may be neither achievable nor even recognizable given sample size limitations. However, this view has been challenged by Lagasse in "Anesthesia Safety: Model or Myth?" in which he finds that anesthesia mortality is closer to 1 in 10,000.⁷² In contrast, a more recent study applying the AHRQ PSIs to an administrative dataset estimated that the number of anesthesia-related deaths was <1 in 100,000.⁷³ However, to our knowledge, the AHRQ PSIs for anesthesia-related complications have not been validated, and it is likely that this latter figure represents a significant underestimate.⁷⁴

Some complications are most likely anesthesia related: patient awareness under general anesthesia, transfusion reaction, neurologic deficit after regional anesthesia, or epidural abscess after neuraxial anesthesia. Death and most major complications, such as acute kidney injury, postoperative MI, respiratory failure, and cerebrovascular accidents, are more difficult to classify as anesthesia related versus surgery related. However, the absence of a validated algorithm to identify anesthesia-related outcomes does not mean that anesthesiologists should escape accountability for these outcomes. After all, if the performance of cardiac surgeons is profiled, why not that of cardiac anesthesiologists? In actuality, anesthesiologists and surgeons are part of a multidisciplinary team, and interact in a complex manner with one another, and with hospital factors, to affect patient outcomes. The goal of performance measurement is first and foremost to improve outcomes, not to assign blame. Multidisciplinary teams of anesthesiologists, surgeons, and nurses working together can use performance measures as a means to motivate and guide targeted quality improvement efforts within their institutions. Process measures should also be used to eliminate unnecessary variability in care, and to accelerate the dissemination of best practices, as long as there is strong evidence that these processes are tightly linked with improved outcomes.

We should enthusiastically promote the growth of the anesthesia registries being created by the AQI⁷ and the Multicenter Perioperative Outcomes Group,⁷⁵ and use them to develop the evidence base to identify best practices. The AQI NACOR will include data on mortality and major complications (e.g., perioperative MI, stroke, cardiac arrest, and unplanned intensive care unit admission), information on comorbidities necessary for risk adjustment, and process measures (anesthesia type, fluid and drug administration, and transfusion), and structural measures (physician training and certification, and hospital and anesthesia group characteristics). NACOR will serve as the data infrastructure for a

national anesthesia perioperative report card to provide anesthesia groups with benchmarking information. The data housed in NACOR will be used as a "learning laboratory"¹⁰ to uncover best practices in perioperative care, which can then be disseminated by the AQI to the anesthesia and surgical community.

THE WAY FORWARD

Performance measures are not perfect and cannot be made perfect. The question is how good do they need to be before we are willing to use them? Where do we set the bar?⁶⁹ How high to set the bar depends on how the report cards are to be used. If performance measures are to be publicly reported, or involve financial penalties to hospitals and physicians, then the bar needs to be much higher than if the performance measures are nonpublic and nonpunitive. Publicly reported measures have much greater potential to do harm than nonpublic reports.

What is the potential downside of public reporting? Public reporting may reduce access to care for the most vulnerable patients: some physicians may avoid high-risk patients to "protect" their outcome measures. In Pennsylvania, which has published a CABG outcomes report card since 1992, 59% of cardiologists found it more difficult to find a surgeon willing to operate on severely ill patients after the release of the report card, and 63% of the cardiac surgeons reported that they were more reluctant to operate on these patients.⁷⁶ In NYS, early evidence that public release of CABG outcomes led to increased out-of-state referrals to the Cleveland Clinic⁷⁷ was countered by a later study showing that the overall percentage of NYS residents undergoing CABG out of state actually declined.⁷⁸ However, >80% of interventional cardiologists in NYS reported that they were less likely to perform interventional procedures on high-risk patients because of public disclosure of physician performance data.⁷⁹ Follow-up studies revealed that high-risk patients in NYS were less likely to undergo percutaneous coronary interventions compared with patients in areas without public reporting,^{80,81} despite the fact that the NYS profiling system is among the oldest and most highly respected profiling efforts in the US and is based on high-quality audited clinical data. In these examples, physicians are "voting with their feet" and indicating that they do not trust risk adjustment to compensate them for caring for high-risk patients who are more likely to have poor outcomes.

Increasing barriers to care are not the only unintended consequence of public reporting. If public outcome reports are inaccurate, patients may actually be directed to lower-quality hospitals. Physicians and hospitals may engage in defensive behaviors that divert resources away from actual quality improvement (i.e., opportunity cost). In these cases, the net result of public reporting may be worse than no reporting at all. Nonpublic reporting (e.g., Northern New England, VA NSQIP, and ACS NSQIP), however, is associated with improved outcomes, and does not seem to have a significant downside.

"How good is good enough"⁶⁹ comes down to the question of data quality and whether the reports are publicly released. In the case of nonpublic reports, even biased reports based on low-quality data may still motivate

hospitals to examine their outcomes and energize quality improvement efforts (the Hawthorne effect). Biased public reports based on low-quality data, however, may distort market forces (patients referred to low-quality physicians and hospitals) with false information, resulting in lower quality care and damage to physicians' reputations.⁸ Although clinical data do not eliminate all the potential limitations of performance measurement, they do reduce bias attributable to data quality. Unfortunately, clinical data are expensive, the cost of participation in ACS NSQIP is \$100,000 per year per hospital,⁶⁹ and the resources for data collection must compete with other needs. However, public reporting responds to the public's demand for transparency and accountability in health care, and is consistent with the move away from a paternalistic care model toward one of shared responsibility.⁸²

This raises the question of whether there is a middle ground that could combine the accuracy and completeness of clinical data with the low expense and widespread availability of administrative data. Electronic health records have the potential to provide the data infrastructure for performance measurement, eliminate the need for expensive chart review-based measures,⁸³ and overcome the data quality limitations of administrative data. Health information technology is the foundation for transforming hospitals into learning laboratories to develop and test, and then disseminate best practices.¹⁰ According to the IOM, a national health information network (NHIN) is critical to accelerating the pace of quality improvement, and will require a national commitment by the federal government, as well as public and private purchasers, to developing an information infrastructure.¹⁴ There are 2 key requirements for an NHIN to serve as the backbone for quality measurement and quality improvement. First is the need for standardized definitions across all vendors so that risk factors and outcomes will be coded identically across all hospitals. Lack of data standardization will severely compromise efforts to use electronic data for quality measurement and benchmarking. Second, the information infrastructure must allow for data exchange within regional networks (interoperability).⁸⁴ Central data management is critical to the production of regional and national benchmarks. Creating an NHIN has been estimated to require an investment equivalent to 2% of annual health care spending over 5 years.⁸⁴ This initial expenditure may be offset by the financial benefit of health information technology, which has been estimated to yield a net value of nearly \$80 billion annually.⁸⁵ To date, fewer than 2% of US hospitals have a comprehensive electronic record.⁸⁶

CONCLUSION

Performance measurement is at a "tipping point." Ten years ago, our beliefs were shaken by evidence that health care in the US was not uniformly safe. With the adoption of a patient safety agenda came the recognition that safety is a quality problem, and that quality problems are best approached using performance measurements to promote and incentivize improvement in the quality of health care. There is now widespread support (federal, state, business groups, physicians, and hospitals) for performance measurement. The ASA and the ACS have created outcome

registries to measure and improve the quality of perioperative care. The STS is partnering with consumer groups to publicly release outcomes data, something that would have been unimaginable even 10 years ago. There are now >600 quality measures endorsed by the National Quality Forum.⁵⁴ According to the Joint Commission, "the proof of concept phase of national quality measurement and public reporting has now been completed."⁵⁴ However, support for performance measurement by major stakeholders has not yet translated into broad-based improvement in the safety and quality of health care.⁸⁷ We still have a long way to go to close the "quality chasm" identified by the IOM.¹⁴ To accelerate the pace of improvement in health care quality, we need comprehensive quality measurement based on an NHIN. We have now reached a critical threshold whereby we must decide if we are willing to make the next leap: investing the necessary resources to create the NHIN envisioned a decade ago by the IOM to serve as the foundation for transforming hospitals, ambulatory centers, and physicians' offices into learning networks capable of delivering the same high quality of care to all patients that we would want for ourselves and our families. ■

AUTHOR AFFILIATIONS

From the *Department of Anesthesiology, University of Rochester School of Medicine, Rochester, New York; †Department of Anesthesiology, University of Pennsylvania, Philadelphia, Pennsylvania; ‡Department of Anesthesiology, Critical Care and Pain Medicine, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts; §Department of Surgery, Anesthesia Services Sub-Section, Saddleback Memorial Medical Center, Laguna Hills, California; and ||Department of Anesthesiology, University of Maryland, Baltimore, Maryland.

REFERENCES

1. Kohn LT, Corrigan JM, Donaldson MS, eds. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press, 2000
2. Young PL, Olsen L. *Healthcare Imperative: Lowering Costs and Improving Outcomes—Brief Summary of the Workshop*. Washington, DC: The National Academies Press, 2010
3. Elmendorf DW. Letter from Congressional Budget Office to Chairman of the Committee on Health, Education, Labor, and Pensions, 2009
4. Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. *JAMA* 2006;296:2694–702
5. Rosenthal MB, Frank RG, Li Z, Epstein AM. Early experience with pay-for-performance: from concept to practice. *JAMA* 2005;294:1788–93
6. Baumann MH, Dellert E. Performance measures and pay for performance. *Chest* 2006;129:188–91
7. Hannenberg AA, Warner MA. The registry imperative. *Anesthesiology* 2009;111:687–9
8. Clancy C. The performance of performance measurement. *Health Serv Res* 2007;42:1797–801
9. Conway PH, Clancy C. Transformation of health care at the front line. *JAMA* 2009;301:763–5
10. Dougherty D, Conway PH. The "3T's" road map to transform US health care: the "how" of high-quality care. *JAMA* 2008;299:2319–21

11. Physician Consortium for Performance Improvement (PCPI) Performance Measure Status Report. Available at <http://www.ama-assn.org/ama1/pub/upload/mm/370/measures.pdf>. Accessed January 24, 2011
12. Cohen NA, Hannenberg AA. Risk stratification index: an important advance in comparing health care apples to oranges. *Anesthesiology* 2010;113:1004-6
13. Donabedian A. Evaluating the quality of medical care: 1966. *Milbank Q* 2005;83:691-729
14. Committee on Quality of Health Care in America, Institute of Medicine. *Crossing the Quality Chasm*. Washington, DC: National Academy Press, 2001
15. Krumholz HM, Normand SL, Spertus JA, Shahian DM, Bradley EH. Measuring performance for treating heart attacks and heart failure: the case for outcomes measurement. *Health Aff (Millwood)* 2007;26:75-85
16. Iezzoni LI. Reasons for risk adjustment. In: Iezzoni LI, ed. *Risk Adjustment*. Chicago: Health Administration Press, 2003:1-16
17. Iezzoni LI. The risks of risk adjustment. *JAMA* 1997;278:1600-7
18. Iezzoni LI. 'Black box' medical information systems: a technology needing assessment. *JAMA* 1991;265:3006-7
19. Shahian DM, Normand SL. Comparison of "risk-adjusted" hospital outcomes. *Circulation* 2008;117:1955-63
20. Glance LG, Osler TM, Dick AW, Mukamel DB, Meredith W. The Survival Measurement and Reporting Trial for Trauma (SMARTT): background and study design. *J Trauma* 2010;68:1491-7
21. Shahian DM, Silverstein T, Lovett AF, Wolf RE, Normand SL. Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards. *Circulation* 2007;115:1518-27
22. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med* 1997;127:666-74
23. Institute of Medicine. *Performance Measurement: Accelerating Improvement*. Washington, DC: National Academies Press, 2006
24. Jollis JG, Ancukiewicz M, DeLong ER, Pryor DB, Muhlbaier LH, Mark DB. Discordance of databases designed for claims payment versus clinical information systems: implications for outcomes research. *Ann Intern Med* 1993;119:844-50
25. Lee DS, Donovan L, Austin PC, Gong Y, Liu PP, Rouleau JL, Tu JV. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care* 2005;43:182-8
26. Jollis JG, Romano PS. Pennsylvania's focus on heart attack: grading the scorecard. *N Engl J Med* 1998;338:983-7
27. Glance LG, Dick AW, Osler TM, Mukamel DB. Accuracy of hospital report cards based on administrative data. *Health Serv Res* 2006;41:1413-37
28. Coffey R, Milenkovic M, Andrews RM. *The Case for the Present-on-Admission (POA) Indicator*. Rockville, MD: U.S. Agency for Healthcare Research and Quality, 2006
29. Hannan EL, Kilburn H Jr, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State: an analysis of risk factors and hospital mortality rates. *JAMA* 1990;264:2768-74
30. Shroyer AL, Coombs LP, Peterson ED, Eiken MC, DeLong ER, Chen A, Ferguson TB Jr, Grover FL, Edwards FH. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. *Ann Thorac Surg* 2003;75:1856-64
31. Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, Chong V, Fabri PJ, Gibbs JO, Grover F, Hammermeister K, Irvin G III, McDonald G, Passaro E Jr, Phillips L, Scamman F, Spencer J, Stremple JF. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. *National VA Surgical Quality Improvement Program*. *Ann Surg* 1998;228:491-507
32. Khuri SF, Henderson WG, Daley J, Jonasson O, Jones RS, Campbell DA Jr, Fink AS, Mentzer RM Jr, Neumayer L, Hammermeister K, Mosca C, Healey N. Successful implementation of the Department of Veterans Affairs' National Surgical Quality Improvement Program in the private sector: the Patient Safety in Surgery Study. *Ann Surg* 2008;248:329-36
33. O'Connor GT, Birkmeyer JD, Dacey LJ, Quinton HB, Marrin CA, Birkmeyer NJ, Morton JR, Leavitt BJ, Maloney CT, Hernandez F, Clough RA, Nugent WC, Olmstead EM, Charlesworth DC, Plume SK. Results of a regional study of modes of death associated with coronary artery bypass grafting. *Northern New England Cardiovascular Disease Study Group*. *Ann Thorac Surg* 1998;66:1323-8
34. PA Health Care Cost Containment Council (PHC4). *Pennsylvania Is a Leader in Quality Outcomes Reporting, 2010*
35. Jha AK, Li Z, Orav EJ, Epstein AM. Care in U.S. hospitals: the Hospital Quality Alliance Program. *N Engl J Med* 2005;353:265-74
36. Lee TH, Meyer GS, Brennan TA. A middle ground on public accountability. *N Engl J Med* 2004;350:2409-12
37. Hurtado MP, Swift EK, Corrigan JM. *Envisioning the National Health Care Quality Report*. Washington, DC: National Academy Press, 2001
38. Drozda JP Jr, Hagan EP, Mirro MJ, Peterson ED, Wright JS. ACCF 2008 health policy statement on principles for public reporting of physician performance data: a Report of the American College of Cardiology Foundation Writing Committee to develop principles for public reporting of physician performance data. *J Am Coll Cardiol* 2008;51:1993-2001
39. Shahian DM, Edwards FH, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, O'Brien SM, Shewan CM, Dokholyan RS, Peterson ED. Quality measurement in adult cardiac surgery. Part 1. Conceptual framework and measure selection. *Ann Thorac Surg* 2007;83:S3-12
40. Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. Regional variations in diagnostic practices. *N Engl J Med* 2010;363:45-53
41. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36
42. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. New York: Wiley, 2000
43. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. *JAMA* 2004;292:847-51
44. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA* 2003;290:1868-74
45. Romano PS, Mull HJ, Rivard PE, Zhao S, Henderson WG, Loveland S, Tsilimingras D, Christiansen CL, Rosen AK. Validity of selected AHRQ patient safety indicators based on VA National Surgical Quality Improvement Program data. *Health Serv Res* 2009;44:182-204
46. Hannan EL, Kilburn H Jr, Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York State. *JAMA* 1994;271:761-6
47. O'Connor GT, Plume SK, Olmstead EM, Morton JR, Maloney CT, Nugent WC, Hernandez F Jr, Clough R, Leavitt BJ, Coffin LH, Marrin CA, Wennberg D, Birkmeyer JD, Charlesworth DC, Malenka DJ, Quinton HB, Kasper JF. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. *JAMA* 1996;275:841-6
48. Hall BL, Hamilton BH, Richards K, Bilimoria KY, Cohen ME, Ko CY. Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: an evaluation of all participating hospitals. *Ann Surg* 2009;250:363-76
49. Hannan EL, Sarrazin MS, Doran DR, Rosenthal GE. Provider profiling and quality improvement efforts in coronary artery bypass graft surgery: the effect on short-term mortality among medicare beneficiaries. *Med Care* 2003;41:1164-72
50. Fleisher LA. Improving perioperative outcomes: my journey into risk, patient preferences, guidelines, and performance measures: Ninth Honorary FAER Research Lecture. *Anesthesiology* 2010;112:794-801
51. Lee TH. Eulogy for a quality measure. *New Engl J Med* 2007;357:1175-7
52. McGlynn EA. Intended and unintended consequences: what should we really worry about? *Med Care* 2007;45:3-5

53. Nolte E. International Benchmarking of Healthcare Quality Technical Report. Santa Monica: RAND, 2010
54. Chassin MR, Loeb JM, Schmaltz SP, Wachter RM. Accountability measures: using measurement to promote quality improvement. *New Engl J Med* 2010;363:683–8
55. Jha AK. Measuring hospital quality: what physicians do? How patients fare? Or both? *JAMA* 2006;296:95–7
56. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC Jr. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009;301:831–41
57. Horlocker TT, Brown DR. Evidence-based medicine: haute couture or the emperor's new clothes? *Anesth Analg* 2005;100:1807–10
58. Shaneyfelt TM, Centor RM. Reassessment of clinical practice guidelines: go gently into that good night. *JAMA* 2009;301:868–9
59. Jha AK, Orav EJ, Li Z, Epstein AM. The inverse relationship between mortality rates and performance in the Hospital Quality Alliance measures. *Health Aff (Millwood)* 2007;26:1104–10
60. Stulberg JJ, Delaney CP, Neuhauser DV, Aron DC, Fu P, Koroukian SM. Adherence to surgical care improvement project measures and the association with postoperative infections. *JAMA* 2010;303:2479–85
61. Hawn MT. Surgical care improvement: should performance measures have performance measures? *JAMA* 2010;303:2527–8
62. Garber AM. Evidence-based guidelines as a foundation for performance incentives. *Health Aff (Millwood)* 2005;24:174–9
63. Romano PS. Improving the quality of hospital care in America. *New Engl J Med* 2005;353:302–4
64. Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista BA, Welch HG, Wennber DE. Hospital volume and surgical mortality in the United States. *N Engl J Med* 2002;346:1128–44
65. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. *N Engl J Med* 2003;349:2117–27
66. Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. *Ann Surg* 2006;243:411–7
67. Meyer GS, Massagli MP. The forgotten component of the quality triad: can we still learn something from "structure"? *Jt Comm J Qual Improv* 2001;27:484–93
68. Hearld LR, Alexander JA, Fraser I, Jiang HJ. Review: how do hospital organizational structure and processes affect quality of care? A critical review of research methods. *Med Care Res Rev* 2008;65:259–99
69. Birkmeyer JD, Kerr EA, Dimick JB. Improving the Quality of Quality Measurement Performance Measurement: Accelerating Improvement. Washington, DC: National Academy of Sciences, 2006:177–203
70. Peterson ED, Coombs LP, DeLong ER, Haan CK, Ferguson TB. Procedural volume as a marker of quality for CABG surgery. *JAMA* 2004;291:195–201
71. Leape LL, Berwick DM, Bates DW. What practices will most improve safety? Evidence-based medicine meets patient safety. *JAMA* 2002;288:501–7
72. Lagasse RS. Anesthesia safety: model or myth? A review of the published literature and analysis of current original data. *Anesthesiology* 2002;97:1609–17
73. Li G, Warner M, Lang BH, Huang L, Sun LS. Epidemiology of anesthesia-related mortality in the United States, 1999–2005. *Anesthesiology* 2009;110:759–65
74. Lagasse RS. Innocent prattle. *Anesthesiology* 2009;110:698–9
75. Kheterpal S, Woodrum DT, Tremper KK. Too much of a good thing is wonderful observational data for perioperative research. *Anesthesiology* 2009;111:1183–4
76. Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care: a survey of cardiovascular specialists. *N Engl J Med* 1996;335:251–6
77. Omoigui NA, Miller DP, Brown KJ, Annan K, Cosgrove D III, Lytle B, Loop F, Topol EJ. Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. *Circulation* 1996;93:27–33
78. Peterson ED, DeLong ER, Jollis JG, Muhlbaier LH, Mark DB. The effects of New York's bypass surgery provider profiling on access to care and patient outcomes in the elderly. *J Am Coll Cardiol* 1998;32:993–9
79. Narins CR, Dozier AM, Ling FS, Zareba W. The influence of public reporting of outcome data on medical decision making by physicians. *Arch Intern Med* 2005;165:83–7
80. Moscucci M, Eagle KA, Share D, Smith D, De Franco AC, O'Donnell M, Kline-Rogers E, Jani SM, Brown DL. Public reporting and case selection for percutaneous coronary interventions: an analysis from two large multicenter percutaneous coronary intervention databases. *J Am Coll Cardiol* 2005;45:1759–65
81. Apolito RA, Greenberg MA, Menegus MA, Lowe AM, Sleeper LA, Goldberger MH, Remick J, Radford MJ, Hochman JS. Impact of the New York State Cardiac Surgery and Percutaneous Coronary Intervention Reporting System on the management of patients with acute myocardial infarction complicated by cardiogenic shock. *Am Heart J* 2008;155:267–73
82. Wang MC, Hyun JK, Harrison M, Shortell SM, Fraser I. Redesigning health systems for quality: lessons from emerging practices. *Jt Comm J Qual Patient Saf* 2006;32:599–611
83. Conway PH, Clancy C. Transformation of health care at the front line. *JAMA* 2009;301:763–5
84. Kaushal R, Blumenthal D, Poon EG, Jha AK, Franz C, Middleton B, Glaser J, Kuperman G, Christino M, Fernandopulle R, Newhouse JP, Bates DW; Cost of National Health Information Network Working Group. The costs of a national health information network. *Ann Intern Med* 2005;143:165–73
85. Walker J, Pan E, Johnston D, Adler-Milstein J, Bates DW, Middleton B. Market watch: the value of health care information exchange and interoperability. *Health Aff (Millwood)* 2005;24:W5–10–W5–18
86. Jha AK, DesRoches CM, Campbell EG, Donelan K, Rao SR, Ferris TG, Shields A, Rosenbaum S, Blumenthal D. Use of electronic health records in U.S. hospitals. *N Engl J Med* 2009;360:1628–38
87. Landrigan CP, Parry GJ, Bones CB, Hackbarth AD, Goldmann DA, Sharek PJ. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med* 2010;363:2124–34