Our New Reality of Public Reporting: Shame Rather Than Blame?

Susan D. Moffatt-Bruce, MD, PhD, Michelle C. Nguyen, MD, James I. Fann, MD, and Stephen Westaby, PhD, FRCS

Ohio State University Wexner Medical Center, Columbus, Ohio; Stanford University, Palo Alto, California; and Nuffield Department of Surgery, Oxford, United Kingdom

The number of organizations issuing reports on hospital and physician quality performance has increased markedly over the past decade. Differences in the measures, data sources, and scoring methodologies produce contradictory results that lead to confusion for the public, providers, and governing boards, and impair the public's ability to make well-informed choices about health care providers [1]. This variability continues today and points to concerns about validity and the ultimate reliability of the measures used by these groups.

The hospital community and surgeons as a whole support the principle of accountability through public reporting of health care performance data. However, performance data that are inappropriately collected, analyzed, and displayed may add more confusion than clarity to the health care quality question [1]. For data to be understood and for results to be comparable, publicly reported data should adhere to a set of guiding principles. With that goal in mind, the Association of American Medical Colleges (AAMC) convened a panel of experts in 2012 and 2013 to develop a set of guiding principles that can be used to evaluate quality reports. The principles were organized into three broad categories: purpose, transparency, and validity.

Under the domain of purpose, the AAMC recognized that public reporting and performance measurement should occur for a variety of reasons, including consumer education, provider quality improvement, and purchaser decision making. Relative to transparency, the AAMC believed that methodologic details should be clearly discerned as they can impact both providers' performance data and the appropriate interpretation of the data. Transparency also requires that all information necessary to understand the data be available to and interpretable by the reader. Limitations in data collection and methodology as well as relevant financial interests should always be disclosed in language that is discernable. Lastly, validity of the data must ensure that the methodology, data collection, scoring, and benchmarks produce an accurate reflection of the characteristic being measured and reflect the care being provided by the hospital or physician. These guiding principles were expanded and proposed by the AAMC to facilitate adherence and to ensure appropriate interpretation of performance as public reporting becomes truly a cottage industry.

Federally Facilitated Quality and Patient Programs

In the recent and final Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long-Term Care Hospital policy for fiscal year 2015, the Centers for Medicare and Medicaid Services (CMS) aimed at promoting high-value and high-quality care using a program that is targeted at a specific set of preventable infections, events, and conditions that occur in the inpatient setting, referred to as hospital-acquired conditions (HAC). Similar to CMS value-based purchasing and readmissions reduction programs, this HAC reduction program has significant implications for academic centers, particularly major teaching hospitals. In its current iteration, the program can only be defined as a penalty program with a current 1% withholding [2]. The program reduces payments to hospitals that rank in the worst performing quartile. The worst performing quartile is identified by calculating the total HAC score, which is based on the hospital's performance on four risk-adjusted quality measures (patient safety indicator 90 composite, central-line associated bloodstream infection, catheter-associated urinary tract infection, and surgical site infection for colon surgery and hysterectomy). Hospitals with a total HAC score above the 75th percentile of the total HAC score distribution are subject to payment reduction [3]. An analysis of the preliminary penalties suggested that major teaching hospitals are 2.9 times more likely to be penalized in this program than nonteaching hospitals [4, 5]. Additionally, CMS estimates that 56% of major teaching hospitals will be penalized [2]. Ultimately, many could face substantial penalties from all three pay-for-performance programs developed by CMS. One has to surmise that the most acutely ill and complex patients will acquire these HACs and that such a hefty penalty program will promote risk-averse behavior, resulting in reluctance of physicians and systems to accept these patients. The results of this HAC program, good or bad, are publicly reported, and this publicly reported score is the basis for both nonpayment and penalties. Not surprisingly, this methodology has been widely scrutinized.

To our knowledge, there is no evidence that this type of administrative composite quality measure is linked to clinical validated risk-adjusted mortality, length of stay,

Address correspondence to Dr Moffatt-Bruce, 130 Doan Hall, 410 W 10th Ave, Columbus, OH 43210; email: susan.moffatt-bruce@osumc.edu.

Abbreviations and Acronyms

AAMC	=	Association of American Medical Colleges
AHRQ	=	Agency for Healthcare Research and Quality
CHSD	=	Congenital Heart Surgery Database
CMS	=	Centers for Medicare and Medicaid Services
HAC	=	hospital-acquired condition
NHS	=	National Health Service
OSUWMC	=	Ohio State University Wexner Medical Center
PSI	=	patient safety indicator
STS	=	The Society of Thoracic Surgeons

and hospital charges. Yet, this information may be linked to hospitals' and health systems' pursuit of what Don Berwick and the Institute for Healthcare Improvement have called the "triple aim"—improving the experience of care, improving health of populations, and reducing per capita cost of health care [6]. A main concern is how health care systems, diverse training programs, and quaternary care centers can use the data to mitigate patient risk and at the same time maintain fiscal survival. Clearly, to stop providing the most specialized care to the sickest patients so as to avoid being penalized by administratively derived data is counterproductive.

In an effort to answer these questions, our group aimed to determine the effects of HACs on mortality, prolonged length of stay, and excessive hospital charges using the Nationwide Inpatient Sample, Healthcare Cost and Utilization Project, Agency for Healthcare Research and Quality (AHRQ), and the largest all-payer publically accessible database of inpatient visits in the United States. In a retrospective, cross-sectional analysis of weighted national estimates from the 2012 Nationwide Inpatient Sample data, we established the effects of at least one HAC on mortality, prolonged length of stay, and excessive hospital charges through univariate and multivariate logistic regression. Prolonged length of stay was defined as a stay longer than 4.5 days (greater than the 75th percentile of all hospital stays in 2012) and excessive hospital charges as a charge greater than $40,448 (greater than the 75th percentile of all hospital charges in 2012). Our findings showed that patients with at least one HAC have a 54% higher likelihood of dying during an inpatient hospital stay than patients without a HAC. Additionally, the odds of patients with at least one HAC having a prolonged hospital stay and excessive charges are 1.64 and 1.85 times that of patients without, respectively (Table 1). In examining the impact of each HAC on mortality, we determined that pressure ulcers stages III and IV, manifestations of poor glycemic control, and vascular catheter-associated infections are the top three drivers of mortality. These findings are consistent with prior studies that have shown these HACs to be major causes of inpatient morbidity and mortality [7–9]. They also underscore the value of national campaigns to reduce their occurrence.

There is no doubt, therefore, that HACs have a negative impact on patient outcomes and provide challenges for hospitals and payers. Currently, hospitals are being penalized and denied payment based on their respective HAC score. Yet, how this score is interpreted remains unclear, and its ability to measure the true performance and quality of an individual hospital is not well defined [10]. So what are these stakeholders to do? Should patients seek care at hospitals with the most favorable HAC score? Does penalizing hospitals promote or hurt their ability to improve performance when resources are so limited?

Resultant Ranking Systems: 700 Top 100 Hospitals

Attempts to assess the quality and safety of hospitals have proliferated, many without the AAMC guiding principles, and some may say have become a cottage industry and include a growing number of consumer-directed hospital rating systems. However, relatively little is known about what these rating systems reveal other than to confirm that there appear to be more than 700 "top 100" hospitals in America. To better understand differences in hospital ratings, Pronovost and colleagues [11] recently published a comparison of four national rating systems, including US News & World Report's "best hospitals" report, Leapfrog, CMS's Hospital Compare, Consumer Reports, and Healthgrades [11]. They designated high and low performers for each rating system and examined the overlap among rating systems and how hospital characteristics corresponded with performance on each. No American hospital was rated as a high performer by all four national rating systems, and only 10% of the 844 hospitals rated as a high performer by one rating system were rated as a high performer by any of the other rating systems. There was a general lack of agreement among the national hospital rating systems, a finding attributed to each system using its own rating methods, having a different focus to its ratings, and stressing different measures of performance. Furthermore, this research group found that differences across hospital ratings add complexity to ascertaining a hospital's actual quality, making it difficult for payers to recognize and reward hospitals for high-quality care, complicating decisions for hospital leadership regarding the focus of their improvement efforts, and most importantly, confounding medical judgment for current and prospective patients and families [11].

Patient Safety Indicator Story: An Example of Futility?

Inherent in many of the ranking systems and public reporting are entities referred to as patient safety indicators (PSIs). To address the need for quality monitoring, the AHRQ established a set of PSIs to assist in monitoring potentially preventable events for patients

Table 1. Effects of At Least One Hospital-Acquired Condition on Mortality, Prolonged Length of Stay, and Excessive Hospital Charges

Effect of HAC	None n (%)	At Least 1 n (%)	Crude OR (95% CI)	Adjusted OR (95% CI)	*p* Value
Mortality	590,845 (1.73)	82,205 (3.67)	2.17 (2.12–2.22)	1.54 (1.51–1.58)	<0.0001
Prolonged LOS	9,304,176 (27.17)	965,170 (43.04)	2.03 (1.99–2.06)	1.64 (1.61–1.66)	<0.0001
Excessive charges	8,012,581 (23.40)	919,405 (41.00)	2.28 (2.22–2.34)	1.85 (1.79–1.90)	<0.0001

Odds ratios are adjusted for age, sex, race, insurance status, household income, Elixhauser comorbidity score, hospital size, hospital type, and hospital region.

CI = confidence interval; HAC = hospital-acquired condition; LOS = length of stay; OR = odds ratio.

treated in hospitals using the International Classification of Diseases, Ninth Revision (soon to be 10th revision), administrative billing codes. The majority of these PSIs affect surgeons and surgical practice. Use of these PSIs has led to an understanding of the flaws in clinical self-reporting and the prohibitive cost of chart abstraction that have been designated to align with patient safety events. The AHRQ PSIs are developed and evaluated by the National Quality Forum on a routine basis and have become hotly debated as they are now within the HAC program as well as the value-based purchasing system, both of which are associated with significant cost gains or reduction based on PSI performance. Several groups have established that PSIs are associated with identifying opportunities to improve patient care such as postsurgical care and readmissions [12, 13]. In addition, the group at Stanford concluded that since their introduction and review, certain PSIs have decreased while patient care and provision of coordinated care have improved [14]. The PSIs were truly designed to be used to flag patient safety events and identify areas for improvement. As they have become part of ranking and gains/penalty programs, the validity of these administratively defined events has been questioned. Ramanathan and colleagues [15] found that at a large academic medical center, the validity of these PSIs was very low, and the most common contributing factors included coding errors (30%), documentation errors (19%), and insufficient criteria for PSI in the chart (16%) [16].

Public Reporting and Its Challenges for Surgeons

The challenge of medical centers and surgeons is to provide medical and surgical care to patients with highly complex injuries and diseases in an evidence-based fashion while maintaining quality and avoiding penalization with reporting systems that use minimally risk adjusted data and may not be truly reflective of the level of care that is provided. Several industries have joined these public reporting efforts by taking advantage of available data. Much controversy has stemmed from the release of several of these types of databases, including ProPublica's Surgeon Scorecard and Consumers' Checkbook.

The investigative journalism company, ProPublica, analyzed the complication rates of 16,827 surgeons operating in 3,575 hospitals. The report used Medicare data from 2009 to 2013 for eight common procedures that were considered to be elective and low risk. Admissions through the emergency room and transfers from other facilities were excluded from the report. Two measures of harm were used: inhospital mortality and readmissions within 30 days with a primary diagnosis identified as a likely complication. A risk-adjustment model was used; however, the model seems unclear and not widely used in existing reports. Importantly, the final adjusted rate that is reported does not reflect past performance. The data attempt to reflect how a surgeon would perform at an average-performing hospital on an average patient population rather than the more complex patient population often treated at large academic institutions. Even surgeons with zero complications were reported with an adjusted rate of at least 1.1%. Aggregate results at the national level show overall low complications (less than 5%), but a wide variation between surgeons. For Ohio State University Wexner Medical Center (OSUWMC) surgeons, five of the eight procedures qualified for publication (volume of remaining procedures was either too low or included too many nonelective cases or transfers). The procedures included for OSUWMC were knee replacement, hip replacement, cholecystectomy, lumbar spinal fusion, and prostatectomy. All of the surgeons at OSUWMC who qualified for a rating were rated as having a medium level of complications even though several had zero complications reported during the study period.

The other surgeon performance reporting agency, Consumers' Checkbook, analyzed Medicare data from 2009 to 2012 for 14 groups of commonly performed procedures. Transfers from other hospitals were excluded from this report. Each surgeon with a sufficient population received a star ranking based on performance in mortality within 90 days of the admission, prolonged hospital stay, and 90-day readmissions. An observed and predicted rate is calculated for each surgeon. The number of stars from three to five reflects how significantly higher the observed rate is compared with the predicted rate. Very few OSUWMC surgeons are reported in several of the categories, likely secondary to high patient transfer rates.

Not unexpectedly, the release of these score cards triggered much discussion on the topic of performance data. While intentions may be positive, care must be taken to ensure that reported data are clinically validated,

reported accurately, and have the buy-in of key stakeholders, the physicians. Everyone supports transparency; however, the value of transparency is often misunderstood. The methodology of the discussed reports and other reports oftentimes lack validity and can be difficult for the general public to understand. One organization can provide conflicting information from another, and the public is left to make decisions with limited information.

With the amount of gains/penalty increase, the viability of the academic enterprise is threatened, providing a real incentive for these systems to become more selective of accepting complex transfers from other institutions. These shifts in penalty and compensation mechanisms not only affect the institution, the provider, but also, most importantly, the patients. What is necessary now is the acceptance of the health care culture change from volume to value and ownership and accountability of our quality as well as our data.

International Experiences With Public Reporting of Cardiac Surgery Outcomes

Most European hospitals and surgical societies collect outcome data for the purposes of quality assurance and internal regulation [17]. With the exception of the National Health Service (NHS) in England, there has been little enthusiasm for public dissemination of the information. Many countries had watched the New York State experience of named surgeon reporting since the 1990s [17]. There has always been awareness that The Society of Thoracic Surgeons (STS [2007]) and American College of Cardiology (2008) had warned of the dangers of public reporting at the individual level [18, 19]. The practice can change focus from patient care to self-preservation as characterized by risk aversion and gaming with risk algorithms [20]. The physiology of risk aversion has been studied in financial traders [21]. Cortisol levels rise by almost 70% in bankers and fund managers after prolonged uncertainty during market volatility. Persistent cortisol elevation leads to heightened anxiety, impaired learning, depression, and unwillingness to take risks. The study from Cambridge University's Judge Business School and Institute of Metabolic Science reproduced risk aversion by administering hydrocortisone to volunteers during financial risk-taking tasks. They performed less effectively. Regardless, publication of surgeon-specific mortality data began in England after repeat NHS institutional failures that resulted in high mortality rates [22, 23]. Cardiac surgeons' risk-adjusted mortality rates were released in 2005, then for other surgical specialties and interventional cardiologists in 2014. The stated aims were to improve standards through openness and transparency and compliance with quality improvement initiatives. No attempt was made to improve NHS facilities or access to care characterized by long waiting lists. There was considerable opposition and criticism of the process. Those who failed to submit their own data had their names supplied to the media. Despite the bell curve of outcomes that applies to any biological system, only three of more than 5,000 surgeons were deemed "outliers." Their pictures appeared on the front pages of the newspapers.

When Donald Berwick was asked by the British prime minister to review the state of the NHS, he reiterated, "Hospitals, not individuals, must be held accountable for poor outcomes" [24, 25]. Nevertheless, the process did reduce mortality to low levels in a system where most cardiac centers were not funded for circulatory support equipment and team consistency was poor through the European Working Time Directive. Failure to rescue rates among the deaths were much higher than reported in the United States. The emphasis throughout the surgical profession switched to avoiding personal liability. That could be regarded as sensible in a system where physical resources are ranked near the bottom of the Organization for Economic Cooperation and Development leagues. As a result, the United Kingdom achieves some of Europe's lowest cardiac surgical death rates but virtually the worst cancer survival rates. Other surgical specialties began to recognize risk-averse practice, and scandals emerged when individual surgeons were discovered to have manipulated their outcome data [20].

A survey of UK cardiothoracic surgeons showed 85% support the transparency agenda, but by reporting center-specific rather than surgeon-specific mortality data. Ninety-five percent recognized risk-averse practice in the profession, and 84% considered that the public were likely to misinterpret the information in its presented form. Seventy-five percent said surgical training had suffered through defensive practice and were unwilling to assist trainees to operate on their patients. Half agreed that risk assessment data were routinely manipulated. There is broad agreement that the individually focused "name and shame" agenda has affected recruitment to cardiothoracic surgery in the United Kingdom. Currently, only 40% of pediatric cardiac surgeons, 54% of thoracic surgeons, and 64% of adult cardiac surgeons had their medical training in the United Kingdom [25]. From the General Medical Council Specialist Register, we know that 68% of trainees entering cardiothoracic surgery at the time of the Bristol children's heart scandal (2000) were UK graduates. In 2013, this figure was 14%. Of those receiving their certificate of completion of training in 2014, only 20% were UK graduates.

Are there reasons why the unintended consequences of public outcome reporting should be more profound in the United Kingdom? Firstly, public outcome disclosure has not improved patient choice, hospital staffing levels, or equipment. When the information disappeared through lack of funding between 2009 and 2012, no one noticed. Secondly, the United States has a fee-for-service system. A surgeon who adopts defensive practice may lose his livelihood. In the United Kingdom, the opposite applies. If a surgeon has a run of failure-to-rescue events with high-risk patients and does not adopt defensive practice, he may lose his livelihood. Equally, in the British NHS, high-risk patients with prolonged intensive care stay increase hospital costs. So risk aversion has economic benefits. In contrast, the US star rating system drives the quality agenda with improved facilities and financial

rewards. The focus has moved away from individual persons because name and shame really does not pay. To quote Berwick, "Measurement is best used for learning rather than for selection, reward or punishment. Real improvement comes from changing systems not change within systems. Concentrate on meeting the needs of patients rather than the needs of organizations. Lastly, effective leaders challenge the status quo by offering clear ideas about superior alternatives" [26].

American Cardiothoracic Surgery and Public Reporting

In January 2015, the STS began to publicly report outcomes of pediatric and congenital cardiac surgery using the 2014 STS Congenital Heart Surgery Database (CHSD) mortality risk model. The STS released the star ratings for congenital heart surgery public reporting, which are based on the STS CHSD mortality risk model. The STS CHSD public reporting initiative continues to grow, increasing from 23% of enrolled participants for the January 2015 data release to 33% in the current data release. The Spring 2015 STS CHSD feedback report includes data from 116 participants in the STS CHSD, including 11 one-star programs, 79 two-star programs, and 6 three-star programs. Twenty participants did not receive a star rating owing to incomplete data, which will also be an issue with public reporting. It is generally believed that because the 2014 STS CHSD mortality risk model adjusts for procedural factors and patient-level factors, centers have to be very aware and actually embrace the resources and rigor to complete all data fields for patient-level factors [27, 28].

The STS, as a society, believes that the public has a right to know the quality of surgical outcomes and considers public reporting an ethical responsibility of the specialty. To that end, STS public reporting online enables STS Adult Cardiac Surgical Database and CHSD participants to voluntarily report to each other and to the public their heart surgery scores and star ratings. This reporting is voluntary at this point, and in the adult world, only includes coronary artery bypass graft surgery, aortic valve replacement, and combined coronary artery bypass graft surgery and aortic valve replacement cases to date. The STS only publishes group or hospital level data for which the participant correctly submitted the required data for score and star rating analysis. Again, as for the congenital data, groups and hospitals that have more than 10% missing data during the reporting period do not receive scores and star ratings and, therefore, are not able to be publicly reported.

Further embracing public reporting, the STS Quality Measurement Task Force is developing a portfolio of composite performance measures for the most commonly performed procedures in adult cardiac surgery and thoracic surgery [28, 29]. The STS composite measure for mitral valve repair/replacement involved all patients undergoing isolated mitral valve repair/replacement who were examined, with or without concomitant performance of tricuspid valve repair, surgical arrhythmia ablation, or repair of atrial septal defect, between July 1, 2011, and June 30, 2014. Star rating classifications included 23 of 867 (2.6%) one-star programs (lower than expected performance), 795 of 867 (91.7%) two-star programs (as expected or average performance), and 49 of 867 (5.7%) three-star programs (higher than expected performance) [28]. The STS has developed its first composite measure for general thoracic surgery for lobectomy, and it is composed of two outcomes: risk-adjusted mortality, and any or none, risk-adjusted major complications. General Thoracic Surgery Database data were included from 2011 to 2014 to provide adequate sample size, and 95% Bayesian credible intervals were used to determine star ratings. Approximately 5% of participants were one-star, 7% were three-star, and 88% were two-star programs. The STS has developed these composite measures to compare programs to be used for quality assessment and provider feedback, and eventually voluntary public reporting [29].

Although these composite measures are new, the STS Adult Cardiac Surgery Database was actually initiated in 1989 and includes more than 1,085 participating centers, representing 90% to 95% of current US-based adult cardiac surgery hospitals. Since its inception, the primary goal of the STS Adult Cardiac Surgery Database has been to use clinical data to track and improve cardiac surgical outcomes. Patients' preoperative risk characteristics, procedure-related processes of care, and clinical outcomes data have been captured and analyzed, with timely risk-adjusted feedback reports to participating providers [30].

To evaluate participant characteristics and outcomes during the first 4 years of the STS public reporting program, one that was initially hotly debated, a detailed analysis of a national, voluntary, cardiac surgery public reporting program using STS clinical registry data and National Quality Forum–endorsed performance measures was completed. Among 8,929 unique observations (approximately 1,000 STS participant centers, 9 reporting periods), 916 sites (10.3%) were classified as low performing, 6,801 (76.2%) were average, and 1,212 (13.6%) were high performing. The STS public reporting participation varied from 22.2% to 46.3% over the nine reporting periods. Risk-adjusted, patient-level mortality rates for isolated coronary artery bypass grafting were consistently lower in public reporting sites versus nonreporting sites (p value range, <0.001 to 0.0077). Interestingly, STS programs that voluntarily participate in public reporting have significantly higher volumes and performance, and no evidence of risk aversion was found [31].

Public Shame and Blame?

In the past decade, much emphasis has been placed on the quality of medical care, and great efforts have been put forward to identify patient safety improvement opportunities. These efforts have generated many different indicators of safety events and low quality of care, including PSIs, HACs, and readmission penalty programs. Additionally, as consumerism has grown,

ranking systems and marketing therein have exploded throughout all domains of health care. It has now become a complicated process to manage these indicators and use these quality and patient safety programs and rankings to actually facilitate better care and meet the expectations of patients, board members, and academic peers. It is imperative that the amount of data generated be validated and utilized to drive well-intended and effective change. The public reporting of data is inevitable and can be of tremendous benefit if the data are accurate. This is the new reality that must be embraced. The actual methodology may not be the issue; rather, the dedication to really embrace accurate reporting and the expenditure to help surgeons reflect the hard work and effort they put into every patient's care must be supported by hospitals and practices. The substantial resources that are required for analysis must not be an excuse to drive risk-averse behavior to avoid negative scores, and there must not be misdirecting of funds away from clinical care and continuous improvement toward data mining and management. Physicians and surgeons need to be at the table in the journey toward transparency so that the data generated are accurate, validated, and patient centric; importantly, such an approach needs to provide an appropriate incentive for improvement in patient care and safety.

So, how can we as surgeons help to minimize what can be perceived as shame and blame? We have to sit at the table. We must not bury our heads in the sand. We need to help set standards at the national level as endorsed by the STS and the American Association of Thoracic Surgeons, and we need to be part of the solution. We need to take ownership of our data and the reporting of our data to ensure its accuracy and validity. We do not want to have things handed down to us. We need to be at the forefront and therefore propose measures we would be proud of as surgeons, and really use the STS database and mutually agreed upon meaningful measures. Doing so will signal our full endorsement of transparency and thereby inform the public that we accept and cultivate accountability. The time is now to hold ourselves accountable, as we have always done, and share our successes with our patients.

The authors wish to thank Darrell Gray, MD, Santino Cua, MS, Susan White, PhD, and Jennifer Hefner, PhD, MPH.

References

1. Rothberg MB, Morsi E, Benjamin EM, Pekow PS, Lindenauer PK. Choosing the best hospital: the limitations of public quality reporting. Health Aff (Millwood) 2008;27:1680–7.
2. Medicare program. Hospital inpatient prospective payment systems for acute care hospitals and the long-term care hospital prospective payment system and fiscal year 2015 rates; quality reporting requirements for specific providers; reasonable compensation equivalents for physician services in excluded hospitals and certain teaching hospitals; provider administrative appeals and judicial review; enforcement provisions for organ transplant centers; and electronic health record (EHR) incentive program. Final rule. Federal Register 2014;79:49853–50536.
3. Medicare. Hospital-acquired condition reduction program. Available at https://www.medicare.gov/hospitalcompare/hac-reduction-program.html. Accessed November 22, 2015.
4. Jha A. Penalizing hospitals for being unsafe. An ounce of evidence; health policy. Boston, MA: WordPress; 2014.
5. Vaz LE, Kleinman KP, Kawai AT, et al. Impact of medicare's hospital-acquired condition policy on infections in safety net and non-safety net hospitals. Infect Control Hosp Epidemiol 2015;36:649–55.
6. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff (Millwood) 2008;27:759–69.
7. Pressure ulcers are increasing among hospital patients. Available at http://archive.ahrq.gov/news/nn/nn120308.htm. Accessed January 13, 2016.
8. Moghissi ES, Korytkowski MT, DiNardo M, et al. American Association of Clinical Endocrinologists and American Diabetes Association consensus statement on inpatient glycemic control. Endocr Pract 2009;15:353–69.
9. Preventing central line-associated bloodstream infections: A global challenge, a global perspective. Available at http://www.jointcommission.org/assets/1/18/clabsi_monograph.pdf. Accessed January 13, 2015.
10. Austin JM, Jha AK, Romano PS, et al. National hospital ratings systems share few common scores and may generate confusion instead of clarity. Health Aff (Millwood) 2015;34:423–30.
11. Rau J. Hospital ratings are in the eye of the beholder. Kaiser Health News 2013. Available at http://khn.org/news/expanding-number-of-groups-offer-hospital-ratings/. Accessed January 13, 2016.
12. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ patient safety indicators (PSIs) on the Veterans Health Administration: the case of readmissions. Med Care 2013;51:37–44.
13. Encinosa WE, Hellinger FJ. The impact of medical errors on ninety-day costs and outcomes: an examination of surgical patients. Health Serv Res 2008;43:2067–85.
14. Downey JR, Hernandez-Boussard T, Banka G, Morton JM. Is patient safety improving? National trends in patient safety indicators: 1998–2007. Health Serv Res 2012;47:414–30.
15. Ramanathan R, Leavell P, Stockslager G, Mays C, Harvey D, Duane TM. Validity of Agency for Healthcare Research and Quality Patient Safety Indicators at an academic medical center. Am Surg 2013;79:578–82.
16. Nelson DW, Simianu VV, Bastawrous AL, et al. Thromboembolic complications and prophylaxis patterns in colorectal surgery. JAMA Surg 2015;150:712–20.
17. Burack JH, Impellizzeri P, Homel P, Cunningham JN. Public reporting of surgical mortality: a survey of New York State cardiothoracic surgeons. Ann Thorac Surg 1999;68:1195–202.
18. Hannan EL, Cozzens K, King SB, et al. The New York State cardiac registries: history, contributions, limitations and lessons for future efforts to assess and publically report healthcare outcomes. J Am Coll Cardiol 2012;59:2309–16.
19. Dehmer GJ, Drozda JP, Brindis RG, et al. Public reporting of clinical quality data. J Am Coll Cardiol 2014;63:1239–45.
20. Westaby S. Publishing individual surgeons' death rates prompts risk averse behavior. BMJ 2014;349:g5026.
21. Kandasamy N, Hardy B, Page L, et al. Cortisol shifts financial risk preferences. Proc Natl Acad Sci 2014;111:3608.
22. Westaby S, Baig K, Pepper J. Publishing surgeon specific mortality data: the risks outweigh the benefits. Bull R Coll Surg Engl 2015;97:155–60.
23. Francis R. Report of the mid Staffordshire NHS Foundation Trust public enquiry 2013. Available at www.midstaffspublicenquiry.com/report. Accessed January 13, 2016.
24. National Advisory Group on the Safety of Patients in England. A promise to learn—a commitment to act: improving the safety of patients in England 2013. Available at

www.england.nhs.uk/tag./berwick-report. Accessed January 13, 2016.
25. Westaby S, Baig K, De Silva R, et al. Recruitment to UK cardiothoracic surgery in the era of public outcome reporting. Eur J Cardiothoracic Surg 2015;47:679–83.
26. Berwick DM. A primer on leading the improvement of systems. BMJ 1996;275:877–8.
27. Jacobs JP, Jacobs ML. Transparency and public reporting of pediatric and congenital heart surgery outcomes in North America. World J Pediatr Congenit Heart Surg 2016;7:49–53.
28. Badhwar V, Rankin JS, He X, et al. The Society of Thoracic Surgeons mitral repair/replacement composite score: a report of The Society of Thoracic Surgeons Quality Measurement Task Force. Ann Thorac Surg 2015 Dec 28; [E-Pub ahead of print].
29. Kozower BD, O'Brien SM, Kosinski AS, et al. The Society of Thoracic Surgeons composite score for rating program performance for lobectomy for lung cancer. Ann Thorac Surg 2016 Jan 16; [E-Pub ahead of print].
30. Winkley Shroyer AL, Bakaeen F, Shahian DM, et al. The Society of Thoracic Surgeons adult cardiac surgery database: the driving force for improvement in cardiac surgery. Semin Thorac Cardiovasc Surg 2015;27:144–51.
31. Shahian DM, Grover FL, Prager RL, et al. The Society of Thoracic Surgeons voluntary public reporting initiative: the first 4 years. Ann Surg 2015;262:526–35.