REVIEWS

Quality improvement through public reporting: The surgeon scorecard – are we there yet?

Afshin A. Anoushiravani[1], Zain Sayeed[2,3,4,5], Muhammad T. Padela[*2,3,4,5], James E. Feng[1], Paul Barach[6], Mouhanad El-Othmani[3], Hussein F. Darwiche[2,3], Khaled J. Saleh[5]

[1]Division of Adult Reconstructive Surgery, NYU Langone Orthopedics, NYU Langone Health, New York, United States

[2]Resident Research Partnership, Detroit, MI, United States

[3]Division of Orthopaedic Surgery and Sports Medicine, Detroit Medical Center, Detroit, MI, United States

[4]Department of Orthopaedic Surgery, Chicago Medical School, Rosalind Franklin University, North Chicago, IL, United States

[5]Michigan Musculoskeletal Institute, Madison Heights, MI, United States

[6]Wayne State University, School of Medicine, Detroit, MI, United States

ABSTRACT

As national healthcare reform continues to place greater emphasis on providing high value care, measures designed to track clinical performance remain relatively overlooked. To that extent, several organizations have attempted to create objective grading systems to evaluate orthopaedic surgeon quality and performance. While attempting to address these issues, ProPublica's Surgeon Scorecard has provoked national debate among patient advocates and healthcare providers. The methodology behind the Scorecard was developed at the Harvard School of Public Health with an aim to provide a more robust means of comparing surgical performance and outcomes for patients and healthcare organizations. Currently, the Scorecard assesses eight elective surgical procedures, including total knee and hip arthroplasty, through the use of the Medicare Claims Dataset. The impact of the Scorecard on orthopaedic practice has yet to be established. In this discussion, we analyze the Scorecard from the perspective of various stakeholders to identify its benefits and shortcomings, as well as offer direction for further improvement.

Key Words: Surgeon Scorecard, Quality, Value, Medicare, HealthGrades, Arthroplasty, Performance measurement

1. INTRODUCTION

As healthcare delivery transitions away from a volume-based reimbursement model to an efficiency-and-quality-oriented model, the balance between high quality, evidence-based and value-based care has posed a challenge for surgeons and healthcare systems. By 2024, Americans will expend an estimated $5.4 trillion on healthcare; a substantial proportion of which will be spent on indispensable surgical procedures such as total joint arthroplasty, spinal fusion, and coronary artery bypass grafting.[1] The Centers for Medicare and Medicaid Services (CMS) responded by developing innovative reimbursement models (e.g., Bundle Payment for Care Improvement [BPCI] and Comprehensive Care of Joint Replacement [CJR]) aimed at controlling healthcare costs while rewarding highly reliable, quality-driven surgeons.[2] With provider reimbursement becoming dependent on postoperative outcomes, public reporting instruments, such as ProPublica's Surgeon Scorecard, have provoked international

debate among patient advocates and healthcare providers.[3] Public and transparent reporting of surgeon performance further engenders the controversy.

Our aim is to provide a historical review of public reporting while also providing a critique of the Surgeon Scorecard. In doing so, we hope to better elucidate the various stakeholder perspectives of the Surgeon Scorecard, highlight its strengths thus far, and identify future areas of improvement for this powerful decision-making tool. We conclude by briefly describing current methodologies used to integrate clinical registries and electronic medical records (EMRs) with reporting instruments to effectively evaluate surgical performance.

1.1 A brief history of publically reporting instruments

Ernest Codman, a 19th-century surgeon and early advocate for assessing clinical outcomes, lobbied the American College of Surgeons to establish an examining committee to enhance hospital efficiency and clinical outcomes.[4] Dr. Codman is better known for his development of the "End Result Idea"; the belief that healthcare practitioners should sufficiently follow all patients under their care, evaluate treatment success, and more importantly, identify treatment failures and how to avoid them. This radical ideology became the foundation for the hospital standardization movement advanced by the American College of Surgeons in 1917, which evolved into, and now is commonly recognized as, the Joint Commission on Accreditation of Hospitals in 1951.[5, 6]

In the 1980s, the Health Care Financing Administration (HCFA), the precursor to the CMS, began publishing the death rates of various hospitals countrywide. The effort was highly controversial and ultimately abandoned due to differences in hospital population demographics which biased comparisons, and was severely pushback by virtually all professional medical societies. In 2006, the Tax Relief and Health Care Act established the Physician Quality Reporting System (PQRS), an initiative that developed a standardized approach to measuring healthcare and physician quality.[7] The PQRS was designed to promote physician accountability for patient outcomes across the entire patient care continuum.[8, 9] However, healthcare organizations are not obligated to participate in the PQRS initiative until the Medicare Access and Children's Health Insurance Program Reauthorization Act (MACRA) of 2015 takes effect in 2019. As part of the Reauthorization Act, all providers will be required to participate in one of two reimbursement models:[10–12]

1) Merit-Based Incentive Payment System (MIPS).

2) An approved alternative payment model (APM).

The MIPS consolidates four performance-indices into a sin-

gle aggregate score, of which the PQRS-equivalent comprises 60% of the total score during the first year of implementation.[10–12] Additionally, all alternative payment models will be required to integrate PQRS metrics into eligible payment arrangements (e.g., BPCI and CJR) on a negotiated basis. Essentially, physician performance will be scored, guiding financial reimbursement and increasing a clinician's professional transparency.

1.2 Physician rating applications

In 2015, CMS launched Physician Compare, a website that links PQRS outcomes and patient experiences to individual providers.[13] The purpose of the website was to provide patients with the information needed make informed decisions regarding patient care.[14] Before Physician Compare, consumers relied on independent, non-validated review and rating websites (see Table 1).[15] However, with the widespread integration of EMRs, it is now economically feasible to collect and disperse large volumes of healthcare data.

Table 1. Instruments for publicly reporting surgeon performance

Instrument and year released	Data Source
HealthGrades, 1998	More than 500 million entries from • Medicare Claims Data • Private Databases • Patient reviews
RateMD.com, 2004	User-submitted reviews
Vitals.com, 2007	Over 1.4 billion entries from • Patient Reviews • Physician publications • State and Federal Medical Databases • Private Databases
Zocdoc.com, 2007	User-submitted reviews
Surgeon Scorecard, 2015	Medicare Claims Dataset

Note. Table 1 shows various instruments used to rate physicians on their performance

To address the demand for transparency within healthcare, a new industry of for-profit organizations provided consumers with information necessary to make informed medical decisions.[15] Corporations such as HealthGrades aggregates and consolidates data from private and public databases, as well as voluntary patient reviews. Patient data is stratified by healthcare organization (HCO) and provider codes obtained from the CMS Medicare Claims Dataset. An undisclosed algorithm is then applied to derive a proprietary, non-validated star-rating system. Moreover, HealthGrades provides ratings on millions of clinicians. As patient participation is voluntary, many physician entries may be empty or contain fewer than 10 entries, making them highly susceptible to outlier patient-submitted ratings.[16] Anecdotally, physician ratings

may also be biased for overly harsh reviews as patients are more likely to submit negative reviews when afflicted by suboptimal outcomes than they are by satisfactory outcomes. Conversely, some practices may focus on these metrics and "cherry-pick" patients to submit reviews through third party electronic perioperative applications.

2. THE SURGEON SCORECARD

ProPublica, a non-profit investigative organization, surpasses the shortcomings of HealthGrades by providing an easy-to-use instrument assessing surgeons who perform common yet costly elective procedures. ProPublica's Surgeon Scorecard, released in July 2015, contains profiles on more than 16,000 surgeons conducting one of eight elective procedures.[17] The methodology behind the Scorecard was developed at the Harvard School of Public Health and was intended to provide patients and HCOs with a reliable means of evaluating and comparing surgical performance and outcomes.[17]

The Surgeon Scorecard avoids many of the errors made in other public reporting instruments by focusing on a limited number of procedures (see Table 2).[18] The data used to evaluate surgeons was extracted from the Medicare Claims Dataset, a reputable and objective administrative database that has proven its reliability over past decade for quality metrics. Additionally, the reporting tool developers avoid rating all surgeons, focusing instead on surgeons who have performed a minimum of 20 procedures.

Table 2. The eight elective surgical procedures included in the Surgeon Scorecard

Procedures
1.   Total Knee Arthroplasty
2.   Total Hip Arthroplasty
3.   Laparoscopic Cholecystectomy
4.   Lumbar and Lumbosacral Arthrodesis of the Posterior Columns
5.   Lumbar and Lumbosacral Arthrodesis of the Anterior Columns
6.   Radical Prostatectomy
7.   Transurethral Prostatectomy
8.   Cervical Arthrodesis of the Anterior Column

*N*ote. Table 2 lists the procedures used by the Surgeon Scorecard to evaluate surgeons

2.1 Surgeon Scorecard methodology

ProPublica acquired the Medicare billing data for in-patient stays from 2009 through 2013. Unlike previous public reporting performance instruments, the Scorecard excluded high-risk surgical candidates, patients admitted through the emergency department, transferred from an outside medical facility, or presented with trauma-related injuries.[19] Thus,

only patients undergoing elective surgeries and meeting strict inclusion and exclusion criteria were used to evaluate surgical performance.

Before rating a surgeon's performance, ProPublica uses a statistical method to adjust for patient age, health, and hospital performance. The mixed-effects statistical modeling adjusts for inter-surgeon case variability. Additionally, Health Scores were calculated to control for variability in comorbidity profiles among patient samples. The Health Score is a continuous coefficient developed using the Van Walraven technique, a modification of the Elixhauser Comorbidity Index.[20] This Index is a unique risk-stratification tool incorporating comorbidities proven to affect clinical outcomes and resources use. Thus, unlike its predecessors, the Surgeon Scorecard accounts for an unprecedented degree of variance among patients and hospitals and is much more reliable as a public-reporting instrument.

2.2 Stakeholder perspectives

2.2.1 *The patient perspective*

The impact of stakeholder transparency on the quality of medical care must be examined through various stakeholder perspectives (see Figure 1). The Surgeon Scorecard represents an objective instrument which patients can review prior to selecting a surgeon. Unlike other physician rating applications, the Surgeon Scorecard is user-friendly and relatively unbiased. However, only 14% of Americans use online information when choosing a surgeon, hospital, and healthcare plan.[16] Such low percentage may reflect the inconsistent formatting and unreliable software often used in public reporting instruments. A 2007 study evaluating five hospital reporting services found that available physician rating applications evaluated different variables and lacked standardization, thus limiting their utility within a large and diverse healthcare ecosystem.

Currently, the Surgeon Scorecard assigns a single adjusted complication rate for each surgeon. In-hospital mortality and 30-day readmissions are the metrics that determine the Surgeon Scorecard's complication coefficient. Although combining these variables simplifies the statistical analysis, they are not equivalent, this can be misleading and they should arguably be evaluated independently. Each variable can then be individually weighed and used to calculate a composite score representing measurable postoperative complications, providing a more prospective patients a more accurate and granular perspective. Such an approach can incorporate additional variables, including 90-day readmissions and patient-reported outcomes (PROs).
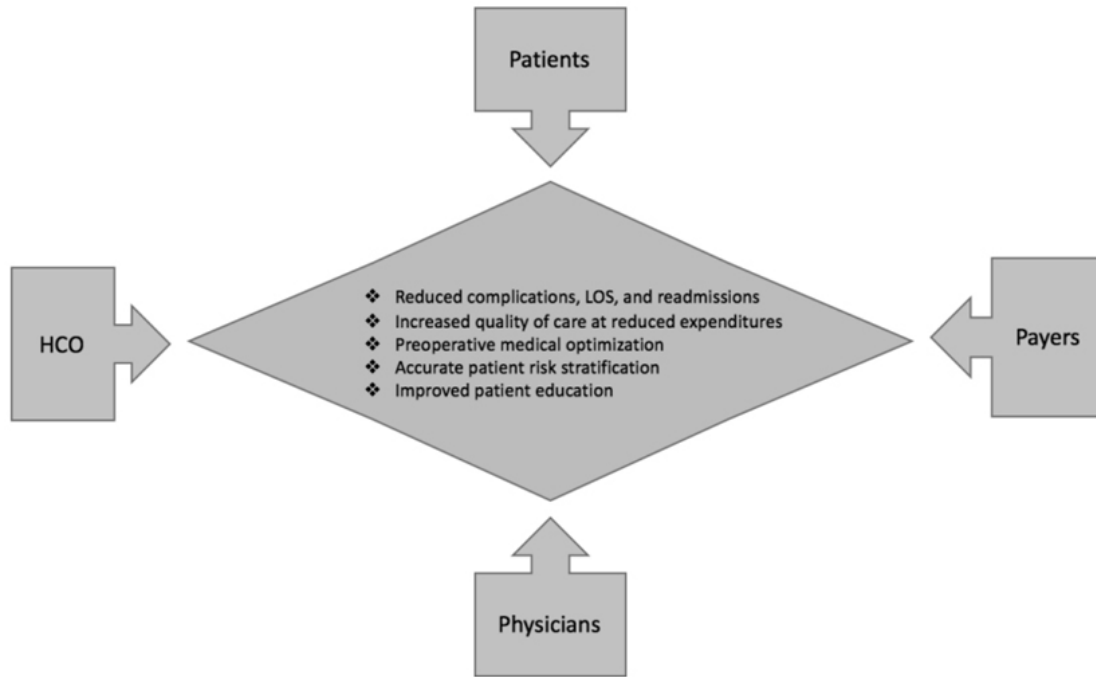
Figure 1. Impact of improved stakeholder transparency on the quality of medical care

2.2.2 *The surgeon perspective*

The Surgeon Scorecard is one of the few public-reporting instruments that does not use subjective patient reviews but instead relies entirely on the Medicare Claims Dataset, providing surgeons with detailed and tangible performance evaluations. The type of information derived from the Medicare Claims Dataset may offer another opportunity to improve the Surgeon Scorecard. The Dataset includes only Medicare beneficiaries, most of whom are more than 65 years old. Therefore, scores may not be based on patients that are representative of a surgeon's practice. Additionally, the administrative nature of the Dataset favors documenting procedures and diagnoses that affect billing rather than outcomes. Consequently, many chronic disorders not managed during the patient's hospitalization may be underreported due to the Medicare Claims Datasets focus on procedures and billing, rather than clinical diagnoses. Both of these components of the Surgeon Scorecard may be improved by newer coding practices implemented in 2015.

The Dataset uses International Classification of Disease (ICD) codes, a complex series of thousands of 3-to-7-digit alphanumeric codes representing diagnoses and procedures.[11] Currently, the vast majority of data comprising the Medicare Claims Dataset utilizes some 17,000 ICD-9 codes, which do not indicate disease severity or progression. Since October 2015, 140,000 ICD-10 codes have been applied nationally, expanding the capabilities for public reporting instruments. However, it will take years to amass enough data to power

a physician rating application such as the Surgeon Scorecard. In addition, the Surgeon Scorecard's reliability can be enhanced if an algorithm is created and validated that linking patients, procedures and healthcare providers. The RAND Corporation reported that more than 60 physicians were linked to procedures out of their scope of practice.[21] Additionally, a 2012 study sponsored by the CMS, reported that the National Provider Identifier of surgeons had been incorrectly associated with Medicare Part A and Part B Claims data in more than 28% of surgical procedures.[22] Although databases are extremely helpful in population health initiatives and quality audits, large administrative databases have inherent limitations which must be appreciated.

The Surgeon Scorecard's developers use well accepted methods to calculate the surgeons' Health Scores. A Health Score coefficient of zero suggests that the prevalence or mix of comorbidities did not vary in the studied population.[19, 21] That said, the Van Walraven modification of the Elixhauser Comorbidity Index may bias the interpretation of surgical outcomes. The method combines 30 individual Elixhauser metrics into a single de-identified composite score. We recommend that ProPublica consider each comorbidity associated with each patient as a unique clinical characteristic that may affect outcomes.[23–27] Public reporting instruments must adequately adjust for baseline characteristics (e.g. age, sex, and comorbidity profiles) in a population. Failure to do so may skew reporting behavior among surgeons.[23–27]

Additionally, the Surgeon Scorecard uses an adjusted com-

plication rate. Although the adjusted complication rate is visually appealing and relatively easy to understand for a given stakeholder, the metric should be validated before being incorporated. Ban and colleagues[28] compared outcomes reported by the Surgeon Scorecard and its proprietary adjusted complications rate for all eight elective procedures with the National Surgical Quality Improvement Program (NSQIP) dataset. After the analysis, the investigators reported that the Surgeon Scorecard's strict criteria excluded 82% of patients, ranging from 42% of patients receiving total knee arthroplasties, to 96% of those undergoing laparoscopic cholecystectomies.[28] Additionally, 84% of the complications associated with the eight elective procedures occurred during hospitalization substantially diluting the number of episodes of care used to power the Surgeon Scorecard.[28] Finally, the authors noted that the Surgeon Scorecard examines only inpatient encounters. This limitation is particularly concerning because many of the episodes of care included in the Surgeon Scorecard are usually done as an outpatient or short-stay procedure. More complex higher morbidity candidates are therefore undergoing surgery as an inpatient further skewing postoperative outcomes.

2.2.3 *The perspective of healthcare organizations*

The Surgeon Scorecard was developed to improve transparency in the healthcare system, particularly among patients and their physicians. Although the Scorecard does not directly affect HCOs, the public's perception of individual providers and their respective HCOs is largely determined by their reputation. Thus, as HCOs continue to merge and acquire physician practices, it is essential that HCOs evaluate the Scorecard as it relates to: 1) value-based care initiatives, 2) the needs of their patients, and 3) the effects of the instrument on the HCOs long-term performance.[29, 30]

Journalists at ProPublica used random-effects methodology to account for variation amongst surgeons and hospitals. In every instance, ProPublica suggested that hospitals are relatively similar and that the bulk of the difference in the adjusted complication rate was more related to the surgeon's performance than to the hospital's performance.[19] Although ProPublica has gone to great lengths to standardize and adjust for variation between hospitals, the Scorecard may nevertheless be overlooking hospital-to-hospital variation.[21, 31] Patients and HCOs must appreciate that a surgeon is limited by the policies and protocols of the HCOs. In addition, successful surgical management of patients requires dozens of treatment protocols (e.g., disposition decisions; infection control) and teams of dedicated staff members (e.g. anesthesia, nursing, physical therapy, surgical, etc.).[21] The standard-of-care implemented by HCOs is embedded in their culture and affects the quality and value of care. Therefore, we recom-

mend that ProPublica consider establishing a more detailed method that relies on mixed methods including qualitative methods to assess an HCO's inherent standard and culture of care for given procedures and to account for variations among protocols.

2.2.4 *The payers perspective*

The interests of payers, in general, align well with those of HCOs because both are trying to provide the best services at the lowest cost. Conversely, patients and physicians are interested in receiving or administering the highest quality of care, whereas payers place a greater importance on the value of care received. ProPublica's Scorecard attempts to objectively evaluate surgeons; however, it does not adjust for the cost of care. In an era of rapidly increasing healthcare expenditures and budget deficits, it is essential that public reporting instruments that attempt to evaluate postoperative outcomes also account for the value of care delivered. Integrating PQRS and PROMs into the Surgeon Scorecard will allow stakeholders to evaluate outcomes from a value-based perspective.

The Scorecard holds potential benefit for all stakeholders because of the increased transparency fostered by public reporting instruments. However, before its widespread implementation, higher-quality data must be used. As suggested by Barach and Lipshultz, public reporting of such data must aim to improve patient safety while avoiding risk-averse environments and providers.[32] They list 11 key principles for improving data collection and public reporting[32] We briefly assesses the clinical registries available to ProPublica should they wish to better evaluate surgical performance.

3. CLINICAL REGISTRIES

Over the past two decades, collecting clinical outcome data has been promoted by various medical societies, including the American Academy of Orthopedic Surgeons.[33] In 1994, the Academy established the Musculoskeletal Outcomes Data Evaluation and Management System (MODEMS), one of the first CDRs collecting functional patient information. Unfortunately, the program was terminated in 1999 due to limited funding and enrollment.[34] Since then, medical societies, including the Academy, have struggled to develop disease-specific databases. Now, more than ever, surgeons must develop standardized disease-specific Qualified Clinical Data Registries (QCDRs) to efficiently collect high-quality patient information to improve transparency and value-based care.[35]

3.1 State and organizational joint registries

In the United States, Virginia, Michigan, and California have developed state-based total joint arthroplasty registries (see

Table 3). The Virginia Joint Registry, established in 2005, is one of the first state-based registries in the United States. It was developed entirely by orthopedic surgeons who valued the importance of monitoring clinical outcomes after total joint arthroplasty.[36] In 2011, a consortium of healthcare providers and insurers developed the Michigan Arthroplasty Registry Collaborative Quality Initiative with the intent of reducing total joint arthroplasty complication and revision rates. Presently, it has documented 73,000 arthroplasty cases at 51 hospitals.[37] The California Joint Replacement Registry was established in 2010, and collects data on more than one-third of the total joint arthroplasty procedures done in the state. The California Registry is also the first total joint arthroplasty registry in the United States to focus on patient-reported outcomes.[38]

Table 3. State-based and institutional joint registries in the United States

Registry and year established	Procedures included	Region
Kaiser Permanente, 2001	TKA, THA, ACLR, hip fracture, TSA, Spine	Predominately West coast & Mid-Atlantic
Virginia Joint Registry, 2005	TKA, THA	Virginia
California Joint Replacement Registry, 2010	TKA, THA	California
Michigan Arthroplasty Registry Collaborative Quality Initiative, 2011	TKA, THA	Michigan
American Joint Replacement Registry, 2011	TKA, THA	National

Note. Table 3 lists different joint registries already in place throughout the nation; TKA = total knee arthroplasty, THA = total hip arthroplasty, ACLR = anterior cruciate ligament reconstruction, TSA = total shoulder arthroplasty

In addition to the state-based and administrative databases, many institutions have established total joint arthroplasty registries for patients treated at their facilities. Kaiser Permanente has established its own orthopaedic CDR using the Swedish Joint Registry as its model (see Table 3). Since 2001, data for more than 192,000 patients receiving total joint arthroplasties, hip fracture care, anterior cruciate ligament reconstructions, and spine procedures have been entered into the registry.[39] Additionally, Kaiser Permanente has begun collaborating with the Norwegian Arthroplasty Registry and serves as a model for inter-registry collaboration.[40]

3.2 The American Joint Replacement Registry

The American Joint Replacement Registry (AJRR) was established in 2011, to standardize and consolidate orthopaedic CDRs. Currently, more than 500 hospitals participate, with about 85% of all arthroplasty cases entered into the registry.[41] Currently, the AJRR is developing methods to expand its data collection capabilities to capture complications, patient-reported outcomes, and radiographic findings (see Table 4). Although the AJRR is the largest and most comprehensive joint registry in the United States, it is limited by its dependence on ICD 9 codes for identifying procedures and comorbidities. Moreover, there is no standardized data-reporting algorithm for level-three (patient-reported outcomes) or level-four (radiographic) data.

To overcome these challenges, surgeons, as well as ProPublica journalists, might study successful disease-specific registries. For example, the Duke Databank for Cardiovascular Disease[42] is the largest cardiovascular database of prospectively collected data in the world. It is considered by many to be the reference standard for clinical data registries because of its size, high-level of organization, and inclusion of long-term follow-up data.[42] These features have allowed investigators to evaluate cardiovascular disease on a large scale. The unique capabilities of these model registries should be considered for incorporation in next-generation surgical registries, such as the AJRR.

3.3 Integrating registries into reporting instruments with use of EMR

To standardize QCDRs that can be used as reporting instruments, surgeons must take the initiative and participate in developing a uniform approach to data collection. As the CMS begins to penalize providers for not complying with reporting requirements, the volume of data available to clinicians will increase exponentially. Moreover, orthopaedic surgeons must follow evidence-based recommendations (see Table 5). Implementing these recommendations will enable orthopaedic surgeons to evaluate surgical performance through clinical outcomes and resource utilization while accounting for unique patient and HCO variables. Additionally, proper implementation of these recommendations will enable clinicians to link multiple disease-specific CDRs to assess the effects of chronic medical conditions on surgical outcomes (e.g., the impact of systemic lupus erythematosus on total hip arthroplasty).

Table 4. Data collected in the American Joint Replacement Registry (Adapted from the AJRR website[41])

Evidence Level	Description	Variables
I	Patient-related data	• Name (first, last) • Date of birth • Social security number • Diagnosis (international classification of disease 10th edition codes) • Sex • Ethnicity • Home address • Hospital related data • National provider identification • Procedure related codes • Implant catalog number
II	Patient risk factors and comorbidities (International Classification of Disease 10th Edition codes)	• American Society of Anesthesiologist Score, operative and postoperative complications • American Joint Replacement Registry risk adjustment analysis
III	Patient-reported outcomes	• Veterans Rand-12/Short Form 12 • Patient Reported Outcomes Measurement Information System-10 • Global Health • Hip Disability and Osteoarthritis Outcome Scores • Modified Western Ontario and McMaster Universities Arthritis Index • Knee Injury and Osteoarthritis Outcomes Score • Oxford Hip and Knee Scores • Knee Society Scoring System • Harris Hip Scores • EuroQol-5D
IV	Radiographs	

Note. Table 4 describes the data that is collected within joint registries

Table 5. Recommendations for developing clinical data registries

Registry Objectives	Description
Disease-specific Registries	• Registries should contain disease-based variables linking clinical outcomes to institutional treatment protocols • Clinicians should take advantage of robust registries currently available (e.g., Duke Databank Cardiovascular Disease; American Joint Replacement Registry)
Patient Identifier	• All patients enrolled in registries should be linked to a single identifier to ensure full traceability (e.g., Social Security Number, National Health Identifier)
Data Organization	• At a minimum, data should be categorized into six groups: • Patient demographics • Disease severity • Treatment protocols • Provider profile • Treatment setting • Resource use
Standardization	• Variables collected during surgical procedures must be measured and assessed in a consistent and standardized manner
Heterogeneity	• Heterogeneity within datasets allows clinicians to examine the effects of different patient specific variables on outcomes • Heterogeneity due to variance in measurement instruments should be minimized
Continuum of care	• Registries should collect longitudinal data reflecting the continuum in care
Electronic Medical Records	• It is essential that Electronic Medical Records be interoperable and locally programmable − enabling them to integrate with databases

Note. Table 5 lists various recommendations that, if implemented, would significantly improve current available joint registries

Table 6. Strengths and weaknesses of the Surgeon Scorecard

Strengths	Weaknesses
Exclusion of trauma patients	Only includes Medicare patients
Exclusion of subjective patient reviews	Excludes 82% of patient entries[28]
Emphasis on eight elective procedures	Does not evaluate in-hospital complications, missing 84% of postoperative[28] complications
	Outpatient procedures not evaluated
	Limited to 30-day readmissions

Note. Table 6 is a summary of strengths and weakness of the Surgeon Scorecard

The CDR's value in evaluating performance is not possible without defining the role of EMRs. Integrating data into disease-specific CDRs should allow millions of data points to be reported at a minimal cost. Thus, it is essential that orthopedic surgeons continuously optimize their EMR systems to improve data documentation capabilities.

4. CONCLUSIONS

ProPublica should be commended for its efforts in improving transparency in healthcare with the Surgeon Scorecard. A summary of its strengths and weaknesses is provided in Table 6. Moreover, it is the first public reporting instrument to use only the Claims Dataset to objectively evaluate the performance of thousands of surgeons nationally. The Scorecard is proof that a public reporting tool may be effectively integrated with QCDRs, potentially providing investigators

and clinicians with the data necessary to power high-quality, population-based studies. However, before achieving such a potential, it is critical that surgeons take the initiative by:

1) Consolidating and standardizing all QCDRs;

2) Using interoperable EMRs, to efficiently organize big data outputs into QCDRs;

3) Coordinating the activities of surgical societies, including the AAOS and ACS, to develop a set of evidence-based guidelines for developing public reporting instruments.

Only after such measures have been taken may surgeons, patients, HCOs, and payers be able to use high-quality data for evaluating surgical competence.

CONFLICTS OF INTEREST DISCLOSURE
The authors declare they have no conflicts of interest.

REFERENCES

[1] Keehan SP, Cuckler GA, Sisko AM, et al. National Health Expenditure Projections, 2014-24: Spending Growth Faster Than Recent Trends. Health affairs. 2015; 34(8): 1407-17. PMid: 26220668. https://doi.org/10.1377/hlthaff.2015.0600

[2] Sanchez JA, Barach PR. High reliability organizations and surgical microsystems: re-engineering surgical care. The Surgical clinics of North America. 2012; 92(1): 1-14. PMid: 22269256. https://doi.org/10.1016/j.suc.2011.12.005

[3] Walker K, Neuburger J, Groene O, et al. Public reporting of surgeon outcomes: low numbers of procedures lead to false complacency. Lancet (London, England). 2013; 382(9905): 1674-7. https://doi.org/10.1016/S0140-6736(13)61491-9

[4] Surgeons ACo. The 1919 "Minimum Standard" document. 2006 [cited 2016 July 20]. Available from: https://www.facs.org/about%20acs/archives/past highlights/minimumhighlight

[5] What is Interoperability?: Health Information and Management Systems Society. 2015 [cited 2016]. Available from: http://www.himss.org/library/interoperability-sta ndards/what-is-interoperability

[6] Neuhauser D. Ernest Amory Codman MD. Quality & Safety in Health Care. 2002; 11(1): 104-5. https://doi.org/10.1136/qhc.11 .1.104

[7] H.R.6111. Tax Relief and Healthcare Act of 2006. P.L. 109-432 D.

[8] Bhowmick K, Varghese VD. Retrograde Intramedullary Nailing for Recurrent Fracture in Congenital Pseudarthrosis of the Tibia. The Journal of Foot and Ankle Surgery. 2015. PMid: 26342666.

[9] CentersforMedicareandMedicaidServices. Physician Quality Reporting System. Available from: www.cms.gov/Medicare/Qua lity-Initiatives-Patient-Assessment-Instruments/P QRS/index.html?redirect=/pqrs

[10] Rosenthal MB. Physician Payment after the SGR–The New Meritocracy. The New England Journal of Medicine. 2015; 373(13): 1187-9. PMid: 26398068. https://doi.org/10.1056/NEJMp1507757

[11] Doherty RB. Goodbye, Sustainable Growth Rate-Hello, Merit-Based Incentive Payment System. Ann Intern Med. 2015; 163(2): 138-9. PMid: 25961185. https://doi.org/10.7326/M15-0992

[12] Bumpass DBSJ. Which path will you take: how SGR repeal will change quality reporting and payment models. AAOSNow. 2016.

[13] Physician Compare: CMS. 2015 [cited 2016 July 24]. Available from: https://www.medicare.gov/physiciancompare/sear ch.html

[14] AMA. 2015 Physician Quality Reporting System. 2015 [cited 2016 July 24]. Available from: http://www.ama-assn.org/ama/pub/ physician-resources/clinical-practice-improvement /clinical-quality/physician-quality-reporting-sys tem.page

[15] Marshall MN, Shekelle PG, Davies HT, et al. Public reporting on quality in the United States and the United Kingdom. Health Affairs. 2003; 22(3): 134-48. PMid: 12757278. https://doi.org/10.1377/hlthaff.22.3.134

[16] Findlay S. Physician Compare. Under the Affordable Care Act, the federal government must post information about physician performance and quality of care on a public website. Will consumers use it? Health Affairs. 2014.

[17] Jha A, Pronovost P. Toward a Safer Health Care System: The Critical Need to Improve Measurement. JAMA. 2016; 315(17): 1831-2. PMid: 27077647. https://doi.org/10.1001/jama.2016.3448

[18] Hofer JK, Ezzet KA. A minimum 5-year follow-up of an oxidized zirconium femoral prosthesis used for total knee arthroplasty. Knee. 2014; 21(1): 168-71. PMid: 24161449. https://doi.org/10.1016/j.knee.2013.08.015

[19] Olga P, Marshall A. Assessing surgeon-level risk of patient harm during elective surgery for public reporting. 2015. Available from: https://static.propublica.org/projects/patient-safety/methodology/surgeon-level-risk-methodology.pdf

[20] van Walraven C, Austin PC, Jennings A, et al. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care. 2009; 47(6): 626-33. PMid: 19433995. https://doi.org/10.1097/MLR.0b013e31819432e5

[21] Friedberg MW PP, Shahian DM, Safran DG, et al. A Methodological Critique of the ProPublica Surgeon Scorecard. Rand Corporation. 2015. https://doi.org/10.7249/PE170

[22] Dowd B KR, Parashuram S, Swenson T, et al. Alternative approaches to measuring physician resource use: Final report. 2012 April 9.

[23] Shahian DM, He X, Jacobs JP, et al. Issues in quality measurement: target population, risk adjustment, and ratings. Ann Thorac Surg. 2013; 96(2): 718-26. PMid: 23816415. https://doi.org/10.1016/j.athoracsur.2013.03.029

[24] Shahian DM, Normand SL. Comparison of "risk-adjusted" hospital outcomes. Circulation. 2008; 117(15): 1955-63. PMid: 18391106. https://doi.org/10.1161/CIRCULATIONAHA.107.747873

[25] Toy JO, Basques BA, Grauer JN. Morbidity, mortality, and readmission after vertebral augmentation: analysis of 850 patients from the American College of Surgeons National Surgical Quality Improvement Program database. Spine (Phila Pa 1976). 2014; 39(23): 1943-9. PMid: 25188603. https://doi.org/10.1097/BRS.0000000000000563

[26] Bilimoria KY, Liu Y, Paruch JL, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. J Am Coll Surg. 2013; 217(5): 833-42. PMid: 24055383. https://doi.org/10.1016/j.jamcollsurg.2013.07.385

[27] Cohen ME, Ko CY, Bilimoria KY, et al. Optimizing ACS NSQIP modeling for evaluation of surgical quality and risk: patient risk adjustment, procedure mix adjustment, shrinkage adjustment, and surgical focus. J Am Coll Surg. 2013; 217(2): 336-46. PMid: 23628227. https://doi.org/10.1016/j.jamcollsurg.2013.02.027

[28] Ban KA, Cohen ME, Ko CY, et al. Evaluation of the ProPublica Surgeon Scorecard "Adjusted Complication Rate" Measure Specifications. Ann Surg. 2016.

[29] Franklin PD, Lewallen D, Bozic K, et al. Implementation of patient-reported outcome measures in U.S. Total joint replacement registries: rationale, status, and plans. J Bone Joint Surg Am. 2014; 96(Suppl 1): 104-9. PMid: 25520425. https://doi.org/10.2106/JBJS.N.00328

[30] Bell R, Krivich MJ, Boyd MS. Charting patient satisfaction. Mark Health Serv. 1997; 17(2): 22-9. PMid: 10170284.

[31] Austin PC, Alter DA, Tu JV. The use of fixed- and random-effects models for classifying hospitals as mortality outliers: a Monte Carlo assessment. Medical Decision Making. 2003; 23(6): 526-39. PMid: 14672113. https://doi.org/10.1177/0272989X03258443

[32] Barach PR, Lipshultz SE. The benefits and hazards of publicly reporting quality outcomes. Progress in Pediatric Cardiology. 2016; 42: 45-9. https://doi.org/10.1016/j.ppedcard.2016.06.001

[33] Brinker MR, O'Connor DP. Stakeholders in outcome measures: review from a clinical perspective. Clin Orthop Relat Res. 2013; 471(11): 3426-36. PMid: 24072624. https://doi.org/10.1007/s11999-013-3265-3

[34] Bulletin A. MODEMS program terminated. 2000.

[35] Endorsement NQFMECaGfEMf. National Quality Forum. Available from: http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=79434

[36] Kalra GD, Agarwal A. Experience with free fibula transfer with screw fixation as a primary modality of treatment for congenital pseudarthosis of tibia in children - Series of 26 cases. Indian journal of plastic surgery : official publication of the Association of Plastic Surgeons of India. 2012; 45(3): 468-77. PMid: 23450675. https://doi.org/10.4103/0970-0358.105951

[37] Dy CJ, Bumpass DB, Makhni EC, et al. The Evolving Role of Clinical Registries: Existing Practices and Opportunities for Orthopaedic Surgeons. J Bone Joint Surg Am. 2016; 98(2): e7. PMid: 26791040. https://doi.org/10.2106/JBJS.O.00494

[38] Chenok K, Teleki S, SooHoo NF, et al. Collecting Patient-Reported Outcomes: Lessons from the California Joint Replacement Registry. EGEMS (Wash DC). 2015; 3(1): 1196. https://doi.org/10.13063/2327-9214.1196

[39] Paxton EW, Kiley ML, Love R, et al. Kaiser Permanente implant registries benefit patient safety, quality improvement, cost-effectiveness. Jt Comm J Qual Patient Saf. 2013; 39(6): 246-52. https://doi.org/10.1016/S1553-7250(13)39033-3

[40] Paxton EW, Furnes O, Namba RS, et al. Comparison of the Norwegian knee arthroplasty register and a United States arthroplasty registry. J Bone Joint Surg Am. 2011; 93(Suppl 3): 20-30. PMid: 22262419. https://doi.org/10.2106/JBJS.K.01045

[41] American Joint Replacement Registry 2016. Available from: http://www.ajrr.net

[42] Bozic KJ, Lau E, Kurtz S, et al. Patient-related risk factors for postoperative mortality and periprosthetic joint infection in medicare patients undergoing TKA. Clin Orthop Relat Res. 2012; 470(1): 130-7. PMid: 21874391. https://doi.org/10.1007/s11999-011-2043-3