Surgery 165 (2019) 985–989

Contents lists available at ScienceDirect

Surgery

journal homepage: www.elsevier.com/locate/surg

Defying public expectations: Publicly reported hospital scores do not always correlate with clinical outcomes☆

Maria S. Altieri, MD[a,*], Jie Yang, PhD[b], Donglei Yin, MS[c], Lisa A. Bevilacqua, MD[d], Konstantinos Spaniolas, MD[e], Mark A. Talamini, MD, FACS, MBA[e], Aurora D. Pryor, MD, FACS[e]

[a] Division of Minimally Invasive Surgery, Washington University School of Medicine, St. Louis, MO
[b] Department of Family, Population and Preventive Medicine, Stony Brook University, Medical Center, NY
[c] Department of Applied Mathematics and Statistics, Stony Brook University, NY
[d] Thomas Jefferson University School of Medicine, Philadelphia, PA
[e] Division of Bariatric, Foregut and Advanced Gastrointestinal Surgery Department of Surgery, Stony Brook University Medical Center, NY

ARTICLE INFO

Article history:
Accepted 3 December 2018
Available online 29 January 2019

ABSTRACT

Background: Publicly reported hospital scores are used by patients to make health care–related decisions; however, their relationship to clinical outcomes is unknown.
Methods: Through the use of the New York Statewide Planning and Research Cooperative System database, the association between two commonly used scores (Healthgrades and Centers for Medicare & Medicaid Services Hospital Compare) and four clinical outcomes was evaluated in several surgical fields (general, colorectal, hepatobiliary, foregut, and bariatric).
Results: After adjusting for patient-level factors, patients from facilities with greater Healthgrades scores were less likely to develop any complication after general surgery operations ($P = .0013$). Also, greater Healthgrades scores were associated with less 30-day readmissions and emergency department visits for general surgery operations only ($P = .0061$ and $P = .0013$, respectively). In addition, greater Healthgrades scores were significantly associated with a lesser hospital length of stay for colorectal, foregut, and general surgery operations. Greater Centers for Medicare & Medicaid Services Hospital Compare scores were significantly associated with less 30-day readmissions and lesser hospital length of stay for specific operative groups.
Conclusion: Although some specialties demonstrated a correlation, there was no consistent relationship between publicly reported hospital scores and surgical outcomes that contributed to clinically meaningful use for patients or operations.

Introduction

Publicly reported hospital scores are being used by patients to make decisions on where to access health care.[1,2] Two such scoring systems are the Healthgrades and the Centers for Medicare & Medicaid Services (CMS) Hospital Compare. Although Healthgrades is an objective scoring system, CMS Hospital Compare relies heavily on subjective patient assessments.

Priorities in the quality of health care have shifted to place greater importance on patient-centered outcomes.[3] Since the implementation of the first national, standardized survey of patients' perspective of health care in 2002, patient-derived measures of satisfaction are now being used publicly to rate hospitals.[4,5] These scores are now also being used by the CMS for reimbursement purposes.[6] In 2014, the value-based purchasing payments were determined by three sets of measures. Indeed, the collective scores of several patients' "experience measures" comprised one of these subsets.[7] Measures of patient experience are now accepted as a core component in the US health care system and are being used to financially incentivize hospitals and physicians to improve the quality of health care.[8,9]

The surgical enterprise of most health care centers represents the "heart" of its mission in terms of investment, complexity of services, risk, and public profile. The surgical leadership is invested in the success of their programs and in ensuring continual improvement by any and all measures. Each year as the increasing sets of publicly reported hospital measures are published, surgeons

☆ Biostatistical consultation and support provided by the Biostatistical Consulting Core at School of Medicine, Stony Brook University, NY.

* Corresponding author: 606 Euclid Ave, Campus Box 8109, St. Louis, MO 63110.
E-mail address: maltieri@wustl.edu (M.S. Altieri).

https://doi.org/10.1016/j.surg.2018.12.006
0039-6060/

often find themselves confused when attempting to correlate the findings with their subjective and objective data sets regarding known programs. This ostensible inability to correlate quality with several of these "scores" led us to explore the correlation between widely accepted outcomes data and high-profile public reports regarding hospital quality.

The relationship between hospital clinical outcomes and patient perspectives is inconsistent.[10–13] This inconsistency is further demonstrated in the surgical arena with even fewer studies, also with conflicting results.[14–16] Because subjective reports can be inconsistent, some groups are publicly reporting hospital quality rating solely on objective measures. Healthgrades is a scoring system that includes objective information about hospitals and health care providers. To score hospitals, CMS inpatient data are analyzed, assessing clinical outcomes for 34 conditions and procedures. Complications and mortality are adjusted for patient demographics and comorbidities. These ratings, however, have been criticized for oversights in their methodology of risk adjustment.[14]

This controversy is important because publicly reported hospital ratings play an increasingly important role in patient decision-making and hospital reimbursement, but it is uncertain whether these scores really do correlate with clinical outcomes. Each year as the increasing sets of publicly reported hospital measures are published, surgeons and administrators often find themselves unable to correlate the findings with their subjective and objective data sets regarding known programs. Our study aimed to evaluate the correlation between clinical outcomes after several groups of operations across specialties and these two publicly available, hospital ranking systems, Healthgrades and CMS Hospital Compare, using data collected by the New York State Department of Health.

Methods

Subjects and enrollment

We performed a cohort study involving patients undergoing several operative procedures during 2010–2014, using the database of the New York (NY) Statewide Planning and Research Cooperative System (SPARCS). The SPARCS database is a comprehensive, administrative database managed and controlled by the NY Department of Health. SPARCS collects data on patient demographics, diagnoses, treatments, services, and charges from each NY hospital inpatient and outpatient visit.[17] Every patient is assigned a unique identifier. Thus, SPARCS is able to track patients longitudinally across encounters, and therefore, accurately capture important outcomes, such as emergency department (ED) visits and readmissions.

Measures

Through the use of ICD-9 procedure codes and CPT codes, all patients undergoing several operative procedures were identified. These procedures included the following: general surgery (appendectomy, cholecystectomy, inguinal hernia, ventral/umbilical hernia); colorectal (colon resection, low anterior resection, abdominoperineal resection); hepatobiliary (HPB) (pancreatectomy, hepatectomy, splenectomy, complex biliary procedures); bariatric (laparoscopic adjustable gastric banding, laparoscopic sleeve gastrectomy, Roux-en-Y gastric bypass); and foregut (fundoplication, esophagogastric myotomy [Heller procedure]).

Additional extracted variables included patient demographics and comorbidities, such as patient age, sex, race/ethnicity, and insurance. Measured clinical outcomes included complication rates, 30-day readmissions, 30-day ED visits, and hospital length of stay (HLOS). Hospital volume for each type of operation (high volume versus low volume) based on yearly mean volume was considered in the analysis. Exclusion criteria included age < 18 years and missing data, such as facility name and hospital score ($n = 64$ for Healthgrades, $n = 41$ for CMS Hospital Compare).

The following 2 publicly available hospital score systems were used: Healthgrades (ranging 0–100, with 100 representing the greatest score) and CMS Hospital Compare (ranging 1–5, with 5 representing the greatest score). The Healthgrades score is based on objective data, examining adjusted complications and mortality, and is grouped into 5 categories: 40–49, 50–59, 60–69, 70–79, 80–89.[14] The CMS Hospital Compare score summarizes up to 57 quality measures across 7 areas of quality into a single star rating for each hospital. The overall hospital rating is calculated using only those measures for which data are available. Thus, this may include as few as 9 or as many as 57 measures. The methodology uses 7 latent variable models, which include mortality, safety of care, readmission, patient experience, effectiveness of care, timeliness of care, and efficient use of medical imaging. A score is then calculated by taking the weighted average of these group scores. If a group is missing, however, the weights are redistributed among the qualifying groups. Hospitals can be eligible for an overall rating if they have at least 3 measures within at least 3 measure groups, including 1 outcome group (mortality, safety, or readmission).[18]

Statistical analysis

For the facility-level analysis, Spearman correlation coefficients were used to investigate the associations between Healthgrades scores and the 4 clinical outcomes identified in SPARCS (facility-level inhospital complication rate, 30-day readmission rate, 30-day ED visit rate, and average HLOS). Kruskal–Wallis tests were used to compare these clinical outcomes among facilities with different CMS Hospital Compare scores (1, 2, 3, 4+) for each surgery group, respectively. Generalized linear mixed models were used to examine the relationship between each of 2 scores and clinical outcomes after adjusting for possible confounding factors. Linear mixed models were used to examine the relationship between each of the 2 scores and HLOS after adjusting for confounding factors. Log transformation was applied on HLOS to satisfy the normality assumption for linear mixed models. The facility was considered as a random effect to take into account the clustering nature of patients from the same facility in all regression models.

Possible confounding factors included age, payment, race, region, sex, hospital volume, comorbidities, any comorbidity, and overall complication. Any factor that was statistically significant in the univariate regression analysis at a significance level of $P < .05$ were further included as covariates in the multivariable models. Odds ratios (ORs) with 95% confidence intervals (CIs) were reported as appropriate. Statistical analysis was performed using SAS 9.4 (SAS Institute Inc., Cary, NC) and the significance level was set at .05.

Results

After applying the inclusion and exclusion criteria, there were 538,134 eligible records in the SPARCS database. Table 1 shows the distribution of operations per year. The analysis included 57,830 bariatric, 10,458 colorectal, 5,984 foregut, 11,471 HPB, and 452,391 general surgery operations. On the facility level, greater Healthgrades scores correlated with lesser complication rates for colorectal operations (Spearman rho = –0.19, $P = .0193$), lesser 30-day readmissions rate for bariatric and general surgery operations (Spearman rho = –0.22, $P = .0455$ and Spearman rho = –0.18, $P = .0251$, respectively), lesser 30-day ED visits rates for bariatric and general surgery operations (Spearman rho = –0.32, $P = .0040$

Table 1
Number of surgery records extracted per year.

Procedure	Year					Total
	2010	2011	2012	2013	2014	
Bariatric	9,699	9,786	10,425	13,345	14,575	57,830
Colorectal resections	2,260	2,196	2,031	2,011	1,960	10,458
Foregut	1,231	1,234	1,208	1,188	1,123	5,984
HPB	2,298	2,329	2,367	2,275	2,202	11,471
General surgery	93,106	93,682	92,586	89,800	83,217	452,391

Table 2
Adjusted Associations for the Healthgrades Scores based on generalized linear mixed models or linear mixed models

Type of operation	Any complication		30-day readmission		30-day ED visit		Hospital length of stay (HLOS)	
	Odds ratio (95% CI)	*P* value*	Odds ratio (95% CI)	*P* value*	Odds ratio (95% CI)	*P* value*	Ratio of HLOS + 1 (95% CI)	*P* value†
Bariatric	0.98 (0.85–1.13)	.7858	0.95 (0.87–1.04)	.2829	0.99 (0.92–1.07)	.8230	1.00 (0.97–1.04)	.8829
Colorectal resections	0.93 (0.86–1.01)	.0942	0.99 (0.91–1.07)	.7783	1.04 (0.97–1.11)	.2891	0.95 (0.93–0.98)	.0002§
Foregut	0.90 (0.77–1.06)	.2061	0.91 (0.78–1.07)	.2483	0.91 (0.78–1.07)	.2483	0.94 (0.90–0.99)	.0136‖
HPB	0.92 (0.82–1.03)	.1287	1.05 (0.95–1.16)	.3341	1.04 (0.95–1.15)	.4051	0.96 (0.93–1.00)	.0709
General surgery	0.82 (0.73–0.93)	.0013¶	0.93 (0.88–0.98)	.0061‖	0.92 (0.88–0.97)	.0013‖	0.98 (0.96–0.99)	.0019§

Note: Odds ratio/ratio were compared based on every 10 units increase in the Healthgrades scores.
* *P* value was based on the Wald test from multivariable generalized linear mixed model with facilities as clustering effect comparing whether there is any significant difference among different levels in one variable.
† P value was based on multivariable linear mixed model accommodating the correlation within the same facility.
§ Covariates include age, sex, insurance type, race, region, hospitals' volume, any comorbidity, any complication.
‖ Covariates include age, sex, insurance type, race, any comorbidity, any complication.
¶ Covariates include age, sex, insurance type, race, region, any comorbidity.

and Spearman rho = –0.33, $P < .0001$, respectively), and lesser average LOS for colorectal (Spearman rho = –0.36, $P < .0001$), HPB (Spearman rho = –0.18, $P = .0312$), and general surgery operations (Spearman rho = –0.26, $P = .0007$). There was no correlation found between CMS Hospital Compare scores and any of the 4 clinical outcomes for any of the operations we evaluated (all *P* values from Kruskal-Wallis tests > 0.05).

After adjusting for patient-level factors, patients from facilities with greater Healthgrades scores were less likely to develop any complication after general surgery procedures (OR for every 10 units increase on Healthgrades: 0.82, 95% CI: 0.73–0.93, $P = .0013$, Table 2). In terms of 30-day readmissions and ED visits, patients from facilities with greater Healthgrades scores were less likely to have 30-day readmissions and ED visits for general surgery operations only (OR for every 10 units increase on Healthgrades: 0.93, 95% CI: 0.88–0.98, $P = .0061$, and OR for every 10 units increase on Healthgrades: 0.92, 95% CI: 0.88–0.97, $P = .0013$, respectively). Greater Healthgrades scores were associated with lesser HLOS for colorectal operations (changing rate in HLOS + 1 for every 10 units increase on Healthgrades: 0.95, 95% CI: 0.93–0.98, $P = .0002$), foregut (changing rate in HLOS + 1 for every 10 units increase on Healthgrades: 0.94, 95% CI: 0.90–0.99, $P = .0136$), and general surgery operations (changing rate in HLOS + 1 for every 10 units increase on Healthgrades: 0.98, 95% CI: 0.96–0.99, $P = .0019$).

No statistically significant association was observed between CMS Hospital scores and any complication or 30-day ED visits for any type of operation. Greater CMS Hospital Compare scores were significantly associated with lesser 30-day readmissions and HLOS only for specific groups of operations (Table 3).

Discussion

The current study examines the association between 2 publicly reported scores, Healthgrades and CMS Hospital Compare, and the clinical outcomes after specific groups of operations as measured by a large longitudinal state government database. Our results suggest that the publicly reported scores, both objective and subjective, do not reliably represent hospital outcomes after specific groupings of operations in terms of postoperative complications or health care utilization, specifically for 30-day readmissions, 30-day ED visits, and HLOS.

The CMS Hospital Compare survey relies heavily on patient satisfaction, such as the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS), which is a score compiled of surveys of recently discharged patients and, thus, is a more subjective score; however, Healthgrades is an objective score adjusted for patient demographics and comorbidities and represents complications and mortality after operations. Although these scores are different, both aim to represent patient experience, safety, and processes of care. In addition, both may be used by patients to make health care decisions and by payers to determine reimbursements; however, neither of these scores represent true surgical quality nor outcomes nor drive change for improvement of health care.

Our study suggests that the CMS Hospital Compare score did not have any association with complications or 30-day ED visits for any type of operation that we surveyed. We did, however, observe a significant association with 30-day readmissions and HLOS, in which greater scores were associated with lesser 30-day readmissions for bariatric and colorectal operations and lesser HLOS for colorectal, HPB, and general surgery operations.

Greater Healthgrades scores were significantly associated with lesser rates of complications and lesser 30-day readmissions and 30-day ED visits for general surgery operations. In addition, greater Healthgrades scores were associated with lesser HLOS for colorectal, foregut, and general surgery operations. Our findings—that these 2 scores were differentially associated with surgical outcomes—is consistent with the literature showing discrepancies between the growing numbers of publicly reported hospital metrics and true outcomes. Austin et al.[19] compared 4 national ranking systems and found that across 844 hospitals, none were rated "high performer" by all 4 metrics. Furthermore, only 10% of the hospitals that received a score of "high" by 1 metric received a second high score by a different metric.[19]

Despite these apparent disparities, publicly reported hospital scores appear to play an important role in today's health care. These scores can be viewed by patients when deciding where to

Table 3
Adjusted correlations for CMS HCAHPS scores based on generalized linear mixed models or linear mixed models.

Surgery type	Medicare gov level	Any complication	30-day readmission		30-day ED visit		Length of stay		
		Odds ratio (95% CI)	*P* value*	Odds ratio (95% CI)	*P* value*	Odds ratio (95% CI)	*P* value*	Changing rate of HLOS+1 (95% CI)	*P* value†
Bariatric	2 vs 1	0.79 (0.57–1.09)	.3301	0.89 (0.73–1.09)	.0417‡	1.00 (0.83–1.20)	.5181	0.98 (0.90–1.06)	.4246
	3 vs 1	0.74 (0.52–1.05)		0.84 (0.68–1.05)		0.92 (0.75–1.13)		0.97 (0.89–1.07)	
	4+ vs 1	0.89 (0.58–1.37)		0.68 (0.52–0.89)		0.85 (0.66–1.10)		1.06 (0.95–1.20)	
Colorectal resections	2 vs 1	1.04 (0.86–1.26)	.1206	0.78 (0.66–0.92)	.0017‡	0.86 (0.73–1.01)	.0821	0.93 (0.87–0.98)	< .0001§
	3 vs 1	0.84 (0.68–1.03)		0.71 (0.59–0.86)		0.79 (0.66–0.95)		0.86 (0.80–0.92)	
	4+ vs 1	0.89 (0.68–1.18)		0.74 (0.60–0.93)		0.89 (0.72–1.10)		0.86 (0.79–0.93)	
Foregut	2 vs 1	1.03 (0.72–1.48)	.2549	0.86 (0.62–1.18)	.7908	1.01 (0.78–1.30)	.1585	0.97 (0.87–1.09)	.4332
	3 vs 1	1.12 (0.75–1.65)		0.85 (0.59–1.21)		0.93 (0.69–1.24)		0.98 (0.86–1.11)	
	4+ vs 1	0.70 (0.44–1.13)		0.87 (0.59–1.30)		0.73 (0.52–1.02)		0.88 (0.75–1.03)	
HPB	2 vs 1	0.99 (0.78–1.27)	.2232	1.02 (0.83–1.25)	.514	1.04 (0.84–1.28)	.5618	0.88 (0.80–0.96)	.0007‖
	3 vs 1	0.93 (0.70–1.24)		0.90 (0.72–1.13)		0.90 (0.71–1.15)		0.86 (0.77–0.95)	
	4+ vs 1	0.72 (0.51–1.01)		0.89 (0.69–1.14)		0.91 (0.69–1.19)		0.80 (0.71–0.90)	
General	2 vs 1	1.00 (0.73–1.38)	.9247	0.95 (0.84–1.06)	.1258	0.93 (0.82–1.04)	.376	0.95 (0.92–0.99)	< .0001§
	3 vs 1	1.09 (0.76–1.57)		0.86 (0.75–0.98)		0.90 (0.79–1.02)		0.93 (0.90–0.97)	
	4+ vs 1	0.95 (0.59–1.53)		0.90 (0.75–1.08)		0.90 (0.76–1.08)		0.90 (0.86–0.95)	

* *P* value was based on the Wald test from a multivariable generalized linear mixed model with facilities as clustering effect comparing whether there is any significant difference among different levels in one variable.
† *P* value was based on the multivariable linear mixed model accommodating the correlation within the same facility.
‡ Covariates include age, insurance type, race, any comorbidity, any complication.
§ Covariates include age, sex, insurance type, race, region, hospitals' volume, any comorbidity, any complication
‖ Covariates include age, sex, insurance type, region, any comorbidity, any complication

pursue their care. In addition, measures of patient experience are believed to correlate with the quality of the health care delivered and, thus, they have become a core component in the evaluation of a hospital and can even affect reimbursement to hospitals. Of note, some studies have found that the implementation of these publicly reported scores are associated with improved outcomes. For example, Werner and Bradlow[20] examined whether the implementation of CMS reporting in the mid-2000s was related to subsequent improvement in hospital performance and to patient outcomes for acute myocardial infarction, heart failure, and pneumonia. They found that, after CMS reporting, hospital process performance was significantly improved, as were mortality rates, HLOS, and readmission rates among patients being treated for acute myocardial infarction.[20]

Both of the studied hospital scores in the present analysis (Healthgrades and CMS Hospital Compare) have been described as controversial elsewhere.[13,14,21,22] In 2014, Sheetz et al.[14] examined the clinical registry of the Michigan Surgical Quality Collaborative 2008–2012 and reported no correlation between patients' perspective and morbidity and mortality for patients undergoing major vascular and general surgery elective operations. Conversely, Kennedy et al.[15] examined patient satisfaction, as measured by the HCAHPS component of the CMS Hospital Compare, and reported an association among hospital size, surgical volume, and low mortality with greater overall patient satisfaction; however, favorable surgical outcomes were not consistently associated with greater HCAHPS scores. More recently, Sacks et al.[16] used the national sample of hospitals and showed a significant association between surgical quality and patient satisfaction scores, but the authors only examined patients older than 65 years of age. To further complicate the analysis, Fenton et al.[13] reported that high satisfaction scores may be associated with greater rates of mortality. Even the Healthgrades scores, which are objective scores, have been questioned as well.[23] Because these scores have great weight in both the public eye and to policy makers, it is important that they represent hospital outcomes accurately.

Among the subjective measures, data on the implications and use of these scores have been mixed. Lyu et al.[24] compared the performance of 31 hospitals that participated in the Patient Satisfaction Survey and the CMS Surgical Care Improvement Program on several metrics of surgical care and found that patient satisfaction was not associated with hospital performance on antibiotic prophylaxis, appropriate hair removal, or deep-vein thrombosis prophylaxis. Conversely, Kaye et al.[25] found that hospitals with greater HCAHPS patient satisfaction scores had modest but significantly lesser rates of more patient-relevant outcomes, including complications, readmissions, HLOS, and mortality for patients undergoing major cancer operations. Similar positive correlations have been reproduced in the neurosurgery[26,27] and general surgery[28] literature. One study looking at spine surgery had more mixed results, finding that greater satisfaction scores were only associated with lesser HLOS, but not with decreased rates of discharge to rehabilitation, mortality, or hospital charges.[29] This study also utilized the SPARCS database to evaluate surgical outcomes with CMS data and had results similar to our findings.

Our study evaluated two publicly reported scoring systems, namely the Healthgrades and CMS and their correlation to outcomes in the state of New York. Objective publicly reported scores, such as Healthgrades, have also had mixed results in their correlations to surgical outcomes. Chau et al.[30] compared the accuracy of multiple metrics in predicting outcomes after pancreatectomy, including hospital volume, *US News and World Report*, Healthgrades, Hospital Compare (HCAHPS), and *Consumer Reports*, and found that Healthgrades was the metric most strongly associated with mortality ($r = 0.50$) versus complications and HLOS. Osborne et al.[31] examined the impact of another objective metric of surgical quality—the National Surgical Quality Improvement Program (NSQIP) of the American College of Surgeons. They found that there was no difference in risk-adjusted 30-day mortality, serious complications, reoperations, or readmissions before and after hospital enrollment in this program, suggesting that the existence of and feedback from these quality databases do not necessarily result in improved quality.[31]

The current state of health care today has many agencies reporting hospital ratings, which can be confusing to the public. This report also shows the inconsistency between hospital ratings and outcomes in the various subspecialties. Perhaps the best measure for hospitals and their outcomes is the one that was created by surgeons and arguably has the best risk-adjustment paradigm, namely the NSQIP. NSQIP is a surgical outcomes database created

by the American College of Surgeons in order to measure risk-adjusted outcomes and compare results between hospitals. Comparisons are achieved with a validated risk-adjustment calculator, thus allowing for unbiased comparison between hospitals of various sizes and populations.[32–35] The Surgical CAHPS (S-CAHPS), also developed by the American College of Surgeons and other surgical specialty societies, assesses patients' experiences before, during, and after an operation. This subjective patient-reported metric, which CMS has endorsed but not implemented, may also provide to patients and payers more clinically relevant types of information, which are surgery-specific and may be consistent with true surgical outcomes at an institution.

These differences in the accuracy of the quality metrics used today by governmental agencies and the public call for a more extensive discussion among state legislation, regulating agencies, and the surgical community to enforce more accurate, single-hospital reporting systems, which we maintain most likely should be ones that are created by the surgical community and which accurately define quality by objective measures but also utilize patient- reported satisfaction appropriately.

Until we can change the metrics that are reported and referenced, however, we need to be aware of how we are reviewed. At our own institution, we track reportable findings and actively work toward minimizing these specific complications. At our department morbidity and mortality conference, for instance, NSQIP metrics are specifically discussed, and strategies to minimize these complications are included in the presentations. We believe these specific measures are essential in today's world of enhanced transparency.

Our study has several limitations. The data used are from operations being performed in New York State. Therefore, these data cannot necessarily be applied nationally. The studied period encompassed 2010–2014, including only the most recent data available from the New York Department of Health. Thus, some hospitals may have either improved or worsened their publicly reported scores. Finally, there are also certain limitations to using a statewide database, which have been described elsewhere.[23]

In conclusion, our study strongly suggests that only the Healthgrades scores, which are objective scores, are appropriately associated with complication rates and 30-day ED visits for some operations, and both Healthgrades and CMS Hospital Compare scores correlate with 30-day readmissions or HLOS for some operations. No association in terms of complication rates were observed for other operations, such as colorectal, bariatric, HPB, and foregut operations for both the Healthgrades and CMS Hospital Compare scores. Although the operations we evaluated in some specialties demonstrated a correlation, there was no broad relationship between hospital scores and outcomes that accurately contribute to clinically meaningful use of these measures for patients or procedures. Specific linked measures add value and clarity and correlate with postoperative outcomes. These two publicly reported hospital metric sets do not correlate overall with established clinical outcomes after certain operations across the broad spectrum of the specialties we evaluated.

References

1. Wang J, Hockenberry J, Chou SY, Yang M. Do bad report cards have consequences? Impacts of publicly reported provider quality information on the CABG market in Pennsylvania. *J Health Econ*. 2011;30:392–407.
2. Schneider EC, Epstein AM. Use of public performance reports: A survey of patients undergoing cardiac surgery. *JAMA*. 1998;279:1638–16342.
3. Selby JV, Beal AC, Frank L. The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. *JAMA*. 2012;307:1583–1584.
4. Goldstein E, Fyock J. Reporting of CAHPS quality information to medicare beneficiaries. *Health Serv Res*. 2001;36:477–488.
5. Hargraves JL, Hays RD, Cleary PD. Psychometric properties of the Consumer Assessment of Health Plans Study (CAHPS) 2.0 adult core survey. *Health Serv Res*. 2003;38:1509–1527.
6. Chatterjee P, Joynt KE, Orav EJ, Jha AK. Patient experience in safety-net hospitals: Implications for improving care and value-based purchasing. *Arch Intern Med*. 2012;172:1204–1210.
7. Rau J. Methodology: How value based purchasing payments are calculated Web site. 2018. https://khn.org/news/value-based-purchasing-medicare-methodology/. Accessed March 4, 2018.
8. Rodriguez HP, von Glahn T, Rogers WH, Safran DG. Organizational and market influences on physician performance on patient experience measures. *Health Serv Res*. 2009;44:880–901.
9. Rodriguez HP, von Glahn T, Elliott MN, Rogers WH, Safran DG. The effect of performance-based financial incentives on improving patient care experiences: A statewide evaluation. *J Gen Intern Med*. 2009;24:1281–1288.
10. Jha AK, Orav EJ, Zheng J, Epstein AM. Patients' perception of hospital care in the United States. *N Engl J Med*. 2008;359:1921–1931.
11. Glickman SW, Boulding W, Manary M, Staelin R, Roe MT, Wolosin RJ, et al. Patient satisfaction and its relationship with clinical quality and inpatient mortality in acute myocardial infarction. *Circ Cardiovasc Qual Outcomes*. 2010;3:188–195.
12. Sequist TD, Schneider EC, Anastario M, et al. Quality monitoring of physicians: linking patients' experiences of care to clinical quality and outcomes. *J Gen Intern Med*. 2008;23:1784–1790.
13. Fenton JJ, Jerant AF, Bertakis KD, Franks P. The cost of satisfaction: A national study of patient satisfaction, health care utilization, expenditures, and mortality. *Arch Intern Med*. 2012;172:405–411.
14. Sheetz KH, Waits SA, Girotti ME, Campbell DA, Englesbe MJ. Patients' perspectives of care and surgical outcomes in Michigan: An analysis using the CAHPS Hospital Survey. *Ann Surg*. 2014;260:5–9.
15. Kennedy GD, Tevis SE, Kent KC. Is there a relationship between patient satisfaction and favorable outcomes? *Ann Surg*. 2014;260:592–600.
16. Sacks GD, Lawson EH, Dawes AJ, et al. Relationship between hospital performance on a patient satisfaction survey and surgical quality. *JAMA Surg*. 2015;150:858–864.
17. New York State Department of Health. Statewide Planning and Research Cooperative System (SPARCS) Web site. https://www.health.ny.gov/statistics/sparcs/. Accessed February 9, 2018.
18. Medicare.gov: Hospital Compare Web site. https://www.medicare.gov/hospitalcompare/Data/Hospital-overall-ratings-calculation.html. Accessed November 23, 2018.
19. 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 Mar;34:423–430.
20. Werner RM, Bradlow ET. Public reporting on hospital process improvements is linked to better patient outcomes. *Health Aff (Millwood)*. 2010;29:1319–1324.
21. Chang JT, Hayes RD, Shekelle PG, et al. Patients' global ratings of their health care are not associated with the technical quality of their care. *Ann Intern Med*. 2006;144:665–672.
22. Weeks JC, Catalano PJ, Cronin A, et al. Patients' expectations about effect of chemotherapy for advanced cancer. *N Engl J Med*. 2012;367:1616–1625.
23. Sullivan T. Editor's commentary: A failing grade. *Ann Vasc Surg*. 2006;20:707–708.
24. Lyu H, Wick EC, Housman M, Freischlag J, Makary MA. Patient satisfaction as a possible indicator of quality surgical care. *JAMA Surg*. 2013;148:362–367.
25. Kaye DR, Richardson CR, Ye Z, Herrel LA, Ellimoottil C, Miller DC. Association between patient satisfaction and short-term outcomes after major cancer surgery. *Ann Surg Oncol*. 2017;24:3486–3493.
26. Bekelis K, Missios S, MacKenzie TA, O'Shaughnessy PM. Does objective quality of physicians correlate with patient satisfaction measured by hospital compare metrics in New York State? *World Neurosurg*. 2017;103:852–858.
27. Bekelis K, Missios S, Coy S, Rahmani R, MacKenzie TA, Asher AL. Correlation of subjective hospital compare metrics with objective outcomes of cranial neurosurgical Procedures in New York State. *Neurosurg*. 2017;80:401–408.
28. Tsai TC, Orav EJ, Jha AK. Patient satisfaction and quality of surgical care in US Hospitals. *Ann Surg*. 2015;261:2–8.
29. Missios S, Bekelis K. How well do subjective Hospital Compare metrics reflect objective outcomes in spine surgery? *J Neurosurg Spine*. 2016;25:264–270.
30. Chau Z, West JK, Zhou Z, et al. Rankings versus reality in pancreatic cancer surgery: A real-world comparison. *HPB (Oxford)*. 2014;16:528–533.
31. Osborne NH, Nicholas LH, Ryan AM, Thumma JR, Dimick JB. Association of hospital participation in a quality reporting program with surgical outcomes and expenditures for medicare beneficiaries. *JAMA*. 2015;313:496–504.
32. Dimick JB, Osborne NH, Hall BL, Ko CY, Birkmeyer JD. Risk adjustment for comparing hospital quality with surgery: How many variables are needed? *J Am Coll Surg*. 2010;210:503–508.
33. Ingraham AM, Richards KE, Hall BL, Ko CY. Quality improvement in surgery: The American College of Surgeons National Surgical Quality Improvement Program approach. *Adv Surg*. 2010;44:251–267.
34. Cohen ME, Liu Y, Ko CY, Hall BL. Improved surgical outcomes for ACS NSQIP hospitals over time: Evaluation of hospital cohorts with up to 8 years of participation. *Ann Surg*. 2016;263:267–273.
35. Montroy J, Breau RH, Cnossen S, Witiuk K, Binette A, Ferrier T, et al. Change in adverse events after enrollment in the National Surgical Quality Improvement Program: A systematic review and meta-analysis. *PLoS One*. 2016;11.