Published in final edited form as:

World Neurosurg. 2017 July ; 103: 852–858.e1. doi:10.1016/j.wneu.2017.04.108.

Does objective quality of physicians correlate with patient satisfaction measured by Hospital Compare metrics in New York State?

Kimon Bekelis, M.D.^{1,2,3}, Symeon Missios, M.D.⁴, Todd A. MacKenzie, Ph.D.^{2,3,5,6}, and Patrick M. O'Shaughnessy, D.O.³

¹Department of Neurosurgery, Thomas Jefferson University Hospitals, Philadelphia, PA

²The Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, NH

³Population Health Research Institute of New York at CHS, Rockville Centre, NY

⁴Center for Neuro and Spine, Cleveland Clinic - Akron General Hospital, Akron, OH

⁵Department of Medicine, Dartmouth-Hitchcock Medical Center, Lebanon, NH

⁶Department of Community and Family Medicine, Dartmouth-Hitchcock Medical Center, Lebanon, NH

Abstract

Background—It is unclear whether publicly reported benchmarks correlate with the quality of physicians and institutions. We investigated the association of patient satisfaction measures from a public reporting platform with the performance of neurosurgeons in New York State.

Methods—We performed a cohort study involving patients undergoing neurosurgical operations from 2009–2013, who were registered in the Statewide Planning and Research Cooperative System (SPARCS) database. This cohort was merged with publicly available data from the CMS Hospital Compare website. A propensity adjusted regression analysis was used to investigate the association of patient satisfaction metrics with neurosurgeon quality, as measured by their individual rate of mortality and average length-of-stay (LOS).

Results—Overall, 166,365 patients underwent neurosurgical procedures during the study. Using a propensity adjusted multivariable regression analysis we demonstrated that undergoing neurosurgical operations in hospitals with a greater percentage of patient-assigned “high” score were associated with higher chance of being treated by a physician with superior performance in terms of mortality (OR 1.90; 95% CI, 1.86 to 1.95), and a higher chance of being treated by a physician with superior performance in terms of length-of-stay (LOS) (OR 1.24; 95% CI, 1.21 to 1.27). Similar associations were identified for hospitals with a higher percentage of patients, who claimed they would recommend these institutions to others.

Corresponding Author: Kimon Bekelis, M.D., Dartmouth-Hitchcock Medical Center, One Medical Center Dr. Lebanon, NH 03756, kbekelis@gmail.com, T: 603-650-5110, F: 603-650-4547.

Conflicts of interest. None

Conclusions—Merging a comprehensive all-payer cohort of neurosurgery patients in New York State with data from the CMS Hospital Compare website, we observed an association of superior hospital-level patient satisfaction measures with the objective performance of individual neurosurgeons in the corresponding hospitals.

Keywords

Hospital Compare; neurosurgery; patient satisfaction; outcomes; SPARCS

INTRODUCTION

Quality measurement and reporting have a central role in the current constantly changing healthcare landscape.² This information can empower all members of the healthcare debate, including patients, peers, payers, and policy makers.^{2,4,21} The Medicare Access and CHIP Reauthorization Act¹⁹ creates the framework for several public reporting avenues for quality metrics. The Hospital Compare website is one such platform, launched by the Centers for Medicare and Medicaid Services (CMS), to inform patients and engage them in the decision-making process.⁶ It mainly focuses on patient satisfaction metrics including provider-patient relationship, communication, care coordination, and efficiency. These measures are subjective and do not necessarily reflect the objective performance of hospitals or providers. However, through heightened public awareness, Hospital Compare is becoming one of the major comparative platforms used by patients to select hospitals for their care.

The results of prior studies on the correlation of such metrics with objective outcomes have been equivocal.^{5,12} Some researchers¹⁴ have demonstrated that improved patient satisfaction was associated with superior hospital-level outcomes in terms of mortality, length-of-stay, and discharge to rehabilitation for patients undergoing spine surgery. The same group³ and others^{17,20,22} failed to identify a similar relationship for different surgical procedures, raising questions about the reliability of this reporting platform. However, no previous study has investigated the correlation of hospital performance in patient satisfaction measures with the objective performance of the individual treating physicians, employed in these institutions.

We used the New York Statewide Planning and Research Cooperative System (SPARCS)¹¹ to study the association of subjective Hospital Compare metrics with the objective performance of individual neurosurgeons in the corresponding hospitals, for patients undergoing neurosurgical procedures. A propensity score adjusted regression model was used to control for confounding.

METHODS

New York Statewide Planning and Research Cooperative System (SPARCS)

Our study was approved by our Institutional Review Board. Informed consent was not required. All patients undergoing neurosurgical procedures who were registered in the SPARCS (New York State Department of Health, Albany, NY)¹¹ database between 2009 and 2013 were included in the analysis. For these years, SPARCS contains patient-level details for every hospital discharge, ambulatory surgery, and emergency department admission in

New York State (and only for this state) as coded from admission and billing records. More information about SPARCS is available at <https://www.health.ny.gov/statistics/sparcs/>.

Hospital Compare Website

Hospital Compare is a public reporting program operated by the CMS, which reports process-of-care, patient satisfaction, and outcome measures for more than 4,000 Medicare-certified hospitals in the United States.⁶ Data between 2009 and 2013 were included in the analysis. The website is designed such that patients can make side-by-side comparisons of hospitals that are geographically close to one another with the stated purpose to “help patients make the right decisions about their healthcare”.⁶ More information is available at <http://www.medicare.gov/hospitalcompare/search.html>.

We merged SPARCS and Hospital Compare data using the hospital names. Two coauthors manually attempted to match hospitals that were unmatched during the initial process. In case of disagreement a third coauthor was involved. Only 5 hospitals in SPARCS were not matched after this process, and were excluded from further analysis.

Cohort Definition

In order to establish the cohort of patients, we used *International Classification of Disease-9-Clinical Modification* (ICD-9-CM) codes to identify patients in the dataset who underwent neurosurgical operations (Table S1) between 2009 and 2013.

Outcome variables

The primary outcome variable was treatment by a physician belonging in the quartile with the lowest mortality for neurosurgical operations. Secondary outcome was treatment by a physician belonging in the quartile with the shortest LOS after neurosurgical operations. Surgeons were identified using the OPERATOR variable in the SPARCS database. Each surgeon was ranked according to their mortality and LOS rates with other surgeons in their field (vascular, tumor or spine surgery). Surgeons with fewer than 10 procedures were excluded from the analysis.

Exposure variables

The primary exposure variable was whether the hospital had a high percentage of patients assigning the institution a satisfaction score of 9 or 10 (on a scale of 1–10). Secondary exposure variable was whether the hospital had a high percentage of patients who would recommend the hospital to others. Both were binary variables. The cutoffs to determine whether the hospital had a “high percentage” of the above satisfaction indices were the respective median satisfaction scores across the state of New York. As part of a sensitivity analysis, we treated the percentage of the above satisfaction indices as a continuous variable and repeated the analysis below. The observed associations did not change and therefore these results are not presented further.

We chose to include these metrics because the rest of the patient satisfaction and process measures were collinear with these two (e.g. if patients scored a hospital highly on “provider communication” they were also likely to recommend it), and the outcomes measures were

not specific to neurosurgery, but were referring to myocardial infarction, pneumonia, and congestive heart failure.

Covariates (Table S1) used for risk-adjustment included age, gender, race (African-American, Hispanic, Asian, Caucasian, other), procedure category (tumor, vascular, spine surgery) and insurance (private, Medicare, Medicaid, uninsured, other). The comorbidities used for risk adjustment were diabetes mellitus (DM), smoking, chronic lung disease, hypertension, hypercholesterolemia, peripheral vascular disease (PVD), congestive heart failure (CHF), coronary artery disease (CAD), history of ischemic stroke, history of transient ischemic attack (TIA), alcohol abuse, obesity, chronic renal failure (CRF), and coagulopathy. Only variables that were defined as “present on admission” were considered part of the patient’s preadmission comorbidity profile.

Statistical analysis

The association of Hospital Compare subjective metrics with our outcome measures was examined in a multivariable setting. A logistic regression was employed for our outcomes (mortality). The covariates used for risk adjustment in these models were: age, gender, race, insurance, and all the comorbidities mentioned previously. Mixed effects methods were employed, with hospital facility used as a random effects variable, in order to account for clustering at the hospital level.

In an alternative method to control for confounding, we used a propensity adjusted (with deciles of propensity score) regression model. Mixed effects with hospital facility key as random effects variable were employed. We calculated the propensity score of our exposure variables with a separate logistic regression model, using all the covariates mentioned previously. Propensity adjustment compares subjects within the same decile of propensity score, balancing potential differences in covariates, while using the entire cohort without any observations being discarded. In sensitivity analysis, we repeated the above analyses in pre-specified subgroups of patients based on the procedural category (tumor, vascular, spine surgery). In additional sensitivity analysis we restricted SPARCS to Medicare patients only, in order to mirror Hospital Compare. The direction and magnitude of the observed associations did not change in these additional analyses, and therefore these results are not reported further.

Regression diagnostics were used for all models. All results are based on two sided tests, and the level of statistical significance was set at 0.05. Only cases without missing values were included in the analysis. This study, based on 166,365 patients, has sufficient power (90%) at a 5% type I error rate to detect differences in mortality, as small as 1.58%. Statistical analyses were performed using the 64-bit version of R.3.1.0 (R Foundation for Statistical Computing), and SPSS version 23 (IBM, Armonk, NY).

RESULTS

Patient characteristics

In the selected study period there were 166,365 patients undergoing neurosurgical procedures (mean age was 55.0 years, with 50.6% females), who were registered in

SPARCS. The characteristic of our cohort at baseline, stratified by the exposure variables can be seen in Table 1a and 1b.

Physician performance-inpatient mortality

Patients treated in hospitals with a higher percentage of patient-assigned high score were associated with superior physician performance in terms of mortality (OR, 2.19; 95% CI, 2.14 to 2.24) in the unadjusted analysis. This persisted (Table 2) in a mixed effects logistic regression model (OR, 1.94; 95% CI, 1.90 to 1.99) and a propensity score adjusted model (OR, 1.90; 95% CI, 1.86 to 1.95).

Likewise, patients treated in hospitals with a higher percentage of patients, who claimed they would recommend these institutions, were associated with superior physician performance in terms of mortality (OR, 1.74; 95% CI, 1.70 to 1.77) in the unadjusted analysis. This persisted (Table 3) in a mixed effects logistic regression model (OR, 1.48; 95% CI, 1.45 to 1.52) and a propensity score adjusted model (OR, 1.45; 95% CI, 1.41 to 1.48).

Physician Performance-Length-of-stay

Patients treated in hospitals with a higher percentage of patient-assigned high score were associated with superior physician performance in terms of LOS (OR, 1.16; 95% CI, 1.13 to 1.18) in the unadjusted analysis. This persisted (Table 2) in a mixed effects logistic regression model (OR, 1.27; 95% CI, 1.24 to 1.30) and a propensity score adjusted model (OR, 1.24; 95% CI, 1.21 to 1.27).

Likewise, patients treated in hospitals with a higher percentage of patients, who claimed they would recommend these institutions, were associated with superior physician performance in terms of LOS (OR, 1.08; 95% CI, 1.06 to 1.11) in the unadjusted analysis. This persisted (Table 3) in a mixed effects logistic regression model (OR, 1.31; 95% CI, 1.28 to 1.35) and a propensity score adjusted model (OR, 1.33; 95% CI, 1.30 to 1.37).

DISCUSSION

Using a comprehensive all-payer cohort of patients undergoing neurosurgical procedures in New York State we identified an association of superior hospital-level patient satisfaction measures with the objective performance of individual neurosurgeons in the corresponding hospitals. Our results were robust when considering several advanced observational techniques to account for confounders. Quality measurement and reporting increasingly guides patient choices and payer reimbursement.⁴ Overall such benchmarking is expected to raise quality standards, lower healthcare cost, and improve population health.^{8,10,21} However, the currently available platforms are faced with significant criticism, given their mostly subjective nature, and the limited consideration of important risk-adjusted objective measures of care (such as mortality, length of stay, or hospitalization charges).^{1,18}

The validity of publicly reported Hospital Compare metrics is debated. Some researchers¹⁴ have demonstrated that hospitals achieving higher patient satisfaction were associated with lower rates of mortality, discharge to rehabilitation, and average LOS for patients undergoing spine surgery. They failed to identify such an association for cranial

neurosurgery.³ Similarly inconsistent results have been observed for other subspecialties. Shafi et al²⁰ demonstrated that the performance of trauma centers on a series of Hospital Compare metrics was not associated with objective trauma outcomes. In their conclusions, they advocated for the development of trauma specific measures. On the contrary, these metrics have shown better performance for general medical conditions. In this setting, Werner et al²² demonstrated that superior disease specific Hospital Compare metrics for pneumonia, myocardial infarction, and congestive heart failure were associated with modestly improved mortality. However, the utility of these measurements in generating hospital rankings has been debated. Paddock et al¹⁷ used comparative rankings provided by the Hospital Compare Website and concluded that these sets of measures make it difficult to distinguish among hospitals that patients are likely to choose from. No prior study attempted to correlate subjective publicly reported measures with objective physician performance.

Despite these conflicting results, hospital leaders pay close attention to Hospital Compare metrics.^{9,12} In a recent survey, executives from 380 hospitals indicated that reported measures significantly influence local planning and quality improvement initiatives.¹² 87.1% of hospitals incorporated performance on publicly reported measures into their annual goals, whereas 90.2% reported regularly reviewing the results with their board of trustees, and 94.3% with senior clinical and administrative leaders.¹² They indicated that patient experience measures, which are undoubtedly the most subjective, are the ones shaping hospital reputations more significantly. However, surveyed leaders expressed concerns about the clinical meaningfulness, unintended consequences, and methods of public reporting.¹²

The present analysis demonstrated an association of superior physician performance with subjective patient satisfaction Hospital Compare metrics. It is unlikely that there is a direct correlation between the two variables. However, this association likely reflects a general “culture of excellence” in institutions with a focus on patient satisfaction. Such facilities are likely to also attract the highest performing physicians.²³ In this regard, patient satisfaction must be viewed as an essential but not sole indicator of surgical quality. Furthermore, the association between subjective and objective surgical outcomes may vary significantly between disease states, patient populations and specific procedures. Alternatively, higher patient satisfaction can be correlated with some desirable patient characteristics, such as compliance with instructions and medications, which we cannot account for.

From a policy perspective, we have no evidence that hospitals with inferior satisfaction scores should be penalized. However, such subjective metrics do not currently allow the side-by-side comparison of different institutions in regards to the quality of care in subspecialties such as neurosurgery. In order for this goal of Hospital Compare to be met, neurosurgery-specific, patient-centered, quality oriented outcomes should be monitored and reported. Recent legislative reform¹⁹ allows a more active role for registries in this public reporting process, avoiding the long and costly task of developing metrics through the National Quality Forum.¹⁵ In this direction, the NeuroPoint Alliance¹⁶ is creating the first neurosurgery specific Qualified Clinical Data Registry Reporting (QCDR),⁷ based on metrics developed as part of the Quality and Outcomes Database (QOD).¹³ These specific measures, in the setting of rigorous risk-adjustment will allow the head-to-head comparison of different hospitals.

Our study has several limitations common to administrative databases. Residual confounding could account for some of the observed associations. However, we used advanced observational techniques to control for confounders. In addition, coding inaccuracies will undoubtedly occur and can affect our estimates. However, coding for procedures is expected to be accurate, since they are a revenue generator, and are under scrutiny by payers. Although SPARCS includes all hospitals from the entire New York State, the generalization of this analysis to the US population at large is uncertain. We were also lacking post-hospitalization, and long-term data on our patients. Quality metrics (i.e. modified Rankin score) are also not available through this database, and therefore we cannot compare our patients on these outcomes. Initiatives, such as the QOD will provide such opportunities in the future. Due to reporting requirements, we cannot present data on physicians from individual hospitals.

Additionally, the two databases compared are not congruent. However, restricting SPARCS to only Medicare patients (in sensitivity analysis) and repeating our analyses yielded similar results. There are numerous differences in patient characteristics between the two groups, although advanced observational techniques are expected to account for such measured confounders. An alternative interpretation of our results is that they do not reflect a culture of excellence, but rather the fact that sicker patients do worse and can be less satisfied with their care. Creating quartiles of care outcomes for neurosurgeons may not be valid if practices are not parallel (e.g. differences between cranial and spinal practices). To address this potential confounder, we repeated all our analyses for every individual procedure. Our threshold for acceptable satisfaction outcomes is arbitrary. However this is the only threshold available through Hospital Compare. Finally, causality cannot be definitively established based on observational data, despite the use of advanced techniques.

Conclusions

It is unclear whether publicly reported benchmarks correlate with the quality of physicians and institutions. To investigate the association of patient satisfaction measures from a public reporting platform with the performance of neurosurgeons in New York State. Merging a comprehensive all-payer cohort of neurosurgery patients in New York State with data from the CMS Hospital Compare website, we observed an association of superior hospital-level patient satisfaction measures with the objective performance of individual neurosurgeons in the corresponding hospitals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding. Supported by grant from the National Center for Advancing Translational Sciences (NCATS) of the NIH (Dartmouth Clinical and Translational Science Institute-UL1TR001086). The funders had no role in the design or execution of the study.

References

1. Austin JM, Jha AK, Romano PS, Singer SJ, Vogus TJ, Wachter RM, et al. National hospital ratings systems share few common scores and may generate confusion instead of clarity. *Health Aff (Millwood)*. 2015; 34:423–430. [PubMed: 25732492]
2. Bekelis, K., Goodney, RP., Dzebisashvili, N., Goodman, DC., Bronner, KK. Practice TDIHPaC. A Dartmouth Atlas of Health Care Series. Lebanon, NH: 2014. Variation in the Care of Surgical Conditions: Cerebral Aneurysms.
3. Bekelis K, Missios S, Coy S, Rahmani R, MacKenzie T, Asher AL. Correlation of subjective Hospital Compare metrics with objective outcomes of cranial neurosurgical procedures in New York State. *Neurosurgery*. 2017 in press.
4. Birkmeyer NJ, Birkmeyer JD. Strategies for improving surgical quality--should payers reward excellence or effort? *N Engl J Med*. 2006; 354:864–870. [PubMed: 16495401]
5. Cassel CK, Conway PH, Delbanco SF, Jha AK, Saunders RS, Lee TH. Getting more performance from performance measurement. *N Engl J Med*. 2014; 371:2145–2147. [PubMed: 25470691]
6. Centers for Medicare and Medicaid Services. Hospital Compare. 2015; 2015
7. Centers for Medicare and Medicaid Services. Physician Compare. 2015; 2015
8. Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Int Med*. 2008; 148:111–123. [PubMed: 18195336]
9. Goff SL, Lagu T, Pekow PS, Hannon NS, Hinchey KL, Jackowitz TA, et al. A qualitative analysis of hospital leaders' opinions about publicly reported measures of health care quality. *Jt Comm J Qual Patient Saf*. 2015; 41:169–177. [PubMed: 25977201]
10. Harder B, Comarow A. Hospital quality reporting by US News & World Report: why, how, and what's ahead. *JAMA*. 2015; 313:1903–1904. [PubMed: 25988456]
11. Health NYSDo. Statewide Planning and Research Cooperative System (SPARCS). 2015; 2015
12. Lindenauer PK, Lagu T, Ross JS, Pekow PS, Shatz A, Hannon N, et al. Attitudes of hospital leaders toward publicly reported measures of health care quality. *JAMA Intern Med*. 2014; 174:1904–1911. [PubMed: 25286316]
13. McGirt MJ, Speroff T, Dittus RS, Harrell FEJ, Asher AL. The National Neurosurgery Quality and Outcomes Database (N2QOD): general overview and pilot-year project description. *Neurosurg Focus*. 2013; 34:E6.
14. Missios S, Bekelis K. How well do subjective Hospital Compare metrics reflect objective outcomes in spine surgery? *J Neurosurg Spine*. 2016; 25:264–270. [PubMed: 26989975]
15. National Quality Forum. Consensus Development Process. 2015; 2015
16. NeuroPoint Alliance. The National Neurosurgery Quality and Outcomes Database (N²QOD). 2015; 2015
17. Paddock SM, Adams JL, Hoces de la Guardia F. Better-than-average and worse-than-average hospitals may not significantly differ from average hospitals: an analysis of Medicare Hospital Compare ratings. *BMJ Qual Saf*. 2015; 24:128–134.
18. Ryan AM, Nallamotheu BK, Dimick JB. Medicare's public reporting initiative on hospital quality had modest or no impact on mortality from three key conditions. *Health Aff (Millwood)*. 2012; 31:585–592. [PubMed: 22392670]
19. Schönenberger S, Möhlenbruch M, Pfaff J, Mundiyanapurath S, Kieser M, Bendszus M, et al. Sedation vs. Intubation for Endovascular Stroke Treatment (SIESTA) - a randomized monocentric trial. *Int J Stroke*. 2015; 10:969–978. [PubMed: 25864723]
20. Shafi S, Parks J, Ahn C, Gentilello LM, Nathens AB, Hemmila MR, et al. Centers for Medicare and Medicaid services quality indicators do not correlate with risk-adjusted mortality at trauma centers. *J Trauma*. 2010; 68:771–777. [PubMed: 20386272]
21. Werner RM, Bradlow ET. Public reporting on hospital process improvements is linked to better patient outcomes. *Health Aff (Millwood)*. 2010; 29:1319–1324. [PubMed: 20606180]
22. Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. *JAMA*. 2006; 296:2694–2702. [PubMed: 17164455]

23. Werner RM, Bradlow ET, Asch DA. Does hospital performance on process measures directly measure high quality care or is it a marker of unmeasured care? *Health Serv Res.* 2008; 43:1464–1484. [PubMed: 22568614]

Patient characteristics stratified by treatment in hospitals with a higher percentage of patient-assigned high score

Table 1a

	All Patients			Patients of hospitals with less than median patient-assigned high scores			Patients of hospitals with more than median patient-assigned high scores		
	N= 166,365			N= 87,394			N= 78,971		
	Mean	SD		Mean	SD		Mean	SD	P-Value
Age	55.05	14.74		54.45	14.59		55.71	14.86	<0.0001
	N	%		N	%		N	%	P-Value
Female gender	84157	50.59		44395	50.80		39762	50.35	0.068
Race	120340	72.54	Caucasian	63218	72.51		57122	72.56	0.986
	15360	9.26	African-American	9795	11.23		5565	7.07	<0.0001
	10726	6.47	Hispanic	6790	7.79		3936	5.00	<0.0001
	3018	1.82	Asian	1237	1.42		1781	2.26	<0.0001
	16462	9.92	Other	6147	7.05		10315	13.10	<0.0001
Insurance	44234	26.64	Medicare	21894	25.11		22340	28.33	<0.0001
	6890	4.15	Medicaid	4429	5.08		2461	3.12	<0.0001
	86226	51.92	Private	43647	50.05		42579	53.99	<0.0001
	4007	2.41	Uninsured	2482	2.85		1525	1.93	<0.0001
	24709	14.88	Other	14748	16.91		9961	12.63	<0.0001
Transient Ischemic Attack	145	0.09		92	0.11		53	0.07	0.008
Ischemic Stroke	162	0.10		106	0.12		56	0.07	0.001
Coronary Artery Disease	16256	9.77		8656	9.90		7600	9.62	0.054
Chronic Obstructive Pulmonary Disease	25440	15.29		14374	16.45		11066	14.01	<0.0001
Congestive Heart Failure	3356	2.02		2045	2.34		1311	1.66	<0.0001
Diabetes Mellitus	25364	15.25		14099	16.13		11265	14.26	<0.0001
Coagulopathy	1801	1.08		976	1.12		825	1.04	0.176
Chronic Renal Failure	3848	2.31		2061	2.36		1787	2.26	0.196
Hypertension	75094	45.14		40073	45.85		35021	44.35	<0.0001
Smoking	25498	15.33		15286	17.49		10212	12.93	<0.0001
Hypercholesterolemia	49582	29.80		25124	28.75		24458	30.97	<0.0001
Obesity	49582	29.80		10304	11.79		8233	10.43	<0.0001
Alcohol	2679	1.61		1673	1.91		1006	1.27	<0.0001

	All Patients		Patients of hospitals with less than median patient-assigned high scores		Patients of hospitals with more than median patient-assigned high scores		P-Value
	Mean	SD	Mean	SD	Mean	SD	
Peripheral Vascular Disease	3495	2.10	1956	2.24	1539	1.95	<0.0001
Top quartile of physicians (Mortality) *	67462	40.55	28048	32.09	39414	49.91	<0.0001
Top quartile of physicians (LOS) §	64239	38.61	32343	37.01	31896	40.39	<0.0001
Mortality	1244	6.35	802	7.91	442	4.67	<0.0001
LOS	7	4-15	8	4-17	7	3-13	<0.0001
	Median	IQR	Median	IQR	Median	IQR	

SD: Standard Deviation; %: Percentage; LOS: Length of Stay; IQR: Interquartile Range

* Includes the quartile of physicians with the lowest inpatient mortality while performing neurosurgical procedures

§ Includes the quartile of physicians with the shortest length-of-stay while performing neurosurgical procedures

Patient characteristics stratified by treatment in hospitals with a higher percentage of patients, who claimed they would recommend these institutions

Table 1b

	All Patients			Patients of hospitals with less than median percentage of patients, who claimed they would recommend these institutions			Patients of hospitals with more than median percentage of patients, who claimed they would recommend these institutions		
	N= 166365			N= 90002			N= 76363		
	Mean	SD		Mean	SD		Mean	SD	P-Value
Age	55.05	14.74		54.33	14.55		55.91	14.92	<0.0001
	N	%		N	%		N	%	P-Value
Female gender	84157	50.59		45656	50.73		38501	50.42	0.209
Race	120340	72.54	Caucasian	65089	72.49		55251	72.59	0.879
	15360	9.26	African-American	9964	11.10		5396	7.09	<0.0001
	10726	6.47	Hispanic	6871	7.65		3855	5.06	<0.0001
	3018	1.82	Asian	1255	1.40		1763	2.32	<0.0001
	16462	9.92	Other	6615	7.37		9847	12.94	<0.0001
Insurance	44234	26.64	Medicare	22276	24.80		21958	28.79	<0.0001
	6890	4.15	Medicaid	4431	4.93		2459	3.22	<0.0001
	86226	51.92	Private	44207	49.22		42019	55.10	<0.0001
	4007	2.41	Uninsured	2563	2.85		1444	1.89	<0.0001
	24709	14.88	Other	16331	18.18		8378	10.99	<0.0001
Transient Ischemic Attack	145	0.09		93	0.10		52	0.07	0.015
Ischemic Stroke	162	0.10		107	0.12		55	0.07	0.002
Coronary Artery Disease	16256	9.77		8835	9.82		7421	9.72	0.501
Chronic Obstructive Pulmonary Disease	25440	15.29		14900	16.56		10540	13.80	<0.0001
Congestive Heart Failure	3356	2.02		2063	2.29		1293	1.69	<0.0001
Diabetes Mellitus	25364	15.25		14443	16.05		10921	14.30	<0.0001
Coagulopathy	1801	1.08		1015	1.13		786	1.03	0.053
Chronic Renal Failure	3848	2.31		2119	2.35		1729	2.26	0.223
Hypertension	75094	45.14		41201	45.78		33893	44.38	<0.0001
Smoking	25498	15.33		16316	18.13		9182	12.02	<0.0001
Hypercholesterolemia	49582	29.80		25650	28.50		23932	31.34	<0.0001
Obesity	49582	29.80		10707	11.90		7830	10.25	<0.0001

	All Patients N= 166365		Patients of hospitals with less than median percentage of patients, who claimed they would recommend these institutions N= 90002		Patients of hospitals with more than median percentage of patients, who claimed they would recommend these institutions N= 76363		P-Value
	Mean	SD	Mean	SD	Mean	SD	
Alcohol	2679	1.61	1744	1.94	935	1.22	<0.0001
Peripheral Vascular Disease	3495	2.10	1952	2.17	1543	2.02	0.036
Top quartile of physicians (Mortality) *	67462	40.55	31295	34.77	36167	47.36	<0.0001
Top quartile of physicians (LOS) §	64239	38.61	34082	37.87	30157	39.49	<0.0001

SD: Standard Deviation

* Includes the quartile of physicians with the lowest inpatient mortality while performing neurosurgical procedures

§ Includes the quartile of physicians with the shortest length-of-stay while performing neurosurgical procedures

Association of top physicians with subjective Hospital Compare metrics (treatment in hospitals with a higher percentage of patient-assigned high score)

Table 2

Model	Crude		Multivariable Regression Adjusted*		Propensity Score Controlled*	
	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
Top 25% (Lowest) Inpatient mortality ^{##}	2.19 (2.14–2.24)	<0.0001	1.94 (1.90–1.99)	<0.0001	1.90 (1.86–1.95)	<0.0001
Top 25% (Lowest) Length-of-Stay ^{##}	1.16 (1.13–1.18)	<0.0001	1.27 (1.24–1.30)	<0.0001	1.24 (1.21–1.27)	<0.0001

OR: Odds Ratio; 95% CI: 95% Confidence Interval

* Mixed effects; Includes treatment hospital as a random effect variable

^{##} Analyses based on logistic regression

Association of top physicians subjective Hospital Compare metrics (treatment in hospitals with a higher percentage of patients, who claimed they would recommend these institutions)

Table 3

Model	Crude		Multivariable Regression Adjusted [*]		Propensity Score Adjusted [*]	
	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
Top 25% Inpatient mortality^{##}	1.74 (1.70–1.77)	<0.0001	1.48 (1.45–1.52)	<0.0001	1.45 (1.41–1.48)	<0.0001
Top 25% Length-of-Stay^{##}	1.08 (1.06–1.11)	<0.0001	1.31 (1.28–1.35)	<0.0001	1.33 (1.30–1.37)	<0.0001

OR: Odds Ratio; 95% CI: 95% Confidence Interval

^{*} Mixed effects; Includes treatment hospital as a random effect variable

^{##} Analyses based on logistic regression