


## Navigating by Stars: Using CMS Star Ratings to Choose Hospitals for Complex Cancer Surgery

Marianna V. Papageorge , MD,<sup>1</sup> Benjamin J. Resio, MD,<sup>1</sup> Andres F. Monsalve, MD,<sup>1</sup> Maureen Canavan, PhD, MPH,<sup>2</sup> Ranjan Pathak, MBBS,<sup>1</sup> Vincent J. Mase, MD,<sup>1</sup> Andrew P. Dhanasopon, MD,<sup>1</sup> Jessica R. Hoag, PhD,<sup>2</sup> Justin D. Blasberg, MD,<sup>1</sup> Daniel J. Boffa, MD<sup>1,\*</sup>

<sup>1</sup>Section of Thoracic Surgery, Department of Surgery, Yale School of Medicine, New Haven, CT, USA and <sup>2</sup>Cancer Outcomes, Public Policy and Effectiveness Research Center, Yale School of Medicine, New Haven, CT, USA

\*Correspondence to: Daniel J. Boffa, MD, Section of Thoracic Surgery, Department of Surgery, Yale School of Medicine, 330 Cedar St, BB205, New Haven, CT 06510, USA (e-mail: daniel.boffa@yale.edu).

### Abstract

**Background:** The Centers for Medicare and Medicaid Services (CMS) developed risk-adjusted “Star Ratings,” which serve as a guide for patients to compare hospital quality (1 star = lowest, 5 stars = highest). Although star ratings are not based on surgical care, for many procedures, surgical outcomes are concordant with star ratings. In an effort to address variability in hospital mortality after complex cancer surgery, the use of CMS Star Ratings to identify the safest hospitals was evaluated.

**Methods:** Patients older than 65 years of age who underwent complex cancer surgery (lobectomy, colectomy, gastrectomy, esophagectomy, pancreaticoduodenectomy) were evaluated in CMS Medicare Provider Analysis and Review files (2013–2016). The impact of reassignment was modeled by applying adjusted mortality rates of patients treated at 5-star hospitals to those at 1-star hospitals (Peters-Belson method). **Results:** There were 105 823 patients who underwent surgery at 3146 hospitals.

The 90-day mortality decreased with increasing star rating (1 star = 10.4%, 95% confidence interval [CI] = 9.8% to 11.1%; and 5 stars = 6.4%, 95% CI = 6.0% to 6.8%). Reassignment of patients from 1-star to 5-star hospitals (7.8% of patients) was predicted to save 84 Medicare beneficiaries each year. This impact varied by procedure (colectomy = 47 lives per year; gastrectomy = 5 lives per year). Overall, 2189 patients would have to change hospitals each year to improve outcomes (26 patients moved to save 1 life). **Conclusions:** Mortality after complex cancer surgery is associated with CMS Star Rating. However, the use of CMS Star Ratings by patients to identify the safest hospitals for cancer surgery would be relatively inefficient and of only modest impact.

The safety of complex cancer surgery is highly variable across hospitals in the United States (1–3), with as much as a fourfold difference in hospital mortality rates, reflecting differences in volume, hospital experience, and surgeon training (4–11). Multiple studies have suggested that a large-scale shift of patients away from high-risk hospitals could meaningfully reduce mortality after complex cancer surgery (12–15). As a result, there is considerable interest in having patients choose hospitals with lower mortality rates for complex cancer surgery.

There are numerous challenges to matching patients with hospitals that have the lowest mortality rates for complex care, not the least of which is identifying hospitals best suited to perform specific procedures. Annual procedural volume has been proposed to align patients with hospitals and has been

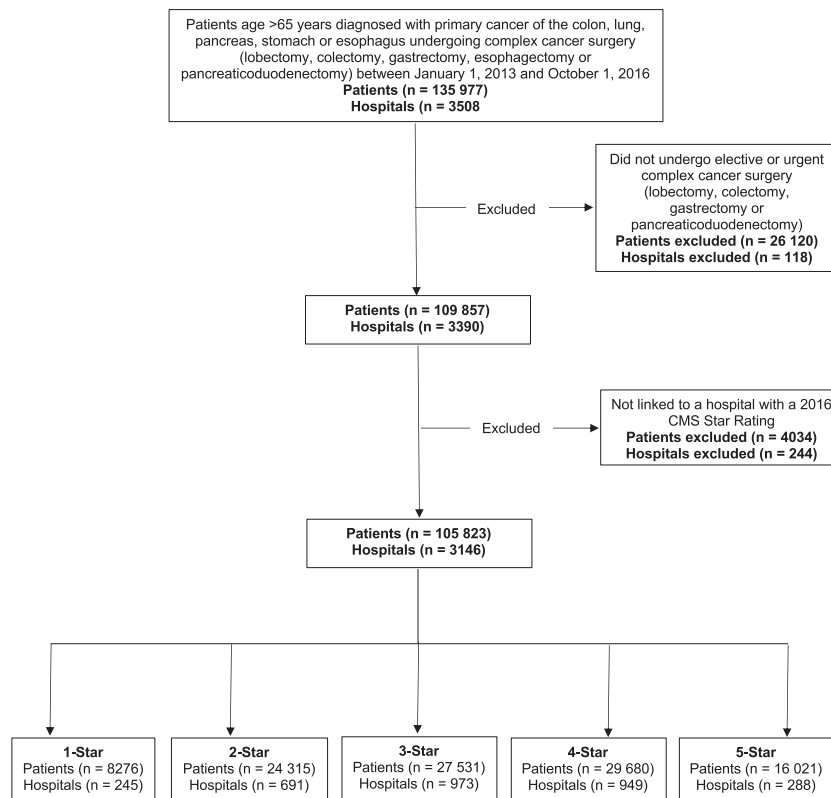
embraced by some payers and clinicians (16–18). On the other hand, “safety performance” (using risk-adjusted standardized mortality ratios) appears to be a more accurate mechanism to match patients with the highest quality hospitals for complex cancer surgery (4). However, neither annual hospital volumes nor procedure-specific safety performance data are readily available to patients and providers. Therefore, a great need exists for an understandable and accessible mechanism to identify the safest hospitals for complex cancer surgery.

The Centers for Medicare and Medicaid Services (CMS) assigned “Star Ratings” in 2015 as a guide for patients to compare hospital quality. By considering 62 measures (eg, mortality, readmissions, and patient experience), CMS assigns each hospital a star rating from 1 (lowest) to 5 (highest) (19,48,49). CMS Star

Received: 25 January 2020; Revised: 30 March 2020; Accepted: 26 June 2020

© The Author(s) 2020. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)



**Figure 1.** Inclusion criteria. Consort diagram of inclusion criteria, including patients older than 65 years with a diagnosis of cancer of the lung, colon, stomach, esophagus, or pancreas who underwent nonemergent cancer surgery, including lobectomy, colectomy, gastrectomy, esophagectomy, and pancreaticoduodenectomy, between January 1, 2013, and October 1, 2016, in a hospital with a Centers for Medicare and Medicaid Services (CMS) Star Rating.

Ratings are not specific to a particular clinical scenario (ie, pneumonia or colectomy) or patient population (ie, transplant, cancer). Despite this fact, CMS Star Ratings have been shown to be associated with quality and safety across many patient care scenarios (20-23). More recently, a study by Kaye et al. (24) demonstrated that star ratings correlate with 30-day mortality after complex cancer surgery. On the other hand, approximately 50% of surgical mortalities occur between 30 and 90 days of surgery (25,26). Furthermore, it is unclear whether the magnitude of mortality difference across star ratings is sufficient to be leveraged to reduce mortality after complex cancer surgery on a national level.

CMS Star Ratings have the potential to influence patient choice because they are not only simple to understand but also readily accessible. For example, in a recent search engine query of “hospital quality,” the first 4 results focused on CMS Star Ratings and the CMS Hospital Compare resource. Furthermore, more than 6000 external websites link to the CMS Star Ratings information page, including those from insurance companies and hospitals, and are included as a component of hospital advertising and hospital ranking mechanisms (27-30). For example, *Becker’s Hospital Review* considers CMS Star Ratings when compiling its yearly “100 great hospitals of America” (31-33).

In an effort to better understand the impact of incorporating CMS Star Ratings into hospital selection for complex cancer surgery, the potential mortality reduction associated with reassigning patients from the least safe (1 star) to the safest (5 star) hospitals was modeled. We hypothesized that selecting hospitals for complex cancer surgery based on CMS Star Ratings could reduce operative mortality.

## Methods

### Data Source

The Medicare Provider Analysis and Review file provided by the Centers for Medicare and Medicaid services consolidates information for Medicare beneficiaries who have used hospital inpatient services (34). The Medicare Provider Analysis and Review files were linked with the Hospital Compare database. These data were then used to isolate hospitals with assigned CMS Star Ratings (35). The analysis was completed in accordance with our institutional review board–approved research protocol, with consent waived.

### Patient Selection

Eligible patients were older than 65 years with a diagnosis of cancer of the lung, colon, stomach, esophagus, or pancreas. All patients had undergone nonemergent surgery, including lobectomy, colectomy, gastrectomy, esophagectomy, or pancreaticoduodenectomy, between January 1, 2013, and October 1, 2016, in a hospital with a CMS Star Rating (Figure 1). This time period, over 4 years, was selected to minimize artifact in mortality assessment in lower volume centers (ie, lower volume centers artificially being pulled toward a mean). These particular cancer surgeries were selected as “complex” because each procedure is associated with a surgical mortality greater than 2%, which has previously been suggested as a criterion for high-risk elective surgery (36). In addition, we evaluated procedures that represented different anatomic regions and likely were managed by

**Table 1.** Patient characteristics<sup>a</sup>

Variable	CMS Star Rating, No. (%)					P <sup>a</sup>
	1	2	3	4	5	
Age, y						
65-69	1292 (15.6)	3641 (15.0)	4121 (15.0)	4396 (14.8)	2343 (14.6)	.04
70-74	2214 (26.8)	6512 (26.8)	7027 (25.5)	7712 (26.0)	4094 (25.6)	.04
75-79	2457 (29.7)	7202 (29.6)	8003 (29.1)	8567 (28.9)	4722 (29.5)	.73
80+	2313 (27.9)	6960 (28.6)	8380 (30.4)	9005 (30.3)	4862 (30.3)	<.001
Sex						
Male	3935 (47.5)	11 702 (48.1)	13 017 (47.3)	14 177 (47.8)	7678 (47.9)	.58
Female	4341 (52.5)	12 613 (51.9)	14 514 (52.7)	15 503 (52.2)	8343 (52.1)	.58
Race						
White	6792 (82.1)	21 117 (86.8)	24 358(88.5)	26 708 (90.0)	14 314 (89.3)	<.001
Black	949 (11.5)	1993 (8.2)	1881 (6.8)	1360 (4.6)	754 (4.7)	<.001
Other or unknown	535 (6.5)	1205 (5.0)	1292 (4.7)	1612 (5.4)	953 (5.9)	.11
Year of surgery						
2013	1812 (21.9)	5230 (21.5)	6009 (21.8)	6406 (21.6)	3260 (20.3)	.005
2014	1980 (23.9)	5572 (22.9)	6438 (23.4)	6747 (22.7)	3799 (23.7)	.71
2015	2164 (26.1)	6476 (26.6)	7464 (27.1)	8104 (27.3)	4339 (27.1)	.12
2016	2320 (28.0)	7037 (28.9)	7620 (27.7)	8423 (28.4)	4623 (28.9)	.18
Admission type						
Elective	7445 (90.0)	2810 (11.6)	24 307 (88.3)	26 922 (90.7)	14 645 (91.4)	<.001
Urgent	831 (10.0)	21 505 (88.4)	3224 (11.7)	2758 (9.3)	1376 (8.6)	<.001
Elixhauser Comorbidity Index						
0	1279 (15.5)	3913 (16.1)	4580 (16.6)	5084 (17.1)	2781 (17.4)	<.001
1-2	3169 (38.3)	9303 (38.3)	10 591 (38.5)	11 243 (37.9)	6198 (38.7)	.55
≥3	3828 (46.3)	11 099 (45.6)	12 360 (44.9)	13 353 (45.0)	7042 (44.0)	.001
Procedure						
Lobectomy	1992 (24.1)	5972 (24.6)	6337 (23.0)	7160 (24.1)	4587 (28.6)	<.001
Colectomy	4945 (59.8)	14 533 (59.8)	17 574 (63.8)	18 049 (60.8)	9068 (56.6)	<.001
Gastrectomy	470 (5.7)	1308 (5.4)	1225 (4.4)	1401 (4.7)	725 (4.5)	<.001
Esophagectomy	245 (3.0)	744 (3.1)	751 (2.7)	1021 (3.4)	537 (3.4)	.10
Whipple	624 (7.5)	1758 (7.2)	1644 (6.0)	2049 (6.9)	1104 (6.9)	.06

<sup>a</sup>1 star vs 5 stars. CMS = The Centers for Medicare and Medicaid Services.

different surgical teams. Overall, this group of procedures captures an estimated 80% of all high-risk cancer surgery (36).

### CMS Star Ratings

CMS Star Ratings have been shown to be associated with surgical outcomes despite the fact they were not designed to consider this patient population. Of the 62 measured variables used by CMS to generate a star rating, none are specific to surgery. However, several studies have found CMS Star Ratings are associated with outcomes after surgery, including complex cancer surgery, and this information is readily accessible to patients and providers (20-24,28-33).

### Data Elements

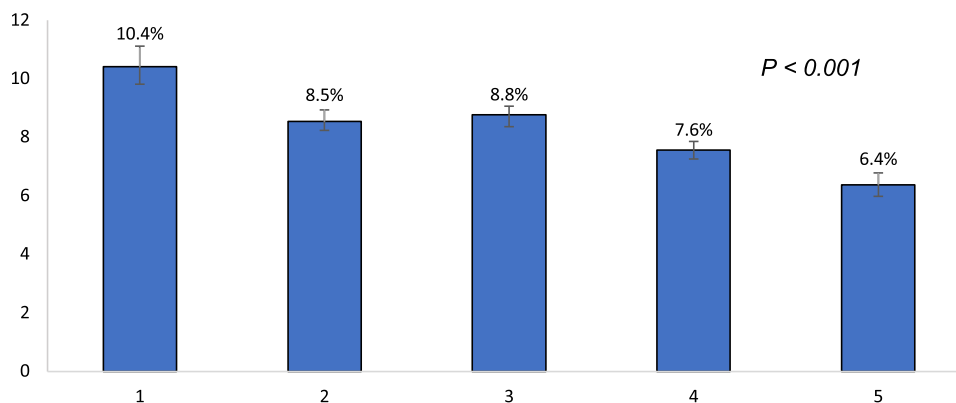
Independent variables in this analysis included patient characteristics (age, sex, race, year of surgery, admission type, Elixhauser Comorbidity Index) and procedure (lobectomy, colectomy, gastrectomy, esophagectomy, pancreaticoduodenectomy; gastrectomy and colectomy were further divided into partial and total resections). Patients were stratified by CMS Star Ratings of the hospital at which they underwent surgery (Table 1). The primary outcome measure was 90-day mortality and was calculated in days following the date of surgery (25). Ninety-day mortality was chosen as the dependent variable

because 30-day mortality may underestimate perioperative mortality risk by approximately 50% (25,26).

### Statistical Analysis

Observed 90-day mortality rates for each of the five procedures were compared across the five CMS Star Rating categories. Secondary outcomes included the potential number of lives saved and the relative efficiency of CMS Star Rating-based regionalization (patients moved per life saved). This study was conducted by considering hospital star-rated cohorts in aggregate (ie, all patients treated at 1-star hospitals), because the objective was to model a generalized impact (ie, if CMS star ratings were broadly used as opposed to changing from one specific hospital to another). Our previous work has demonstrated the challenges of studying safety performance on the individual hospital level (ie, deriving individual standardized mortality ratios for each hospital) given that low-volume hospitals tend to be pulled toward the mean (4). Therefore, the aggregate data approach was more relevant and likely more reflective of the current landscape. Confidence intervals and statistical significance were calculated with the Cochran-Armitage trend test.

The movement of patients from 1-star to 5-star hospitals was studied to evaluate the maximum potential impact that may be achieved (ie, greatest differential in mortality rate). In addition, a sensitivity analysis was performed in which patients from both 1-star and 2-star hospitals were modeled to change



**Figure 2.** Overall observed 90-day mortality across The Centers for Medicare and Medicaid Services (CMS) Star Rating cohorts. The unadjusted 90-day mortality for each of the CMS Star Rating rated cohorts is represented on the y-axis. The whiskers atop each bar represent the 95% confidence interval, and the reported  $P$  value is for the Cochran-Armitage trend test.

their care to 5-star hospitals. In an attempt to account for differences in patients across 1-star and 5-star hospitals (eg, differences in health or age between patients treated at 1-star and 5-star hospitals), the Peters-Belson approach was used to model the predicted mortality of 1-star patients being treated at 5-star hospitals (37,38). Logistic regression was used to estimate the effects of covariates on mortality at 5-star hospitals, which were then applied to the population who underwent surgery at 1-star hospitals. This method was chosen to demonstrate the impact of reassigning patients. Patients originally from 1-star hospitals, once moved to 5-star hospitals, would be affected by the regression coefficients from the 5-star hospital rather than the original 1-star hospital (regression coefficients are available on request). An estimated mortality difference was calculated between the observed and estimated mortality had the population that was treated at 1-star hospitals instead been treated at 5-star hospitals. Using a primary outcome of 90-day mortality, individual effect sizes and 95% confidence intervals were determined for each procedure. To adjust for the potentially elevated type I error rate due to multiple comparisons, the Bonferroni correction was used. A family-wise error rate of less than 0.05 and a 2-sided  $P$  value of .01 for the five individual cancer surgery types (lobectomy, colectomy, gastrectomy, esophagectomy, and pancreaticoduodenectomy) were considered to be statistically significant for each individual test. This accounted for the total number of 90-day mortality outcomes assessed. All data analysis was conducted with SAS (Cary, NC). All statistical tests were 2-sided.

## Results

### Patient Characteristics

A total of 105 823 patients underwent complex cancer surgery at 3146 hospitals. Overall, patient characteristics across star categories were similar (Table 1). However, a greater proportion of Black patients (11.5% vs 4.6% at 1-star vs 5-star hospitals) and urgent cases (10.0% vs 8.6% at 1-star vs 5-star hospitals) were treated at hospitals with lower star ratings. Overall, the distribution of patients across star categories was heterogeneous, with fewer patients at the extremes (only 7.8% of complex cancer surgeries performed at 1-star hospitals and 15.1% at 5-star hospitals).

### Observed 90-Day Mortality

Unadjusted 90-day mortality decreased with increasing CMS Star Rating, with the greatest difference observed comparing 1-star and 5-star hospitals (1 star = 10.4%, 2 stars = 8.5%, 3 stars = 8.8% 4 stars = 7.6% 5 stars = 6.4%,  $P < .001$ ) (Figure 2). However, the difference in mortality between 1-star and 5-star hospitals varied by surgery type (9.1% for pancreaticoduodenectomy and 2.8% for lobectomy) (Supplementary Figure 1, available online).

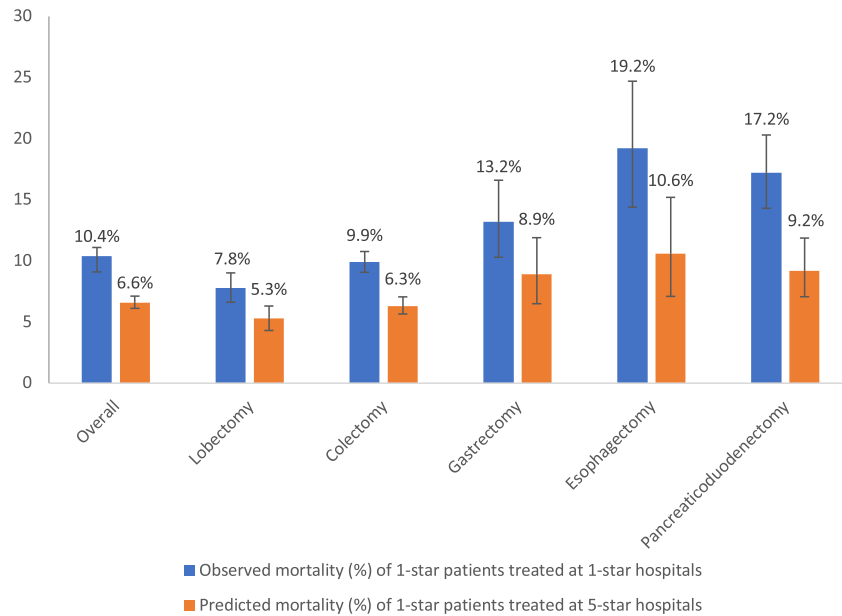
### Modeled Impact of Reassignment Based on CMS Star Rating

In an attempt to estimate the potential impact of using CMS Star Ratings to align patients for complex cancer surgery, an adjusted maximally compliant (all patients changed care) model was created that assumed all patients treated at 1-star hospitals had instead been treated at 5-star hospitals (Figure 3). This resulted in an estimated 84 lives saved per year. The potential number of lives saved per year was variable across procedures, ranging from 5 for gastrectomy to 47 for colectomy (Figure 4). In a regionalization model that moved 1-star hospital patients to 5-star hospitals, 26 patients would need to change hospitals to save a single life. This represents 2189 patients changing hospitals each year to improve outcomes. The efficiency was variable across procedures, ranging from 12 for esophagectomy to 40 for lobectomy (Figure 4).

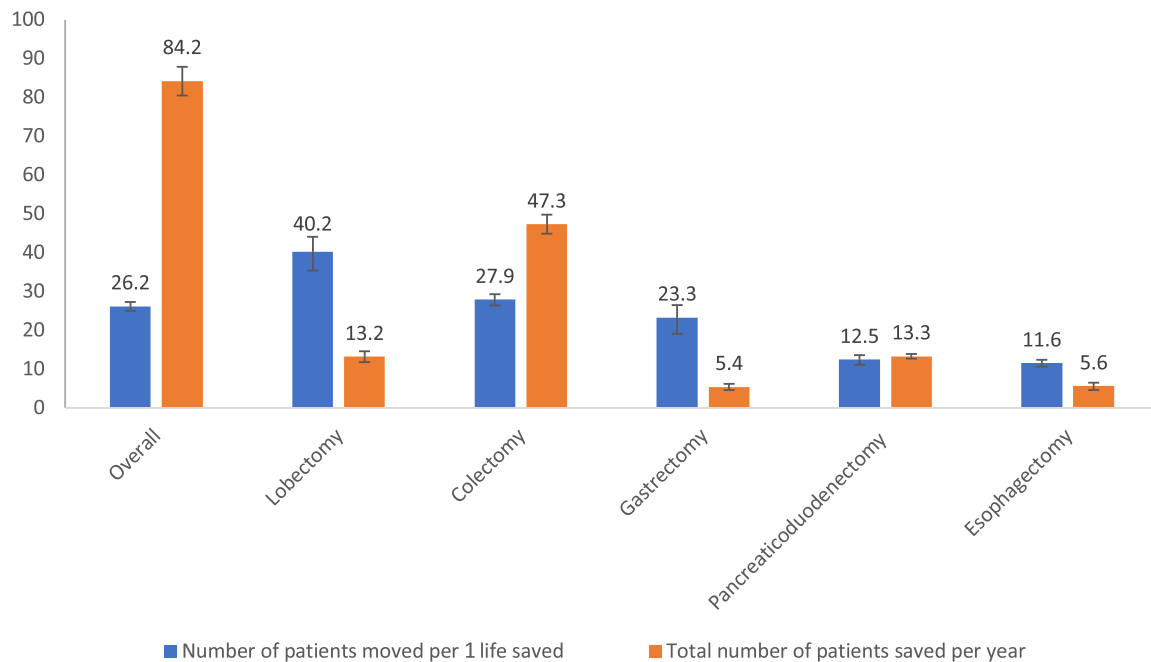
Realizing that a minority (7.8%) of patients were treated at 1-star hospitals, a sensitivity analysis was performed in which patients from both 1-star and 2-star hospitals (capturing 30.8% of study patients) changed their care to 5-star hospitals. An estimated 208 lives per year would be saved. This ranged from 7 per year for esophagectomy to 140 per year for colectomy. Overall, this model required 42 patients to be moved to save 1 life, ranging from 57 for lobectomy to 37 for colectomy (Supplementary Figure 2, available online).

## Discussion

These study findings indicate there is an association between lower all-cause 90-day mortality after complex cancer surgery and increasing CMS Star Rating. This finding is consistent with previous studies that have found CMS Star Ratings to be associated with surgical mortality and safety across a range of clinical



**Figure 3.** Modeled impact of reassignment on mortality. The observed mortality at 1-star hospitals is represented as the blue bar for each of the procedures. The modeled “reassigned mortality” (orange bar) was intended to represent the mortality that would have been observed had the 1-star population been cared for at the 5-star hospitals. The reassigned mortality was determined by applying the risk coefficients for each independent variable achieved at the 5-star hospitals (using logistic regression) to the population of patients treated at 1-star hospitals (see Peters-Belson approach in “Methods”). The whiskers represent 95% confidence intervals.



**Figure 4.** The number of patients moved (1 star to 5 stars) to save 1 life. The number of patients that would need to be moved from 1-star to 5-star hospitals to save a single life is represented by the blue bar. The total number of patients saved in each group per year as a result of this move is represented by the orange bar. The whiskers represent 95% confidence intervals.

scenarios, including cancer surgery (20-24,50-52). These previous studies evaluated factors such as hospital characteristics, including type (community vs academic), size, and location, and patient outcomes, including mortality rates and readmissions. The study by Kaye et al. examined outcomes, including 30-day mortality, complications, length of stay, and readmission, after complex cancer surgery (24). Our study builds on this literature in several important ways. Only one-half of mortalities occur in

the first 30 days; therefore, by studying 90-day mortality, our study could be painting a more complete picture of hospital differences. More important, we then modeled impact of regionalization based on star ratings, which to our knowledge is novel. We combined our findings as an estimate of the potential effectiveness, efficiency, and feasibility of using this rating system to influence mortality through the movement of patients for complex cancer surgeries.



The effectiveness or the overall impact of using CMS Star Rating system to choose hospitals for complex cancer surgery (84 lives saved per year) appears to be modest relative to other proposed realignment strategies. For example, moving patients away from low-volume hospitals is estimated to save 400 lives per year, and using “safety performance” (ie, avoiding hospitals with high standardized mortality ratios) could save 700 lives per year (4). Even if one were to double the lives saved by avoiding 1-star hospitals (because the current study only considers patients older than 65 years), the number of lives saved ( $84 \times 2 = 168$ ) is far less than what is predicted by avoiding the least safe hospitals. Although our sensitivity analysis of avoiding both 1- and 2-star hospitals could save more than twice as many lives (compared with avoiding 1-star hospitals alone), it would require moving 7000 patients each year.

The efficiency (number of patients needed to move to save a single life) of using CMS Star Ratings to identify hospitals with the lowest mortality rates is less than other proposed models. Our study suggests that 26 patients would have to change hospitals (1 star to 5 stars) to save 1 life. Other reassignment models, based on volume or safety performance, predict a larger impact and require as few as 15 patients to change hospitals to save 1 life (4). Given the many barriers that patients face when choosing hospitals, the likelihood of a reassignment strategy working is almost certainly tied to the number of patients required to move (39). On the other, the movement of higher risk patients (including those underdoing esophagectomy and pancreaticoduodenectomy) may actually be more reasonable (but again, this is less effective compared with other models). A critical consideration with any regionalization model would be the potential access implications of preventing a group of hospitals from performing surgery. A number of barriers prevent subsets of patients from being able to travel for complex care, as such, restricting a subset of hospitals from performing procedures could take away a patient’s only realistic option for surgical care. Any attempt to regionalize must include a strategy to mitigate barriers to travel to prevent exaggerating existing disparities in treatment access (39).

The feasibility of reassignment depends on the number of “destination” hospitals the reassignment metric identifies. Not only does the number of destination hospitals affect accessibility (travel time and distance), but each hospital likely has limits on the number of patients they can manage. It would be unrealistic to expect hospitals to handle massive increases in their complex surgical volume. Therefore, the greater number of hospitals that are identified as destination hospitals, the more likely a reassignment strategy is feasible. In the CMS model, 288 hospitals were identified as the “safest” (ie, 5 stars), representing 9.2% of all hospitals. This is in direct comparison with previously discussed performance-based regionalization models, which have identified more than 500 destination hospitals (4,17). This would suggest that the feasibility of reducing mortality using CMS Star Ratings may be lower than expected with performance-based models.

This study has several limitations in addition to those traditionally associated with an observational study design. A key assumption in this analysis is that patients moved from 1-star to 5-star hospitals will experience the same (superior) outcomes as the 5-star cohort. However, there may be differences in patient-specific surgical risk (general health, support) that are not captured by the database but that affect surgical outcomes. We attempted to account for population differences by using the Peters-Belson method as used in other comparative studies in oncology and surgery (37,40,41). Interestingly, the difference

between the adjusted mortality reduction and an unadjusted mortality reduction (ie, observed mortality differential) appeared to be similar, suggesting the patient populations were similar across cohorts. However, potential bias may be due to variables not captured by claims data. This study focuses on the Medicare population (>65 years of age), and it is possible that our findings would be different if a broader range of ages was considered (42). There is incomplete overlap between the study period (2013-2016) and the first consideration of hospital activity for star ratings (2014). This was done to give a better sense of performance at lower volume hospitals (4 years of activity), but it is possible that there were differences in the CMS star metric measures during 2013 that would have caused hospitals to be reclassified as a different star category. This highlights a challenge in that the ratings may not represent real-time performance and safety because there is no guarantee that past performance will predict future performance. In addition, the studied mortality reduction is not the only surgical outcome that is relevant. Important endpoints such as disabling complications (eg, stroke, supplemental oxygen) and long-term survival were not studied and could influence preference for patient treatment. Finally, our study did not evaluate which specific factors contribute to postoperative mortality. Further study into why rates are lower at higher rated centers could provide guidance for lower star-rated hospitals and ultimately improve outcomes without shifting patients.

The current study models a maximally compliant movement of care from 1-star to 5-star hospitals. This perspective is relevant because it demonstrates the total gains from using the CMS Star Rating to inform patients. On the other hand, several important assumptions should be considered in any study that models regionalization. First is the assumption that “destination” hospitals (in this case, the 5-star hospitals) have the capacity to care for additional patients. The additional case-load could overwhelm hospital infrastructure and jeopardize both the patients that move as well as patients who were destined to receive care there initially. Next, patients are assumed to travel to a destination hospital, but in reality, specific barriers (eg, transportation resources, costs) may restrict patients from participating in regionalization. Fortunately, most patients are willing to travel if these barriers are addressed (39,43-47). Finally, an important assumption is that 1-star hospital patients would be motivated by the mortality differences to change hospitals. Based on previous work, 80%-90% of patients would be motivated to travel based on a potential mortality reduction (39). There are many other complex considerations in how a patient chooses a hospital for care, including personal motivators and clinical variables such as physician and therapy availability (39,43-47). Ultimately, patients likely differ in the extent that travelling to a particularly safe hospital fits with their overarching priorities and goals of care.

CMS Star Ratings are associated with mortality among Medicare recipients undergoing complex cancer surgery. However, using CMS Star Ratings to identify hospitals for complex cancer surgery is not as effective, efficient, or feasible as other proposed models for identifying the safest hospitals. Additional efforts to inform patients regarding the safest choice for complex cancer surgery and align hospital choice with goals of care are needed.

## Funding

None.

## Notes

**Disclosures:** There are no conflicts of interest to disclose.

**Author contributions:** MVP: Study design, data acquisition, data analysis, data interpretation, manuscript drafting, critical revisions. BJR: Study design, data interpretation, manuscript drafting, critical revisions. AFM: Study design, data interpretation, manuscript drafting, critical revisions. MC: Data analysis, data interpretation, manuscript drafting, critical revisions. RP: Study design, data interpretation, critical revisions. VJM: Study design, data interpretation, critical revisions. APD: Study design, data interpretation, critical revisions. JRH: Study design, data acquisition, data analysis, data interpretation. JDB: Study design, data interpretation, critical revisions. DJB: Study design, data interpretation, manuscript drafting, critical revisions.

## References

- Fisher E, Wennberg D, Stukel T, et al The implications of regional variations in Medicare spending. *Ann Intern Med.* 2003;138(4):273–287.
- Birkmeyer JD, Sun Y, Goldfaden A, et al Volume and process of care in high-risk cancer surgery. *Cancer.* 2006;106(11):2476–2481.
- Etzioni DA, Young-Fadok TM, Cima RR, et al Patient survival after surgical treatment of rectal cancer: impact of surgeon and hospital characteristics. *Cancer.* 2014;120(16):2472–2481.
- Chiu AS, Arnold BN, Hoag JR, et al Quality versus quantity: the potential impact of public reporting of hospital safety for complex cancer surgery. *Ann Surg.* 2019;270(2):281–287.
- Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA.* 1998;280(20):1747–1751.
- Birkmeyer JD, Finlayson SR, Tosteson AN, et al Effect of hospital volume on in-hospital mortality with pancreaticoduodenectomy. *Surgery.* 1999;125(3):250–256.
- Hannan EL, Radzyner M, Rubin D, et al The influence of hospital and surgeon volume on in-hospital mortality for colectomy, gastrectomy, and lung lobectomy in patients with cancer. *Surgery.* 2002;131(1):6–15.
- Schrag D, Panageas KS, Riedel E, et al Hospital and surgeon procedure volume as predictors of outcome following rectal cancer resection. *Ann Surg.* 2002;236(5):583–592.
- Sosa JA, Bowman HM, Tielsch JM, et al The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Ann Surg.* 1998;228(3):320–330.
- Porter GA, Soskolne CL, Yakimets WW, Newman SC. Surgeon-related factors and outcomes in rectal cancer. *Ann Surg.* 1998;227(2):157–167.
- Sahni NR, Dalton M, Cutler DM, et al Surgeon specialization and operative mortality in United States: retrospective analysis. *BMJ.* 2016;354:i3571.
- Arnold BN, Chiu AS, Hoag JR, et al Spontaneous regionalization of esophageal cancer surgery: an analysis of the National Cancer Database. *J Thorac Dis.* 2018;10(3):1721–1731.
- Birkmeyer JD, Siewers AE, Marth NJ, Goodman DC. Regionalization of high-risk surgery and implications for patient travel times. *JAMA.* 2003;290(20):2703–2708.
- Schlottmann F, Strassle PD, Charles AG, Patti MG. Esophageal cancer surgery: spontaneous centralization in the US contributed to reduce mortality without causing health disparities. *Ann Surg Oncol.* 2018;25(6):1580–1587.
- Birkmeyer JD, Siewers AE, Finlayson EV, et al Hospital volume and surgical mortality in the United States. *N Engl J Med.* 2002;346(15):1128–1137.
- Austin JM, D'Andrea G, Birkmeyer JD, et al Safety in numbers: the development of Leapfrog's composite patient safety score for US hospitals. *J Patient Saf.* 2014;10(1):64–71.
- Birkmeyer JD, Dimick JB. Potential benefits of the new Leapfrog standards: effect of process and outcome measures. *Surgery.* 2004;135(6):569–575.
- Urbach D. Pledging to eliminate low-volume surgery. *N Engl J Med.* 2015;373(15):1388–1390.
- CMS. Overall hospital quality star rating on Hospital Compare methodology report. CMS.gov. December 2017. 2019. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Overall-Hospital-Quality-Star-Rating-on-Hospital-Compare-Public-Input-Period.pdf>. Accessed April 20, 2020.
- Wang DE, Tsugawa Y, Figueroa JF, Jha AK. Association between the Centers for Medicare and Medicaid Services Hospital star rating and patient outcomes. *JAMA Intern Med.* 2016;176(6):848–850.
- Bilimoria KY, Barnard C. The new CMS hospital quality star ratings: the stars are not aligned. *JAMA.* 2016;316(17):1761–1762.
- Delancey JO, Softcheck J, Chung JW, et al Associations between hospital characteristics, measure reporting, and the centers for Medicare and Medicaid services overall hospital quality star ratings. *JAMA.* 2017;317(19):2015–2017.
- Chatterjee P, Joynt Maddox K. Patterns of performance and improvement in US Medicare's Hospital Star Ratings, 2016–2017. *BMJ Qual Saf.* 2019;28(6):486–494.
- Kaye DR, Norton EC, Ellimoottil C, et al Understanding the relationship between CMS' Hospital Compare Star rating, surgical case volume, and short-term outcomes after major cancer surgery. *Cancer.* 2017;123(21):4259–4267.
- In H, Palis BE, Merkow RP, et al Doubling of 30-day mortality by 90 days after esophagectomy: a critical measure of outcomes for quality improvement. *Ann Surg.* 2016;263(2):286–291.
- Adam MA, Turner MC, Sun Z, et al The appropriateness of 30-day mortality as a quality metric in colorectal cancer surgery. *Am J Surg.* 2018;215(1):66–70.
- Amazon. medicare.gov Competitive Analysis, Marketing Mix and Traffic. Alexa. [https://www.alexa.com/siteinfo/medicare.gov#section\\_traffic](https://www.alexa.com/siteinfo/medicare.gov#section_traffic). Accessed January 4, 2020.
- BlueCross BlueShield. BlueCross BlueShield's star rating increases for Medicare Advantage plan. <https://www.bcbs.com/news/press-releases/blue-cross-blueshields-star-rating-increases-medicare-advantage-plans>. Accessed January 4, 2020.
- Healthpayer Intelligence. 5 strategies for 5 stars: Cigna's approach to CMS Star Ratings. <https://healthpayerintelligence.com/features/5-strategies-for-5-stars-from-cigna>. Accessed January 4, 2020.
- Selecthealth physician. Medicare Advantage Star Ratings. <https://intermountainphysician.org/selecthealth/fehb/advantagereference%20manual/Pages/Stars-Ratings.aspx>. Accessed January 4, 2020.
- Cape Cod HealthCare. Cape Cod hospital earns Becker's Top 100 honor for the ninth straight year. <https://www.capecodhealth.org/about/press-room/news/2019/cch-earns-beckers-top-100-for-the-9th-straight-year/>. Accessed January 4, 2020.
- Vanderbilt University Medical Center. VUMC named among "100 great hospitals in America." <http://news.vumc.org/2019/04/23/vumc-named-among-100-great-hospitals-in-america/>. Accessed January 4, 2020.
- Becker's Hospital Review. The 293 hospitals with 5 stars from CMS. <https://www.beckershospitalreview.com/rankings-and-ratings/the-293-hospitals-with-5-stars-from-cms.html>. Accessed January 4, 2020.
- CMS.gov. MEDPAR. <https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicarefeeforsvcpartsab/medpar.html>. Accessed July 27, 2019.
- http://Medicare.gov. About Hospital Compare data. <https://www.medicare.gov/hospitalcompare/data/about.html>. Accessed July 27, 2019.
- Schwarze ML, Barnato AE, Rathouz PJ, et al What is high risk surgery? Development of a list of high risk operations for patients age 65 and older. *JAMA Surg.* 2015;150(4):325–331.
- Graubard BI, Sowmya Rao R, Gastwirth JL. Using the Peters-Belson method to measure health care disparities from complex survey data. *Stat Med.* 2005;24(17):2659–2668.
- Li Y, Graubard BI, Huang P, Gastwirth JL. Extension of the Peters-Belson method to estimate health disparities among multiple groups using logistic regression with survey data. *Stat Med.* 2015;34(4):595–612.
- Resio BJ, Chiu AS, Hoag JR, et al Motivators, barriers, and facilitators to traveling to the safest hospitals in the United States for complex cancer surgery. *JAMA Netw Open.* 2018;1(7):e184595.
- Rao SR, Breen N, Graubard BI. Trends in Black-White disparities in breast and colorectal cancer screening rates in changing screening environment: the Peters-Belson approach using United States National Health Interview Surveys 2000–2010. *Med Care.* 2016;54(2):133–139.
- Sathianathan NJ, Jarosek SL, Fan Y, Krishna SR, Konety BR. Racial disparities in surgical outcomes among males following major urologic cancer surgery. *Am J Prev Med.* 2018;55(5):S14–21.
- Massarweh NN, Legner VJ, Symons RG. Impact of advancing age on abdominal surgical outcomes. *Arch Surg.* 2009;144(12):1108–1114.
- Smith AK, Shara NM, Zeymo A, et al Travel patterns of cancer surgery patients in a regionalized system. *J Surg Res.* 2015;199(1):97–105.
- Kahn JM, Asch RJ, Iwashyna TJ, et al Perceived barriers to the regionalization of adult critical care in the United States: a qualitative preliminary study. *BMC Health Serv Res.* 2008;8(1):239.
- Finlayson SR, Birkmeyer JD, Tosteson AN, Nease RF. Patient preferences for location of care: implications for regionalization. *Med Care.* 1999;37(2):204–209.
- Koster ES, Schmidt A, Philbert D, et al Health literacy of patients admitted for elective surgery. *J Public Health.* 2017;25(2):181–186.
- Wu RR, Myers RA, Buchanan AH, et al Effect of sociodemographic factors on uptake of a patient-facing information technology family health history risk assessment platform. *Appl Clin Inform.* 2019;10(02):180–188.
- Centers for Medicare and Medicaid Services (CMS). *Overall Hospital Quality Star Ratings on Hospital Compare Methodology Report (v2.0)*. Baltimore, MD: Centers for Medicare and Medicaid Services; 2016.

49. <http://CMS.gov>. First release of the overall quality star rating on Hospital Compare. <https://www.cms.gov/newsroom/fact-sheets/first-release-overall-hospital-quality-star-rating-hospital-compare>. Accessed July 27, 2019.
50. Koh CY, Inaba CS, Sujatha-Bhaskar S, Nguyen NT. Association of Centers of Medicare and Medicaid Services overall hospital quality star rating with outcomes in advanced laparoscopic abdominal surgery. *JAMA Surg*. 2017;152(12):1113–1117.
51. Courtney PM, Frisch NB, Bohl DD, Della Valle CJ. Improving value in total hip and knee arthroplasty: the role of high volume hospitals. *J Arthroplasty*. 2018; 33(1):1–5.
52. Chung JW, Dahlke AR, Barnard C, DeLancey JO, Merkow RP, Bilimoria KY. The Centers of Medicare and Medicaid Services hospital ratings: pitfalls of grading on a single curve. *Health Aff*. 2019;38(9):1523–1529.