

# Properties of the Overall Hospital Star Ratings and Consumer Choice

Jianhui Hu, PhD; and David R. Nerenz, PhD

**T**he CMS Overall Hospital Quality Star Ratings<sup>1</sup> (referred to subsequently as the “Star Ratings”) were created to provide consumers with a simple, easy-to-understand summary of the measures available in the Hospital Compare measure set.<sup>2</sup> Under the current method, a set of 57 individual measures of quality are categorized into 7 quality groups or domains. Scores are calculated using a statistical model for each of the 7 groups and are weighted and combined to yield an overall hospital rating of 1 to 5 stars. Although concerns have been raised about technical aspects of the Star Ratings<sup>3</sup> and the accuracy and fairness of the ratings,<sup>4</sup> a variety of consumer advocacy and large purchaser groups have reported that they find the Star Ratings useful.<sup>5</sup>

An overall hospital rating or ranking system could be used in either of 2 ways. It can be used by purchasers or health plans that are designing limited provider networks and whose interest in hospital quality spans the entire range of clinical specialties and programs. An overall rating that brings together scores on multiple individual measures of quality simplifies the decision about inclusion in the network, and it allows a plan or purchaser or patient advocacy group to make recommendations to its members or clients about “good” vs “bad” hospitals for purposes of their choices.<sup>6</sup>

An overall rating system could also be used by individual patients or consumers for choices of hospitals for specific needs—elective surgical procedures being a common example. In this case, the value of any overall rating system can be questioned because it inevitably includes information on aspects of care that are not relevant to the patient’s specific condition. Available data on the small amount of correlation among specific Hospital Compare measures<sup>7</sup> suggest that an overall rating may not be informative about relative hospital performance for specific conditions or procedures.

Existing literature examining associations between Star Ratings and outcomes of cancer surgery,<sup>8-10</sup> advanced laparoscopic abdominal surgery,<sup>11</sup> and percutaneous coronary intervention (PCI)<sup>12</sup> demonstrates mixed results. For example, Star Ratings seemed to be significantly associated with cancer surgery outcomes,<sup>8-10</sup> although the efficiency, effectiveness, and feasibility of patients’ use of Star Ratings to identify high-quality hospitals for complex

## ABSTRACT

**OBJECTIVES:** To examine characteristics of the CMS Overall Hospital Quality Star Ratings related to their use by consumers for choosing hospitals.

**STUDY DESIGN:** Observational study using secondary data analyses.

**METHODS:** Hospital Star Rating data reported in February 2019 and additional quality data from California and New York were used, with a mix of analytical approaches including descriptive statistics, correlational analysis, and Poisson regression models.

**RESULTS:** The distribution of hospitals’ Star Rating summary scores was tightly compressed, with no hospitals at or near the scores that would be obtained if a hospital were either best or worst across all quality domains. Hospitals did not consistently perform well or poorly across the range of measures and quality groups included in the Star Ratings. On average, for a given quality measure included in the Star Rating program, 12% of 1-star hospitals received top-quartile scores and 16% of 5-star hospitals received bottom-quartile scores. No significant associations were found between hospitals’ overall Star Ratings and their performance on a set of condition-specific quality measures for hospitals in California and New York State.

**CONCLUSIONS:** Hospitals’ overall scores clustered in the middle of the potential distribution of scores; no hospitals were either best at everything or worst at everything. The Star Ratings did not predict hospital quality scores for separate quality measures related to specific medical conditions or health care needs. These 2 observations suggest that the Star Ratings are of limited value to consumers choosing hospitals for specific care needs.

*Am J Manag Care.* 2021;27(5):203-210. doi:10.37765/ajmc.2021.88634

## TAKEAWAY POINTS

The CMS Star Ratings were created to provide a simple guide to patients or consumers for choices of hospitals. For patients making a choice for a specific medical condition or health care need:

- ▶ the distribution of overall scores was very compressed, and hospitals did not consistently perform well or poorly across the range of measures and quality groups in the Star Rating program; and
- ▶ Star Ratings did not predict scores for other quality measures related to patients' specific medical conditions or health care needs.

cancer surgeries was questioned.<sup>10</sup> On the other hand, no significant associations were found between Star Ratings and mortality in laparoscopic abdominal surgeries, serious morbidity for bariatric or hiatal hernia surgeries, or 30-day risk-standardized mortality after PCI. However, serious morbidity for colorectal surgery and 30-day post-PCI readmission rate were lower at high-star hospitals compared with low-star hospitals.<sup>11,12</sup>

We attempted to address 3 questions:

1. Given the potential range of scores for a hospital best at everything and a hospital worst at everything, how different are hospitals from each other?
2. What percentage of 5-star hospitals receive low scores on either specific quality groups or individual measures, and what percentage of 1-star hospitals receive high scores on either specific quality groups or individual measures?
3. To what extent do the Star Ratings predict measures of quality for specific clinical programs or services that are not represented in the Star Ratings but might be relevant to patient decisions about where to seek care?

## METHODS

### Data

We used data from the CMS Overall Hospital Quality Star Rating Program and additional quality data from the states of California and New York. We ran the publicly available SAS Pack (SAS Institute) for the February 2019 Star Rating report period to obtain hospitals' overall Star Ratings, scores on 7 quality groups (Mortality, Readmission, Safety of Care, Patient Experience, Effectiveness of Care, Timeliness of Care, and Efficient Use of Medical Imaging), and scores on 57 individual quality measures.<sup>13</sup> (Quality groups, quality measure identifiers, and quality measure names can be found in [eAppendix A](#) [eAppendices available at [ajmc.com](http://ajmc.com)].)

Additional data in the quality domains of mortality and surgical site infection (SSI) were obtained from California and New York. We selected quality measures that were not included in the February 2019 Star Ratings. For hospitals in California, mortality rates for 13 medical conditions and procedures were obtained from California's Office of Statewide Health Planning and Development<sup>14</sup>; SSIs for 17 individual operative procedures as well as SSIs for all

types of procedures were obtained from the California Health and Human Services Open Data Portal.<sup>15</sup> For the state of New York, mortality rates for 7 medical conditions and procedures, SSIs for 3 types of operative procedures, and 4 carbapenem-resistant Enterobacteriaceae (CRE)-related measures were obtained from Health Data NY, a State of New York Open Data Portal.<sup>16</sup> All mortality and SSI rates were risk adjusted. (See [eAppendix B](#) for details of the state-level quality measures.)

### Analytical Approaches

To answer the initial research question, we first created a distribution of hospitals' summary scores, which were the sum of the weighted group scores for each hospital. We included hospitals with group scores for all 7 quality groups. We also calculated 2 "benchmark" summary scores using the lowest and the highest group scores for each quality group, respectively, to represent a "potential lowest" summary score and a "potential highest" summary score that a hypothetical hospital could have obtained, had it performed the worst or the best on all quality groups. For example, if a hospital had the best score in all 7 groups, it would receive the high benchmark score. On the other hand, if a hospital had the worst score in all 7 groups, it would receive the low benchmark score. The best or worst measure group score was the best or worst group score actually obtained by a real hospital—not a hypothetical best or worst score based on performance on individual measures within groups.

To address the second question, we first calculated percentages of hospitals receiving top-quartile and bottom-quartile scores on 7 quality groups, by their Star Rating. We then calculated percentages of hospitals receiving top-quartile and bottom-quartile scores on 50 quality measures, by Star Rating. Seven measures that could not be categorized into quartiles due to their distributions (HAI-4, OP-4, OP-30, OP-33, OP-11, PC-01, and VTE-6, defined in [eAppendix A](#)) were excluded. Differences between the 5 Star Rating groups were compared using the  $\chi^2$  test.

To address the third question, we examined the relationship between hospitals' Star Ratings and their performance on individual quality measures, not in Hospital Compare and not included in the Star Rating system, from New York and California. Any hospitals without a Star Rating or a quality measure value were excluded.

We took 3 analytical approaches: First, a correlational analysis was conducted using the Spearman rank correlations between overall Star Ratings and individual quality measures. Second, we compared mean quality measure rates for 3 groups of hospitals, categorized by their Star Ratings: hospitals receiving 1 or 2 stars, hospitals receiving 3 stars, and hospitals receiving 4 or 5 stars. Differences between the 3 resulting groups were tested using the Kruskal-Wallis test. We reported the raw *P* values and *P* values adjusted for multiple comparisons using the Sidak-Holm method. Third, to take into

consideration the different case volumes for each quality measure at different hospitals, we ran Poisson regression models with hospitals' Star Ratings (1-5) as the dependent variable and hospitals' quality measure rates and case volumes as independent variables.

All data management and statistical analyses were performed using SAS version 9.4 (SAS Institute) and Stata/SE version 13 (StataCorp).

## RESULTS

Numbers of included hospitals in different parts of the analyses are presented in [eAppendix C](#), with a range of 467 to 4408 hospitals in national-level analyses and a range of 20 to 313 hospitals in state-level analyses. We included both acute care and critical access care hospitals in all analyses, although additional analyses including only acute care hospitals yielded almost identical results.

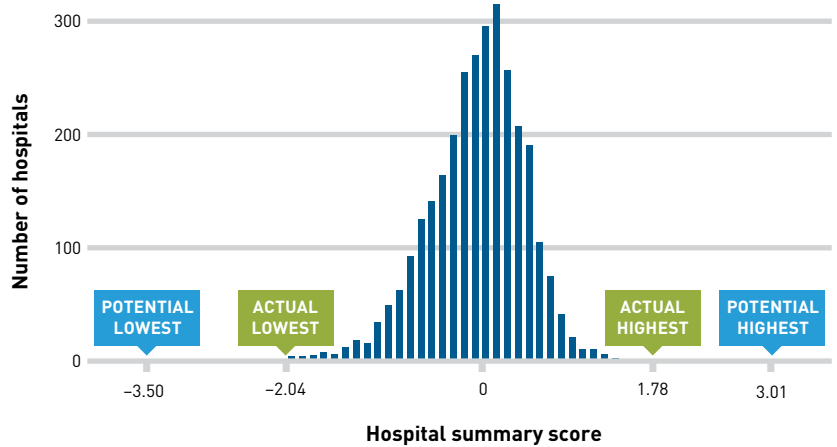
Hospitals' summary score distribution is presented in [Figure 1](#). The score that would be obtained by a hospital with the actual observed worst score on every quality group is identified as "potential lowest" at -3.50. The score that would be obtained by a hospital with the actual observed best score on every quality group is identified as "potential highest" at 3.01.

The range of actual scores was much narrower, with the vast majority of hospitals clustered at or near the middle of the score distribution (mean actual score, -0.05; lowest actual score, -2.04; highest actual score, 1.78). Only a handful of high-scoring hospitals scored at or beyond half of the potential highest score above the midpoint of 0, and there were also only a handful of low-scoring hospitals that were at or beyond half of the potential lowest score.

There were no obvious break points or clusters in the distribution, which was smooth and continuous throughout the range.

The percentage of hospitals with top- and bottom-quartile scores at the quality group level, for hospitals in each of the 5 overall Star Rating groups, is shown in [Table 1](#). As expected because of the design of the Star Rating system, hospitals with high overall Star Ratings were more likely to have group scores in the top quartile, but the strength of that association varied by quality group. The relationship was clear in 3 of the 4 highly weighted measure groups (Readmission, Safety of Care, and Patient Experience), but less strong in the other 3 measure groups and in the Mortality group. In the Mortality group, the largest difference was between 5-star and 4-star hospitals; there were much smaller differences among the 3 other Star Rating groups. In the Effectiveness of Care group, only 43% of 5-star hospitals had scores in the top quartile, whereas 23% of 1-star hospitals had scores in the top quartile. For Efficient Use of Medical Imaging, there was no relationship between overall star rating and scores in the top quartile for the measure group.

**FIGURE 1.** Distribution of Hospitals' Star Rating Summary Scores



**TABLE 1.** Percentages of Hospitals Receiving Top-Quartile and Bottom-Quartile Scores on 7 Quality Groups, by Their Stars

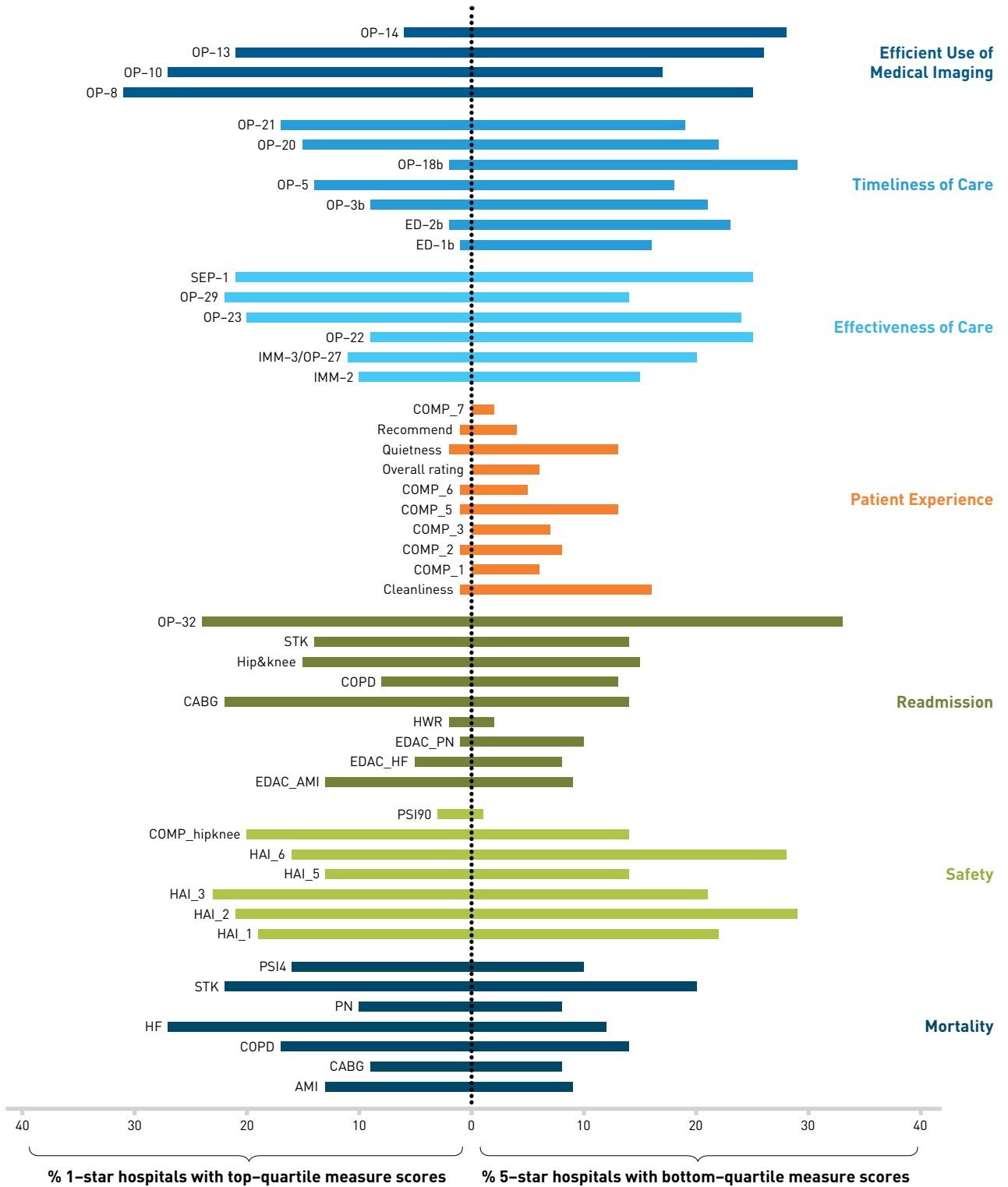
Percentage of hospitals receiving top-quartile scores						
Quality groups	Stars					P
	1	2	3	4	5	
Mortality	19	25	23	30	61	<.001
Safety of Care	5	12	20	34	70	<.001
Readmission	2	8	19	43	77	<.001
Patient Experience	0	3	13	44	60	<.001
Effectiveness of Care	23	27	27	33	43	<.001
Timeliness of Care	1	11	22	27	22	<.001
Efficient Use of Medical Imaging	30	27	25	27	27	.44

Percentage of hospitals receiving bottom-quartile scores						
Quality groups	Stars					P
	1	2	3	4	5	
Mortality	46	40	28	17	7	<.001
Safety of Care	82	46	23	8	1	<.001
Readmission	82	55	20	8	2	<.001
Patient Experience	79	52	19	6	2	<.001
Effectiveness of Care	41	33	27	17	11	<.001
Timeliness of Care	66	44	23	15	17	<.001
Efficient Use of Medical Imaging	31	32	28	20	16	<.001

The results of a similar analysis at the individual measure level are shown in [Figure 2](#). Figure 2 shows the percentage of 1-star hospitals with a top-quartile measure score (left side of figure), and the percentage of 5-star hospitals with a bottom-quartile measure score (right side of figure). There is clear variability from

**FIGURE 2.** Percentages of 1-Star Hospitals Receiving Top-Quartile Scores and Percentages of 5-Star Hospitals Receiving Bottom-Quartile Scores on 50 Quality Measures



**TABLE 2.** Correlation Associations Between Hospitals' Star Ratings and Mortality and Surgical Site Infection Rates in California and New York

California	rho	P		New York	rho	P		
		(raw)	(adjusted) <sup>a</sup>			(raw)	(adjusted) <sup>a</sup>	
<b>Mortality</b>								
AAA repair: unruptured endovascular	0.17	.06	.38	AAA repair: unruptured endovascular	0.25	.28	.73	
Acute stroke	-0.08	.19	.66	AAA repair	0.04	.85	.85	
Acute stroke hemorrhagic	-0.12	.08	.44	Carotid endarterectomy	-0.10	.53	.78	
Acute stroke subarachnoid	-0.22	.02	.23	Craniotomy	-0.39	.01	.08	
Carotid endarterectomy	0.00	.96	.97	Gastrointestinal hemorrhage	-0.19	.02	.09	
Craniotomy	-0.27	<.01	.01	Hip fracture	-0.10	.26	.73	
Esophageal resection	0.36	.04	.35	Hip replacement	-0.12	.23	.73	
Gastrointestinal hemorrhage	-0.08	.16	.66					
Hip fracture	-0.01	.84	.97					
Percutaneous coronary intervention	-0.11	.19	.66					
Pancreatic cancer	0.08	.55	.91					
Pancreatic other	-0.27	.04	.35					
Pancreatic resection	-0.26	.02	.23					
<b>Surgical site infections</b>								
All	-0.01	.84	>.99	CRE admission prevalence rate: all body sites	-0.25	<.01	.01	
Appendix surgery	-0.06	.43	.99	CRE admission prevalence rate: bloodstream infections	-0.25	<.01	.01	
Bile duct, liver, or pancreatic surgery	0.09	.33	.99	CRE hospital onset infection rate: all body sites	-0.31	<.001	<.001	
Cardiac surgery	0.17	.18	.95	CRE hospital onset infection rate: bloodstream infections	-0.14	.08	.27	
Cesarean section	0.14	.06	.65	Hip prosthesis	-0.03	.76	.94	
Coronary bypass: chest and donor incisions	0.26	.01	.12	Coronary artery bypass graft: chest site	-0.27	.11	.30	
Exploratory abdominal surgery	0.07	.34	.99	Coronary artery bypass graft: donor site	-0.05	.80	.94	
Gallbladder surgery	0.03	.67	>.99					
Gastric surgery	0.10	.22	.96					
Hip prosthesis	0.03	.61	>.99					
Knee prosthesis	0.03	.71	>.99					
Laminectomy	0.14	.18	.95					
Open reduction of fracture	0.02	.80	>.99					
Pacemaker surgery	0.14	.28	.98					
Rectal surgery	0.18	.06	.66					
Small bowel surgery	-0.06	.34	.99					
Spinal fusion	-0.04	.58	>.99					
Thoracic surgery	0.04	.75	>.99					

AAA, abdominal aortic aneurysm; CRE, carbapenem-resistant Enterobacteriaceae.

<sup>a</sup>P values were adjusted for multiple comparisons using the Sidak-Holm method.

measure to measure, with some measures having essentially no “crossover” ratings (high Star Rating and scores in lowest quartile or vice versa), and others with more than 30% of hospitals with crossover ratings. On average across the 50 quality measures, there were 12% of 1-star hospitals with ratings for a given measure in the top quartile and 16% of 5-star hospitals with ratings for a given measure in the bottom quartile. (Appendix D reports all data on all 5 hospital star groups.)

Table 2 and Table 3 present results of the relationships between overall Star Ratings and individual quality measures using California and New York quality data. Table 2 reports the results of the correlational analysis. Of the 31 measures available from California, there were significant associations ( $P < .05$ ) in the unadjusted values for 6 measures, but only 1 of those associations (craniotomy mortality) remained significant in the adjusted analysis. Of the 14 measures available for New York, there were 5 significant associations in the

**TABLE 3.** Mean Mortality and Surgical Site Infection Rates by Hospital Star Rating Groups: California and New York

California	1-2 stars	3 stars	4-5 stars	P (raw)	P (adjusted) <sup>a</sup>	New York	1-2 stars	3 stars	4-5 stars	P (raw)	P (adjusted) <sup>a</sup>
<b>Mortality</b>											
AAA repair: unruptured endovascular	2.3	4.1	4.3	.17	.74	AAA repair: unruptured endovascular	3.2	6.3	1.9	.40	.92
Acute stroke	6.7	7.1	7.0	.57	.86	AAA repair	6.4	5.7	5.7	.94	.95
Acute stroke hemorrhagic	17.9	15.6	16.8	.12	.72	Carotid endarterectomy	0.3	0.3	0.0	.63	.95
Acute stroke subarachnoid	19.8	15.2	16.7	.14	.74	Craniotomy	8.9	3.5	8.1	<.01	.02
Carotid endarterectomy	0.7	0.2	0.6	.14	.74	Gastrointestinal hemorrhage	4.6	3.6	4.8	.06	.30
Craniotomy	11.1	7.8	5.8	.02	.20	Hip fracture	6.1	5.8	6.9	.73	.95
Esophageal resection	0.0	4.5	16.7	.18	.74	Hip replacement	0.2	0.2	0.0	.47	.92
Gastrointestinal hemorrhage	2.2	1.6	1.6	.39	.86						
Hip fracture	3.1	2.7	2.1	.98	.98						
Percutaneous coronary intervention	3.8	3.9	3.6	.44	.86						
Pancreatic cancer	0.7	4.9	1.9	.31	.85						
Pancreatic other	7.7	0.5	1.8	.08	.62						
Pancreatic resection	11.5	3.1	2.1	.08	.62						
<b>Surgical site infections</b>											
All	0.9	0.9	0.9	.94	>.99	CRE admission prevalence rate: all body sites	0.5	0.3	0.2	.06	.28
Appendix surgery	1.2	1.0	0.7	.56	>.99	CRE admission prevalence rate: bloodstream infections	0.1	0.0	0.0	.02	.13
Bile duct, liver, or pancreatic surgery	1.1	0.8	1.3	.57	>.99	CRE hospital onset infection rate: all body sites	0.6	0.3	0.2	.02	.12
Cardiac surgery	0.3	0.4	0.8	.14	.91	CRE hospital onset infection rate: bloodstream infections	0.1	0.1	0.0	.36	.83
Cesarean section	0.5	1.3	0.9	.01	.18	Hip prosthesis	1.1	0.9	0.7	.65	.91
Coronary bypass: chest and donor incisions	1.0	0.9	1.5	.01	.17	Coronary artery bypass graft: chest site	1.4	1.1	0.8	.56	.91
Exploratory abdominal surgery	1.1	1.0	1.1	.59	>.99	Coronary artery bypass graft: donor site	0.3	0.2	0.3	.92	.92
Gallbladder surgery	1.1	1.2	1.0	.88	>.99						
Gastric surgery	0.7	0.7	0.6	.72	>.99						
Hip prosthesis	1.0	1.1	1.3	.44	>.99						
Knee prosthesis	0.9	1.3	0.9	.30	.99						
Laminectomy	1.0	0.9	1.3	.18	.95						
Open reduction of fracture	1.0	0.8	1.0	.74	>.99						
Pacemaker surgery	0.8	1.3	1.3	.47	>.99						
Rectal surgery	0.5	0.2	0.6	.26	.98						
Small bowel surgery	0.9	0.8	0.7	.37	>.99						
Spinal fusion	1.0	0.8	0.9	.71	>.99						
Thoracic surgery	0.9	0.7	0.5	.94	>.99						

AAA, abdominal aortic aneurysm; CRE, carbapenem-resistant Enterobacteriaceae.

<sup>a</sup>P values were adjusted for multiple comparisons using the Sidak-Holm method.

unadjusted values; 3 were still significant in the adjusted analyses. All 3 of those measures were for care related to CRE.

The same relationships are shown in Table 3, but in this case the table entries are the mean values on specific measures for 3 sets of hospitals: those with 1- or 2-star CMS ratings, those with 3-star ratings, and those with 4- or 5-star ratings. As in the correlational analysis, there were essentially no significant relationships found in the analyses adjusted for multiple comparisons. The only relationship that was found in the adjusted analysis (craniotomy mortality in New York) showed an interesting pattern in which the lowest mortality was found in 3-star hospitals rather than in 4- or 5-star hospitals.

Results from the Poisson regressions controlling for case volume are presented in [Appendix E](#). Similar to the above state-level analysis, relationships between star ratings and individual quality measures were all nonsignificant, the only exception being the “CRE Hospital Onset Infection Rate—all body sites” measure (incidence rate ratio, 0.79;  $P = .017$ ).

## DISCUSSION

Given the range of possible total summary scores in the Star Ratings, the range of actual scores was narrow. The shape of the curve is to be expected as the Star Ratings methodology ensures a normal distribution of hospital ratings. However, the extent to which the curve is compressed within the potential range of scores vs spread across that whole range would not seem to be an intentional design feature, as a wider distribution across the full possible scale range would be more useful for consumer choice. CMS has been making changes to different components of the methods over the years, however, trying to “flatten” the distribution bell curve and assign more hospitals in the 1-star and 5-star categories.<sup>17</sup>

The large gaps between the actual highest/lowest scores and the potential highest/lowest scores seemed noteworthy. The actual best hospital total score was 1.78; a hospital with the best score on all 7 quality groups would have scored 3.01. On the other side of the distribution, the actual worst hospital score was  $-2.04$ ; a hospital with the worst score on all 7 quality groups would have scored  $-3.50$ . The median 5-star hospital score was 0.73, which is much closer to the overall hospital mean of  $-0.05$  than to either the best actual hospital score (1.78) or the “best possible” score for a hypothetical hospital with the best scores in all quality groups at 3.01. Similarly, the median 1-star hospital score was  $-0.95$ , which is closer to the overall mean than to either the score of the actual worst-scoring hospital ( $-2.04$ ) or the “worst possible” hospital score of  $-3.50$ . Choosing a “high-performing” hospital or avoiding a “poor-performing” hospital just on the basis of the Star Ratings alone, then, is likely to yield a choice that is not much different from average in terms of total score.

Many hospitals at the top of the summary score distribution had low scores in some measure groups, and many hospitals at the bottom of the score distribution had high scores in some measure groups. A consumer choosing a 4-star hospital on the basis of that

rating, for example, will not find consistent “4-star quality” in all measure groups or in all specific individual quality measures.

There were also no obvious break points in the distribution of overall scores. Although CMS uses a method to divide hospitals into 5 groups and assign Star Ratings to those groups (“k-means clustering”),<sup>18</sup> hospitals do not naturally cluster into groups in terms of the overall scores.

There clearly are hospitals with relatively good performance on multiple quality groups, which thereby earn a 5-star rating. One characteristic of those hospitals is that they report on relatively fewer measures and do not have scores in all quality groups, including the important quality groups of Safety of Care and Mortality.<sup>19</sup>

We find, however, that many 1-star hospitals have top scores in important areas such as Safety of Care, Mortality, or Patient Experience. Likewise, some 5-star hospitals have low scores in 1 or more of the 7 quality groups. When we examine correlations between the overall Star Ratings and external, independent measures of quality of care for specific clinical conditions (the California and New York measures), we find no correlations, so that the Star Ratings cannot be used to predict quality on those specific clinical conditions either.

This pattern of findings would seem to raise concerns not only about the CMS Star Rating system, but about the value of any global hospital rating system that attempts to combine scores on a large number of disparate measures of performance into a single overall score or categorical rating.<sup>20</sup> The lack of correlation among individual measures of hospital quality by domain (eg, readmission, infection) or clinical program (eg, cardiology, orthopedics) means that any global rating system will have little or no ability to predict quality of care outside the set of measures used to construct the global rating. If a global rating system does not include, for example, measures of quality for elective knee replacement surgery, there is no basis to presume that the global rating will be useful in choosing a hospital for elective knee replacement surgery.

Health plans, purchaser groups, patient advocacy groups, and others seeking to make broad judgments on overall performance or quality may still find it useful to use an overall 5-category rating system for the choices or recommendations they make. Any process that involves some yes/no decision about a hospital (eg, inclusion in a limited provider network) has to have some basis for the decision, and the Star Ratings may be as good as any other available. It has to be noted, however, that major global hospital rating programs including the CMS Star Ratings, U.S. News & World Report, and Healthgrades do not generally agree with each other on their hospital ratings<sup>20</sup>; they do not even agree on hospitals’ performance on specific quality domains for a particular condition.<sup>21</sup>

For individual consumers, the problem is perhaps more difficult, in that the overall Star Ratings do not seem to predict a patient’s likely experience in specific clinical domains, or with specific elective surgical procedures. This problem has been noted previously.<sup>12</sup> Patients who are interested in finding the best hospital for heart



surgery or for elective hip replacement will not be able to use the Star Ratings to do that.

### Limitations

Our study had limitations. First, state-level quality measure reporting periods were not perfectly aligned with those of the CMS Star Ratings. We used 2015 mortality data and 2018 SSI data from both California and New York (the most recent years available for risk-adjusted rates), and Star Ratings for the February 2019 reporting period. Given that hospitals' performance on quality indicators usually does not change substantially over time, we did not expect that the different reporting periods would have affected our key findings. Second, in the state-level analysis, some member hospitals in multihospital health systems reported their quality data for Star Ratings in a consolidated way under the same CMS Certification Number. In this case, these member hospitals would have the same "consolidated" stars (ie, they were not rated separately in Star Ratings), even though they reported separately, to their states, the quality measure data we included in the study. Although we cannot fix the underlying reporting issue, it was not a common issue in our data set, and we did not expect it to affect our key findings in any substantial manner. Third, our study did not explore the possible impact of the minimum thresholds for a hospital to receive a Star Rating (at least 3 quality groups, one of which must be the outcome group—Mortality, Safety of Care, or Readmission—and at least 3 measures in each group).<sup>18</sup> Our overall score distribution analysis excluded hospitals with missing scores in any of the 7 quality groups. We do not know how this exclusion could have affected the relationship between Star Ratings and individual quality measure or changed the summary score distribution. Lastly, hospital sample sizes in some parts of the analyses were small, especially in the state-level analysis, limiting our ability to detect significant relationships that we might have otherwise found.

## CONCLUSIONS

No hospitals were either best in all quality groups or worst in all quality groups. Because the overall Star Ratings are not correlated with other quality measures for specific clinical conditions or procedures, the global Star Ratings may not provide useful guidance to consumers for choices of their specific medical conditions or needs. ■

**Author Affiliations:** Center for Health Policy & Health Services Research, Henry Ford Health System (JH, DRN), Detroit, MI.

**Source of Funding:** None.

**Author Disclosures:** Drs Hu and Nerenz are employed by Henry Ford Health System; the 5 acute care hospitals in the Henry Ford Health System are scored every year in the Star Rating System.

**Authorship Information:** Concept and design (JH, DRN); acquisition of data (JH); analysis and interpretation of data (JH, DRN); drafting of the manuscript (JH, DRN); critical revision of the manuscript for important intellectual content (JH, DRN); statistical analysis (JH); obtaining funding (DRN); and supervision (DRN).

**Address Correspondence to:** Jianhui Hu, PhD, Center for Health Policy & Health Services Research, Henry Ford Health System, 1 Ford Pl, Ste 3A, Detroit, MI 48202. Email: jhu1@hfhs.org.

## REFERENCES

- Overall hospital quality star rating. CMS. Accessed April 5, 2021. <https://data.cms.gov/provider-data/topics/hospitals/overall-hospital-quality-star-rating/>
- Jha AK. The stars of hospital care: useful or a distraction? *JAMA*. 2016;315(21):2265-2266. doi:10.1001/jama.2016.5638
- Bilimoria KY, Barnard C. The new CMS hospital quality star ratings: the stars are not aligned. *JAMA*. 2016;316(17):1761-1762. doi:10.1001/jama.2016.13679
- Chatterjee P, Joynt Maddox K. Patterns of performance and improvement in US Medicare's Hospital Star Ratings, 2016-2017. *BMJ Qual Saf*. 2018;28(6):486-494. doi:10.1136/bmjqs-2018-008384
- National Quality Forum releases multistakeholder recommendations for strengthening the overall hospital star rating system. News release. National Quality Forum. November 6, 2019. Accessed April 26, 2020. [http://www.qualityforum.org/News\\_And\\_Resources/Press\\_Releases/2019/National\\_Quality\\_Forum\\_Releases\\_Multistakeholder\\_Recommendations\\_for\\_Strengthening\\_the\\_Overall\\_Hospital\\_Star\\_Rating\\_System.aspx](http://www.qualityforum.org/News_And_Resources/Press_Releases/2019/National_Quality_Forum_Releases_Multistakeholder_Recommendations_for_Strengthening_the_Overall_Hospital_Star_Rating_System.aspx)
- Haeder SF, Weimer DL, Mukamel DB. California hospital networks are narrower in Marketplace than in commercial plans, but access and quality are similar. *Health Aff (Millwood)*. 2015;34(5):741-748. doi:10.1377/hlthaff.2014.1406
- Hu J, Jordan J, Rubinfeld I, Schreiber M, Waterman B, Nerenz D. Correlations among hospital quality measures: what "Hospital Compare" data tell us. *Am J Med Qual*. 2017;32(6):605-610. doi:10.1177/1062860616684012
- Kaye D, Norton E, Elimoottil C, et al. Understanding the relationship between the Centers for Medicare and Medicaid Services' Hospital Compare star rating, surgical case volume, and short-term outcomes after major cancer surgery. *Cancer*. 2017;123(21):4259-4267. doi:10.1002/cncr.30866
- Mehta R, Paredes A, Tsilimigras D, et al. CMS Hospital Compare system of star ratings and surgical outcomes among patients undergoing surgery for cancer: do the ratings matter? *Ann Surg Oncol*. 2020;27(9):3138-3146. doi:10.1245/s10434-019-08088-y
- Papageorge MV, Resio BJ, Monsalve AF, et al. Navigating by stars: using CMS Star Ratings to choose hospitals for complex cancer surgery. *JNCI Cancer Spectr*. 2020;4(5):pkaa059. doi:10.1093/jncics/pkaa059
- Koh CY, Inaba CS, Sujatha-Bhaskar S, Nguyen NT. Association of Centers for Medicare & Medicaid Services Overall Hospital Quality Star Rating with outcomes in advanced laparoscopic abdominal surgery. *JAMA Surg*. 2017;152(12):1113-1117. doi:10.1001/jamasurg.2017.2212
- Khatana SA, Groeneveld P, Giri JS. Association between New York State hospital post-percutaneous coronary intervention mortality and readmissions and CMS Hospital Star Ratings. *J Am Coll Cardiol*. 2018;71(suppl 11):A95.
- Statistical Analysis System (SAS) package. QualityNet. 2020. Accessed May 4, 2020. <https://www.qualitynet.org/inpatient/public-reporting/overall-ratings/sas>
- California hospital inpatient mortality rates and quality ratings. California Health and Human Services Open Data Portal. Accessed May 4, 2020. <https://data.chhs.ca.gov/dataset/california-hospital-inpatient-mortality-rates-and-quality-ratings>
- Surgical site infections (SSIs) for operative procedures in California hospitals. California Health and Human Services Open Data Portal. Accessed April 5, 2021. <https://data.chhs.ca.gov/dataset/surgical-site-infections-ssis-for-28-operative-procedures-in-california-hospitals>
- Hospital-acquired infections: beginning 2008. Health Data NY. Updated November 14, 2019. Accessed May 4, 2020. <https://health.data.ny.gov/Health/Hospital-Acquired-Infections-Beginning-2008/utrz-zdsi>
- Castellucci M. CMS unveils updated hospital star ratings formula. *Modern Healthcare*. December 21, 2017. Accessed August 11, 2020. <https://www.modernhealthcare.com/article/20171221/NEWS/171229968/cms-unveils-updated-hospital-star-ratings-formula>
- Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHC/CORE). Overall hospital quality star rating on Hospital Compare: public input request. CMS. February 2019. Accessed April 5, 2021. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Overall-Hospital-Quality-Star-Rating-on-Hospital-Compare-Public-Input-Period.pdf>
- Chung JW, Dahlke AR, Barnard C, DeLancey JD, Merkow RP, Bilimoria KY. The Centers for Medicare and Medicaid Services hospital ratings: pitfalls of grading on a single curve. *Health Aff (Millwood)*. 2019;38(9):1523-1529. doi:10.1377/hlthaff.2018.05345
- Bilimoria KY, Birkmeyer JD, Burstin H, et al. Rating the raters: an evaluation of publicly reported hospital quality rating systems. *NEJM Catalyst*. August 14, 2019. Accessed April 5, 2021. <https://catalyst.nejm.org/doi/full/10.1056/CAT.19.0629>
- Austin JM, Derk JM, Kachalia A, Pronovost PJ. Assessing the agreement of hospital performance on 3 national mortality ratings for 2 common inpatient conditions. *JAMA Intern Med*. 2020;180(6):904-905. doi:10.1001/jamainternmed.2020.0450

Visit [ajmc.com/link/88634](http://ajmc.com/link/88634) to download PDF and eAppendix



**eAppendix A. Quality measure groups, IDs, and measure names**

Measure Group	Measure ID	Measure Name
Mortality	mortality_AMI	Acute Myocardial Infarction (AMI) 30-Day Mortality Rate
Mortality	mortality_CABG	Coronary Artery Bypass Graft (CABG) 30-Day Mortality Rate
Mortality	mortality_COPD	Chronic Obstructive Pulmonary Disease (COPD) 30-Day Mortality Rate
Mortality	mortality_HF	Heart Failure (HF) 30-Day Mortality Rate
Mortality	mortality_PN	Pneumonia (PN) 30-Day Mortality Rate
Mortality	mortality_STK	Acute Ischemic Stroke (STK) 30-Day Mortality Rate
Mortality	mortality_PSI4	Death Rate Among Surgical Inpatients with Serious Treatable Complications
Readmission	EDAC_AMI	Excess Days in Acute Care after Hospitalization for Acute Myocardial Infarction
Readmission	readmission_CABG	Coronary Artery Bypass Graft (CABG) 30-Day Readmission Rate
Readmission	readmission_COPD	Chronic Obstructive Pulmonary Disease (COPD) 30-Day Readmission Rate
Readmission	EDAC_HF	Excess Days in Acute Care after Hospitalization for Heart Failure
Readmission	readmission_hipknee	Hospital-Level 30-Day All-Cause Risk- Standardized Readmission Rate (RSRR) Following Elective Total Hip Arthroplasty (THA)/Total Knee Arthroplasty (TKA)
Readmission	EDAC_PN	Excess Days in Acute Care after Hospitalization for Pneumonia (PN)
Readmission	readmission_STK	Stroke (STK) 30-Day Readmission Rate
Readmission	readmission_HWR	HWR Hospital-Wide All-Cause Unplanned Readmission
Readmission	OP-32	Facility Seven-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy
Safety of Care	HAI-1	Central-Line Associated Bloodstream Infection (CLABSI)
Safety of Care	HAI-2	Catheter-Associated Urinary Tract Infection (CAUTI)
Safety of Care	HAI-3	Surgical Site Infection from Colon Surgery (SSI-colon)
Safety of Care	HAI-4	Surgical Site Infection from Abdominal Hysterectomy (SSI-abdominal hysterectomy)
Safety of Care	HAI-5	MRSA Bacteremia
Safety of Care	HAI-6	Clostridium Difficile (C.difficile)
Safety of Care	COMP-HIP-KNEE	Hospital-Level Risk-Standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and Total Knee Arthroplasty (TKA)
Safety of Care	PSI-90-Safety	Patient Safety and Adverse Events Composite
Patient Experience	H-clean	Cleanliness of Hospital Environment
Patient Experience	H-COMP_1	Nurse Communication
Patient Experience	H-COMP_2	Doctor Communication
Patient Experience	H-COMP_3	Responsiveness of Hospital Staff
Patient Experience	H-COMP_5	Communication About Medicines
Patient Experience	H-COMP_6	Discharge Information
Patient Experience	H-hosp	Overall Rating of Hospital
Patient Experience	H-quiet	Quietness of Hospital Environment
Patient Experience	H-COMP_7	HCAHPS 3 Item Care Transition Measure
Patient Experience	H-recommend	Willingness to Recommend Hospital
Efficient Use of Medical Imaging	OP-8	MRI Lumbar Spine for Low Back Pain
Efficient Use of Medical Imaging	OP-10	Abdomen CT Use of Contrast Material

Efficient Use of Medical Imaging	OP-11	Thorax CT Use of Contrast Material
Efficient Use of Medical Imaging	OP-13	Cardiac Imaging for Preoperative Risk Assessment for Non-Cardiac Low-Risk Surgery
Efficient Use of Medical Imaging	OP-14	Simultaneous Use of Brain Computed Tomography (CT) and Sinus CT
Timeliness of Care	ED-1b	Median Time from ED Arrival to ED Departure for Admitted ED Patients
Timeliness of Care	ED-2b	Admit Decision Time to ED Departure Time for Admitted Patients
Timeliness of Care	OP-3b	Median Time to Transfer to Another Facility for Acute Coronary Intervention
Timeliness of Care	OP-5	Median Time to ECG
Timeliness of Care	OP-18b	Median Time from ED Arrival to ED Departure for Discharged ED Patients
Timeliness of Care	OP-20	Door to Diagnostic Evaluation by a Qualified Medical Professional
Timeliness of Care	OP-21	ED-Median Time to Pain Management for Long Bone Fracture
Effectiveness of Care	IMM-2	Influenza Immunization
Effectiveness of Care	OP-4	Aspirin at Arrival
Effectiveness of Care	IMM-3/OP-27	Healthcare Personnel Influenza Vaccination
Effectiveness of Care	OP-22	ED-Patient Left Without Being Seen
Effectiveness of Care	OP-23	ED-Head CT or MRI Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke who Received Head CT or MRI Scan Interpretation Within 45 Minutes of Arrival
Effectiveness of Care	OP-29	Endoscopy/Polyp Surveillance: Appropriate Follow-up Interval for Normal Colonoscopy in Average Risk Patients
Effectiveness of Care	OP-30	Endoscopy/Polyp Surveillance: Colonoscopy Interval for Patients with a History of Adenomatous Polyps – Avoidance of Inappropriate Use
Effectiveness of Care	OP-33	External Beam Radiotherapy for Bone Metastases
Effectiveness of Care	PC-01	Elective Delivery Prior to 39 Completed Weeks Gestation: Percentage of Babies Electively Delivered Prior to 39 Completed Weeks Gestation
Effectiveness of Care	SEP-1	Severe Sepsis and Septic Shock
Effectiveness of Care	VTE-6	Hospital Acquired Potentially-Preventable Venous Thromboembolism

**eAppendix B. State-level quality measures (California and New York)**

California		New York	
Measures	Measure Units	Measures	Measure Units
<b>Mortality (report year: 2015)</b>			
Abdominal Aortic Aneurysm Repair: Unruptured	Risk-adjusted mortality rate	Abdominal Aortic Aneurysm Repair	Risk-adjusted mortality rate
Acute Stroke	Risk-adjusted mortality rate	Abdominal Aortic Aneurysm Repair: Unruptured	Risk-adjusted mortality rate
Acute Stroke Hemorrhagic	Risk-adjusted mortality rate	Carotid Endarterectomy	Risk-adjusted mortality rate
Acute Stroke Subarachnoid	Risk-adjusted mortality rate	Craniotomy	Risk-adjusted mortality rate
Carotid Endarterectomy	Risk-adjusted mortality rate	Gastrointestinal Hemorrhage	Risk-adjusted mortality rate
Craniotomy	Risk-adjusted mortality rate	Hip Fracture	Risk-adjusted mortality rate
Esophageal Resection	Risk-adjusted mortality rate	Hip Replacement	Risk-adjusted mortality rate
GI Hemorrhage	Risk-adjusted mortality rate		
Hip Fracture	Risk-adjusted mortality rate		
PCI	Risk-adjusted mortality rate		
Pancreatic Cancer	Risk-adjusted mortality rate		
Pancreatic Other	Risk-adjusted mortality rate		
Pancreatic Resection	Risk-adjusted mortality rate		
<b>Surgical Site Infections (report year: 2018)</b>			
All procedures	Standardized Infection Ratio (SIR)	Carbapenem-resistant Enterobacteriaceae Admission Prevalence Rate- all body sites	# admission prevalent CRE infections per 1,000 admissions
Appendix surgery	Standardized Infection Ratio (SIR)	Carbapenem-resistant Enterobacteriaceae Admission Prevalence Rate- Bloodstream Infections	# admission prevalent CRE BSIs per 1,000 admissions
Bile duct, liver or pancreatic surgery	Standardized Infection Ratio (SIR)	Carbapenem-resistant Enterobacteriaceae Hospital Onset Infection Rate- all body sites	# hospital onset CRE infections per 10,000 patient days
Cardiac surgery	Standardized Infection Ratio (SIR)	Carbapenem-resistant Enterobacteriaceae Hospital Onset Infection Rate- Bloodstream Infections	# hospital onset blood stream CRE infections per 10,000 patient days
Cesarean section	Standardized Infection Ratio (SIR)	Surgical Site Infections: Hip	# SSI per 100 procedures, risk-adjusted
Coronary bypass, chest and donor incisions	Standardized Infection Ratio (SIR)	Surgical Site Infections: CABG chest site	# SSI per 100 procedures, risk-adjusted
Exploratory abdominal surgery (laparotomy)	Standardized Infection Ratio (SIR)	Surgical Site Infections: CABG donor site	# SSI per 100 procedures, risk-adjusted
Gallbladder surgery	Standardized Infection Ratio (SIR)		
Gastric surgery	Standardized Infection Ratio (SIR)		
Hip prosthesis	Standardized Infection Ratio (SIR)		
Knee prosthesis	Standardized Infection Ratio (SIR)		
Laminectomy	Standardized Infection Ratio (SIR)		
Open reduction of fracture	Standardized Infection Ratio (SIR)		

Pacemaker surgery	Standardized Infection Ratio (SIR)		
Rectal surgery	Standardized Infection Ratio (SIR)		
Small bowel surgery	Standardized Infection Ratio (SIR)		
Spinal fusion	Standardized Infection Ratio (SIR)		
Thoracic surgery	Standardized Infection Ratio (SIR)		

**eAppendix C.** Numbers of included hospitals in different parts of the analyses

Quality Groups (or Measures)	N
<b>Figure 1.</b>	
Hospitals with all 7 group scores	3000
<b>Figure 2. Quality Measures</b>	
mortality_AMI	2329
mortality_CABG	1014
mortality_COPD	3556
mortality_HF	3539
mortality_PN	4054
mortality_STK	2578
mortality_PSI4	1708
HAI_1	1979
HAI_2	2283
HAI_3	1874
HAI_5	1683
HAI_6	3078
COMP_hipknee	2740
PSI90	3227
EDAC_AMI	2116
EDAC_HF	3606
EDAC_PN	4043
readmission_HWR	4408
readmission_CABG	1010
readmission_COPD	3611
readmission_hipknee	2771
readmission_STK	2517
OP-32	2919
H-clean	3444
H-COMP_1	3444
H-COMP_2	3444
H-COMP_3	3444
H-COMP_5	3444
H-COMP_6	3444
H-hosp	3444
H-quiet	3444
H-recommend	3444
H-COMP_7	3444
IMM-2	4217
IMM-3/OP-27	4253
OP-22	3630
OP-23	1523
OP-29	2857

SEP-1	3071
ED-1b	3983
ED-2b	3954
OP-3b	467
OP-5	2626
OP-18b	3819
OP-20	3821
OP-21	3606
OP-8	1428
OP-10	3761
OP-13	2351
OP-14	2954
<b>Table 1. Quality Groups</b>	
Mortality	3646
Safety	3366
Readmission	3723
Patient experience	3317
Effectiveness of care	3709
Timeliness of care	3614
Efficient use of image	3558
<b>Tables 2&amp;3. CA Quality Measures</b>	
<b>Mortality</b>	
AAA Repair Unruptured	132
Acute Stroke	287
Acute Stroke Hemorrhagic	211
Acute Stroke Subarachnoid	106
Carotid Endarterectomy	179
Craniotomy	147
Esophageal Resection	32
GI Hemorrhage	289
Hip Fracture	281
PCI	154
Pancreatic Cancer	53
Pancreatic Other	56
Pancreatic Resection	77
<b>SSI</b>	
All	313
Appendix surgery	162
Bile duct, liver or pancreatic surgery	115
Cardiac surgery	61
Cesarean section	180
Coronary bypass,chest and donor incisions	104
Exploratory abdominal surgery (laparo..	208
Gallbladder surgery	209



Gastric surgery	150
Hip prosthesis	230
Knee prosthesis	204
Laminectomy	89
Open reduction of fracture	229
Pacemaker surgery	59
Rectal surgery	112
Small bowel surgery	240
Spinal fusion	159
Thoracic surgery	71
<b>Tables 2&amp;3. NY Quality Measures</b>	
<b>Mortality</b>	
Abdominal Aortic Aneurysm (AAA) Repair	25
Abdominal Aortic Aneurysm (AAA) Repair - Unruptured Endovascular	20
Carotid Endarterectomy	46
Craniotomy	42
Gastrointestinal Hemorrhage	154
Hip Fracture	129
Hip Replacement	100
<b>SSI</b>	
CRE Admission Prevalence Rate- all body sites	163
CRE Admission Prevalence Rate- Bloodstream Infections	163
CRE Hospital Onset Infection Rate- all body sites	163
CRE Hospital Onset Infection Rate- Bloodstream Infections	163
SSI Hip	138
SSI CABG chest site	35
SSI CABG donor site	35

## eAppendix D. Percentages of best and worst performing hospitals by stars

Percentages of hospitals performing top quartile (**BEST**) on individual quality measures, by their stars

Quality Measures	Hospital Stars				
	1	2	3	4	5
mortality_AMI	13	19	22	29	44
mortality_CABG	9	18	24	26	39
mortality_COPD	17	23	23	28	42
mortality_HF	27	23	21	25	43
mortality_PN	10	15	22	33	55
mortality_STK	22	21	23	29	37
mortality_PSI4	16	18	26	29	44
HAI_1	19	24	25	30	24
HAI_2	21	26	24	27	23
HAI_3	23	24	25	26	27
HAI_4	31	32	34	34	37
HAI_5	13	21	27	32	30
HAI_6	16	25	24	27	27
COMP_hipknee	20	18	21	27	47
PSI90	3	10	17	38	69
EDAC_AMI	13	16	26	35	42
EDAC_HF	5	11	25	35	42
EDAC_PN	1	9	25	34	33
readmission_HWR	2	8	19	42	76
readmission_CABG	22	17	23	25	39
readmission_COPD	8	15	22	29	46
readmission_hipknee	15	16	20	24	45
readmission_STK	14	18	22	32	42
OP-32	24	24	23	24	28
H-clean	1	4	15	31	29
H-COMP_1	0	3	10	28	42
H-COMP_2	1	4	11	28	39
H-COMP_3	0	2	13	32	30

Domain

Mortality

Safety

Readmission

Patient Experience

Percentages of hospitals performing lowest quartile (**WORST**) on individual quality measures, by their stars

Quality Measures	Hospital Stars				
	1	2	3	4	5
mortality_AMI	40	32	31	20	9
mortality_CABG	38	33	27	21	8
mortality_COPD	41	33	30	22	14
mortality_HF	27	30	28	24	12
mortality_PN	44	38	28	17	8
mortality_STK	39	29	26	20	20
mortality_PSI4	40	31	24	18	10
HAI_1	34	27	24	20	22
HAI_2	27	25	24	25	29
HAI_3	29	26	24	25	21
HAI_4	35	39	29	33	32
HAI_5	36	31	22	19	14
HAI_6	22	24	27	24	28
COMP_hipknee	34	32	28	23	14
PSI90	79	44	21	7	1
EDAC_AMI	40	32	24	17	9
EDAC_HF	63	40	20	13	8
EDAC_PN	64	41	23	13	10
readmission_HWR	83	56	22	9	2
readmission_CABG	36	31	24	22	14
readmission_COPD	48	36	23	17	13
readmission_hipknee	42	38	34	26	15
readmission_STK	57	35	24	17	14
OP-32	26	26	26	25	33
H-clean	78	57	31	17	16
H-COMP_1	80	57	30	11	6
H-COMP_2	75	51	32	15	8
H-COMP_3	76	51	26	10	7

H-COMP_5	1	4	13	32	39
H-COMP_6	1	2	10	25	32
H-hosp	0	1	5	26	48
H-quiet	2	8	14	28	34
H-recommend	1	3	10	33	63
H-COMP_7	0	1	6	24	42
IMM-2	10	12	12	15	16
IMM-3/OP-27	11	13	22	23	25
OP-4	31	25	30	37	34
OP-22	9	16	20	29	30
OP-23	20	25	25	23	25
OP-29	22	23	25	22	21
OP-30	27	31	31	33	33
OP-33	27	33	32	34	28
PC-01	56	54	57	58	56
SEP-1	21	20	24	28	27
VTE-6	60	59	65	64	72
ED-1b	1	11	22	27	22
ED-2b	2	10	23	26	18
OP-3b	9	20	21	28	40
OP-5	14	16	17	19	17
OP-18b	2	12	24	26	16
OP-20	15	19	20	26	25
OP-21	17	21	24	28	28
OP-8	31	24	25	26	22
OP-10	27	23	22	25	22
OP-11	29	25	28	34	31
OP-13	21	27	27	23	19
OP-14	6	12	18	21	18

Effectiveness of Care

Timeliness of Care

Efficient Use of Medical Imaging

H-COMP_5	77	56	31	16	13
H-COMP_6	67	45	25	12	5
H-hosp	84	63	35	12	6
H-quiet	63	45	25	15	13
H-recommend	76	58	34	10	4
H-COMP_7	74	52	24	7	2
IMM-2	27	23	23	19	15
IMM-3/OP-27	34	32	24	21	20
OP-4	40	43	37	26	34
OP-22	63	51	42	30	25
OP-23	25	25	27	22	24
OP-29	33	26	28	22	14
OP-30	36	38	38	31	22
OP-33	46	34	34	33	32
PC-01	44	46	43	42	44
SEP-1	38	27	25	20	25
VTE-6	40	41	35	37	28
ED-1b	65	42	22	15	16
ED-2b	64	40	22	16	23
OP-3b	50	24	25	21	21
OP-5	41	35	30	22	18
OP-18b	53	36	22	20	29
OP-20	48	34	27	19	22
OP-21	46	36	26	18	19
OP-8	26	26	24	25	25
OP-10	29	31	27	20	17
OP-11	36	39	36	28	21
OP-13	29	23	26	30	26
OP-14	41	35	28	21	28

**eAppendix E.** Poisson regressions results (incidence rate ratios): relationship between hospital star ratings and individual quality measures, controlling for case volume

**California**

	<b>IRR</b>	<b>p-value</b>
<b>Mortality</b>		
AAA Repair Unruptured	1.00	0.98
Acute Stroke	1.00	0.96
Acute Stroke Hemorrhagic	1.00	0.62
Acute Stroke Subarachnoid	0.99	0.24
Carotid Endarterectomy	1.00	0.96
Craniotomy	0.99	0.07
Esoophageal Resection	1.00	0.43
GI Hemorrhage	0.98	0.26
Hip Fracture	1.00	0.60
PCI	0.99	0.66
Pancreatic Cancer	1.00	0.86
Pancreatic Other	0.98	0.16
Pancreatic Resection	0.99	0.07
<b>SSI</b>		
all	0.98	0.65
Appendix surgery	0.98	0.51
Bile duct, liver or pancreatic surgery	1.01	0.65
Cardiac surgery	1.00	0.96
Cesarean section	1.03	0.39
Coronary bypass,chest and donor incis..	1.03	0.34
Exploratory abdominal surgery (laparo..	1.00	1.00
Gallbladder surgery	1.00	0.93
Gastric surgery	1.01	0.86
Hip prosthesis	1.02	0.34
Knee prosthesis	1.00	0.88
Laminectomy	1.02	0.61
Open reduction of fracture	1.00	0.94
Pacemaker surgery	1.01	0.76
Rectal surgery	1.01	0.70
Small bowel surgery	0.98	0.59
Spinal fusion	0.98	0.57
Thoracic surgery	0.98	0.75

**New York**

	<b>IRR</b>	<b>p-value</b>
<b>Mortality</b>		
Abdominal Aortic Aneurysm (AAA) Repair	1.00	0.99
Abdominal Aortic Aneurysm (AAA) Repair - Unruptured Endovascular	1.03	0.44
Carotid Endarterectomy	0.93	0.61
Craniotomy	0.98	0.41
Gastrointestinal Hemorrhage	0.98	0.22
Hip Fracture	0.99	0.66
Hip Replacement	0.92	0.52
<b>SSI</b>		
CRE Admission Prevalence Rate- all body sites	0.84	0.08
CRE Admission Prevalence Rate- Bloodstream Infections	0.37	0.13
CRE Hospital Onset Infection Rate- all body sites	0.79	0.02
CRE Hospital Onset Infection Rate- Bloodstream Infections	0.63	0.20
SSI Hip	0.97	0.52
SSI CABG chest site	0.85	0.24
SSI CABG donor site	0.77	0.48