BMJ Open

# Identifying high-value care for Medicare beneficiaries: a cross-sectional study of acute care hospitals in the USA

Jeph Herrin ,[1,2] Huihui Yu,[3,4] Arjun K Venkatesh,[5] Sunita M Desai,[6] Cassandra L Thiel,[6] Zhenqiu Lin,[1,4] Susannah M Bernheim,[4,7] Leora I Horwitz [6,8]

**To cite:** Herrin J, Yu H, Venkatesh AK, *et al*. Identifying high-value care for Medicare beneficiaries: a cross-sectional study of acute care hospitals in the USA. *BMJ Open* 2022;**12**:e053629. doi:10.1136/bmjopen-2021-053629

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-053629).

Received 19 May 2021
Accepted 04 March 2022



For numbered affiliations see end of article.

**Correspondence to**
Dr Jeph Herrin;
jeph.herrin@yale.edu

## ABSTRACT

**Objectives** High-value care is providing high quality care at low cost; we sought to define hospital value and identify the characteristics of hospitals which provide high-value care.

**Design** Retrospective observational study.

**Setting** Acute care hospitals in the USA.

**Participants** All Medicare beneficiaries with claims included in Center for Medicare & Medicaid Services Overall Star Ratings or in publicly available Medicare spending per beneficiary data.

**Primary and secondary outcome measures** Our primary outcome was value defined as the difference between Star Ratings quality score and Medicare spending; the secondary outcome was classification as a 4 or 5 star hospital with lowest quintile Medicare spending ('high value') or 1 or 2 star hospital with highest quintile spending ('low value').

**Results** Two thousand nine hundred and fourteen hospitals had both quality and spending data, and were included. The value score had a mean (SD) of 0.58 (1.79). A total of 286 hospitals were classified as high value; these represented 28.6% of 999 4 and 5 star hospitals and 46.8% of 611 low cost hospitals. A total of 258 hospitals were classified as low value; these represented 26.6% of 970 1 and 2 star hospitals and 49.3% of 523 high cost hospitals. In regression models ownership, non-teaching status, beds, urbanity, nurse to bed ratio, percentage of dual eligible Medicare patients and percentage of disproportionate share hospital payments were associated with the primary value score.

**Conclusions** There are high quality hospitals that are not high value, and a number of factors are strongly associated with being low or high value. These findings can inform efforts of policymakers and hospitals to increase the value of care.

## Strengths and limitations of this study

- This study incorporates a comprehensive measure of overall hospital quality derived from five distinct domains.
- This study uses payments as proxy for actual costs; however, these payments do include postdischarge utilisation.
- The quality and payment measures include only patients enrolled in a single insurance programme, Medicare Fee for Service, and excludes most patients under 65 years of age.

## INTRODUCTION

The past two decades have seen an increasing focus by US policymakers and caregivers on improving hospital quality of care. The movement began in the early 2000s with public reporting of process measures, moved on to public reporting of outcomes measures such as readmissions, complications and mortality and, particularly since the passage of the Affordable Care Act in 2010, has since evolved to tying quality of care to payment.[1] There are now more than 100 quality measures on the Center for Medicare & Medicaid Services' (CMS) Care Compare, and a summary of over 50 of these measures into an Overall Hospital Quality Star Rating.[2] Patients and policymakers thus have available a range of quality metrics for comparing hospitals.

However, only recently have policymakers begun to take on the challenge of driving both improved quality and lower costs in tandem. The Value Based Purchasing Programme (for hospitals) and the Quality Payment Programme (for clinicians), for instance, both incorporate some measures of cost along with quality measures into their overall scoring system. Yet despite prior research indicating that people often perceive higher quality to be associated with higher costs and policymakers' concerns that cost reduction efforts will lower quality, the relationship between hospital quality and hospital cost, and in turn value, remains empirically underexplored. While in some cases high cost care is clearly of low quality (ie, excess use of unnecessary care, higher complication rates), it may be that higher costs produce better outcomes in other situations. For example, higher spending may focus on high-value care interventions or services such as postdischarge planning or expensive, evidence-based

treatments such as percutaneous coronary intervention and implantable cardioverter defibrillators.[3–5] Moreover, even if there is no systemic association between quality and cost, it is plausible that among hospitals that provide high quality care, some will be more cost effective at providing this care than others. Characteristics of hospitals that produce high-value healthcare, therefore, as defined by quality accounting for cost, may not be the same as those that produce high quality alone.

We therefore sought to identify hospitals that provide 'high value' care and to examine how they differ from other hospitals, including specifically those which provide 'low value' care. Using the CMS Overall Hospital Quality Star Ratings to identify high and low quality care hospitals and the Medicare spending per beneficiary (MSPB) total cost of care scores to identify high and low cost hospitals, we created a value score as the difference between quality and cost scores. Using those hospitals that were particularly high or low value, we analysed the characteristics of those hospitals, including anticipated drivers of costs (total expenses, supply costs) and factors historically associated with quality (beds, nurse staffing, teaching status, ownership, geographic region and urban status).

## METHODS

### Data and cohort

We used publicly reported Star Ratings data from Hospital Compare for 2018. The Star Rating system assigns each hospital an integer from 1 to 5, with 1 reflecting the lowest quality and 5 reflecting the highest quality. The Star Rating is a composite that reflects hospital performance across 57 measures, grouped into any of seven quality domains mortality, readmission, safety, patient experience, efficiency, effectiveness of care and timeliness of care. Measures in each domain are used to estimate a latent variable model with a single latent variable; and the latent variables (group scores) for each domain are combined using a weighted average to generate an overall summary score. Finally, k-means clustering with 5 means is used to group the summary scores into five star categories.[6] A hospital was assigned an overall Star Rating if it reported at least three measures in at least three domains, one of which was safety, mortality or readmission.

Hospital-level Star Ratings data were linked to 2018 MSPB data using hospitals' CMS Certification numbers. The MSPB score is calculated by the CMS using Medicare Part A and Part B payments for all care provided from 3 days prior to, to 30 days following an inpatient hospital stay (defined as an 'episode' of care). MSPB episodes span all conditions, and MSPB scores are calculated for all Medicare fee-for-service hospital stays, excluding stays in psychiatric, rehabilitation, cancer, children's, critical access or long-stay hospitals; excluding hospitalisations resulting in transfers to another acute facility and excluding readmissions that are within 30 days of an eligible hospitalisation. Episodes in which the patient dies during the measurement period are also excluded, as are those with $0 payments or involving transfers. The measure adjusts for the Medicare Severity-Diagnosis Related Group of the index hospitalisation, age, use of long-term care and 79 comorbidities. All payments are standardised across geographies for comparability; additional payments such as for indirect medical education or disproportionate share payments are omitted.[7] The MSPB score is constructed as a ratio so that a value of 1 is 'average' spending, and values lower or higher than 1 represent lower or higher spending than average.

Our study cohort comprised all hospitals which had an overall Star Rating; none were missing MSPB scores.

### Value

We examined two related measures of value. For our first value outcome, we constructed a linear metric of value by subtracting the standardised MSPB score from the standardised Star Rating hospital summary score (used to create the 5 star categories). Though we initially considered a ratio of quality to spending, we were concerned that a ratio could allow hospitals to appear high value while having low quality, if their spending were low enough. Therefore, we used a difference measure, which ensures that hospitals categorised as high value were always of high quality. To facilitate interpretation, both spending and quality scores were standardised to range from 0 to 10; thus, higher scores indicate higher value, with the maximum value score being 10 and minimum −10.

For our second value outcome, we classified hospitals as 'high value' if they received a 4 or 5 star rating on Hospital Compare and had an MSPB score in the lowest quintile of all hospitals, and classified hospitals as 'low value' if they received a 1 or 2 star rating on Hospital Compare and had an MSPB score in the highest quintile of all hospitals; all other hospitals were classified as 'average' value. For this classification, we selected quintiles to categorise hospitals by their MSPB scores as providing meaningful cost distinctions while identifying adequate numbers of high and low value hospitals for analysis.

### Other variables

Hospital characteristics were linked from the American Hospital Association (AHA) survey for 2018[8] and the CMS Provider of Services file of 2018. We examined hospital characteristics that have previously been associated with quality: geographic region (New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain and Pacific); teaching status (teaching, residency programme, non-teaching); number of beds (<100, 100–199, 200–299, 300–399 and 400+); nurse to bed ratio (≤0.75, 0.75–1, 1–1.5, >1.5); urban location (urban, rural); ownership status (for profit, private not for profit, public, federal); percentage of Medicare admissions that are for patients dually eligible for both Medicare and Medicaid (≤10%, 11%–20%, 21%–40% and >40%); total expenses (quintiles); total cost of supplies (quintiles); ratio of supply expenses to total expenses (quintiles) and

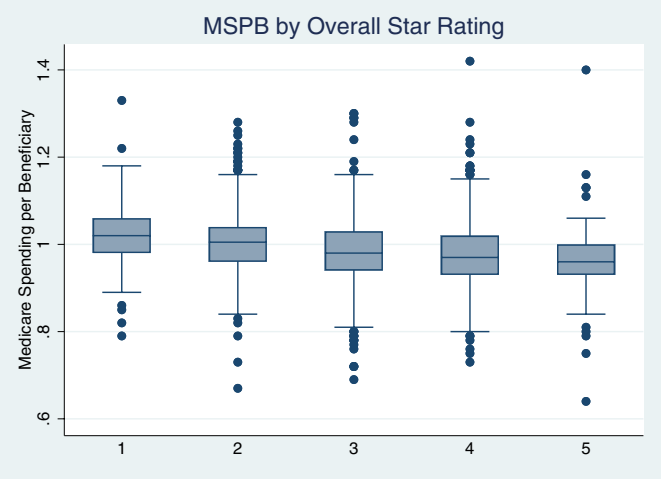


**Figure 1** Medicare spending per beneficiary by overall star rating. MSPB, Medicare spending per beneficiary.

disproportionate share hospital (DSH) patient percentage (quintiles). DSH is a Medicare payment adjustment based on the proportion of Medicare admissions at a hospital that are for patients enrolled in Medicaid.[9]

### Patient and public involvement

There was no patient or public involvement in the design, conduct or reporting of this study.

### Statistical analysis

We summarised value score (mean and SD) and value category (frequency and per cent) overall and by hospital characteristics, testing for differences in category characteristic using analysis of variance for the value score and chi-square tests for the value category. To assess the overall relationship between cost and quality of care, we graphed the distribution of MSPB scores over each Star Rating using a box plot (figure 1), and then plotted the value scores against the Star Rating summary scores to visualise the association of value with quality (figure 2). Then, to identify the hospital characteristics independently associated with value, and to identify contrasts in associations with quality, we estimated a series of models. We first used bivariate linear regression models to estimate the relationship between each hospital characteristics and value, where the outcome was our continuous value score; all factors were then included in a final multivariate model. For comparison with factors associated with quality alone, we estimated identical models where the dependent variable was the Star Ratings summary score, rather than our value score. We next estimated a multinomial logit model with the three-category value as the outcome using 'average' as the reference group, including all factors. We used a multinomial rather than ordinal model to avoid assuming that associations were monotone; that is, we anticipated that some factors might be associated with low value and not high, and some with high value but not low. Both analyses included the same set of hospital characteristics, and all models included indicators for each domain, equal to 1 if the domain was included in the hospital's overall star rating score. We reported overall Wald test p values for each characteristic. We report numbers and per cent of missing values, and use multiple imputation (with 20 imputations) to account for missing values in all models.

All analyses were done using SAS V.9.4 and Stata V.16.1 (2020, StataCorp, College Station, Texas, USA). All statistical tests were two-tailed, and we used $p<0.05$ to determine statistical significance.


**Figure 2** Value score versus star rating summary score.

## RESULTS

Our final study sample included 2914 hospitals that received both a Star Rating and an MSPB score, all of which matched the AHA Survey file for 2018 (table 1). The value score had a mean (SD) of 0.58 (1.79); see figures 1 and 2. A total of 286 were classified as high value and 258 as low value by the categorical definition (table 1). The 286 high-value hospitals represented 28.6% of 999 4 and 5 star hospitals and 46.8% of 611 low cost (bottom quintile MSPB) hospitals; the 258 low value hospitals represented 26.6% of 970 1 and 2 star hospitals and 49.3% of 523 high cost hospitals.

In biviariate analyses (table 1), we found not-for-profit private and government owned hospitals to have much higher values scores than for profit private hospitals, with score of 0.39 and 1.38 versus 0.08. Non-teaching hospitals and those with residency programmes both had three times the value score of teaching hospitals, while hospitals with fewer than 100 beds (1.48) and less than 10% dual eligible patients (1.74) had the highest values scores. Hospitals in rural areas and smaller percentages of DSH payments also had higher value scores (all p values <0.05). For the categorical bivariate comparisons (table 1), results were similar: high-value hospitals were more likely to be not-for-profit, have no teaching programme, have more beds and a higher nurse to bed ratio and have a lower percentage of dual-eligible admissions compared with low-value and average-value hospitals (all p values <0.001). Regionally, lower value hospitals were more

Table 1 Characteristics of low, average and high value hospitals

| Characteristic | Value score Mean (SD) | P value* | Low value n (%) or mean (SD) | Average n (%) or mean (SD) | High value n (%) or mean (SD) | P value |
|---|---|---|---|---|---|---|
| n | 0.58 (1.79) | | 258 (100.0) | 2370 (100.0) | 286 (100.0) | |
| Star rating | | <0.001 | | | | |
| 1 | −2.38 (1.26) | | 89 (34.5) | 179 (7.6) | 0 (0.0) | |
| 2 | −0.66 (1.02) | | 169 (65.5) | 533 (22.5) | 0 (0.0) | |
| 3 | 0.71 (1.00) | | 0 (0.0) | 945 (39.9) | 0 (0.0) | |
| 4 | 1.75 (1.00) | | 0 (0.0) | 528 (22.3) | 213 (74.5) | |
| 5 | 3.16 (1.07) | | 0 (0.0) | 185 (7.8) | 73 (25.5) | |
| Ownership | | <0.001 | | | | <0.001 |
| Public | 0.39 (1.80) | | 43 (16.7) | 356 (15.0) | 33 (11.5) | |
| Not for profit | 0.77 (1.69) | | 128 (49.6) | 1521 (64.2) | 215 (75.2) | |
| For profit | 0.08 (1.96) | | 87 (33.7) | 478 (20.2) | 36 (12.6) | |
| Government | 1.38 (1.60) | | 0 (0.0) | 15 (0.6) | 2 (0.7) | |
| Teaching status | | <0.001 | | | | <0.001 |
| Non-teaching | 0.87 (1.72) | | 22 (8.5) | 381 (16.1) | 48 (16.8) | |
| Residency | 0.72 (1.75) | | 124 (48.1) | 1130 (47.7) | 168 (58.7) | |
| Teaching | 0.25 (1.82) | | 112 (43.4) | 859 (36.2) | 70 (24.5) | |
| Beds | | <0.001 | | | | <0.001 |
| <100 | 1.48 (1.57) | | 21 (8.1) | 575 (24.3) | 155 (54.2) | |
| 100–199 | 0.65 (1.65) | | 58 (22.5) | 645 (27.2) | 67 (23.4) | |
| 200–299 | 0.21 (1.73) | | 54 (20.9) | 399 (16.8) | 33 (11.5) | |
| 300–399 | 0.08 (1.86) | | 58 (22.5) | 274 (11.6) | 17 (5.9) | |
| 400+ | −0.12 (1.74) | | 67 (26.0) | 477 (20.1) | 14 (4.9) | |
| Urbanity | | <0.001 | | | | <0.001 |
| Rural | 1.08 (1.52) | | 29 (11.2) | 579 (24.4) | 110 (38.5) | |
| Urban | 0.41 (1.84) | | 229 (88.8) | 1791 (75.6) | 176 (61.5) | |
| Region | | <0.001 | | | | <0.001 |
| New England | 0.70 (1.48) | | 3 (1.2) | 115 (4.9) | 9 (3.1) | |
| Mid Atlantic | −0.27 (1.92) | | 62 (24.0) | 260 (11.0) | 13 (4.5) | |
| E North Central | 1.02 (1.58) | | 22 (8.5) | 382 (16.1) | 57 (19.9) | |
| W North Central | 1.60 (1.48) | | 1 (0.4) | 179 (7.6) | 50 (17.5) | |
| South Atlantic | 0.40 (1.71) | | 43 (16.7) | 411 (17.3) | 42 (14.7) | |
| E South Central | 0.15 (1.46) | | 23 (8.9) | 236 (10.0) | 11 (3.8) | |
| W South Central | 0.17 (1.80) | | 60 (23.3) | 341 (14.4) | 20 (7.0) | |
| Mountain | 1.13 (1.78) | | 12 (4.7) | 155 (6.5) | 31 (10.8) | |
| Pacific | 0.81 (1.91) | | 32 (12.4) | 291 (12.3) | 53 (18.5) | |
| Nurse/beds | | <0.001 | | | | <0.001 |
| ≤0.75 | 0.51 (1.88) | | 86 (33.3) | 536 (22.6) | 79 (27.6) | |
| 0.75–1 | 0.35 (1.73) | | 52 (20.2) | 405 (17.1) | 47 (16.4) | |
| 1–1.5 | 0.41 (1.72) | | 86 (33.3) | 783 (33.0) | 65 (22.7) | |
| 1.5–2 | 0.79 (1.81) | | 25 (9.7) | 381 (16.1) | 52 (18.2) | |
| 2+ | 1.24 (1.67) | | 9 (3.5) | 265 (11.2) | 43 (15.0) | |
| % Duals | | <0.001 | | | | <0.001 |
| ≤10% | 1.74 (1.93) | | 7 (2.7) | 176 (7.4) | 33 (11.5) | |
| 11%–20% | 0.80 (1.58) | | 67 (26.0) | 863 (36.4) | 91 (31.8) | |

Continued

Table 1 Continued

| Characteristic | Value score Mean (SD) | P value* | Low value n (%) or mean (SD) | Average n (%) or mean (SD) | High value n (%) or mean (SD) | P value |
|---|---|---|---|---|---|---|
| 21%–40% | 0.51 (1.71) | | 118 (45.7) | 1057 (44.6) | 151 (52.8) | |
| >40% | −0.56 (1.94) | | 66 (25.6) | 274 (11.6) | 11 (3.8) | |
| Total expenses | | <0.001 | | | | <0.001 |
| Quintile 1 | 1.09 (1.66) | | 28 (10.9) | 459 (19.4) | 95 (33.2) | |
| Quintile 2 | 0.79 (1.77) | | 54 (20.9) | 446 (18.8) | 81 (28.3) | |
| Quintile 3 | 0.44 (1.73) | | 57 (22.1) | 475 (20.0) | 50 (17.5) | |
| Quintile 4 | 0.31 (1.81) | | 71 (27.5) | 473 (20.0) | 37 (12.9) | |
| Quintile 5 | 0.24 (1.83) | | 48 (18.6) | 511 (21.6) | 22 (7.7) | |
| Missing | 0.68 (1.61) | | 0 (0.0) | 6 (0.3) | 1 (0.3) | |
| Supplies $ | | <0.001 | | | | <0.001 |
| Quintile 1 | 1.11 (1.63) | | 23 (8.9) | 301 (12.7) | 72 (25.2) | |
| Quintile 2 | 0.83 (1.64) | | 31 (12.0) | 311 (13.1) | 53 (18.5) | |
| Quintile 3 | 0.50 (1.82) | | 35 (13.6) | 326 (13.8) | 35 (12.2) | |
| Quintile 4 | 0.59 (1.72) | | 37 (14.3) | 334 (14.1) | 24 (8.4) | |
| Quintile 5 | 0.27 (1.81) | | 34 (13.2) | 349 (14.7) | 12 (4.2) | |
| Missing | 0.40 (1.86) | | 98 (38.0) | 749 (31.6) | 90 (31.5) | |
| Supplies/total % | | 0.042 | | | | 0.002 |
| Quintile 1 | 0.83 (1.79) | | 29 (11.2) | 311 (13.1) | 56 (19.6) | |
| Quintile 2 | 0.75 (1.69) | | 23 (8.9) | 329 (13.9) | 43 (15.0) | |
| Quintile 3 | 0.51 (1.75) | | 38 (14.7) | 317 (13.4) | 41 (14.3) | |
| Quintile 4 | 0.53 (1.56) | | 28 (10.9) | 344 (14.5) | 23 (8.0) | |
| Quintile 5 | 0.68 (1.92) | | 42 (16.3) | 320 (13.5) | 33 (11.5) | |
| Missing | 0.40 (1.86) | | 98 (38.0) | 749 (31.6) | 90 (31.5) | |
| % DSH | | <0.001 | | | | <0.001 |
| Quintile 1 | 1.35 (1.74) | | 32 (12.4) | 468 (19.7) | 84 (29.4) | |
| Quintile 2 | 0.85 (1.62) | | 33 (12.8) | 480 (20.3) | 70 (24.5) | |
| Quintile 3 | 0.63 (1.59) | | 42 (16.3) | 488 (20.6) | 52 (18.2) | |
| Quintile 4 | 0.30 (1.60) | | 55 (21.3) | 484 (20.4) | 44 (15.4) | |
| Quintile 5 | −0.26 (1.96) | | 96 (37.2) | 450 (19.0) | 36 (12.6) | |

Source: Center for Medicare and Medicaid Services Overall Star Ratings Score and Medicare spending per beneficiary; American Hospital Association Survey.
*P value based on analysis of variance model estimated using multiple imputation for missing values.
DSH, disproportionate share hospital.

often in the Southern and Mid-Atlantic regions, while high-value hospitals were in Northern Central and Pacific Regions (p<0.001).

In the multivariable model, all were significantly associated with value except those related to total and supply expenses (table 2). The largest value effects were again for non-teaching, not-for-profit and government hospitals, hospitals with fewer than 100 beds and those with fewer than 10% dual eligible admissions. In equivalent model using the quality score instead of value as an outcome, there was no association with urban status, while, unlike with value, higher supply expenses were associated with higher quality; other patterns of effects were very similar.

In multinomial regression assessing the associations between characteristics and low-value, average-value and high-value classification (table 3), we collapsed the region categories to account for sparse cells. In the final model Government owned and not-for-profit hospitals, hospitals with fewer than 100 beds, hospitals in the Western region and those with lower percentages of dual eligible patients and DSH patients were more likely to be high value than average value (all p values <0.05). Correspondingly, public hospitals, those with the more beds, those in the Northeast region and lowest nurse to bed ratio all had higher odds of being low-value (all p values <0.05).

Table 2 Results of multivariable linear models with value score and star ratings summary score as dependent variables

| | Value | | Quality | |
|---|---|---|---|---|
| Characteristic | Coeff (SE) | P value | Coeff (SE) | P value |
| Ownership | | <0.001 | | <0.001 |
| Public | Ref | | Ref | |
| Not for profit | 0.40 (0.21 to 0.59) | | 0.34 (0.20 to 0.48) | |
| For profit | −0.04 (−0.29 to 0.21) | | 0.14 (−0.05 to 0.33) | |
| Government | 0.96 (−0.23 to 2.15) | | 0.47 (−0.42 to 1.36) | |
| Teaching status | | 0.008 | | 0.006 |
| Non-teaching | Ref | | Ref | |
| Residency | −0.30 (−0.50 to −0.10) | | −0.19 (−0.34 to −0.05) | |
| Teaching | −0.27 (−0.47 to −0.07) | | −0.24 (−0.39 to −0.09) | |
| Beds | | <0.001 | | <0.001 |
| <100 | Ref | | Ref | |
| 100–199 | −0.55 (−0.77 to −0.33) | | −0.31 (−0.47 to −0.14) | |
| 200–299 | −0.99 (−1.29 to −0.70) | | −0.57 (−0.79 to −0.36) | |
| 300–399 | −1.24 (−1.57 to −0.90) | | −0.67 (−0.93 to −0.42) | |
| 400+ | −1.45 (−1.81 to −1.08) | | −0.80 (−1.07 to −0.52) | |
| Urbanity | | 0.015 | | 0.894 |
| Rural | Ref | | Ref | |
| Urban | −0.23 (−0.42 to −0.04) | | −0.01 (−0.15 to 0.13) | |
| Region | | <0.001 | | <0.001 |
| New England | Ref | | Ref | |
| Mid Atlantic | −0.67 (−1.03 to −0.31) | | −0.56 (−0.83 to −0.29) | |
| E North Central | 0.35 (0.01 to 0.70) | | 0.27 (0.02 to 0.53) | |
| W North Central | 0.57 (0.20 to 0.95) | | 0.03 (-0.25 to 0.31) | |
| South Atlantic | −0.09 (−0.45 to 0.27) | | −0.25 (−0.52 to 0.02) | |
| E South Central | −0.70 (−1.13 to −0.28) | | −0.55 (−0.86 to −0.23) | |
| W South Central | −0.62 (−0.99 to −0.26) | | −0.15 (−0.42 to 0.12) | |
| Mountain | 0.23 (−0.19 to 0.65) | | −0.02 (−0.34 to 0.29) | |
| Pacific | 0.83 (0.45 to 1.22) | | 0.35 (0.06 to 0.64) | |
| Nurse/beds | | 0.006 | | 0.001 |
| ≤0.75 | Ref | | Ref | |
| 0.75–1 | −0.19 (−0.40 to 0.03) | | −0.21 (−0.37 to −0.05) | |
| 1–1.5 | −0.14 (−0.34 to 0.05) | | −0.16 (−0.30 to −0.01) | |
| 1.5–2 | 0.05 (−0.18 to 0.28) | | 0.03 (−0.14 to 0.20) | |
| 2+ | 0.22 (−0.04 to 0.48) | | 0.10 (−0.10 to 0.29) | |
| % Duals | | <0.001 | | <0.001 |
| ≤10% | Ref | | Ref | |
| 11%–20% | −0.48 (−0.78 to −0.18) | | −0.26 (−0.48 to −0.03) | |
| 21%–40% | −0.81 (−1.14 to −0.48) | | −0.68 (−0.93 to −0.43) | |
| >40% | −1.66 (−2.10 to −1.23) | | −1.20 (−1.53 to −0.88) | |
| Total expenses | | 0.821 | | 0.79 |
| Quintile 1 | Ref | | Ref | |
| Quintile 2 | −0.01 (−0.31 to 0.29) | | −0.09 (−0.32 to 0.13) | |
| Quintile 3 | −0.15 (−0.54 to 0.24) | | −0.15 (−0.44 to 0.14) | |
| Quintile 4 | −0.12 (−0.60 to 0.35) | | −0.17 (−0.52 to 0.18) | |

Continued

Table 2 Continued

| Characteristic | Value | | Quality | |
|---|---|---|---|---|
| | Coeff (SE) | P value | Coeff (SE) | P value |
| Quintile 5 | −0.24 (−0.81 to 0.34) | | −0.27 (−0.71 to 0.16) | |
| Supplies $ | | 0.034 | | 0.008 |
| Quintile 1 | Ref | | Ref | |
| Quintile 2 | −0.06 (−0.37 to 0.24) | | 0.07 (−0.16 to 0.30) | |
| Quintile 3 | −0.06 (−0.47 to 0.34) | | 0.01 (−0.29 to 0.31) | |
| Quintile 4 | 0.30 (−0.20 to 0.79) | | 0.34 (−0.03 to 0.71) | |
| Quintile 5 | 0.50 (−0.09 to 1.10) | | 0.50 (0.05 to 0.94) | |
| Supplies/total % | | 0.424 | | 0.479 |
| Quintile 1 | Ref | | Ref | |
| Quintile 2 | −0.10 (−0.32 to 0.12) | | 0.03 (−0.14 to 0.19) | |
| Quintile 3 | −0.21 (−0.45 to 0.03) | | 0.06 (−0.12 to 0.24) | |
| Quintile 4 | −0.23 (−0.49 to 0.03) | | 0.02 (−0.17 to 0.22) | |
| Quintile 5 | −0.24 (−0.54 to 0.06) | | −0.09 (−0.31 to 0.13) | |
| % DSH | | 0.003 | | <0.001 |
| Quintile 1 | Ref | | Ref | |
| Quintile 2 | −0.17 (−0.39 to 0.04) | | −0.27 (−0.43 to −0.10) | |
| Quintile 3 | −0.32 (−0.55 to −0.09) | | −0.40 (−0.57 to −0.23) | |
| Quintile 4 | −0.38 (−0.63 to −0.14) | | −0.37 (−0.55 to −0.18) | |
| Quintile 5 | −0.56 (−0.85 to −0.27) | | −0.66 (−0.88 to −0.44) | |

Source: Center for Medicare and Medicaid Services Overall Star Ratings Score and Medicare spending per beneficiary; American Hospital Association Survey.
DSH, disproportionate share hospital.

## DISCUSSION

In this examination of hospital value of care, we identified hospitals that provided relatively higher quality of care to Medicare beneficiaries at lower comparative costs, as well as hospitals that provided lower quality care at a high level of spending. In bivariate analyses, both high and low value hospitals differed from other hospitals strongly in ownership, urban location, bed size and region, nurse-to-bed ratio, amount spent on supplies relative to total expenses and percentage of dual-eligible patients. Most associations were significant in the multivariable primary analysis using the value score; of these, all but those related to urbanicity and expenses persisted in the multinomial logit analysis using a categorical outcome of low, average or high value.

The most notable finding, however, may not be the associations of hospital characteristics with value, but that value and quality of care are not synonymous: though generally, the MSPB declined with increasing Star Ratings (figure 1), less than a third of 4 and 5 star hospitals were high value, while more than a fourth of 1 and 2 star hospitals were lowest value. The observed trend in spending across overall quality (figure 1) as well as the pattern of value classification is very similar to that found previously for patient experience Star Ratings[10 11]; though the patient experience Star Rating is a component of the overall Star Rating, it accounts for only 22% of the weight of the overall summary score, so it is unlikely to drive the current findings. This finding was also reflected in our model of quality score, which found urban status was not related to quality while quality increased with more spending on supplies. That quality and value are not identical indicates that it will be possible to reduce spending among even the highest quality hospitals; however, there is clearly a tension between the two, for if quality improvements driven by increased supply spending (or other additional resources) that translate to greater spending, then only offsets in other areas can increase value.

Most of the associations we found for value are similar to those seen for hospital characteristics and quality, both here with the Star Ratings summary score and in prior research; for example, DeLancey *et al* found that hospitals with fewer beds, no medical school affiliation and lower proportions of DSH patients were more likely to have higher Star Ratings.[12] This was similar to the results of our analysis of the star summary quality score (table 2). However, in our categorical analysis, we found no relationship between teaching status and high value. Geographic region also explained some of the difference between low value and average value hospitals across all of our analyses, with hospitals in East South Central and Mid-Atlantic regions having lowest value and being more

Table 3 Results of multilogit model for low and high value versus average value hospitals

| | Low value | | High value | |
|---|---|---|---|---|
| Characteristic | OR (95% CI) | P value | OR (95% CI) | P value |
| Ownership | | <0.001 | | 0.002 |
| Public | Ref | | Ref | |
| Not for profit | 0.4 (0.2 to 0.6) | | 2.1 (1.2 to 3.5) | |
| For profit | 0.8 (0.5 to 1.5) | | 0.8 (0.4 to 1.8) | |
| Government | 0.0 (0.0 to 0.0) | | 2.0 (0.3 to 13.3) | |
| Teaching status | | 0.012 | | 0.698 |
| Non-teaching | Ref | | Ref | |
| Residency | 2.9 (1.4 to 5.8) | | 1.1 (0.7 to 1.7) | |
| Teaching | 2.4 (1.2 to 4.9) | | 1.2 (0.7 to 2.1) | |
| Beds | | <0.001 | | 0.003 |
| <100 | Ref | | Ref | |
| 100–199 | 2.9 (1.2 to 6.8) | | 0.5 (0.3 to 0.7) | |
| 200–299 | 5.9 (2.2 to 16.1) | | 0.4 (0.2 to 0.8) | |
| 300–399 | 11.9 (4.1 to 34.5) | | 0.3 (0.1 to 0.8) | |
| 400+ | 10.0 (3.2 to 31.7) | | 0.2 (0.1 to 0.5) | |
| Urbanity | | 0.3 | | 0.758 |
| Rural | Ref | | Ref | |
| Urban | 1.4 (0.8 to 2.5) | | 1.1 (0.7 to 1.6) | |
| Region* | | <0.001 | | <0.001 |
| West | Ref | | Ref | |
| Midwest | 0.5 (0.2 to 1.2) | | 0.4 (0.2 to 0.6) | |
| Northeast | 2.7 (1.4 to 5.3) | | 0.1 (0.1 to 0.3) | |
| South | 1.3 (0.7 to 2.4) | | 0.3 (0.2 to 0.4) | |
| Nurse/beds | | 0.025 | | 0.673 |
| ≤0.75 | Ref | | Ref | |
| 0.75–1 | 1.0 (0.6 to 1.7) | | 1.0 (0.6 to 1.7) | |
| 1–1.5 | 0.8 (0.5 to 1.3) | | 0.8 (0.5 to 1.2) | |
| 1.5–2 | 0.5 (0.2 to 0.9) | | 1.1 (0.6 to 1.9) | |
| 2+ | 0.3 (0.1 to 0.7) | | 1.0 (0.5 to 1.8) | |
| % Duals | | 0.427 | | 0.009 |
| ≤10% | Ref | | Ref | |
| 11%–20% | 2.2 (0.6 to 7.8) | | 1.1 (0.6 to 2.2) | |
| 21%–40% | 2.7 (0.7 to 9.8) | | 1.9 (0.9 to 4.0) | |
| >40% | 3.2 (0.8 to 13.2) | | 0.6 (0.2 to 2.0) | |
| Total expenses | | 0.536 | | 0.626 |
| Quintile 1 | Ref | | Ref | |
| Quintile 2 | 1.7 (0.7 to 4.6) | | 1.4 (0.8 to 2.6) | |
| Quintile 3 | 1.2 (0.4 to 4.0) | | 1.1 (0.5 to 2.4) | |
| Quintile 4 | 1.3 (0.3 to 5.1) | | 1.2 (0.4 to 3.5) | |
| Quintile 5 | 0.9 (0.2 to 4.4) | | 0.9 (0.2 to 4.1) | |
| Supplies $ | | 0.66 | | 0.845 |
| Quintile 1 | Ref | | Ref | |
| Quintile 2 | 0.6 (0.2 to 1.4) | | 0.7 (0.4 to 1.4) | |
| Quintile 3 | 0.4 (0.1 to 1.3) | | 0.6 (0.2 to 1.4) | |
| Quintile 4 | 0.4 (0.1 to 1.4) | | 0.5 (0.2 to 1.7) | |
| Quintile 5 | 0.3 (0.1 to 1.5) | | 0.5 (0.1 to 2.6) | |

Continued

Table 3 Continued

| Characteristic | Low value | | High value | |
|---|---|---|---|---|
| | OR (95% CI) | P value | OR (95% CI) | P value |
| Supplies/total % | | 0.01 | | 0.273 |
| Quintile 1 | Ref | | Ref | |
| Quintile 2 | 1.0 (0.5 to 2.0) | | 0.9 (0.5 to 1.4) | |
| Quintile 3 | 2.0 (1.0 to 4.0) | | 1.2 (0.7 to 2.1) | |
| Quintile 4 | 1.4 (0.7 to 3.1) | | 0.6 (0.3 to 1.3) | |
| Quintile 5 | 3.0 (1.3 to 6.8) | | 1.1 (0.5 to 2.3) | |
| % DSH | | 0.048 | | 0.006 |
| Quintile 1 | Ref | | Ref | |
| Quintile 2 | 0.7 (0.4 to 1.4) | | 0.8 (0.5 to 1.4) | |
| Quintile 3 | 1.1 (0.6 to 2.0) | | 0.5 (0.3 to 0.9) | |
| Quintile 4 | 0.6 (0.3 to 1.3) | | 0.4 (0.2 to 0.7) | |
| Quintile 5 | 1.6 (0.7 to 3.3) | | 0.5 (0.2 to 1.0) | |

Source: Center for Medicare and Medicaid Services Overall Star Ratings Score and Medicare spending per beneficiary; American Hospital Association Survey.
*Region groups were collapsed due to sparse cells under original groupings.
DSH, disproportionate share hospital.

likely to be low value than those in any other regions, and Pacific region hospitals being consistently high value in all models. Since our spending measure is adjusted for local wage and cost variations, this geographic variation is likely driven substantially by variations in quality.

Of most relevance to hospitals and policymakers are the associations with factors that are within the control of the hospital. A key factor that was related to low value in all analyses was nurse to bed ratio, with lower ratios being more often found at the lowest value hospitals—only 9 (3.5%) low value hospitals had a ratio of 2 or more, while 43 (15.0%) of high-value hospitals had a ratio of at least 2. This finding suggests that high levels of nurse staffing, while increasing the expense of inpatient care, may reduce overall Medicare payments while promoting high quality care. Importantly, this finding also highlights a perennial challenge for policymakers: the wrong pocket problem. Since the costs of increased staffing are borne by hospitals but the benefits of lower spending are reaped by payers and patients, there is no intrinsic incentive for hospitals to pursue such interventions. Payment policies seeking to improve value should explicitly take these challenges into account.

Conversely, 66 (25.6%) of low value hospitals reported at least 40% of their Medicare patients were also eligible for Medicaid, while only 11 (3.8%) of high-value hospitals reported such high rates of dual eligibility, with proportion of DSH patients exhibiting a similar but weaker pattern. This may indicate that lower income patients have unmeasured comorbidities or care complexities that are not captured by the risk adjustment used in Star Ratings' underlying measures or that these patients drive costs in ways unrelated to quality of care (eg, limited social support may lead to greater postdischarge costs), that these patients may have higher cost needs that are not completely adjusted for, or possibly that DSH patients are more likely to receive care at lower quality hospitals.

This study has the limitations of any observational study, including that no causal inferences can be drawn; for example, it may be that lower value hospitals have increased costs of care because of lower quality, rather than independent of it. In addition, we have only examined a few key factors that may be associated with quality, in order to better focus on value of care. And, we have focused only on quality and spending metrics largely related to Medicare fee-for-service beneficiaries; other, broader measures of quality and costs that included the entire patient population could plausibly produce different results. Moreover, Medicare payments are distorted somewhat by the Inpatient Prospective Payment System. However, CMS Overall Star Ratings captures quality of care across seven domains, and the MSPB is a measure of spending with cost-of-living adjustments that plausibly represents the overall spending per patient resulting from a hospitalisation.

It is worth considering that the MSPB does not include admissions that lead to transfers, nor admissions in which the patient does not survive 30 days. Both exclusions are likely to confound the relationship between the spending score and the mortality domain, which accounts for 22% of the Star Rating summary score, since publicly reported mortality measures typically measure outcomes within the same 30 days, and attribute them to the first hospital in the event of transfers. Excluding patients who do not survive and/or are transferred likely reduces the MSPB score (assuming these are more complex patients), and more so for hospitals with higher mortality or transfer rates. However, this would tend to bias our current findings toward the null, attenuating the relationship between spending value.

## CONCLUSION

We have shown that care quality and care value are not identical; there are high quality hospitals that are not high value. In addition, there are a number of structural factors, some under the control of hospitals such as nurse to bed ratio, that are strongly associated with being low or high value. These findings can inform efforts of policymakers and hospitals to increase the value of care.

**Author affiliations**
[1]Section of Cardiovascular Medicine, Yale University School of Medicine, New Haven, Connecticut, USA
[2]Flying Buttress Associates, Charlottesville, Virginia, USA
[3]Yale School of Medicine, New Haven, Connecticut, USA
[4]Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, Connecticut, USA
[5]Department of Emergency Medicine, Yale University, New Haven, Connecticut, USA
[6]Department of Population Health, NYU Grossman School of Medicine, New York City, New York, USA
[7]Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut, USA
[8]Center for Healthcare Innovation and Delivery Science, NYU Langone Health, New York City, New York, USA

**Twitter** Arjun K Venkatesh @arjunvenkatesh and Leora I Horwitz @leorahorwitzmd

**Contributors** JH, AKV, SMB and LIH contributed to the conception and design of the study. JH and HY performed all analyses. JH drafted the manuscript. AKV, SMD, CLT, ZL, SMB and LIH contributed to critical additions and conceptual revisions to the manuscript. JH serves as guarantor of the study.

**Funding** This work was funding by grant R01HS022882 from the Agency for Healthcare Quality and Research. SMD was also supported by grant number K01HS026980 from the Agency for Healthcare Research and Quality.

**Competing interests** SMB, JH, ZL and AKV recieve salary support from the Centers for Medicare and Medicaid Services to develop, implement and maintain hospital performance outcome measures, including the methodology for the Overall Hospital Star Ratings, that are publicly reported. LIH and HY have worked under contract to the Centers for Medicare and Medicaid Services to develop quality measures, including some used in the Overall Hospital Star Ratings program.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** The study was approved by the Yale and NYU Institutional Review Boards.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Star Ratings and Medicare spending data are publicly available, and an analytical file including those data can be provided by the authors. Data collected by the American Hospital Association are proprietary, and will not be available for open access sharing.



**ORCID iDs**
Jeph Herrin http://orcid.org/0000-0002-3671-3622
Leora I Horwitz http://orcid.org/0000-0003-1800-6040

## REFERENCES

1 U.S. Centers for Medicare & Medicaid Services. Hospital readmissions reduction program (HRRP). Available: https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html [Accessed 08 Dec 2020].
2 U.S. Centers for Medicare and Medicaid Services. Hospital compare overall ratings. Available: https://qualitynet.cms.gov/inpatient/public-reporting/overall-ratings [Accessed 08 Dec 2020].
3 Doyle JJ, Graves JA, Gruber JA, *et al*. Do high-cost hospitals deliver better care? Evidence from ambulance referral patterns. National Bureau of economic research. Available: http://www.nber.org/papers/w17936.pdf?new_window=1 [Accessed 19 Dec 2013].
4 Silber JH, Kaestner R, Even-Shoshan O, *et al*. Aggressive treatment style and surgical outcomes. *Health Serv Res* 2010;45:1872–92.
5 Virtual mentor. *American Medical Association Journal of Ethics* 2014;16:124–30.
6 QualityNet. Available: https://www.qualitynet.org/inpatient/public-reporting/overall-ratings/resources
7 QualityNet. Available: https://www.qualitynet.org/inpatient/measures/mspb/methodology
8 American Hospital Association. American Hospital association (AHA) annual survey database 2018.
9 Disproportionate share Hospital (DSH). Available: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/dsh
10 Liao JM, Zhou L, Navathe AS. Nationwide Hospital performance on publicly reported episode spending measures. *J Hosp Med* 2020;15:E1–7.
11 Trzeciak S, Gaughan JP, Bosire J, *et al*. Association between Medicare StAR ratings for patient experience and Medicare spending per beneficiary for us hospitals. *J Patient Exp* 2017;4:17–21.
12 DeLancey JO, Softcheck J, Chung JW, *et al*. Associations Between Hospital Characteristics, Measure Reporting, and the Centers for Medicare & Medicaid Services Overall Hospital Quality Star Ratings. *JAMA* 2017;317:2015–7.