

An Examination of CMS's Dialysis Star-Rating System

By Alex Brill

MARCH 2015

In January 2015, CMS expanded its star-rating system to include U.S. dialysis facilities. This article offers a preliminary analysis of the nearly 6,000 rated facilities and explores trends and variations in scores based on facility characteristics and local demographic factors. Statistical tests show variation across states and by demographic factors. Regression analysis indicates that communities with higher incomes have facilities that perform better on some metrics of the star-rating system. This finding suggests that the system, which seeks to measure facility quality, not patient characteristics, needs further revision.

BACKGROUND

In a [blog post](#) last June, Patrick Conway, CMS's Chief Medical Officer and Deputy Administrator for Innovation and Quality and now also the Acting Principal Deputy Administrator, explained the objective of CMS's star ratings, which also apply to nursing homes, hospitals, home health agencies, and some physician practices: "The star ratings empower consumers with information to make more informed health care decisions, encourage providers to strive for higher levels of quality, and drive overall health system improvement." In its [announcement](#) of the dialysis star-rating program, CMS further explained, "These ratings summarize performance data, making it easier for consumers to use the information on the website. These ratings also spotlight excellence in health care quality."

However, the new dialysis star-rating program is not without controversy. MedPAC, the commission that advises Congress on Medicare reimbursement matters, [expressed concern](#) last summer that CMS intended to establish the dialysis star-rating system without considering public comment. MedPAC urged a delay, noting that "an open and transparent process will give beneficiaries, providers, and other members of the public the opportunity to submit comments

to the agency's proposal." Many in the provider community objected as well. Kidney Care Partners, an umbrella coalition, [wrote](#) to CMS about the lack of consultation, the risk that the new measure will create patient confusion, the validity of certain component factors used to generate the final scores, and the bell curve assumption that forces the facility rating into a preset distribution.

Recently, CMS [announced](#) plans to establish an End Stage Renal Disease (ESRD) Star Rating Technical Experts Panel (TEP) to review the methodology developed to produce the star ratings. The TEP will include experts from a variety of perspectives able to offer advice and recommendations for improving the statistical methodology, the clinical measures underlying the star-rating calculation and public reporting, and the readability and presentation of the star-rating results.

Hooper, Lundy & Bookman provides public policy and government relations services to Dialysis Clinics, Inc. (DCI) and the Nonprofit Kidney Care Alliance (NKCA). Members of NKCA include Centers for Dialysis Care, DCI, Independent Dialysis Foundation, Northwest Kidney Centers, and The Rogosin Institute. NKCA provided financial support for the preparation of this analysis.

How Star Ratings Are Constructed

The dialysis star-rating system is based on a composite score for seven facility-specific quality measures from the CMS Dialysis Facility Compare (DFC) Database:

1. Standardized hospitalization ratio (SHR)
2. Standardized mortality ratio (SMR)
3. Standardized transfusion ratio (STrR)
4. Share of adult patients who received treatment through arteriovenous fistula (Fistula)
5. Share of adult patients who had a catheter (tube) left in a vein longer than 90 days for regular hemodialysis treatment (Catheter)
6. Share of adult patients who had hypercalcemia (that is, an average calcium level greater than 10.2 mg/d over the past three months) (Calcium)
7. Share of patients who had adequate Kt/V levels (that is, enough wastes removed from their blood during dialysis)² (Kt/V)

CMS “standardizes” each metric by translating the raw scores into percentile ranks.

CMS [explains](#), “To further differentiate facilities that performed exceptionally well or poorly, these percentile ranks (pRanks) were ‘normalized’ or mapped from the uniform percentile rank distribution to a normal distribution (nRanks).”

As a result of this forced normal distribution, the difference in the value of a facility with a percentile rank of 99 for a given metric and a facility with a rank of 98 is greater than the difference between a facility with a percentile rank of 51 and a facility with a rank of 50. This

effort to differentiate scores in the tails of the distribution is an artificial distinction imposed on the data by CMS and has been criticized by industry representatives, as noted above.

When the seven factors are combined into a star rating, each is not weighted equally. Instead, a statistical method known as factor analysis is used to ensure that individual factors that are correlated are not disproportionately weighted. (For example, the share of adult patients who received treatment through arteriovenous fistula is correlated with the share of patients who had a catheter for longer than 90 days; $\rho = 0.45$.) The factor analysis creates three groups, each equally weighted in determining the final facility score. The first group, “Standardized Outcomes,” comprises the variables for SHR, SMR, and STrR. The second group, “Other Outcomes 1 (Fistula, Catheter),” comprises the variables for patients who received treatment through arteriovenous fistula and those who had a catheter left in a vein longer than 90 days. The final group, “Other Outcomes 2 (Kt/V, Calcium),” comprises the variables for hypercalcemia and Kt/V.

CMS averages the scores within each group and then averages across the three groups to arrive at a final score, which is then used to divide the facilities into five star-rating categories.

Facilities with the top 10% of final scores are given a star rating of 5. Facilities with the next 20% of scores are rated 4. Facilities within the middle 40% of final scores are given a star rating of 3. Facilities with the next 20% are rated 2. And facilities with the bottom 10% of final scores are rated 1.

² This variable is a weighted average derived from three separate measures for each of three patient types: adult hemodialysis, pediatric hemodialysis, and adult peritoneal dialysis.

Data Review

CMS reports quality metrics for 6,307 dialysis facilities, but some facilities are missing data for certain metrics. For the star-rating system, CMS excludes any facility that lacks data for all metrics in one of the three groups. If a facility is not missing all values in a group, CMS assigns the median value, 50, to any missing measure.

As Table 1 shows, 5,143 facilities (82%) have data for all seven measures, while 321 (5%) lack data for all seven. In total, 727 facilities (12%) are not rated.

Table 1. Facilities with Missing Measures

Missing Measures	Facilities		Facilities	
	Facilities	%	Unrated	%
0	5,143	82	27	4
1	334	5	40	6
2	180	3	59	8
3	135	2	109	15
4	75	1	60	8
5	53	1	46	6
6	66	1	65	9
7	321	5	321	44
Total	6,307		727	12

Source: January 2015 DFC data.

The most common missing measure is STTrR, with 952 facilities (16%) not reporting this metric (see Table 2). For the other measures, non-reporting rates ranged from 8% to 10%.

Table 2. Missing Data by Measure

Measures	Facilities with Missing Data	%
STTrR	952	16
SHR	551	9
SMR	599	10
Kt/V	511	8
Calcium	524	9
Fistula	585	10
Catheter	585	10

Source: January 2015 DFC data.

The seven individual metrics are reported on different scales. STTrR, SHR, and SMR have mean scores of 1.0. For these metrics, a lower score is better. Calcium and Catheter have mean scores of 2.3 and 10.5, respectively, and lower scores also reflect better outcomes. For Kt/V and Fistula, higher scores are better, and the average facility's results are 87.9 and 63.8, respectively. Table 3 reports the average score for each metric among facilities within each of the five star-rating categories.³

Table 3. Mean Values by Star Rating

	1	2	3	4	5
STTrR	1.5	1.2	1.0	0.8	0.6
SHR	1.3	1.1	1.0	0.9	0.8
SMR	1.3	1.1	1.0	0.9	0.8
Kt/V	79.2	85.0	88.5	91.3	93.0
Calcium	4.4	3.4	2.3	1.2	0.8
Fistula	51.1	58.7	63.9	69.3	74.7
Catheter	19.5	13.7	9.9	7.1	5.1

Source: January 2015 DFC data.

The distribution of rated facilities is very close to CMS's stated objective: 9.8% are 1-star, 19.8% are 2-star, 40.1% are 3-star, 20.2% are 4-star, and 10.1% are 5-star.

³ CMS has previously reported Tables 1–3 based on January 2014 data. However, the final published star ratings are based on newer data, available as of January 29, 2015. This analysis reports results from that new dataset. The January 2015 dataset bases all quality measures on calendar year 2013 data except SMR, which is based on the previous four years of data (January 2010–December 2013).

ANALYSIS

Distribution Analysis

In order to better understand the relationship between the star ratings associated with different types of facilities, census tract-level demographic data from the U.S. Census Bureau are matched with the CMS star-rating data using facility zip codes. The Census Bureau data include information about median incomes, educational attainment, race, and age. These data are also matched with facility-level data from the 2011 Dialysis Facility Report (DFR) dataset, which reflects data from calendar year 2010 and allows for the incorporation of information about facility staffing levels.⁴

Table 4 on the following page details the distribution of star ratings for various subgroups. After reporting the share of facilities in each star-rating category, the table reports the total number of facilities within the given subgroup, the average score, and the results of a t-test and Pearson's chi-squared test. The t-test determines if the average score for that subgroup is statistically different from the mean of the rest of the dialysis facilities. Pearson's chi-squared test assesses whether the distribution of scores in a subpopulation is statistically independent from the star rating. P-values of less than 0.05 indicate statistically significant differences in the mean (t-test) and the distribution (chi-squared test).

Table 4 illustrates variations in star ratings across subgroups of facilities. Facilities in census tracts with a black population above the national median tend to have lower star ratings—just 6.6% are 5-star, and the average score is 2.87. Conversely, tracts with a white population above the national median have an average score of 3.05, and 11.6% are 5-star.

Facilities in low-income communities clearly perform poorly. Facilities located in census tracts where the median income is in the bottom 20% average 2.8 stars and are twice as likely to have 1 star (18.4%) compared to facilities with median incomes near the national average (9.1%). States with poverty rates below the median rate have dialysis facilities that average 3.09 (not shown in Table 4).

There is considerable variation in star ratings across states. The three states with the most facilities (Texas, California, and Florida) have mean scores of 3.19, 3.29, and 2.60, respectively. New York, Ohio, and Washington, three diverse states spread across the country, have mean scores of 2.82, 2.60, and 3.57, respectively. (See Figure 1 on the next page.)

In Puerto Rico, 62% of facilities are 1-star, while 38% are 2-star. Dialysis Patient Citizens, a patient advocacy group, [reports](#) that healthier states tend to have higher-rated facilities.

⁴ At the time of publication, the author was aware that more recent DFR public data may be available, and future work will hopefully utilize that more recent data. However, a comparison of comparable variables in both the star-rating dataset and the DFR suggest that most facilities are highly similar in 2013 and 2010.

Table 4. Distribution of Star Ratings for a Range of Facility Subpopulations

	Star Rating (%)					No. of facilities	Avg. score	T-test (p-value)	Chi-squared test (p-value)
	1	2	3	4	5				
All facilities	9.8	19.8	40.1	20.2	10.1	5,580	3.0	N/A	N/A
All nonprofit	10.4	18.7	39.3	18.1	13.5	758	3.1	0.24	0.01
Black % above national median	11.3	21.0	43.4	17.7	6.6	2,060	2.9	0.00	0.00
White % above national median	9.0	20.9	38.1	20.4	11.6	3,104	3.0	0.01	0.00
Median income in bottom 20%, by census tract	18.4	19.7	35.0	19.7	7.2	223	2.8	0.00	0.00
Median income in middle 20%, by census tract	9.1	19.8	40.3	19.8	10.9	2,383	3.0	0.16	0.35
Median income in top 20%, by census tract	7.1	17.7	39.0	24.8	11.3	141	3.2	0.11	0.54

Source: January 2015 DFC data and U.S. Census Bureau.

Figure 1. State Variation in Star Ratings



Regression Analysis

The distribution analysis above suggests that community demographics and geography impact a facility's star rating. Given that the objective of the star-rating program is to measure facility quality, geographic and demographic variation in scores is suggestive of potential measurement bias. Specifically, if patient and community demographic data predict a facility's star rating, this may indicate that patient characteristics influence a system intended to measure facility quality. Regression analysis can further examine this question by estimating the impact of various explanatory variables on star ratings.

Predicting Star Ratings

Table 5 reports ordinary least squares (OLS) regression results. I regress star ratings on facility traits (column 1), add patient and community demographics (column 2), and add state dummy variables (column 3). Column 4 excludes variables from the 2011 DFR. Independent variables include dummy variables to indicate if a facility is owned by one of the three large dialysis organizations (LDOs) or by another chain (compared to independent facilities); patient/staff ratios; the share of patients on Medicare; average patient age; and, by census tract, the log of the median income and the percent of the population with a college degree or above.

Table 5. Effect of Facility and Demographic Factors on Star Rating

	(1)	(2)	(3)	(4)
LDO: Davita	0.796** (0.048)	0.803** (0.048)	0.840** (0.049)	0.829** (0.046)
LDO: Fresenius	-0.217** (0.047)	-0.211** (0.048)	-0.152** (0.049)	-0.134** (0.047)
LDO: DCI	0.361** (0.080)	0.393** (0.080)	0.466** (0.081)	0.479** (0.079)
Non-LDO chains	0.154** (0.051)	0.169** (0.051)	0.150** (0.053)	0.173** (0.050)
Number of stations	-0.007** (0.002)	-0.004* (0.002)	-0.005** (0.002)	-0.006** (0.002)
Patient/staff ratio, 2010	-0.035** (0.009)	-0.027** (0.009)	-0.013 (0.009)	
% not on Medicare, 2010		-0.002 (0.003)	0.001 (0.003)	
Average patient age, 2010		0.008* (0.004)	0.016** (0.004)	
% white patients, 2010		0.002** (0.001)	0.002* (0.001)	
Log median income		0.193** (0.041)	0.099* (0.042)	0.140** (0.040)
% bachelors degree or above		-0.006** (0.001)	-0.005** (0.001)	-0.005** (0.001)
Observations	4,978	4,946	4,946	5,560
R-squared	0.156	0.168	0.239	0.221
State dummies	No	No	Yes	Yes

Standard errors in parentheses. ** p<0.01, * p<0.05, + p<0.1

Regression results indicate the following:

- Census tract–level median income is a statistically significant factor in predicting star rating, after controlling for other observable factors. For example, a 10% increase in median wage increases the star rating by at least 0.1 (depending on the particular regression specification). A more significant change in median income—from \$40,000 to \$60,000—would result in a 0.5 or more increase in the rating.
- Relative to being an independent facility, the marginal impact of being Davita-owned is 0.8 stars; DCI-owned, 0.4–0.5 stars; and another chain, 0.2 stars. The marginal effect of being a Fresenius facility is negative.

Predicting Score Components

Three of the seven score components (STrR, SHR, and SMR) are risk-adjusted factors that reflect variations in patient population; the other four are not. Table 6 reports regressions for each factor.

- SHRs and SMRs are largely independent of the observable factors.
- LDOs and other chains achieve significantly better STrRs than independent facilities.
- For Kt/V, Calcium, Fistula, and Catheter, ownership has a significant effect. For Catheter and Fistula, census tract–level median income has a powerful positive effect.

Table 6. Effect of Facility and Demographic Factors on Quality Measures

	(1) SHR	(2) SMR	(3) STrR	(4) Kt/V	(5) Calcium	(6) Catheter	(7) Fistula
LDO: Davita	0.037** (0.014)	-0.006 (0.013)	-0.219** (0.027)	5.212** (0.441)	-1.643** (0.134)	-5.750** (0.308)	4.940** (0.522)
LDO: Fresenius	0.047** (0.014)	-0.026* (0.013)	-0.209** (0.027)	-1.277** (0.444)	1.606** (0.134)	-3.443** (0.310)	2.829** (0.525)
LDO: DCI	-0.042+ (0.023)	-0.093** (0.021)	-0.298** (0.044)	3.575** (0.724)	0.045 (0.216)	-3.059** (0.503)	3.088** (0.852)
Non-LDO chains	0.036* (0.015)	0.011 (0.014)	-0.134** (0.029)	2.372** (0.477)	0.183 (0.144)	-2.064** (0.333)	1.349* (0.564)
Number of stations	0.000 (0.001)	-0.002** (0.000)	-0.002* (0.001)	0.121** (0.017)	-0.004 (0.005)	0.015 (0.012)	-0.050* (0.020)
Patient/staff ratio, 2010	0.012** (0.002)	-0.008** (0.002)	0.008+ (0.005)	0.129+ (0.077)	-0.027 (0.023)	-0.123* (0.054)	0.125 (0.092)
% not on Medicare, 2010	0.004** (0.001)	-0.000 (0.001)	-0.005** (0.002)	-0.142** (0.029)	-0.004 (0.009)	-0.043* (0.020)	0.079* (0.035)
Avg. patient age, 2010	-0.003** (0.001)	-0.005** (0.001)	0.000 (0.002)	0.408** (0.035)	-0.016 (0.011)	0.015 (0.025)	-0.091* (0.042)
% white patients, 2010	-0.001** (0.000)	0.000 (0.000)	0.001* (0.000)	-0.043** (0.006)	-0.008** (0.002)	0.006 (0.004)	0.090** (0.007)
Log median income	0.029* (0.012)	-0.016 (0.011)	-0.046* (0.023)	0.565 (0.379)	-0.122 (0.114)	-0.956** (0.264)	1.409** (0.447)
% bachelors or above	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.001)	-0.071** (0.009)	0.006* (0.003)	0.033** (0.007)	-0.029** (0.011)
Observations	4,938	4,946	4,765	4,917	4,882	4,894	4,894
R-squared	0.183	0.099	0.125	0.165	0.255	0.179	0.180
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. ** p<0.01, * p<0.05, + p<0.1.

CONCLUSION

CMS created the dialysis facility star-rating system to help patients find the best quality of care in the vital area of in-center dialysis treatment. Improving consumer awareness and promoting health care quality are worthy goals, but the star-rating system as currently constructed may not be accurately directed toward those ends.

In addition to legitimate concerns raised by other stakeholders about the preset distribution of the star ratings, and given the importance of patient demographics in predicting patient outcomes and health status, it is necessary to consider refinements to the star-rating system. This preliminary analysis suggests that the current system may be predicting local demographic factors in addition to facility quality metrics.

Specifically, there are statistically significant differences in the average star rating for facilities located in lower-income census tracts and facilities in census tracts with higher-than-average black populations.

Regression analysis results indicate that patient age and median income are significant predictors of some of the seven individual underlying metrics even after controlling for facility ownership, facility size, patient/staff ratios, and state.

In many respects, the finding that patient characteristics influence health outcomes should come as no surprise as it is a common finding in many aspects of health care. To ensure that the star-rating system accurately informs patients about the quality of care they can expect from a given facility, star-rating scores could be adjusted to account for underlying community characteristics.

ABOUT THE AUTHOR

Alex Brill is an economic policy advisor to Hooper, Lundy & Bookman. Previously, he served as policy director and chief economist to the House Ways and Means Committee and as an economist at the White House Council of Economic Advisers. Currently also a research fellow at the American Enterprise Institute, he has a BA in Economics from Tufts University and an MA in Mathematical Finance from Boston University.

ABOUT THE FIRM

Founded in 1987, Hooper, Lundy & Bookman is the largest full service law practice in the country dedicated solely to the representation of health care providers and suppliers. With offices in Los Angeles, San Francisco, San Diego and Washington, D.C., and clients in all 50 States, we meet the business, litigation, regulatory and government relations needs of a broad array of health care providers—ranging from the largest national health care organizations to community hospitals and individual physician practices. For more information, please visit our website at www.health-law.com.