RESEARCH ARTICLE

Evaluating national trends in outcomes after implementation of a star rating system: Results from dialysis facility compare

Stephen Salerno MS[1] | Claudia Dahlerus PhD[1] | Joseph Messana MD[1] |
Karen Wisniewski MPH[1] | Lan Tong MPH[1] | Richard A. Hirth PhD[1] |
Jordan Affholter BA[1] | Garrett Gremel MS[1] | YiFan Wu MPH[1] | Ji Zhu PhD[1] |
Jesse Roach MD[2] | Elena Balovlenkov RN[2] | Joel Andress PhD[2] | Yi Li PhD[1]

[1]Kidney Epidemiology and Cost Center, University of Michigan, Ann Arbor, Michigan, USA

[2]The Centers for Medicare & Medicaid Services, Baltimore, Maryland, USA

Correspondence
Yi Li, PhD, Kidney Epidemiology and Cost Center, University of Michigan, Ann Arbor, MI, USA.
Email: yili@umich.edu

Abstract

Objective: To examine which factors are driving improvement in the Dialysis Facility Compare (DFC) star ratings and to test whether nonclinical facility characteristics are associated with observed longitudinal changes in the star ratings.

Data Sources: Data were collected from eligible patients in over 6,000 Medicare-certified dialysis facilities from three annual star rating and individual measure updates, publicly released on DFC in October 2015, October 2016, and April 2018.

Study Design: Changes in the star rating and individual quality measures were investigated across three public data releases. Year-to-year changes in the star ratings were linked to facility characteristics, adjusting for baseline differences in quality measure performance.

Data Collection: Data from publicly reported quality measures, including standardized mortality, hospitalization, and transfusion ratios, dialysis adequacy, type of vascular access for dialysis, and management of mineral and bone disease, were extracted from annual DFC data releases.

Principal Findings: The proportion of four- and five-star facilities increased from 30.0% to 53.4% between October 2015 and April 2018. Quality improvement was driven by the domain of care containing the dialysis adequacy and hypercalcemia measures. Additionally, independently owned facilities and facilities belonging to smaller dialysis organizations had significantly lower odds of year-to-year improvement than facilities belonging to either of the two large dialysis organizations (Odds Ratio [OR]: 0.736, 95% Confidence Interval [CI]: 0.631-0.856 and OR: 0.797, 95% CI: 0.723-0.879, respectively).

Conclusions: The percentage of four- and five-star facilities has increased markedly over a three-year time period. These changes were driven by improvement in the specific quality measures that may be most directly under the control of the dialysis facility.

KEYWORDS
public reporting, quality measures, star ratings, dialysis, medicare

1 | INTRODUCTION

In 2000, Dialysis Facility Compare (DFC) emerged as one of the earliest public reporting programs implemented by the Centers for Medicare & Medicaid Services (CMS) and has since served as a template for reporting the quality of other providers in the Medicare program.[1] For the past decade, CMS has expanded its public reporting efforts by establishing Five-Star Quality Rating Systems in response to consumers and other stakeholders. The star rating systems represent enhancements by CMS to increase the utility of publicly reported quality data on Medicare providers and health plans.[2,3] Commensurate with these initiatives, reporting of dialysis facility quality on DFC expanded in 2015 to include the DFC Quality of Patient Care Star Rating system. The goal of the DFC star rating was to give patients and other consumers a balanced, easy-to-use tool for comparing the overall quality of care provided by dialysis facilities.

The DFC star rating represents a global summary of the quality measures reported on DFC, rating facilities from one to five stars. As originally implemented, the dialysis star rating system assigned a fixed proportion of facilities to each of the five-star categories based on performance measured in each reporting period relative to other facilities. A consequence of this methodology is that even if facilities, on average, improve their performance over time, the distribution of star ratings would remain unchanged. In response to consumer and stakeholder feedback, this original methodology was updated in October 2016 to score facilities against absolute quality standards set in a baseline year. The establishment of a fixed baseline allowed consumers to compare year-over-year changes in the star ratings. Since implementing this update in 2016, the star rating distribution has shifted upward: The proportions of four- and five-star facilities have increased progressively, while those of one- and two-star facilities have consistently decreased.

Our study examines which measures in the DFC star rating drove the observed trend in star rating improvement over three public star rating releases on the DFC site: October 2015, October 2016, and April 2018. Our objective is to better understand the extent to which quality in these measured areas improved, in addition to whether specific features of the measures and methodology may have inadvertently placed greater provider focus on certain quality measures versus others. For example, dialysis providers may be able to more easily impact performance on certain intermediate outcome measures tied directly to dialysis care than broader outcome measures like mortality that could be influenced by multiple providers. Our second objective is to investigate the associations between year-over-year changes in star rating with several nonclinical metrics and to examine whether these facility or organizational characteristics are associated with greater rating improvement over time. We conclude by discussing the implications of our results on the public reporting of dialysis facility quality.

What is Known on This Topic

- Dialysis Facility Compare (DFC) is one of the earliest public reporting programs implemented by the Centers for Medicare & Medicaid Services (CMS)
- The DFC star rating represents a global summary of the quality measures reported on DFC, rating facilities from one to five stars
- Since 2016, the star rating distribution has shifted upward: The proportions of four- and five-star facilities have increased progressively, while those of one- and two-star facilities have consistently decreased

What this Study Adds

- This trend was attributed to a rapid improvement in two intermediate outcomes, dialysis adequacy and hypercalcemia, which may be directly influenced by facility practices
- Facilities from large dialysis organizations had higher odds of improving their star rating over the study period compared to independent facilities
- Our study provides a systematic approach to understanding the sources underlying changes in the star rating, which may aid dialysis patients in making informed decisions about their care

---

2 | METHODS

2.1 | Data

We utilized CMS clinical and administrative data on eligible chronic dialysis patients in Medicare-certified dialysis facilities. The data were collected from three annual star rating and individual measure updates, publicly released on DFC in October 2015, October 2016, and April 2018 (postponed release from October 2017).

2.2 | DFC quality measures

Seven DFC quality measures, which broadly represent either primary or intermediate outcomes, are used in this analysis. The three primary outcome measures are the Standardized Mortality Ratio, the Standardized Hospitalization Ratio, and the Standardized Transfusion Ratio. The Standardized Transfusion Ratio serves as a measurable marker for the quality of a facility's anemia management (the primary outcome) in ESRD patients. These three measures report the ratio of the number of observed events (deaths, hospitalizations, or blood transfusions, respectively) to the number of events that would be expected based on the characteristics of each facility's

patients. Lower measure values indicate better performance. The national average value for each of the standardized measures is approximately 1.0, interpreted as a facility performing "as expected."

The four intermediate outcome measures include total Kt/V—a summary of three DFC Kt/V measures that apply to different patient subpopulations (adult hemodialysis, adult peritoneal dialysis, and pediatric hemodialysis), vascular access for hemodialysis (fistula or long-term catheter), and hypercalcemia (blood serum or plasma calcium >10.2 mg/dL). Kt/V reports the time-averaged small solute clearance of urea nitrogen, a marker for hydrophilic waste products of intermediary metabolism that accumulate in patients with kidney failure. Essentially, higher Kt/V serves as a marker of the delivery of a higher "dose" of dialysis via the removal of more urea nitrogen, and the measure is an indicator of the adequacy of the dose. The preferred type of vascular access for delivering hemodialysis treatment is an arteriovenous (AV) fistula. An AV fistula is associated with the lowest risk of infection and other complications that put patients at higher risk for hospitalization. In contrast, a long-term central venous catheter is the least preferred vascular access for most patients. Finally, the hypercalcemia measure assesses one component of the facility's management of bone and mineral disease. Elevated calcium levels place patients at higher risk for cardiovascular events and associated morbidity and mortality. Individually, these four intermediate quality measures represent the percent of patients within a dialysis facility meeting each of the above criteria. Higher percentage values for the total Kt/V and fistula measures indicate better performance, while the opposite is true for the hypercalcemia and catheter measures.

To calculate the star ratings, all measure values are first transformed into either probit (risk-adjusted standardized measures) or z-scores (percentage measures). This transformation is done relative to the data in the baseline period.[4] The transformed measures are then grouped into three quality domains. Domain 1 consists of the three standardized primary outcome measures, Domain 2 consists of the two vascular access measures, and Domain 3 consists of total Kt/V and hypercalcemia. The measure scores within each domain are averaged and the three resulting domain scores are averaged, with equal weight, to produce a final score for each facility. Note that facilities that treat only peritoneal dialysis patients are exempt from scoring on Domain 2. Cutoff values defining the star rating categories are then calculated using data from a fixed baseline period to determine a facility's star rating in subsequent releases. Star ratings are assigned by comparing the final scores to these cutoffs, thereby allowing the distribution of star ratings to shift toward higher ratings if average facility performance improves over time. Additional methodological details concerning the standardization of the quality measures can be found in the supplemental materials. Further details on how the star ratings were calculated can be found in the Technical Notes on the Updated Dialysis Facility Compare Star Rating Methodology.[4]

2.3 | Trends analysis

The star ratings compare performance in a current period of evaluation, or the annual DFC data release, to star category cutoffs established from a fixed baseline period. This allows consumers to track changes in facility performance over time against a pre-determined standard. For the three consecutive data releases utilized in this study, the baseline period data were released in October 2015.

To study what aspects of care quality drove the observed upward shift in the star ratings, average final scores, domain scores, and quality measure values were calculated by averaging each metric across all facilities with available data for each release. The summary statistics presented in the results for the percentage measures (total Kt/V, hypercalcemia, fistula, and catheter) reflect the means of the raw reported measure values. Due to the standardization of the ratio measures for mortality, hospitalizations, and transfusions, reporting mean values directly would not inform a trend in these data, as they inherently have a mean value of approximately 1.00 each year. Therefore, the standardized ratios reported in the results section were first multiplied by an adjustment factor, to adjust for differences in event rates between each evaluation period and the baseline period. This adjustment factor is also used in the star rating system.[4] To further examine potential differences in trends across provider operational structures, we stratified these results by organizational (chain) affiliation: independently owned facilities, facilities owned by small chain organizations (2-1000 affiliated facilities), and facilities owned by large chain organizations (1000+ affiliated facilities). We note that the large chain group is comprised of two large dialysis organizations, which together constitute ownership for approximately 70% of all facilities nationwide.

2.4 | Associations between quality improvement and nonclinical facility characteristics

We further modeled the year-to-year change in facility star rating as a function of several facility organizational and locational characteristics: facility size, urbanicity, regional socioeconomic disadvantage, and chain affiliation. Change in star rating was classified into three categories: a decrease in rating, maintaining the same rating, or an increase in rating between consecutive DFC releases. Facility size was defined by classifying facilities into the smallest, middle, or largest tertile for the number of reported patients. Based on known facility locations, facility urbanicity was defined as urban or rural based on the July 2016 Office of Management and Budget's definition of core-based statistical areas.[5] Additionally, each facility was assigned an Area Deprivation Index (ADI) percentile rank. The ADI was constructed by the University of Wisconsin-Madison Neighborhood Atlas Project using 5-year data from the American Community Survey.[6] Percentile ranks were created by grouping neighborhoods nationally into 1% ADI intervals. A ranking of 1 indicates the least disadvantaged neighborhoods, while a ranking of 100 indicates the most disadvantaged. This metric was scaled by a factor of 10 in our model. Lastly, we categorized chain affiliation into three levels: independently owned facilities, facilities owned by small dialysis organizations, and facilities owned by the two largest dialysis organizations.

To mitigate potential multicollinearity with other facility characteristics of interest, information such as profit status and hospital affiliation were excluded from our analytical model (see Tables S5 and S6). In addition, control variables were added to account for facilities' baseline final scores and facilities with a one- or five-star baseline star rating, as these facilities could only maintain their rating or experience a change in one direction. To control for potential confounding from measure definition changes and differences in the population of facilities between releases, the star ratings were retrospectively recalculated with consistent measure definitions and a common population of facilities across all three DFC releases. Thus, all facilities modeled were both open and eligible to receive a star rating, under the same criteria, at the time of each release. Lastly, to account for correlations among the ratings for the same facility over years, we fit a proportional odds cumulative logit model using a generalized estimating equation (GEE) approach to compute robust standard errors for the parameter estimates.[7,8] When implementing GEE, we used a compound symmetric (exchangeable) covariance structure. All analyses were carried out using SAS software, version 9.4 of the SAS System for Windows.[9]

3 | RESULTS

Relevant data were extracted from each of the three DFC releases used in this study, resulting in $n_1 = 6418$, $n_2 = 6606$, and $n_3 = 7047$ eligible facility records for the October 2015, October 2016, and April 2018 releases, respectively. These facilities all contributed quality measure data based on the measure definitions in place at the time of each release. Due to the star rating guidelines for data quality, 546, 545, and 844 of these facilities, respectively, were not eligible to receive a star rating. These ratings were suppressed

primarily if a facility did not have enough patients or was not open for a sufficient portion of the evaluation period for a particular DFC release.[4]

Since the establishment of the star category cutoffs based on the October 2015 baseline year data, the percentage of four- and five-star facilities increased from 30.0% in year 2016 to 53.4% in 2018, whereas the percentage of one- and two-star facilities decreased from 30.0% to 11.4% (see Figure 1). Figure 2 reports mean measure values for the individual clinical quality measures used in the star ratings (a), their corresponding domain-averaged measure sores (b), and resulting average final scores (c), stratified by chain affiliation. Mean performance values for the total Kt/V (dialysis adequacy) and hypercalcemia quality measures improved most notably over the three DFC release periods (increasing values for total Kt/V and decreasing values for hypercalcemia) and most rapidly among facilities belonging to either of the large chain organizations. Mortality rates, transfusion rates, and arteriovenous fistula utilization also improved, while average performance on the catheter utilization and hospitalization measures declined. However, these changes were less pronounced than the changes observed for total Kt/V and hypercalcemia (see Tables S1 for average measure values, standardized scores, and standard deviations).

Mean quality measure domain scores more than doubled in Domain 3, which includes the dialysis adequacy and hypercalcemia measures. This suggests that the Domain 3 measures were the main drivers in the improvement of facility-level final scores and the upward shift in the star rating distribution over this period. Scores in Domain 2 (vascular access type) improved slightly as of the April 2018 release, while average Domain 1 scores (standardized mortality, hospitalizations, and transfusions) improved in the October 2016 release then declined slightly in the April 2018 release. These trends in mean domain scores persisted in a sensitivity analysis of domain
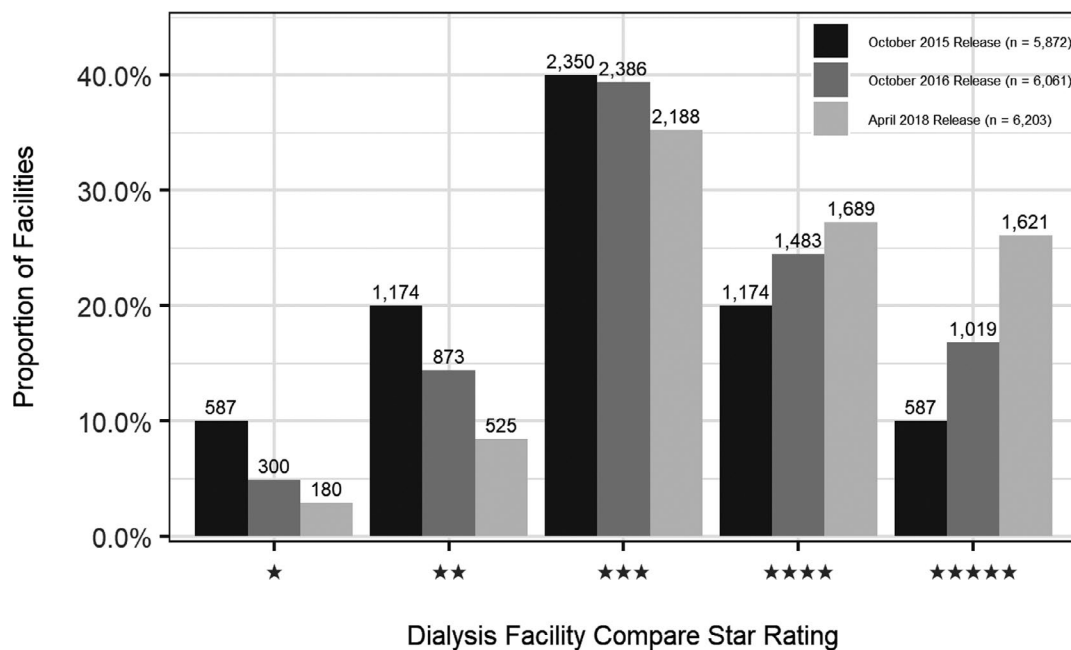

FIGURE 1 Distribution of the Dialysis Facility Compare Star Rating by data release year
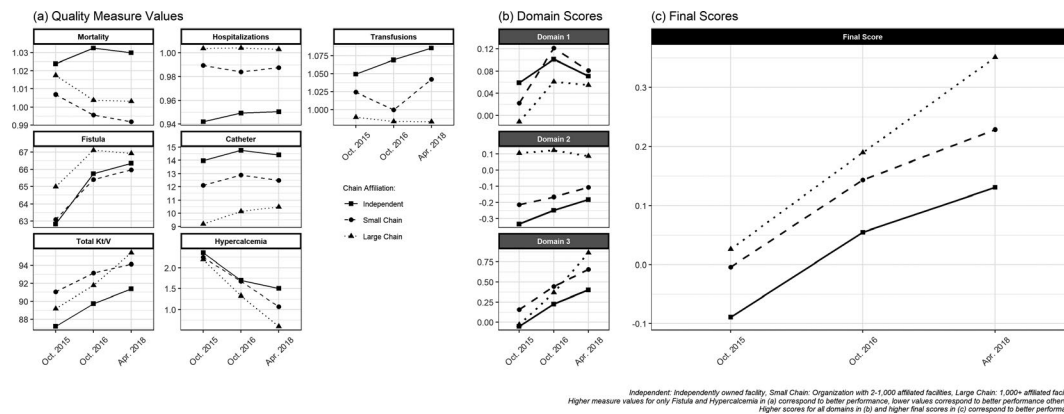
FIGURE 2 Trends in clinical measure values and standardized scores over time. (A) Mean measure values for the individual clinical quality measures used in calculating the Dialysis Facility Compare Star Rating; (B) their corresponding domain-averaged measure sores; and (C) the resulting average final scores, stratified by chain affiliation

TABLE 1 Descriptive statistics for nonclinical characteristics of the n = 5914 common population of dialysis facilities in our analytical sample, stratified by facility chain affiliation

Characteristic	Overall	Stratified by chain affiliation[a]			P-value[b]
		Independent	Small chain	Large chain	
n	5914	602	994	4318	
Urbanicity, n (%)					
Rural	401 (6.8)	55 (9.1)	74 (7.4)	272 (6.3)	<.001
Urban	5354 (90.5)	412 (68.4)	899 (90.4)	4043 (93.6)	
Missing	159 (2.7)	135 (22.4)	21 (2.1)	3 (0.1)	
Facility size, n (%)					
Smallest tertile	1778 (30.1)	234 (38.9)	315 (31.7)	1229 (28.5)	<.001
Middle tertile	2064 (34.9)	159 (26.4)	347 (34.9)	1558 (36.1)	
Largest tertile	2072 (35.0)	209 (34.7)	332 (33.4)	1531 (35.5)	
National ADI rank, mean (SD)	54.78 (27.10)	54.44 (29.75)	54.29 (27.38)	54.93 (26.66)	.759
Baseline final score, mean (SD)	0.17 (0.48)	0.07 (0.64)	0.15 (0.50)	0.19 (0.44)	<.001
One-star at baseline, n (%)					
Yes	322 (5.4)	81 (13.5)	71 (7.1)	170 (3.9)	<.001
No	5592 (94.6)	521 (86.5)	923 (92.9)	4148 (96.1)	
Five-star at baseline, n (%)					
Yes	961 (16.2)	114 (18.9)	163 (16.4)	684 (15.8)	.154
No	4593 (83.8)	488 (81.1)	831 (83.6)	3634 (84.2)	

Abbreviations: National ADI Rank, National Area Deprivation Index Percentile Rank; SD, standard deviation.

[a]Independent: Independently owned facilities; Small Chain: Facilities owned by small chain organizations (2-1000 affiliated facilities); Large Chain: Facilities owned by large chain organizations (1000+ affiliated facilities).

[b]Unadjusted P-values for differences in the distributions of the nonclinical characteristic between chain affiliations (chi-squared tests for discrete variables, Kruskal-Wallis tests for continuous variables).

scores derived by retrospectively applying consistent measure definitions in each release year, with improvements in Domain 3 being slightly more pronounced (see Tables S2, S3 and S4).

To address the second aim of our study, individual clinical quality measure values, facility final scores, and star ratings were retrospectively recalculated, applying consistent measure definitions to a common population of facilities across the three DFC releases. Under these measure specifications (the most current as of the April 2018 release), n = 5914 facilities were eligible to receive a star rating in all three releases. Nonclinical characteristics for this common

population of facilities used in our analytical model are presented in Table 1. The majority of facilities were in urban areas (90.5%), owned by one of the two large dialysis organizations (73%), and were rated neither one nor five stars at baseline (78.4%). Considering chain affiliation, we found that a higher proportion of chain facilities were located in urban areas and were larger in size. No significant differences were found in area deprivation between chain affiliation groups.

Results from the proportional odds cumulative logit model are presented in Table 2. As shown, independently owned facilities and facilities belonging to smaller dialysis organizations had significantly lower odds of year-over-year improvement when compared to facilities belonging to either of the two large dialysis organizations (Odds Ratio [OR]: 0.736, 95% Confidence Interval [CI]: 0.631-0.856 and OR: 0.797, 95% CI: 0.723-0.879, respectively). Facilities in the smallest size tertile had 29% higher odds of improvement when compared to facilities in the largest tertile (OR: 1.291, 95% CI: 1.177-1.417). The national area deprivation index had a small, but significant, effect with higher area deprivation being associated with lower odds of improvement (OR: 0.981, 95% CI: 0.967-0.994). However, no significant differences in improvement were found between facilities located in a rural area *vs* an urban area, after adjusting for all other facility characteristics.

4 | DISCUSSION

A substantial upward shift was observed in the proportions of four- and five-star facilities between October 2015 and April 2018. This trend was attributed to a rapid improvement of Domain 3, consisting of two intermediate outcomes, total Kt/V and hypercalcemia, which may be directly influenced by facility practices (delivery of dialysis and provision of medications, respectively). This result persisted when using consistent measure definitions across the DFC releases, as compared to those implemented for each release, suggesting improvement in quality performance was driven by improved outcomes rather than changes in measure definitions. We also noted improved reporting in the hypercalcemia measure over time (see Tables S4). We did not observe substantial changes in national performance for mortality, hospitalizations, transfusions (Domain 1), or type of vascular access (Domain 2).

The goal of public reporting for quality measures is to provide information to patients that can help them choose where to receive health care or to know how well their current providers perform on care delivery across a range of important outcomes.[1,2] The DFC star ratings were implemented to provide objective and easy-to-use information to patients, their families, and other stakeholders for comparing the quality of dialysis facility care. The rather rapid increase in four- and five-star facilities may reduce the utility of the star rating, particularly when there is greater variation in facility performance within the larger four- and five-star categories. Additionally, the main drivers for high performance are measures most directly under the control of the dialysis facility. Despite this, a potential benefit to public reporting of facility outcomes is the creation of improvement incentives which facilitate provider marketing opportunities. The rapid improvement in those metrics under the direct control of dialysis providers may be an indicator of the power of those incentives.

The progressive improvement in both the total Kt/V and hypercalcemia measures over time led to the increase of facilities with higher star ratings. This may be, in part, a consequence of the star rating design paradigm, which placed equal weight on all three

TABLE 2 Estimated log-odds ratios (Estimate), standard errors (SE) and 95% confidence intervals (95% CI) from the proportional odds cumulative logit regressing change in star rating (decrease, no change, or increase) on facility characteristics

Parameter	Estimate	SE	95% CI	Z-value	P-value
Intercept 1	0.0606	0.0525	(−0.0423, 0.1636)	1.1539	.2485
Intercept 2	0.5888	0.0529	(0.4851, 0.6925)	11.1258	<.0001
Urbanicity (Reference = Urban)					
Rural	0.1219	0.0809	(−0.0367, 0.2804)	1.5064	.1329
Facility size (Reference = Largest Tertile)					
Smallest tertile	0.2557	0.0473	(0.1630, 0.3483)	5.4095	<.0001
Middle tertile	0.0814	0.0434	(−0.0037, 0.1664)	1.8753	.0607
Chain affiliation (Reference = Large chain)					
Independent	−0.3078	0.0776	(−0.4598, −0.1557)	−3.9680	<.0001
Small chain	−0.2265	0.0500	(−0.3244, −0.1286)	−0.5334	<.0001
National ADI rank	0.0196	0.0069	(−0.0332, −0.0061)	−2.8357	.0046
Baseline final score	−0.0057	0.0007	(−0.0072, −0.0043)	−7.7648	<.0001
One-star rating at baseline (Reference = No)					
Yes	0.6830	0.0788	(0.5287, 0.8374)	8.6722	<.0001
Five-star rating at baseline (Reference = No)					
Yes	−1.6874	0.0621	(−1.8091, −1.5657)	−27.1743	<.0001

Abbreviation: National ADI Rank, National Area Deprivation Index Percentile Rank.

measure domains. In effect, this attributed equal weight to Domain 3, a domain whose outcomes could be directly influenced by dialysis facility practices, and Domains 1 and 2, comprised of measures (mortality, hospitalization, and blood transfusion; and creation and maintenance of vascular access) that require more resources and coordination of care in order to improve performance. Dialysis adequacy and hypercalcemia are generally considered to be both more directly attributed to, and actionable by, dialysis facilities.[10,11] It is possible that facilities' efforts to improve on these measures, combined with the star rating weighting scheme, had the unintended consequence of focusing provider efforts on intermediate outcomes, which can be improved over the shorter term but may ultimately be less important to patients than primary outcomes like mortality, hospitalizations, or vascular access. In hindsight, the DFC star ratings might have been more impactful on improving these primary outcomes if proportionally greater weight had been assigned to Domains 1 and 2. For future DFC releases, CMS has announced an update to the methodology which will include a 50% reduction in the weight of Domain 3, relative to Domains 1 and 2. However, the reweighting approach will require a tradeoff with one of the primary stated goals of the star ratings: to provide a simple, balanced summary of all DFC results.

Our study also found that facilities from large dialysis organizations had higher odds of improving their star rating over the study period compared to independent facilities. This result is best understood in context. The structure of dialysis care in the United States is dominated by medium and large dialysis chain organizations, with well over half the market made up of for-profit facilities owned by two large chains. Incentives are strong for these organizations to have high star ratings in order to maintain or grow their market share and attract new dialysis patients, something that has also been seen within the nursing home setting.[12] Large dialysis organizations also enjoy economies of scale, potentially allowing them to more efficiently monitor and track their quality data across facilities and readily implement quality improvement programs to maintain or improve outcomes. It is possible that the high achievement on the quality indicators most directly under the control of dialysis facilities is, in part, influenced by incentives to focus on those measures in the star rating that can be directly impacted by facilities in relatively short periods of time. This is borne out in studies showing that for-profit facilities tend to have better performance improvement on intermediate outcomes like dialysis adequacy and anemia management, with mortality outcomes being similar between nonprofit and for-profit facilities.[13-15]

Some in the dialysis community have questioned the value of including the dialysis adequacy and hypercalcemia measures in public reporting and the star ratings because they are already at high levels of achievement and not as meaningful as outcomes like mortality and hospitalization.[10,11,16,17] Nissenson has argued that, while performance on certain intermediate outcomes has been high, it has not resulted in demonstrable improvements in primary outcomes. Moreover, Nissenson notes that while strong performance is necessary on these intermediate outcomes, it cannot be the only focus

in order to achieve excellent primary outcomes. Better outcomes in these areas can be achieved if the dialysis community focuses on more complex quality metrics that require coordination of care beyond the dialysis facility.[18] Moving the needle on primary outcomes thus requires a paradigm shift and realignment of incentives to focus more on metrics like mortality. If this is the case, weighting of quality measure domains in the DFC star rating needs to reflect outcomes that the quality program wants to prioritize because of their importance to patients, providers, and payers. This will have implications for interpreting changes in star rating trends over time, where improvement may be slower.

In other sectors of CMS public reporting, such as for nursing homes and hospitals, there is strong interest in demonstrating clinically meaningful associations between the star ratings and quality performance. Several studies on the Nursing Home Compare Star Rating have found minimal or no clear association between star ratings and patient safety and patient-reported outcomes.[19-22] It was found that the association between the Nursing Home Compare star ratings and preventable 30-day hospitalizations was slightly diminished for postacute care stays after the implementation of the Nursing Home Compare Star Rating System in 2008.[23] However, the authors suggest nursing homes may have been focusing their efforts more on improving performance on the reported quality measures at the expense of other care areas that could prevent hospitalizations. Similarly, Brauner et al. suggest there is a generally inconsistent relationship between nursing home quality reflected in the Nursing Home Compare star ratings and some key indicators of patient safety.[24] In contrast, one study reported that the Hospital Compare star ratings were associated with a better experience of care, lower risk-adjusted mortality, and lower readmission rates.[25] Clearly there is a fair amount of variability within and across star rating programs, something that has been reported across the Hospital, Nursing Home, Dialysis Facility, and Home Health CMS Compare quality reporting programs.[26] The authors found very few geographical markets achieved high-quality ratings across all five provider settings. These differences could be attributed to a variety of factors such as those considered in our study, including the types of measures in each of the star rating systems, the star rating methodology, and the structure of each health care sector's provider market.

This study has several limitations. We only examined trends in two DFC releases after the updated star rating methodology was implemented. Since the time of this study, more data have been released, albeit with further methodological updates and changes to the star rating measure set that may make direct comparisons more difficult. We also did not contrast the results for the DFC star ratings to other CMS initiatives, such as the ESRD Quality Incentive Program (QIP). The ESRD QIP value-based purchasing program may also have helped drive the focus on dialysis adequacy and hypercalcemia, which are also in the ESRD QIP. Further, while there was a small decline in mortality trends during the study period, ESRD dialysis patient mortality overall has been on the decline since prior to the implementation of the DFC star ratings.[27] Any attribution to the star rating warrants further investigation. Lastly,

it should be noted that many factors influence patient choice of facility beyond quality. These include but are not limited to, recommendations by one's primary nephrologist, issues related to access and travel burden, and considerations surrounding modality selection (e.g., home dialysis), and quality of life.[28-30] While some initial improvement trends of star ratings were attributed to specific measures in the DFC star rating, further examination will be needed to assess the longer term association of the DFC star ratings with dialysis facility quality outcomes and patient selection of a facility. Evaluation of future trends will also be important when there are changes to the star rating measure set and methodology.

Our analysis suggests that the upward trend in star rating performance was driven by improvement in the specific quality measures that may be most directly under the control of the dialysis facility. Many facilities also have high rates of achievement on these measures. In addition, equal weighting of the star rating measure domains may have focused facility efforts to maintain performance in these areas versus other potentially more meaningful outcomes. We also found that facilities belonging to large dialysis organizations had significantly higher odds of year-over-year improvement. These facilities may more efficiently monitor and track their quality data and readily implement quality improvement programs for high achievement on the quality measures most directly under their control.

In summary, our study provides a systematic approach to understanding the sources underlying the changes in the star rating over time, which may aid dialysis patients in making informed decisions about their care. The results can further inform design of future summary reporting statistics for public reporting sites that could result in greater impact on program outcomes.

ORCID

Stephen Salerno https://orcid.org/0000-0003-2763-0494

REFERENCES

1. Frederick PR, Maxey NL, Clauser SB, Sugarman JR. Developing dialysis facility-specific performance measures for public reporting. *Health Care Financ Rev*. 2002;23(4):37–50.

2. *Health Care Transparency: Actions Needed to Improve Cost and Quality Information for Consumers*. In: Office USGA, ed. Vol GAO-15-11. Washington DC: October 2014.

3. Gerteis M, Thomas C, Blatt L, et al. *Quality Reporting on Medicare's Compare Sites: Lessons Learned from Consumer Research, 2001–2014*. Princeton: Mathematica Policy Research; 2015.

4. University of Michigan Kidney Epidemiology and Cost Center. *Technical Notes on the Updated Dialysis Facility Compare Star Rating Methodology*. 2016. https://dialysisdata.org/sites/default/files/content/Methodology/UpdatedDFCStarRatingMethodology.pdf. Accessed August 24, 2018.

5. Ratcliffe MR. Creating metropolitan and micropolitan statistical areas. *Measuring Rural Diversity*. 2006;3(1):1–13. Accessed August 24, 2018

6. Kind AJ, Buckingham WR. Making neighborhood-disadvantage metrics accessible—the neighborhood atlas. *N Engl J Med*. 2018;378(26):2456.

7. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 1986;42(1):121–130.

8. Heagerty PJ, Zeger SL. Marginal regression models for clustered ordinal measurements. *J Am Stat Assoc*. 1996;91(435):1024–1036.

9. SAS Institute. *SAS SJC*. Cary, NC: SAS Institute. 9.4 [Computer Program]; 2013.

10. Weiner DE. Assessing quality care in kidney disease: the double-edged sword versus the Gordian knot. *Semin Dial*. 2020;33(1):10–17.

11. Gupta N, Wish JB. Do current quality measures truly reflect the quality of dialysis? *Semin Dial*. 2018;31(4):406–414.

12. Perraillon MC, Brauner DJ, Konetzka RT. nursing home response to nursing home compare: the provider perspective. *Med Care Res Rev*. 2019;76(4):425–443.

13. Szczech LA, Klassen PS, Chua B, et al. Associations between CMS's Clinical Performance Measures project benchmarks, profit structure, and mortality in dialysis units. *Kidney Int*. 2006;69(11):2094–2100.

14. Zhang Y. The association between dialysis facility quality and facility characteristics, neighborhood demographics, and region. *Am J Med Qual*. 2016;31(4):358–363.

15. Foley RN, Fan Q, Liu J, et al. Comparative mortality of hemodialysis patients at for-profit and not-for-profit dialysis facilities in the United States, 1998 to 2003: a retrospective analysis. *BMC Nephrol*. 2008;9:6.

16. Fuller DS, Robinson BM. Facility practice variation to help understand the effects of public policy: insights from the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Clin J Am Soc Nephrol*. 2017;12(1):190–199.

17. Pozniak A, Pearson J. The dialysis facility compare five-star rating system at 2 years. *Clin J Am Soc Nephrol*. 2018;13(3):474–476.

18. Nissenson AR. Improving outcomes for ESRD patients: shifting the quality paradigm. *Clin J Am Soc Nephrol*. 2014;9(2):430–434.

19. Grabowski DC, Town RJ. Does information matter? Competition, quality, and the impact of nursing home report cards. *Health Serv Res*. 2011;46(6pt1):1698–1719.

20. Williams A, Straker JK, Applebaum R. The nursing home five star rating: how does it compare to resident and family views of care? *Gerontologist*. 2016;56(2):234–242.

21. Konetzka RT, Grabowski DC, Perraillon MC, Werner RM. Nursing home 5-star rating system exacerbates disparities in quality, by payer source. *Health Aff (Millwood)*. 2015;34(5):819–827.

22. Konetzka RT, Perraillon MC. Use of nursing home compare website appears limited by lack of awareness and initial mistrust of the data. *Health Aff (Millwood)*. 2016;35(4):706–713.

23. Ryskina KL, Konetzka RT, Werner RM. Association between 5-star nursing home report card ratings and potentially preventable hospitalizations. *Inquiry*. 2018;55:46958018787323.

24. Brauner D, Werner RM, Shippee TP, Cursio J, Sharma H, Konetzka RT. Does nursing home compare reflect patient safety in nursing homes? *Health Aff (Millwood).* 2018;37(11):1770–1778.

25. Wang DE, Tsugawa Y, Figueroa JF, Jha AK. Association between the centers for medicare and medicaid services hospital star rating and patient outcomes. *JAMA Intern Med.* 2016;176(6):848–850.

26. Figueroa J, Feyman Y, Blumenthal D, Jha A. Do the stars align? Distribution of high-quality ratings of healthcare sectors across US markets. *BMJ Qual Saf.* 2018;27(4):287–292.

27. Saran R, Robinson B, Abbott KC, et al. US renal data system 2016 annual data report: epidemiology of kidney disease in the United States. *Am J Kidney Dis.* 2017;69(3):A7–A8.

28. Stephens JM, Brotherton S, Dunning SC, et al. Geographic disparities in patient travel for dialysis in the United States. *J Rural Health.* 2013;29(4):339–348.

29. Moist LM, Bragg-Gresham JL, Pisoni RL, et al. Travel time to dialysis as a predictor of health-related quality of life, adherence, and mortality: the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Am J Kidney Dis.* 2008;51(4):641–650.

30. Chanouzas D, Ng KP, Fallouh B, Baharani J. What influences patient choice of treatment modality at the pre-dialysis stage? *Nephrol Dialysis Transplant.* 2012;27(4):1542–1547.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.