DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Medicare & Medicaid Services
7500 Security Boulevard, Mail Stop S3-02-01
Baltimore, Maryland 21244-1850

CENTERS FOR MEDICARE & MEDICAID SERVICES
CENTER FOR CLINICAL STANDARDS & QUALITY

Jackson Williams
Director of Government Affairs
Dialysis Patient Citizens
1012 14th Street, NW, Suite #905
Washington, DC 20005

Dear Mr. Williams:

Thank you for your "Information Quality Request for Correction" letter, dated December 1, 2014, which was submitted under the HHS\CMS Guidelines for Ensuring the Quality of Information Disseminated to the Public. We appreciate the opportunity to respond to the Dialysis Patient Citizens (DPC) concerns regarding the Star Ratings consumer information on Dialysis Facility Compare (DFC).

You requested correction out of concern that the Star Rating methodology effective as of the January 2015 release of DFC is biased, and that the Star Ratings do not comply with the Objectivity and Utility requirements of the Health and Human Services Information Quality Guidelines. Specifically, you were concerned that:

A. The DFC Star Ratings do not consider regional disparities in mortality due to socio-demographic factors, and therefore are systematically biased so as to disseminate artificially low scores in places where, in general, mortality is higher than average, and artificially high scores in places where mortality is lower than average.

B. In promulgating the DFC Star Ratings, CMS did not follow processes nor apply performance metrics necessary to ensure the usefulness of the information to its intended users.

C. The DFC Star Ratings as promulgated do not achieve usefulness as that standard is defined for presentation of health care quality information.

We appreciate the opportunity to respond to each of your concerns separately below.

The DFC Star Ratings do not consider regional disparities in mortality due to socio-demographic factors, and therefore are systematically biased so as to disseminate artificially low scores in places where, in general, mortality is higher than average, and artificially high scores in places where mortality is lower than average.

You are correct that we do not account for regional disparities in mortality, or for some related socio-demographic factors. While we understand the concern you have raised, we do not believe that the lack of risk adjustment for these factors introduces systematic bias in the Star Ratings assessment. Rather, it identifies systematic differences in facility performance from region to region.

Several measures used in the construction of the final facility score have been risk adjusted, which to some degree have accounted for the differences in the underlying health of populations

served by particular facilities. Specifically, the standardized mortality ratio, one key measure for the rating, has been adjusted for variation of mortality at the state level.

Risk adjustment for socio-demographic factors, however, requires careful consideration of which factors may be appropriate for adjustment. Such adjustments have the potential to forgive or conceal disparities in care, if, for example, higher mortality may be in part attributed to facility practices or aspects of care that are under the facility's control.

Even with adjustments accounting for variation in mortality as well as select patient characteristics, further analyses detected some possible geographic variations (across states) in the distribution of star ratings. This finding is expected as there are wide heterogeneities in quality of care across states and the star ratings are capturing this variation in quality. However regional variations do not systematically affect the relative rankings of nearby facilities. Because almost all patients are selecting from facilities that are proximate to each other, the relative rankings are more important to achieving the goal of aiding consumer decision-making, than whether all facilities in an area have slightly higher or slightly lower risk-adjusted scores than they would if they were located in another region. Finally, seeing low scores of facilities in a local area will encourage the dialogue between patients and care-providers in order to achieve a better quality of care in that region, and reflects an opportunity for improvement so that patients can receive the best possible care and have accurate information to use in making informed decisions about their care. We also note that CMS has invested in providing technical assistance to the facilities in need of improvement through the ESRD Network Program.

There are differing opinions regarding whether measures should be risk adjusted for socio-demographic variables. While some stakeholders argue that socio-demographic variable risk adjustments should be made, others believe that risk adjusting for these variables may mask the potential disparities of care provided by dialysis facilities and other providers and could potentially result in lower quality care for vulnerable populations, including Medicare beneficiaries.

We apply a risk adjustment methodology to our outcomes measures to account for co-morbidities and other factors, including age, sex, and markers of severity to ensure that dialysis facilities are not penalized for serving populations that are sicker or have higher incidences of chronic disease. In addition, we monitor the impact of socioeconomic status on dialysis facility and other provider results as we continuously refine the measures used in our quality reporting and payment programs and to address stakeholder recommendations and concerns, including socio-demographic status and risk adjustment.

We are currently collaborating with the Office of the Assistant Secretary of Planning and Evaluation (ASPE) to assess the impact of risk adjusting for socio-demographic characteristics at both the measure and program level. This work is required as part of the IMPACT Act of 2014 and may influence future methodological changes to the Star Ratings. We are also participating in a trial period with the National Quality Forum (NQF) to consider under what conditions risk adjustment by socio-demographic characteristics may be appropriate. Until these efforts reach some conclusion, we do not believe these issues are sufficiently settled for us to substantially alter our approach.

In promulgating the DFC Star Ratings, CMS did not follow processes nor apply performance metrics necessary to ensure the usefulness of the information to its intended users.

You stated that you believe that the DFC Star Ratings lacks usefulness to consumers, the website's intended users. You argue that appropriate processes were not followed, and that CMS has not applied performance metrics necessary to provide useful information.

We would like to note that the display of the measures on DFC, prior to the implementation of the Star Ratings, was difficult for many users to understand. Our experience with other CMS websites that use star ratings indicates that display of quality measures in this fashion is much easier for consumers to understand. We believe that the star ratings represent a significant improvement in patient usability over individual measures presented as odds ratios or performance rates and are consistent with congressional mandate (Affordable Care Act, Section 399JJ).

All of the quality measures underlying the Star Ratings reflect the same rigorous testing and assessment to which we subject quality measures used in quality reporting, public reporting, and value-based purchasing programs for hospitals, nursing homes, rehabilitation facilities, and other providers. With the exception of the Standardized Transfusion Ratio (STrR), all measures used to calculate the star ratings are NQF-endorsed. We submitted the STrR for endorsement in early 2015 and have subjected it to the same rigorous testing for scientific acceptability as other outcomes measures later endorsed by NQF, such as the Standardized Hospitalization Ratio (SHR) and the Standardized Mortality Ratio (SMR). As a result of that testing, we concluded that it meets the necessary threshold for scientific acceptability, and this was reflected when the STrR was proposed and finalized for implementation in the PY 2018 Quality Incentive Program (QIP).

In developing the DFC star ratings methodology, we reviewed the existing methodologies used for star ratings reported by CMS for nursing homes, health plans and physicians. We sought to be consistent with other methodologies, but to also learn from their implementation. We also considered the implications of having a more limited set of available measures derived from different data sources than could be found for nursing homes or health plans, for example. Following the rollout of the DFC Star Ratings methodology, we conducted an exhaustive comment process at the request of the dialysis community. This process included multiple Open Door Forums that served as opportunities to ask us questions related to the methodology, and an ongoing response to questions and concerns submitted to us in writing. As part of this process, our methodology has been made publicly available, including analyses we conducted in development to ensure the Star Ratings met the goal of providing patients with a meaningful and interpretable summary of available quality data.

We carefully considered alternative scoring methodologies, both as part of the development process, and in response to concerns raised by the ESRD community. After considering those concerns, we determined that the star ratings methodology was appropriate to the task of providing patients and other consumers with reliable and valid summary data on the quality of care received at dialysis facilities and made the decision to move forward with program implementation in January 2015.

We conducted consumer testing on the presentation of star ratings to dialysis patients, and concluded that patients were able to understand the relevance of the star ratings. They understood that more stars indicated better performance and were willing to make use of the star ratings for their own decisions. The ratings did not appear to overshadow other concerns, such as proximity to the facility. As additional data become available it will be necessary to consider the role of new measures to address gaps in the current Star Ratings methodology. In an effort to continuously improve, we are convening a Technical Expert Panel that will be responsible for

reviewing the methodology and measures included in the star ratings, and making recommendations regarding changes to the methodology and measure selection.

In summary, we have incorporated quality data tested to endorsement standards, and in all cases, either endorsed by NQF or implemented in the QIP. We have received comments on the specific rating methodology and responded to those comments, and we have tested the presentation of the ratings to dialysis patients to ensure that they are useful to consumers who visit Dialysis Facility Compare.

The DFC Star Ratings as promulgated do not achieve usefulness as that standard is defined for presentation of health care quality information.

You expressed concern about CMS' decision to use a single composite score for the star ratings. Section 399JJ of the Affordable Care Act mandates that the "Secretary shall make available to the public, through standardized Internet websites, performance information summarizing data on quality measures." The use of scores for individual measures or measure domains was considered for this and other websites implementing the star ratings. In some cases, multiple sub-domain ratings were considered appropriate, because of the nature of available data. Hospital Compare, for instance, publicly reports on dozens of quality measures across a broad array of quality domains and topics. Nursing Home Compare includes multiple domains, including patient staffing, because those data are available, but are not necessarily considered quality measure data.

By contrast, we determined that DFC would use only nine quality measures for the initial star ratings, in large part due to our decision to only use data previously reported on DFC. In many cases, only a single quality measure is relevant to a particular domain due to missing data, small facility size, or similar reasons. Factor analyses indicated that the measures available fell within three statistically coherent domains, and we considered reporting using those groupings independently, but elected not to do so for the following reasons.

1. A non-trivial proportion of facilities include no data for at least one measure, meaning that these sub-domains would often be based on only a single indicator.
2. The statistically coherent domains were statistical constructs, and did not necessarily fit within a coherent theme easily explainable to patients.
3. Dialysis facilities generally have relatively few patients, which creates an added risk of uncertainty in assessments based only on a single indicator. This is not a flaw in the methodology, but a characteristic of the population. By incorporating all quality measures in a single indicator, we believe the overall composite star rating is a more robust assessment of facility quality.

Additionally, you suggested that DFC Star Ratings lack utility because they are counter-intuitive, they do not use the type of measures consumers expect, the cut-points do not convey information about the magnitude of differences or the degree of variation, and they use national rather than local reference points. We assessed facilities on their performance using a number of quality metrics. We believe the resultant rating is intuitive and easy to understand: a higher rating reflects better performance and a lower rating reflects worse performance. This presentation reflects an advance beyond the individual measure percentages (without adequate context) and odds ratios (hard for patients to interpret), which were previously the only quality data available on DFC and similar sites. Moreover, we consumer tested the rating system prior to the release of the star ratings, and it appeared sufficiently intuitive for consumers to use.

Your concern regarding the magnitude of differences and the degree of variation not being accurately reflected in the cut points we selected is a fair one, and worthy of consideration here. Since the final score was continuous, we exercised discretion in choosing cut points to categorize facilities into star rating categories. The cut points were chosen and resulted in 10% 1-star, 20% 2-stars, 40% 3-stars, 20% 4-stars, and 10% 5-stars for the following reasons:

1. The distribution of final scores was symmetric, which justifies assigning equal percentages to the two tails of the distribution (1 and 2 stars vs 4 and 5 stars).
2. These cut-offs distinguish facilities based on the final score and individual measure components. Specifically, the analysis showed that each star tier distinguishes itself from the other tiers. Each tier has mean final scores (Table 1) and individual measures (Table 2) that differ from other tiers, and these differences are highly statistically significant with p-values less than 0.0001 in all comparisons. Higher DFC Star Quality Ratings are able to indicate better quality of both clinically, and based on high statistical significance.
3. The rating methodology allows for similar differences in final scores (roughly 9 points) between adjacent star ratings (Table 1) and yields percentiles that should be easily interpretable to users (5-star facilities are in the top 10 percent, 4-star facilities are the next 20%, etc.).
4. Selecting a pre-set distribution is consistent with the quality measure methodology employed by the Nursing Home star ratings, and in our case, the distribution is reflective of the performance of facilities throughout the country.

Table 1. Average Measure Values Within Star Rating Categories

Star Rating 1	Mean Final Score (Range)	Star Rating 2	Mean Final Score (Range)	T-test P-value
★	32.4 (8.6, 37.5)	★★	41.7 (37.5, 44.92)	<.0001
★★	41.7 (37.5, 44.9)	★★★	50.3 (44.9, 55.4)	<.0001
★★★	50.3 (44.9, 55.4)	★★★★	58.4(55.4, 62.0)	<.0001
★★★★	58.4(55.4, 62.0)	★★★★★	66.7 (62.0, 84.4)	<.0001

Using a national basis of comparison allows patients to observe how a particular facility performs against the universe of available facilities. If many of the local facilities have low ratings, we believe that information should be made available to the dialysis patients and should be used as an opportunity for improvement. Using regional reference points may obscure that information, and a patient at a regionally referenced 3-star facility may never know that it would receive only a single star in comparison to all facilities. The ratings for many of the facilities in regions with a higher proportion of 1 and 2 star facilities represent a great opportunity for improvement so that patients can receive the best possible care and have accurate information to use in making informed decisions about their care. Ultimately, making this information public is the very purpose of the star ratings, and we are committed to ensuring that patients and families have accurate information about facilities in their region. That is not to say that understanding regional variation has no value; we may certainly entertain how it may factor into public reporting efforts in the future.

We do recognize that the current iteration of Star Ratings has room for improvement. CMS' practice in all of its public reporting sites has been to make changes based upon lessons learned

from implementation and stakeholder feedback. We readily agree with you that additional measures of patient safety and patient engagement, among others, are appropriate for consideration in future iterations of the star ratings. We are, however, limited by the measures currently available to us. A Star Rating TEP was held in April of 2015 and the patient and consumer advocacy workgroup did express strong interest in patient reported and patient centered outcome measures. We will continue to explore the possibilities of adding these types of measures to Dialysis compare and as additional measures and information become available, we may present them to a Star Rating TEP convened for its consideration. We believe the decisions that we made regarding the selection of measures and methodology are reasonable through analysis of the available evidence, but also recognize that different decisions can be made and remain willing to hear the suggestions presented to us by you and other stakeholders in the ESRD community. We believe that the Star Ratings TEP and our preview period will give the community ample opportunity to provide us with feedback in an ongoing fashion.

We want to thank you for including recommendations for how you believe we can improve the Star Ratings. We believe that many of the issues you raise are ones that we can consider and evaluate through the TEP process. Indeed, the focus of one of the TEP workgroups was to review and discuss public reporting approaches and patient/consumer needs and goals.

In summary, we believe that the Star Ratings constitute quality information collected, calculated, and reported in compliance with the HHS\CMS Guidelines for Ensuring the Quality of Information Disseminated to the Public. We sincerely appreciate your comments and hope the information provided helps clarify the status of our work. If you do not agree with our decision you may send a written hard copy or electronic request for reconsideration within 30 days of receipt of this letter. The appeal must state the reasons why you consider our decision or correction within the scope of the Act to be insufficient or inadequate. You must also attach a copy of your original request and our response to it, clearly mark the appeal with the words, "Information Quality Appeal," and send the appeal to the following address:

Information Quality
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244-1850

Alternatively, your appeal can be e-mailed to infoquality@cms.hhs.gov. Please note that appeals should contain the word "appeal" on the envelope or in the subject line.

Sincerely yours,

Mary Pratt
Division of Chronic and Post-Acute Care
Center for Clinical Standards and Quality
Centers for Medicare and Medicaid Services