



March 16, 2016

Kate Goodrich, M.D.  
Acting Director  
Center for Clinical Standards and Quality  
Centers for Medicare & Medicaid Services  
7500 Security Boulevard  
Baltimore, MD 21244

Dear Dr. Goodrich,

On behalf of Kidney Care Partners (KCP), I want to thank you and your team for providing the kidney care community with the opportunity to provide comments on the “Planned Changes to the DFC Star Rating Methodology” (Planned Changes Report). As you know, addressing methodology concerns with the ESRD Star Rating program is a top priority for the members of KCP. We especially appreciate that CMS has established a technical expert panel (TEP) to review the issues and has developed proposed modifications.

Specifically, we support the decision to use fixed year-to-year benchmarks for the scoring of the performance measures included in the Star Rating program. These benchmarks will allow dialysis facilities to demonstrate annual improvement in the quality of care they deliver to their patients, which will more accurately convey the commitment to quality of the industry. We also applaud the proposal to use a z-score methodology for many of the measures; z-scores will more accurately reflect the underlying performance distribution of facilities than the previous percentile scoring model did. Yet, as noted below, applying the proposed rebasing policy would eliminate the improvement that would result from using the z-score methodology.

Other aspects of the proposed changes also represent progress, but still raise concerns for KCP’s membership. KCP supports the decision to allow the distribution of Star Ratings to shift over time to show improvement, but has concerns about the continued use of the 10-20-40-20-10 distribution as a baseline. The rebasing process also raises several questions about the criteria that will trigger rebasing, the frequency of rebasing, and the re-basing methodology. We note that the “Planned Changes to the DFC Star Rating Methodology” document was ambiguous on these critical issues. If the ESRD Five Star is frequently rebased, then in effect the program will retain the forced distribution. If the re-basing significantly changes Star Ratings, even though underlying performance has not shifted, then consumers may be confused about what the changes mean.

The 10-20-40-20-10 predetermined distribution of DFC Star Ratings remains our chief concern with the program. KCP fully supports the position articulated so well during the ESRD Star Rating TEP that the current methodology is difficult to understand for patients and inconsistent with other online rating systems (including other CMS Star Rating Programs). Performance on measures should determine the number of stars a facility receives, not a pre-determined distribution that may not accurately reflect the actual distribution of quality results. The methodology should not create artificial distinctions among facilities. Distinctions should be identified through the selection of measures that matter to patients and accurately reflect the quality of dialysis care being provided by a facility, and every facility that performs well should have the opportunity to achieve a high Star Rating.

The standardized ratio measures remain another source of concern. As we have discussed previously, rates are preferable to the use of standardized ratios. Whether CMS ultimately adopts the rates or maintains the current standardized ratio measures, we believe that a Z-score method could also be used with standardized ratios. Using the Z-score method consistently across all measures in ESRD Five Star would make the program easier for patients and consumers to understand, as well as make it more internally consistent.

The Planned Changes Report includes some important steps forward, but more needs to be done before the next roll out of the ESRD Five Star ratings in the fall of 2017. To that end, we offer the following recommendations:

- For the upcoming star ratings (released in the Fall of 2016) eliminate the 10-20-40-20-10 distribution and assign stars based upon the star definitions of:
  - The facility's performance in every quality domain is better than average (5 stars)
  - The facility's overall performance is better than average (4 stars)
  - The facility's overall performance is close to average (3 stars)
  - The facility's overall performance is well below average (2 stars)
  - The facility's performance in each quality domain is well below average (1 star)
  - The facility has insufficient data in one of the measure domains (no stars).
- Use fixed year-to-year benchmarks for the scoring of the performance measures to allow dialysis facilities to demonstrate annual improvement in the quality of care they deliver to their patients.
- Do not force rebasing using the triggers outlined and shifting back to an artificial 10-20-40-20-10 distribution. Rather allow rebasing to occur organically as new measures are added and others are eliminated; shifts

in stars should be the result of changes in actual performance as determined by the measures.

- If, for example, a measure does not show a distinction in performance, it could be eliminated from ESRD Five Star, but remain as an individual measure publicly reported on Dialysis Facility Compare (DFC) assuming it remains an important piece of information for patients.
- Allow individual measure benchmarks to be rebased without rebasing the entire program.
- Use the z-score methodology for scoring all of the measures in ESRD Five Star.
  - Ideally CMS would use the rates rather than maintain the standardized ratio and apply the z-score methodology to the rates (which are calculated as part of the current standardized ratios).
  - If CMS cannot shift to the rates for the next round of ESRD Five Star, it can still use the z-score methodology as described below.

#### **I. Overall Star Distribution and Rebasing**

KCP is pleased that CMS has proposed a new methodology for calculating facilities' overall star ratings that allows facilities to demonstrate improvement over time. The new methodology will present patients and consumers with a more accurate representation of facilities' performance than the current methodology, which requires that facilities be assigned stars based on a rigid 10-20-40-20-10 distribution. We appreciate that CMS has recognized KCP's concern with the forced distribution of stars under the current methodology and made an effort to address this problem.

However, while the proposed methodology seeks to address the problem of forcing a normal distribution on the assignment of stars to facilities, the rebasing policy appears to result in little actual movement away from the rigid 10-20-40-20-10 distribution. This outcome is due to: (1) using 2014 and the current methodology that relies upon the forced distribution as the baseline year, and (2) using rebasing triggers that seem likely to result in the rebasing the star ratings every year, especially in the near term. These two aspects of the Planned Changes Report would result in maintaining the current methodology contrary to the intent expressed in the document to move away from it.

**A. Retaining the forced distribution as the baseline is inconsistent with the stated purpose of providing patients and consumers with an accurate view of how facility quality changes over time.**

Consistent with the views of from the patient members of the ESRD Star Rating TEP, KCP recommends that CMS move away from a pre-determined 10-20-40-20-10 distribution for the DFC Star Ratings. Instead of setting the cut points at the normalized bell curve percentiles, the Agency could set performance criteria defining each star level<sup>1</sup> and allow all qualifying facilities to achieve that level.

One option for assigning star ratings would be to use the following categorical approach.

**Table 1: Description of Recommended Categories Used Determining Overall Star Ratings**

<b>5 Stars</b>	The facility's performance in <u>every quality domain</u> is better than average.
<b>4 Stars</b>	The facility's <u>overall performance</u> is better than average.
<b>3 Stars</b>	The facility's overall performance is close to the average.
<b>2 Stars</b>	The facility's <u>overall performance</u> is well below average.
<b>1 Star</b>	The facility's performance in <u>every quality domain</u> is well below average.
<b>Not rated</b>	The facility has insufficient data in one of the measure domains.

Five stars would mean that a facility's actual performance is above average in every domain. Facilities with four stars would have above average performance, but not in every domain. Three stars would mean that a facility's overall performance is as expected. Facilities with two stars would have below average performance, but not in every domain. One star would mean that a facility's actual performance in every domain is worse than expected.

---

<sup>1</sup>These criteria at the star level should not be confused with setting absolute benchmarks at the individual measure level. We appreciate the concerns raised by the methodology TEP that setting absolute benchmarks at the individual measure level could be different. However, that problem does not exist when setting specific performance criteria for each star level.

We understand CMS's concern that the methodology should not result in the vast majority of facilities being 4 or 5 stars. The recommendations we are making would not lead to such an outcome either. On the other hand, our methodology would not force a set percentage of facilities into the lowest categories when in fact they are providing quality that is comparable to their peers. The point that patients, consumers, and KCP have continually stressed is that if too many facilities are in the top or bottom star rating categories, then it is the measures that may need to change. The methodology should not drive artificial distinctions that do not reflect actual quality. As Table 2 shows, the recommended performance categories would not change significantly, but would provide patients and consumers with performance information that has not be distorted by the methodology.

**Table 2: Star Ratings Determined Using Recommended Methodology  
(2014 DFC Data)**

Rating	Number of Facilities	Percent of Facilities
<b>5 Stars</b>	846	15%
<b>4 Stars</b>	1692	29%
<b>3 Stars</b>	2606	45%
<b>2 Stars</b>	540	9%
<b>1 Star</b>	56	1%

In our view, this overall distribution of results is preferable to the current 10-20-40-20-10 distribution. Facilities are able to show high performance, with the lower ratings reserved for facilities whose performance is significantly below par. This reflects a more useful profile of facility performance than a forced symmetrical distribution.

These definitions are internally consistent and represent meaningful performance differences for consumers. Moreover, these definitions could be maintained over time, even as benchmarks for underlying measures are periodically updated based on increasing performance. This approach provides patients and consumers with an easy to understand representation of facilities' quality performance, empowering their decision-making.

**B. If designed properly, the star ratings will organically adjust themselves over time, making a forced rebasing unnecessary.**

We agree that star ratings must evolve over time. If designed correctly, the methodology will allow an ongoing shift in star ratings to happen without artificially reinstating the forced 10-20-40-20-10 distribution to assign star ratings.<sup>2</sup> Rebasing

---

<sup>2</sup>Given the lack of a definition of rebasing in the Planned Changes Report, it appears that in a rebasing year the star ratings would be determined using the artificial 10-20-40-20-10 distribution. As noted

periodically to the forced 10-20-40-20-10 distribution creates its own problems, not the least of which is that it will result in distorting signals to consumers about the quality of facilities, since rebasing would change star ratings without any actual change in performance. It is possible a facility that maintains quality could still drop from a five to a three star rating without any actual change in performance. This would be inaccurate and confusing to patients and consumers and not serve them well.

The Planned Changes Report suggests that rebasing should occur when one of the following criteria are met:

- Measures are added or retired;
- TEP recommends the baselines should be re-evaluated
- When the Star Rating distribution “obscures differences between facilities”; “obscuring” would be determined using the following criteria:
  - Greater than 50 percent of facilities achieve 4 or 5 stars or greater than 50 percent of facilities achieve 1 or 2 stars;
  - Differences between 4 and 5 star facilities are not statistically significant for more than half of the individual measures; or
  - Differences between 1 and 2 star facilities are not statistically significant for more than half of the individual measures.

KCP is concerned that these criteria make it extremely likely that the star ratings would be rebased each year. Given the strong interest to continually adding measures to ESRD quality programs, the criterion of rebasing whenever measures are added or retired makes it likely rebasing would occur annually, especially in the near terms. Having a TEP recommend rebasing is concerning because it leaves the decision to the discretion to a small group of individuals without any other criteria to evaluate the decision. They too could decide to rebase every year and maintain the 10-20-40-20-10 distribution for assigning stars. Finally, the “obscuring” criteria are concerning because they assume that only a certain percentage of facilities should be allowed to achieve four or five stars and conversely that only a certain number of facilities should be allowed to be rated as one or two stars. This approach once again establishes an arbitrary cut off instead of allowing the actual performance of facilities to determine the star ratings.

Rebasing is a term of art used in economic programs to adjust for changes in input over time. Medicare traditionally rebases payment systems to address changes in inputs that have lead to the payment rates being inconsistent with the

---

earlier, this rebasing would essentially eliminate the ability for patients and consumers to see improvement over time and return to a methodology about which patients, consumers, and KCP have continually raised concerns.

costs incurred by providers to serve patients.<sup>3</sup> While we understand that CMS has “rebased” other star programs, we do not believe it is necessary or methodologically sound to use this economic concept in a quality program.

To avoid the situation where there is no distinction among facilities in terms of quality performance, we recommend that CMS allow star ratings to shift organically as new measures are added and topped out measures are retired. If measures are added or retired, the distribution of star ratings will naturally change. This process should be transparent and open to comment from all stakeholders. Measures that are driving higher scores could be determined to be topped out and removed from the rating program. When new measures are added, we would assume they meet the NQF criterion of Importance, meaning there is need to measure the area because there is a clinically relevant gap in performance. Using the criteria we suggest in this letter to establish the star rating cut points in Table 1 would allow for the stars to shift over time based on measures rather than the methodology. This method would avoid a complicated methodology that would mask the actual performance of the facilities. Most importantly, patients and consumers would understand the shifts because they could see the changes in the actual measures being used.

## **II. Use of Z-Score versus Probit Methodology**

KCP supports the proposal to use the truncated z-score methodology for the percentile measures. As we have noted in the past and the report indicates, a truncated Z-score allows for “greater precision in scores,” “eliminates the need to make a decision on when to use different scoring methods,” and “eliminates the possibility that an outlier on a single measure would completely determine the Star Rating.” We agree that the truncated Z-score is superior to the probit methodology.

We urge CMS to consistently use the z-scores for every star rating measure. There are two options for using a Z-score methodology with the current standardized ratio measures. First, CMS could move forward with shifting from the standardized ratios to rate measures for evaluating hospitalization, transfusions, and mortality. We understand and are pleased that CMS is interested in moving in this direction, and we request that CMS expedite this process.

KCP continues to support the use of rate measures because they allow patients and facilities to see year-over-year differences between normalized rates (deaths per 100 patient years) for mortality and hospitalization. Including the year-over-year rate difference at this time to allow patients, consumers, and the program to acknowledge improvement as well as attainment. These rates are currently

---

<sup>3</sup>See, e.g., Medicare Payment Advisory Commission, *Report to the Congress*, 182 (March 2015).

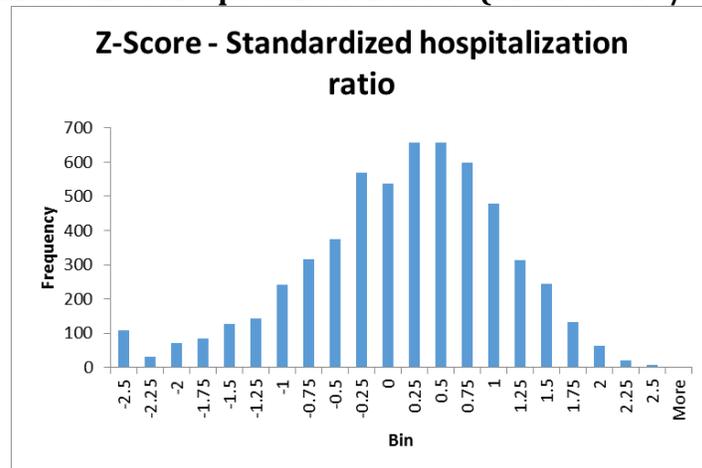
available from Dialysis Facility Reports data and should be used in DFC/ESRD Five Star. More recently, CMS's proposed changes to the SMR, SHR, and STrR indicated the measures could be calculated as risk standardized rates, and we have commented on those models.

If CMS were to shift to these rate measures, which already exist, it could easily use the truncated z-score methodology as well. This approach would not only create consistency and make DFC Five Star easier to understand, but it also would ensure that patients have more precise and accurate data on hospitalization and mortality, which they have repeatedly indicated are important measures for evaluating dialysis facilities.

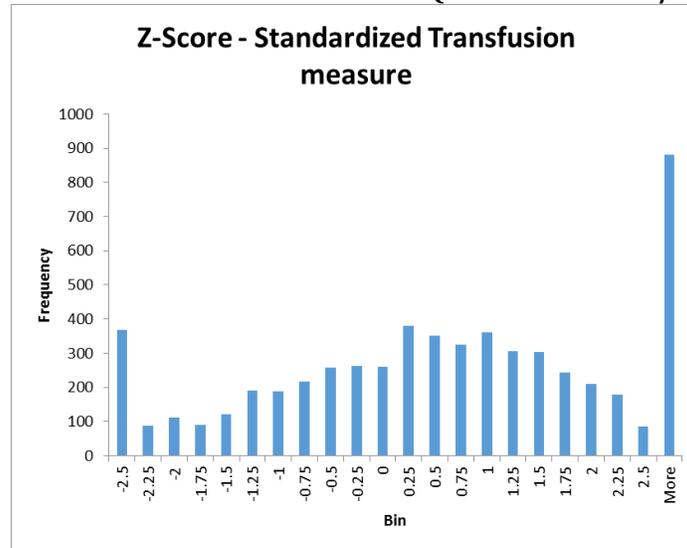
If for some reason, CMS is not able to immediately shift to rate measures, it could still use the z-score methodology for the standardized ratio measures. While we understand that some of the statisticians indicated during the TEP that a z-score could not be used for the standardized ratio measures, in reality the results for the DFC standardized ratio measures are amenable to the z-score model, because the actual distribution of those results is tightly clustered and symmetrical around the average.

To demonstrate how z-scores can work for the DFC standardized ratio measures, we applied z-scoring to the current DFC data for those measures. For each of the three standardized ratio measures we calculated z-scores with truncation at +/- 2.5 standard deviations. The results are illustrated below.

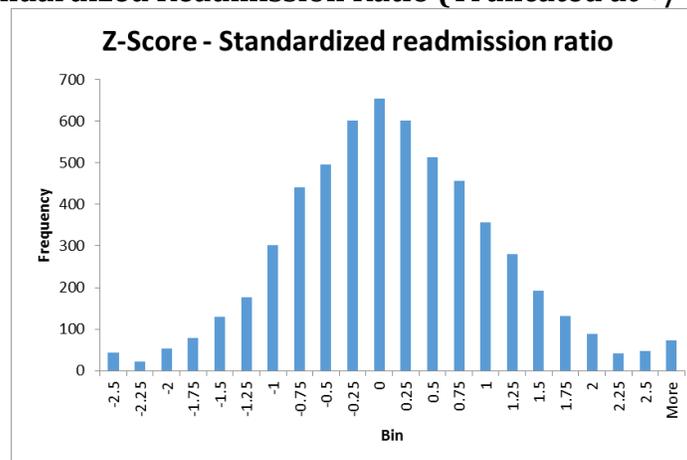
### Standardized Hospitalization Ratio (Truncated +/- 2.5)



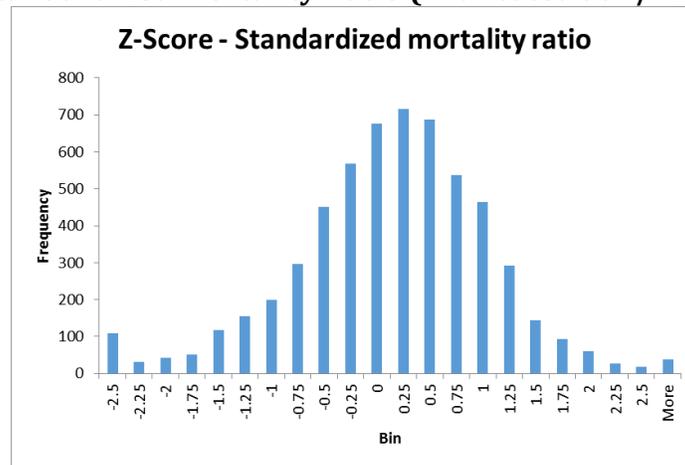
### Standardized Transfusion Ratio (Truncated at +/- 2.5)



### Standardized Readmission Ratio (Truncated at +/- 2.5)



### Standardized Mortality Ratio (Truncated at +/- 2.5)



As can be observed from the graphs, the distribution of performance for all three measures is generally symmetrical, and the vast majority of values fall within 2.5 standard deviations of either side of the distribution (which is very similar to the z-score results for the other measures). Using z-scores for all the measures will make the Five Star methodology more internally consistent, easier for stakeholders to understand, and more reflective of the quality of care provided.

### III. Transparency

KCP appreciates the opportunity to provide comments on the proposal, but remains concerned that the Planned Changes Report does not include all of the information necessary to sufficiently understand the proposal. While we appreciate that the Agency provided answers to the questions we raised after the release of the Planned Changes Report, we want to emphasize the importance of providing a complete proposal at the release date so that all stakeholders have a full understanding of the proposals and the entire comment period to analyze them. Having all of the information at the outset is particularly important given the extremely short comment period that has been provided to the kidney care community. For example, while we can provide our general comments on rebasing, not understanding if rebasing means the scoring returns to the 10-20-40-20-10 distribution, or to a distribution more precisely related to actual performance, makes it extremely difficult to assess the rebasing proposal.

Given the perennial nature of this problem, we recommend that when CMS releases a proposal in the future it provide an opportunity for the community to submit clarification questions. The Agency should provide answers to these questions within a week or two of the submission deadline and then provide a full 30-day comment period once it releases the answers. This would allow for a full

Dr. Kate Goodrich

March 16, 2016

Page 11 of 12

and fair review of proposals and establish a more collaborative approach to the comment period.

#### **IV. Conclusion**

KCP appreciates the efforts CMS has made to address the concerns raised by patients, consumers, and our members. We encourage you to adopt the additional modifications suggested in this letter to avoid the proposed modifications from becoming meaningless. We look forward to working with you on these changes to make ESRD Five Star a program that all patients, consumers, and the kidney care community can support and rely upon.

Sincerely,

A handwritten signature in black ink that reads "Frank Maddux, M.D." with a stylized flourish at the end.

Frank Maddux, M.D.

Chairman

Kidney Care Partners

**Appendix: KCP Members**

AbbVie  
Akebia Therapeutics, Inc  
American Kidney Fund  
American Nephrology Nurses' Association  
American Renal Associates, Inc.  
American Society of Nephrology  
American Society of Pediatric Nephrology  
Amgen  
AstraZeneca  
Baxter Gambro Renal  
Board of Nephrology Examiners and Technology  
Centers for Dialysis Care  
DaVita Healthcare Partners Inc.  
Dialysis Clinic, Inc.  
Dialysis Patient Citizens  
Fresenius Medical Care North America  
Fresenius Medicare Care Renal Therapies Group  
Greenfield Health Systems  
Keryx Biopharmaceuticals, Inc.  
Kidney Care Council  
National Kidney Foundation  
National Renal Administrators Association  
Nephrology Nursing Certification Commission  
Northwest Kidney Centers  
NxStage Medical, Inc.  
Renal Physicians Association  
Renal Support Network  
Rogosin Institute  
Sanofi  
Satellite Health Care  
U.S. Renal Care