Public Comment Summary Report

Project Title:

End Stage Renal Disease Dialysis Facility Compare (DFC) Star Ratings Technical Expert Panel (TEP)

Dates:

- The Call for Public Comment ran from February 16, 2016 to March 16, 2016.
- The Public Comment Summary was made available on May 20, 2016.

Project Overview:

The Centers for Medicare & Medicaid Services (CMS) has contracted with The University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) to review the methodology developed to produce the DFC Star Ratings. The contract name is the ESRD Quality Measure Development, Maintenance, and Support contract. The contract number is HHSM-500-2013-13017I. CMS has requested interested parties to submit comments on the Dialysis Facility Compare Star Rating methodology.

Project Objectives:

A TEP was convened to evaluate and make recommendations on the DFC Star Rating methodology and display. Specific objectives included:

1. Review of the statistical methodology behind the Star Rating calculations
2. Review of the measures used in the Star Ratings
    a. Consider measures for retirement
    b. Consider measures for future implementation
3. Review the readability and presentation of the Star Ratings on the DFC website

As a result of TEP deliberations several updates to the methodology have been proposed. CMS requested public comments on these methodological updates.

Additionally, in parallel CMS is reviewing and updating the content on DFC to increase readability and comprehension of quality data and the meaning of the Star Ratings.

Information About the Comments Received:

- Public comments were solicited by email.
- Six public comments were received on this topic.

Stakeholder Comments—General and Measure-Specific

General Comments

Several commenters supported the proposed changes informed by the Star Rating Technical Expert

Panel (TEP) deliberations. Several commenters stated the changes made to the measure scoring and setting a baseline year were improvements to the original DFC Star Rating methodology. One commenter recommended redefining Star Ratings based on setting performance criteria for each star level. For example, a facility would receive five stars when a "facility's performance in every quality domain is better than average."

***Response**: Thank you for your comments. The Star Rating category definitions have been released twice to public consumers of the DFC website. For the present time maintaining consistency and clear definitions of the current Star Rating categories will help consumers use the DFC Star Ratings. A limitation of setting performance criteria at the star (or domain) level is that information is lost. The loss of information and accuracy is not consistent with the recommendations that emerged during the TEP deliberations on the follow-up teleconference calls.*

Comments on Measure Scoring

Several commenters supported using z-scores on the intermediate outcome (percentage) measures. A suggestion was also made to use z-scores for the standardized ratio measures rather than probit scoring. Another commenter supported using probit scores on the standardized ratio measures.

***Response**: Analyses demonstrated that continuous scoring techniques (such as using probit scores and z-scores) are more accurate and have more power in distinguishing facilities than categorizing measure performance into groups (such as passes and fails). As the Star Rating TEP held accuracy as a high priority, we advocate for continuous scoring of the measures, including the probit scores and truncated z-scores.*

Among the percentage measures, we used truncated z-scores since they maintained the shape of the distribution of the measures, without allowing scores in the tails of a measure distribution to be too influential on the rating. Truncated z-scores are more appropriate because some of the percentage measures have highly skewed distributions.

Among the ratio measures, we used probit scoring, a rank based approach since the quality associated with a unit change in a ratio measure is not likely to be consistent along the range of the measure value. The planned changes document that was posted for public comment provides additional detail on the measure scoring.

Comments on the Baseline Year

Several commenters supported establishing a fixed baseline year as it allows facilities to show improvement over time, and achieve higher Star Ratings. There was also general support of moving away from the current fixed distribution

***Response**: Thank you for your comments.*

Comments on the Rebaselining

Several commenters wanted more detail on the criteria for rebasing. One commenter did not support

rebasing the Star Ratings and stated the concern that rebasing would result in a return to the fixed distribution of 10-20-40-20-10 for the DFC Star Ratings. Several commenters expressed concern that the Star Ratings would be rebased too frequently if it was rebased whenever measures were added or retired, while other comments suggested other or additional criterion, such as rebasing when measures are added or removed; using set time intervals; or only rebase individual measures as needed. One commenter recommended only rebasing the individual measures as needed and not the Star Ratings.

***Response**: As described in the planned methodological updates, draft technical criteria for rebasing were described. The planned rebasing criteria do not affect the upcoming October 2016 release of the Star Ratings.*

Re-baselining, or establishing a new baseline distribution for the DFC Star Ratings when the distribution of facility Star Ratings no longer optimally supports the site's goal of informing dialysis consumers, is critically important. One potential unintended consequence of re-baselining is that historical improvements and declines in facility outcomes are obscured by the re-baselining. Given the importance placed by many commenters on the importance of this historical or longitudinal trend, re-baselining should be performed judiciously and with this potential consequence in mind. Nonetheless, the experience with other CMS Compare sites demonstrates the importance of re-baselining to ensure that facility rating information remains both contemporary and relevant to consumers.

The comments received related to re-baselining highlight the complexity of the issue. CMS, as sponsor of the DFC Compare site, has responsibility for ensuring that the information contained in Star Ratings is current and relevant. In order to fulfill that responsibility, intermittent re-baselining will be necessary, if dialysis providers continue to improve care over time. When re-baselining occurs, some resetting of the facility Star Rating distribution is inevitable. That is the goal of re-baselining. Given that the proposed re-baselining criteria do not affect the October 2016 release of Star Ratings, CMS has additional time to consider the broad range of options included in comments, including 1) how many facilities should be included in each Star Rating category in the new baseline year (Star Rating distribution), and 2) when and how to re-baseline in order to optimize the Star Rating efficacy while recognizing that re-baseline will obscure some historical information about dialysis facility performance.

Comments on Other Recommendations

Public Commenters offered other suggestions such as (1) incorporating more patient-reported outcomes into the DFC Star Ratings, (2) aligning the DFC Star Ratings distribution with the distribution of other compare sites, and (3) including CKD patients (CKD moderate stage 3, and CKD severe stage 4) in the measures used for the DFC Star Ratings.

***Response**: We thank you for your comments. The quality measures used in the Star Ratings only apply to dialysis facilities and the care they are providing to patients. As this does not include care for CKD-ND patients it would not be appropriate to include these patients in the measure and Star Ratings calculations. Available patient-reported outcome measures for the dialysis population are limited, but we agree that pursuing their implementation should be an important focus for a public reporting program intended to provide transparent quality data that meaningfully supports patient decisions about care. We intend to turn our efforts in this direction in the future. We also agree that alignment with other Star Ratings programs at CMS is important, but we also believe that the distribution of performance should as closely reflect the performance of the facilities in our program. Our intention is that using a baseline approach will allow us to provide a point of comparison against which facility*

performance may be compared without constraining facility performance to a particular distribution.

Preliminary Recommendations

The developers recommend that the DFC Star Ratings move forward with the proposed updated DFC Star Rating methodology that was submitted for public comment. The updated methodology will be applied to the Star Ratings released in the October 2016 refresh on DFC.

A detailed finalized methodology document providing key components of the methodology will be provided on June 1, 2016 (prior to the DFC Star Rating Preview Period).

Overall Analysis of the Comments and Recommendations

CMS and UM-KECC appreciate the time dedicated to reviewing and providing comments on the DFC Star Ratings.

Public Comment Verbatim Report

Date Posted	Measure Set or Measure	Text of Comments	Name, Credentials, and Organization of Commenter	Type of Organization	Recommendations/Actions Taken
May 20, 2016	Star Ratings	See appendix	Allen R. Nissenson, MD, FACP, FASN, FNKF, Chief Medical Officer, DaVita Healthcare Partners Inc.	Provider Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.
May 20, 2016	Star Ratings	See appendix	Jackson Williams, Director of Government Affairs, Dialysis Patient Citizens (DPC)	Patient Advocacy Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.
May 20, 2016	Star Ratings	See appendix	Frank Maddux, M.D., Chairman, Kidney Care Partners (KCP)	Patient Advocacy Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.
May 20, 2016	Star Ratings	See appendix	John Wagner, MD, MBA, President, National Forum of ESRD Networks, Inc.	ESRD Network	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.

May 20, 2016	Star Ratings	See appendix	Kerry Willis, PhD, Chief Scientific Officer, National Kidney Foundation (NKF)	Patient Advocacy Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.
May 20, 2016	Star Ratings	See appendix	Joyce F. Jackson, President & Chief Executive Officer, Northwest Kidney Centers	Provider Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.

DaVita OFFICE OF THE
Chief Medical Officer

601 Hawaii Street, El Segundo, CA 90245 : : 800.313.4872 : : davita.com/physicians

March 15, 2016

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

Dear Kate,

On behalf of DaVita HealthCare Partners, the 176,000 patients with end-stage renal disease (ESRD) that we serve, and our 65,000 teammates dedicated to their care, we are pleased to respond to the Planned Changes to the DFC Star Rating Methodology. We share a common goal with the Centers for Medicare and Medicaid Services (CMS): to strengthen the ESRD quality programs and payment system so that beneficiaries with kidney disease have access to information regarding the highest quality care.

We share a common goal with the Centers for Medicare and Medicaid Services (CMS): to strengthen the ESRD quality program and payment system so that beneficiaries with kidney disease have access to information regarding the highest quality care.

Overview

We applaud CMS for listening to feedback from the dialysis community and exploring the use of defined thresholds to assign star ratings versus the current method based on the relative distribution. However, the proposal provided to us regarding the changes to the DFC Star Rating Methodology lacks significant detail needed to fully evaluate the changes and spawns several questions. We fear this lack of detail and clarity will result in missing the objective of observing improvement in clinical quality over time. As a result, we respectfully request a consolidated document detailing the DFC Star Rating Methodology and an extension of the comment deadline to 21 days after such documentation is made available.

Baselines for standardized ratio measures

In general we agree with the proposal to adjust standardized ratio performance based on improvement or deterioration using rate data, but as mentioned in prior commentary we oppose the use of standardized ratios. Additionally, we express concern the data used to adjust the standardized ratio measures as referenced in the "Planned Changes…" document is currently not available publicly and therefore does not allow CMS' calculation to be replicated. The example provided in the documentation uses transfusion transfusions per patient year for 2013 and 2014 to adjust the Standardized Transfusion Ratio. We request micro-specifications for the

transfusions per patient year measure along with any other event rate data used to adjust the Standardized Ratio Measures, along with the specific numerator and denominator data for each facility used to calculate the event rates.

Summary of DFC Star Rating Modifications – Update of Scores Relative to Baseline Year

We reviewed the proposal to define thresholds using a baseline year, starting with performance from 2014. While we agree in principal with leveraging a prior year to define thresholds, which is relatively consistent with the Quality Incentive Program (QIP) methodology, we are concerned when and how measures or Five-Star cutoffs will be rebased. Under the current proposal a facility could theoretically be rated a '5', but then fall to a '3' star rating (or vice versa) due solely to rebasing. We request further detail on the rebasing methodology and to implement guard rails prohibiting a facility from a significant change in star rating solely due to rebasing.

Revision of Individual Measure Scoring

We concur binary scoring to define thresholds is not an appropriate method and agree with the use of z-scores. Unfortunately the proposal lacks requisite details on the next steps to calculate star ratings once the measure scoring is complete. The proposal does not provide detail on the following which are required for us to meaningfully comment:

1) how thresholds for each measure are defined;
2) how performance relative to an individual measure threshold will translate into a facility overall star rating;
3) use of baseline data in the calculation of performance year z-scores;
4) how the truncation/winzorization point is determined after calculated z-scores are calculated

We request detailed specifications to address these concerns, most importantly how z-scores will translate into a star rating.

Additional comments

Upon review of the "Planned Changes…" document it is unclear whether thresholds will be set at the measure level, at the overall score level, or both. It is also unclear whether domains will be used going forward star rating calculation. We also request clarity on which specific measures will be included in the Five-Star program and when, if any, new measures will be adopted into the Star rating calculation. The process for adding and removing measures in QIP and Dialysis Facility Compare is clearly defined. However, for the star rating program the process is very ambiguous. We ask CMS implement a similar process to review and comment on any measure proposals for the star rating calculation.

In summary, the following list represents our current open questions:

- How will measure scores translate into actual facility star ratings?
- How specifically will measure thresholds be defined?
- Is the threshold specific to the measure, the overall score, or both?
- Will domains continue to be used or will some other weighting apply?
- What constitutes "statistically significant" when rebasing measures?

DaVita. OFFICE OF THE
Chief Medical Officer
601 Hawaii Street, El Segundo, CA 90245 : : 800.313.4872 : : davita.com/physicians

- How does CMS intend to handle when a facility is not eligible for a measure (i.e., missing values)?
- What data is used to calculate mortality, hospitalization and transfusion rates to adjust year over year changes (i.e., transfusions per patient year as provided as an example in the proposal)?
- What measures will be added or removed as part of the Star Rating? If measures are added or removed, what measures and when will changes be implemented?
- When will the proposed change to the Five-Star program go into effect?

Recommendations

We respectfully recommend and request the following:

1) Provide detailed, comprehensive micro-specifications on the entire DFC Star Rating Methodology that addresses the questions and concerns provided in this letter;
2) Extend the timeline for comments to 21 days after a comprehensive methodology is made available;
3) Maintain the use of domains to prevent any one measure from being weighted to heavily;

Conclusion

We thank CMS for the opportunity to comment on this matter. Once again we commend the Agency for its increased transparency and willingness to work with us and the entire ESRD community. This mutual collaboration ensures that ESRD beneficiaries receive the best possible care, and that these patients are presented with sufficient publicly-reported data to allow for meaningful conversations with their caregivers in facilities and their physicians to assess the quality of care they currently receive or wish to receive. Programs such as the DFC Star Rating are clear examples how to represent clinical quality, but can only be strengthened by attention to the concerns we have raised and the policy improvements we have proposed.

Sincerely,

Allen R. Nissenson, MD, FACP, FASN, FNKF
Chief Medical Officer, DaVita Healthcare Partners Inc.
David Geffen School of Medicine at UCLA

FORUM OF END STAGE RENAL DISEASE NETWORKS

March 15, 2016

Re: ESRD DFC Star Rating TEP - Call for Public Comment Methodology Recommendations

The Kidney Patient Advisory Council (KPAC) of the Forum of ESRD Networks is pleased to comment on the planned changes to the DFC Star Rating Methodology. The KPAC is composed of patient representatives from each of the 18 Networks. Our focus is to improve patient's quality of care and quality of life by supporting the ESRD Networks System and the ESRD patient communities throughout the United States and our Territories.

The KPAC applauds CMS for ensuring patient involvement and the patient voice have a large role in the End Stage Renal Disease Dialysis Facility Compare (DFC) Star Ratings Technical Expert Panel (TEP). We certainly recognize and appreciate the focus and changes made to make the Dialysis Facility Compare web-site more patient and family centered. Although this is a highly technical recommendation, the KPAC offers the following comments:

1. In Table 3, we see how using the new methodology will improve patients' ability to compare facilities by better representing relative changes in facility quality over time.
2. In the case that Star distributions shift towards extremes, we appreciate the proposed suggestion to "rebase" thresholds as per the listed rebasing criteria. After all, these metrics represent quality of care that will ideally improve with time.
3. We recognize the benefits in precision and mathematical convenience of using truncated z-scores to represent the highly skewed (hypercalcemia and Kt/V) metrics. However, it is not obvious from the 2013 DFC data in Table 2 whether or not these changes result in appreciable differences in Star Distributions.

It is our hope that DFC 5 Star will be a living, ever-changing, tool for patients. We urge that future Measure Development Activities continue to focus on quality outcomes that matter most to patients, and the KPAC welcomes the opportunity to partner with CMS in this effort.

PO Box 203 • Birchwood, WI 54817 • (715) 354-3735 • Fax: 1 (888) 571-2065
email: forumcoord@centurytel.net • http://www.esrdnetworks.org

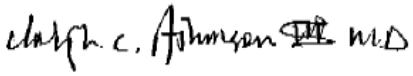
The KPAC feels the DFC 5 Star can add much value to ESRD patients making a choice between facilities and seeing the various modalities available in their area. We also feel that the DFC 5 Star should be expanded and shared with CKD patients in the CKD moderate stage 3 to look at available modalities choices and CKD severe stage 4 to choose a facility to be better prepared for ESRD and treatment.

Sincerely,

Derek Forfang
Vice-Chair, Kidney Patient Advisory Council, National Forum of ESRD Networks, Inc.

Maggie Carey

Maggie Carey
Chair, Kidney Patient Advisory Council, National Forum of ESRD Networks, Inc.



Ralph Atkinson III, MD
Chair, Medical Advisory Council, National Forum of ESRD Networks, Inc.

Donald Molony, MD
Past-Chair, Medical Advisory Council, National Forum of ESRD Networks, Inc.

John Wagner, MD, MBA
President, National Forum of ESRD Networks, Inc.

National Kidney Foundation™

30 E. 33rd Street
New York, NY 10016

Tel 212.889.2210
Fax 212.689.9261
www.kidney.org

March 14, 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,

The National Kidney Foundation appreciates the changes CMS has proposed to the methodology for the Dialysis Five Star program. The proposed methodology facilities large scale quality improvement among dialysis facilities over the existing methodology, which limits the number of dialysis facilities that can achieve above average ratings. This change aligns with NKF's past recommendations and those of the Consumer Technical Expert Panel (TEP) convened by CMS in April of last year. NKF also appreciates the substantial consideration CMS gave to patient input on the TEP and the additional time and meetings to discuss how the patients' recommendations could be included in the methodology used to assign star ratings.

However, we are concerned with the proposal that broadly allows for rebasing anytime more than 50% of facilities achieve 4 to 5 stars. NKF believes facilities performance on the measures should drive the star ratings and that the methodology should reflect that performance. In the TEP report, the patient workgroup reported being less interested in using the star ratings to identify small, insignificant differences between dialysis facilities and instead preferred to see the star ratings reflect actual performance on quality care. Those recommendations were in line with NKF's August 2014 survey of dialysis patients' views of quality care, where 66% of patients (out of 860 respondents) stated that facilities should be rated on their individual performance rather than compared to other facilities. This was the impetus for patient advocates to call for a change in the methodology away from assigning

National Kidney Foundation
30 E. 33rd Street
New York, NY 10016

Tel 212.889.2210
Fax 212.689.9261
www.kidney.org

stars based on a bell curve distribution. Rebasing as a result of high performance without a change in the measures could artificially facilitate a drop in facility performance, which may mislead patients into thinking their facility's quality of care dropped when performance on the measure is unchanged. This would not achieve the stated goals of the program.

Rebasing as new measures are added and others are retired is expected. However, rebasing back to a methodology that requires facilities to fit into different percentile buckets would not be an appropriate approach. Keeping the methodology consistent to ensure that measures are driving the assignment of stars is critical to ensuring beneficiaries can rely on the stars to reflect on how facilities are performing on the quality measures.

In addition, NKF encourages CMS to incorporate measures that are more meaningful to patients into the star ratings. The Consumer TEP had many suggestions for new areas of performance measurement. In addition NKF's quality survey found that patients' highest priority for determining value in care was the attentiveness of the facility staff. The Consumer TEP also encouraged the agency to develop an updated, interactive website that would allow patients to pick the quality measures and facility attributions that matter most to them and see star ratings in relation to those selections.

NKF hopes future iterations of the star ratings program will be more responsive to patients' interests. In the interim it is vital that patients be able to rely on the ratings as reflective of performance on the measures and for that reason we encourage CMS to only undergo rebasing when measures are removed, added or changed.

NKF has been engaging kidney patients in education, science, research, and advocacy for over 60 years. We share the agency's goals to empower patients to make informed decisions about their care and we would like the opportunity to work closely with you on further improving and testing the Five Star program with patients to ensure its success as a useful tool in patient informed decision making.

Sincerely,

Kerry Willis

Kerry Willis, PhD
Chief Scientific Officer

TO: Joel Andress, Phd
Centers for Medicare and Medicaid Services
University of Michigan Epidemiology and Cost Center
dialysisdata@umich.edu

February 22, 2016

The Northwest Kidney Centers is a non-profit, community-based organization with a mission to promote the optimal health, quality of life and independence of people with kidney disease through patient care, education and research. We would like to thank CMS for the opportunity to provide comments on the planned changes to the DFC Star Rating Methodology.

Overall, we support the planned changes to the DFC Star rating methodology. Our comments are organized in the same order as the publicly-provided document that explains the proposed changes, available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Planned-Changes-to-DFC-Star-Rating-Methodology.pdf. The sections are:

1. Baseline Year for Scoring Measures and Rating Facilities
2. Revision of Individual Measure Scoring
3. Summary of Recommendations

1. Baseline Year for Scoring Measures and Rating Facilities
NKC strongly supports use of a fixed baseline year.

- NKC would like to commend the Technical Expert Panel (TEP) for emphasizing the importance of tracking a facility's improvement or decline over time, and suggesting the use of a fixed baseline year for comparison.

2. Revision of Individual Measure Scoring
NKC offers comments on the planned revisions.

- NKC agrees that using a binary scoring methodology would not be appropriate, given the limitations discussed. Namely, the insufficient empirical evidence to define single thresholds for skewed measures and the arbitrary nature of setting cut-points.
- NKC supports the use of the z-scoring methodology for scoring percentage measures. In order to provide complete transparency of the number of facilities removed, we would encourage CMS to provide information about how many

facilities are excluded from these calculations due to the truncated methodology described in this document.

- NKC supports the use of the probit distribution for standardized measures. We would recommend that CMS provide information on how the number of percentiles (199) was determined, or if it was arbitrary.

3. Summary of Recommendations

NKC would like to thank CMS for providing Tables 1 – 3.

- These tables demonstrate the impact of the planned changes. In particular, the impact Star Ratings for facilities in each of the five categories.

NKC again thanks you for the opportunity to comment on this important work.

Sincerely,

Joyce F. Jackson

President & Chief Executive Officer

Administration

Northwest Kidney Centers

700 Broadway Seattle WA 98122

Tel:206 720 8500| Fax: 206 860 5821

Joyce.Jackson@nwkidney.org | www.nwkidney.org

KIDNEY CARE PARTNERS

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 previou 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. Th 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 progra 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 <u>every quality domain</u> is better than average (5 stars)
    - The facility's <u>overall performance</u> is better than average (4 stars)
    - The facility's overall performance is close to average (3 stars)
    - The facility's <u>overall performance</u> is well below average (2 stars)
    - The facility's performance in <u>each quality domain</u> 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[1] 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

5 Stars	The facility's performance in every quality domain is better than average.
4 Stars	The facility's overall performance is better than average.
3 Stars	The facility's overall performance is close to the average.
2 Stars	The facility's overall performance is well below average.
1 Star	The facility's performance in every quality domain is well below average.
Not rated	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.

[1]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
5 Stars	846	15%
4 Stars	1692	29%
3 Stars	2606	45%
2 Stars	540	9%
1 Star	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.[2] Rebasing

[2]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.[3] 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

[3] *See, e.g.,* Medicare Payment Advisory Commission, *Report to the Congress,* 182 (March 2015).

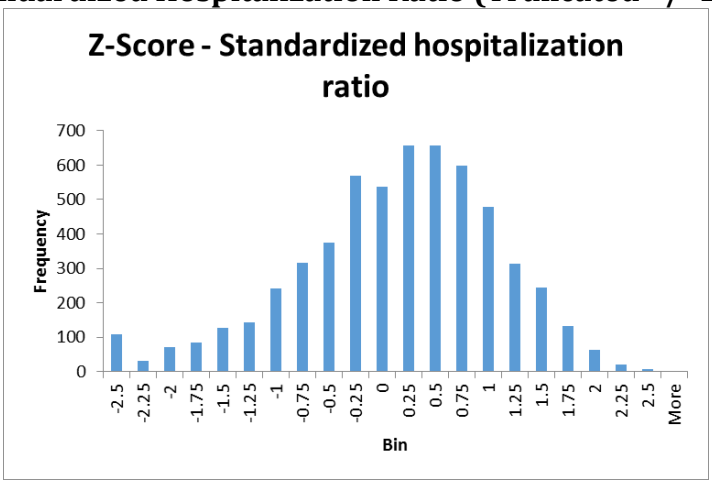
available from Dialysis Facility Reports data and should be used in DFC/ESRD Five Star. More recently, CMS's proposed changes to the SMR, SHR, and STrR indicated the measures could be calculated as risk standardized rates, and we have commented on those models.

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.

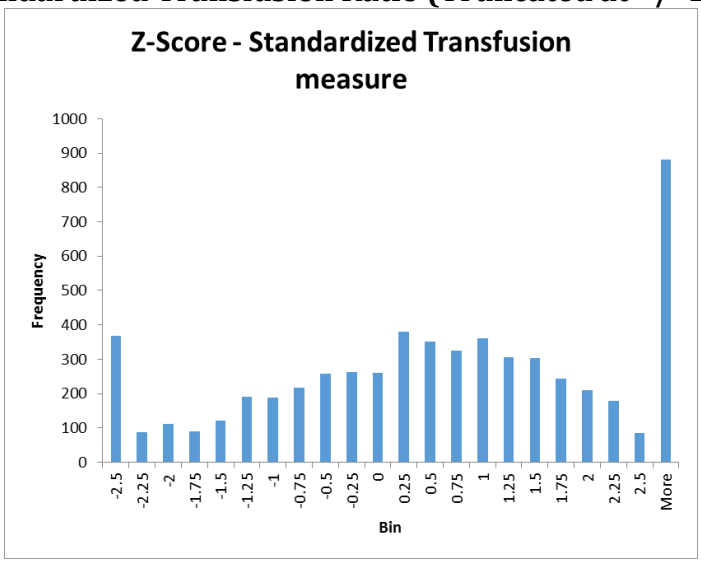
If for some reason, CMS is not able to immediate shift to rate measures, it could still use the z-score methodology for the standardized ratio measures. While we understand that some of the statisticians indicated during the TEP that a z-score could not be used for the standardized ratio measures, in 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.

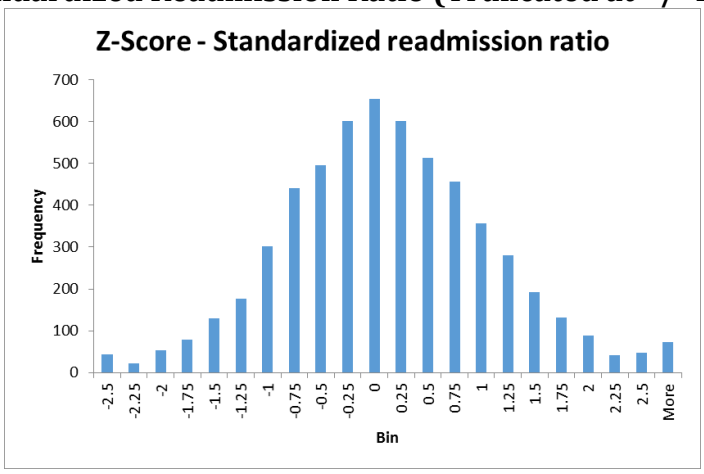
Standardized Hospitalization Ratio (Truncated +/- 2.5)



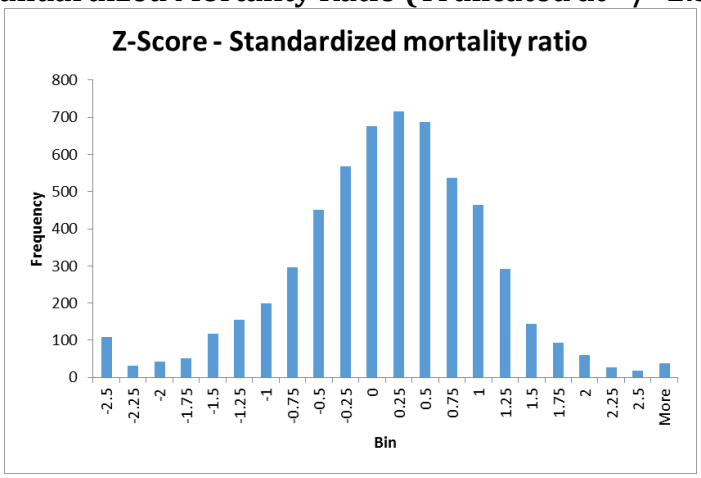
Standardized Transfusion Ratio (Truncated at +/- 2.5)



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). Usin 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

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,

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

DPC
DIALYSIS PATIENT CITIZENS

Improving Life Through Empowerment

March 16, 2016

Kate Goodrich, M.D.
Chief Medical Officer
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244

Re: Planned Changes to the DFC Star Rating Methodology

Dear Dr. Goodrich:

We are pleased to see that, after much back and forth, progress is being made in incorporating stakeholder concerns into the design of the DFC star ratings program. The immediate outcome proposed in the Planned Changes—using the bell-curve thresholds for 2014 as a fixed baseline and reassigning star ratings to encompass improved performance since that time—is acceptable to us as a next step. However, given the need to periodically re-set the baseline, we are concerned that the Planned Changes document is silent on how standards will be set upon re-basing. Many in the kidney community have interpreted the silence as an indication that DFC would return to the stack-ranking method when re-setting the cut-offs. We are perhaps more optimistic that accord can be reached on this question, but the because the document refers to possibly re-basing upon the addition of new measures, and potential new measures will be ready for inclusion in the very near future, we must regrettably temper our approbation with a sense of urgency that a new system be ready to supplant the bell curve should that option be chosen.

Maintaining a degree of stability in the star ratings must be a key consideration. There were changes in star rating awards to 2500 dialysis facilities in the latest iteration of DFC. Per research by Mark Stephens reported in NNI, it appears that many facilities whose scores were near the previous version's cutpoints were repositioned. We have two concerns about such changes. First, year-to-year volatility of the ratings—with 45 percent of facilities changing places—may undermine consumer confidence in their authoritativeness. Second, for patients in the 1250 facilities whose ratings were downgraded, the change may be alarming, conveying a sense that their facility is in some way troubled. We believe downgrading of facilities should be limited to those that actually experienced an absolute drop in their quality performance that reflects poorer quality and we interpret the Planned Changes document to mean that CMS agrees with us on this principle.

1012 14th St. NW, Suite 905 • Washington, D.C. 20005 • Toll Free Number 1.866.877.4242 • Fax 1.888.423.5002
www.dialysispatients.org • Email: dpc@dialysispatients.org

As the Planned Changes document recognizes, another important consideration is avoiding a lopsided distribution, as where more than half of facilities are rated as "above average." We agree that new baselines must be set in such circumstances in order for the star ratings to retain legitimacy.

CMS must appropriately balance these competing considerations. In the recent case of Nursing Home Compare, star ratings were re-set after about five years. This is not, perhaps, the best case to cite, given that the rebasing occurred after advocates and journalists called into question the appropriateness and usefulness of the prior system. DFC is different in that we don't believe the gaming that nursing homes engaged in is possible in DFC. But a four-to-five year run before rebasing would seem to strike the proper balance between consumer preferences for stability and the need to periodically raise the bar on quality expectations. HHS' *Healthy People 2020* initiative sets 2020 as the target year for achieving certain health outcomes so it seems fitting to us that 2020 be the target year for re-setting star rating cut-offs.

We have expressed a number of concerns about the Dialysis Facility Compare star ratings program, but at present our policy priority is aligning its rating criteria with those used in other CMS programs. We are able to support the next step described in the Planned Changes document because it reallocates DFC's star awards so that its distribution will resemble those on the other CMS sites that beneficiaries and caregivers are likely to visit. In the new distribution, fewer patients will be alarmed by one- and two-star ratings that are identified in the public's mind with substandard quality

The table below compares the current distributions of stars across CMS' consumer-facing sites. On Dialysis Facility Compare, thirty percent of facilities are awarded one or two stars; by contrast, fewer than one percent of Medicare Advantage plans and twelve percent of home health agencies receive such ratings. If one considers three stars to represent "average," then 49 percent of nursing homes and home health agencies, and 60 percent of MA plans are portrayed as "above average," while only thirty percent of dialysis facilities are so portrayed.

Star Rating	Percentage in Rating Range				
	Dialysis Facility	**Nursing Facility**	**Home Health**	**Medicare Advantage**	**Hospital**
1	10	13	.06	0	3
1 ½			1.95	0	
2	20	*	9	1	17
2 ½			17	4	
3	30	*	23	12	40
3 ½			23	27	
4	20	49*	15	18	34
4 ½			8	13	
5	10		3	2	7

* denotes detailed breakdown could not be found.

We realize that this can seem somewhat abstract, so we have prepared two additional charts that graphically depict the star ratings that a consumer is likely to see while searching for providers and plans in two localities. Imagine an ESRD patient's caregiver, who may be a spouse over 65 or a son or daughter, who has also searched for or assisted a beneficiary in search of a Medicare Advantage plan. We expect that, owing to the complex needs of ESRD patients, this caregiver may also have searched for hospital and post-acute care provider ratings prior to or contemporaneous with a search for dialysis facility ratings.

The first graph depicts the distribution of star ratings a consumer will find in Dayton Ohio. Because of the greater weight given to health outcomes on DFC relative to other CMS rating systems, Ohio has disproportionate number of 1- and 2-star dialysis facilities; but this is not the case with other rated entities. The graph depicts the ratings of the first 12 entities that are listed on five CMS sites. (Home health ratings on our graphs are for patient experience—Home Health Compare issues separate star ratings for clinical quality and patient experience.)

Of the first 60 entities listed, 12 are below 3 stars and 29 are above 3 stars. Of the 12 entities with 2 stars or below, eight are dialysis facilities. Of the 29 entities with 3 ½ stars or above, none are dialysis facilities. To the beneficiary who has been exposed to the other CMS star systems (or for that matter, systems like safercars.gov or Yelp) the sight of solely 2- and 3-star facilities nearby will be jarring. While 4- and 5-star hospitals, post-acute providers and health plans are readily accessible in Dayton, the nearest 4-star dialysis facility is 21 miles north in Troy, Ohio.

The second graph depicts the distribution of star ratings in Springfield, Missouri. Because fewer than twelve rated home health agencies and MA plans are available in Springfield, this graph shows ten entities in each provider category and nine in the health plan category.

Missouri is a state in which dialysis facilities are symmetrically distributed, and this pattern is seen in Springfield. However, the distributions of other entities skew to the high end as in Dayton: overall, 5 entities are below 3 stars while 28 are above. The effect of the distribution is to make 3-star entities look mediocre or laggard. This impression would be accurate with regard to most categories, since those star systems assign ratings above 3 stars to 41, 49, 49, and 60 percent of rated entities. But the impression is inaccurate with regard to dialysis facilities since 3 stars purports to mean "average" and only 30 percent of facilities can earn ratings above that.

To be clear, we expect such disparities in star rating systems to be a problem only in two thirds of the country. In the northwest, mountain west, upper Midwest and New England, the superior health of local populations skew DFC star ratings upwards so as to roughly match the skewing of the other rating systems.

We believe CMS' other star rating systems appropriately convey a note of caution in identifying a small number of truly poor performers with one or two stars. They follow the conventions of movie reviews, Yelp, airline ratings, and other star systems familiar to the average consumer. It remains unclear to us why DFC was given a unique design nor why CMS would see a benefit to retaining its unusual design since dialysis patients and their caregivers are arguably the demographic group most likely to have reason to search multiple "Compare" websites and therefore experience confusion. We are therefore relieved that the Planned Changes document indicates that CMS will, at least for the immediate future, step away from this system.

We want to emphasize that we do not dispute the normative considerations behind DFC's bell-curve distribution. We agree that ranking providers among their peers harnesses positive competitive dynamics and can provoke self-critical analysis among practitioners who all too typically believe themselves to be "above average." We believe, and we suspect that the DFC team agrees, that awarding 4- and 5-star ratings to 40 percent of facilities, and a greater percentage in the future, is a less-than-ideal outcome.

Our concern has always been the disconnect between these idealized notions of rankings and the public's experience with and understanding of star ratings. Since we raised that initial objection 18 months ago, CMS has rolled out new star ratings or modified old star ratings that adhere to the common-understanding paradigm rather than to the idealized paradigm manifested in DFC's first iterations. As such, awarding four or five stars to 40 percent or more of dialysis facilities can be justified for purposes of harmonizing the various CMS programs, as well as conveying to consumers a sense, which we believe is accurate, that overall care is improving.

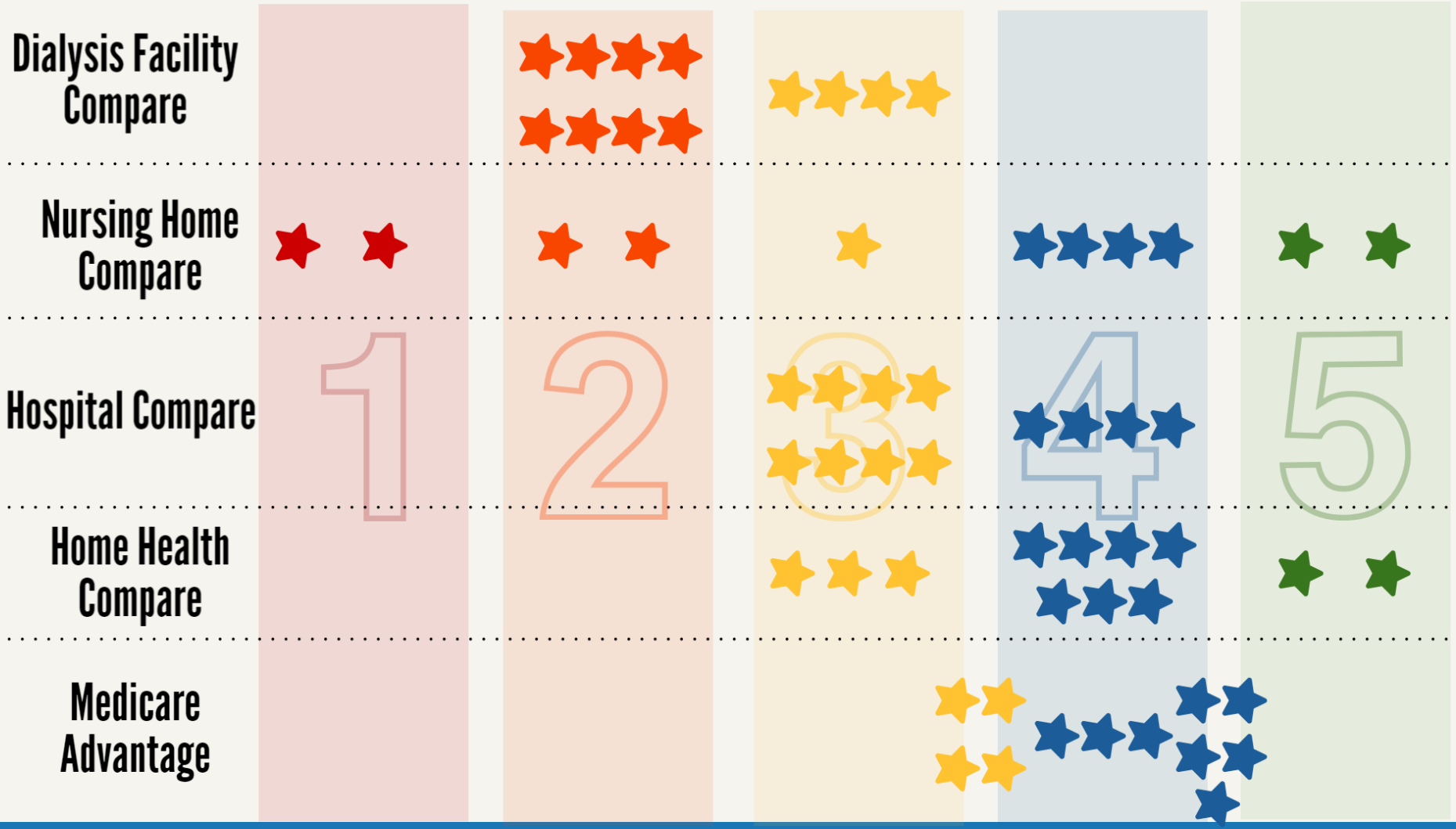
Thank you for your consideration of our comments. As you'll recall, we had requested a meeting with you or Dr. Conway to discuss this matter in detail. We no longer feel an immediate need to hold such a discussion. However, if there are plans to revert to the bell-curve distribution in the near future we believe it would be helpful to gather and confer with the CMS personnel overseeing the various star rating systems to go over the disparities illustrated on the attached charts and seek a more uniform experience for consumers who visit CMS's website.

Respectfully submitted,

Jackson Williams
Director of Government Affairs

DISTRIBUTION OF STAR RATINGS ON CMS "COMPARE" AND "PLAN FINDER" WEB PAGES

DAYTON, OH



DISTRIBUTION OF STAR RATINGS ON CMS "COMPARE" AND "PLAN FINDER" WEB PAGES

SPRINGFIELD,MO