

End-Stage Renal Disease Dialysis Facility Compare Star Rating Technical Expert Panel

Summary Report

October 4, 2019

Prepared By:

The Kidney Epidemiology and Cost Center
University of Michigan, School of Public Health
1415 Washington Heights, Suite 3645 SPH I
Ann Arbor, MI 48109-2029

CMS Contract Number:

75FCMC18D0041

Task Order Number:

75FCMC18F0001

Table of Contents

List of Figures	3
List of Tables.....	3
1. Introduction	4
Technical Expert Panel Objectives.....	4
Technical Expert Panel Meeting	4
TEP Membership	4
2. Background	5
3. Summary of the Technical Presentation	7
Trends in DFC Star Rating Distribution	8
Goals and Options for Resetting the DFC Star Rating.....	9
Goals and Options for Reweighting the DFC Star Rating Measures or Domains	11
Follow-Up to Questions from the Teleconference Calls	13
4. TEP Discussion.....	14
What techniques would help users interpret the DFC Star Rating after a reset?.....	14
Is a reset of the DFC Star Rating needed now?	16
What distribution should the DFC Star Rating be reset to?.....	17
Discussion of Re-Weighting the DFC Star Rating Measure Domains	18
TEP Vote on Resetting the DFC Star Rating	19
5. TEP Summary	20
Recommendations for Resetting the DFC Star Rating	20
Additional Themes from TEP Member Comments.....	20
Outstanding Issue.....	21
Additional TEP Input on Topics Discussed Beyond the Scope of the TEP Charter	21
6. Public Comment Period	25
7. Appendices.....	26
Appendix A. List of TEP Members and Contractor Staff	27
Appendix B. Technical Expert Panel Charter	31
Appendix C. DFC Star Rating TEP Teleconference Call #1 Minutes.....	35
Appendix D. DFC Star Rating TEP Teleconference Call #2 Minutes	41

Appendix E. DFC Star Rating Post-TEP Teleconference Call Minutes.....	47
Appendix F. DFC Star Rating TEP Supporting Materials List.....	52
Appendix G. TEP Composition Form.....	53
Appendix H. Public Comment Letter from Kidney Care Partners (KCP).....	60
Appendix I. Teleconference Call 1 Presentation.....	75
Appendix J. Teleconference Call 2 Presentation.....	116
Appendix K. In-Person Meeting Presentation.....	155
Appendix L. Post-TEP Teleconference Call Presentation	217

List of Figures

Figure 1: Trends in the DFC Star Rating Distribution, October 2015 – October 2018	8
Figure 2: Distribution of Facility Final Scores for the October 2018 DFC Release	9

List of Tables

Table 1: Mean (SD) Standardized Domain Score by DFC Release.....	8
--	---

1. Introduction

The Centers for Medicare & Medicaid Services (CMS) has contracted with the University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) to act as quality measure developer and DFC technical content support contractor under the *Kidney Disease Quality Measure Development, Maintenance, and Support* contract. The contract number is 75FCMC18D0041, task order number 75FCMC18F0001. As part of this contract, UM-KECC convenes technical expert panels (TEPs) to provide valued consumer and provider input for both quality measure development and the Dialysis Facility Compare (DFC) Quality of Patient Care Star Rating (DFC Star Rating). The End Stage Renal Disease (ESRD) DFC Star Rating TEP, hereby referred to as ‘the DFC Star Rating TEP’ or ‘the TEP,’ was convened in response to CMS’ request for TEP recommendations on options for a future reset of the DFC Star Rating baseline distribution.

Technical Expert Panel Objectives

The DFC Star Rating TEP charter was distributed to the TEP members for review prior to the in-person meeting and was approved by the TEP members during the first of two pre-TEP teleconference calls. The TEP charter is included in Appendix B. Technical Expert Panel Charter. As per the TEP charter, the 2019 DFC Star Rating TEP will provide recommendations on options to inform the development of a methodology for resetting the DFC Star Rating distribution. The final methodology developed is intended to allow the DFC Star Rating to continue to reflect meaningful performance differences among facilities.

Technical Expert Panel Meeting

The DFC Star Rating TEP met via teleconference on May 6 and May 24, 2019. An in-person meeting was held in Baltimore, Maryland on June 6, 2019.

TEP Membership

A public call for nominations opened on March 6, 2019 and closed on April 5, 2019. The DFC Star Rating TEP membership represents many diverse perspectives and backgrounds. Members were selected based on their experience and/or expertise in one or more of the following areas: patients, caregivers, dialysis providers, or methodological expertise. Given that the intended audience for the DFC Star Rating is primarily patients, selection of TEP members aimed to achieve ample representation from patients and patient advocates. A comprehensive list of the TEP membership is provided in Appendix A. List of TEP Members and Contractor Staff.

2. Background

CMS developed the DFC Star Rating to help consumers, including patients and caregivers, understand CMS quality measures and more easily identify differences in overall quality when selecting dialysis facilities. The DFC Star Rating was developed as part of CMS' broader initiative for all of the Medicare Compare sites to make quality information more accessible to patients, caregivers, providers and policymakers.

The DFC Star Rating was first released in 2014 and scored dialysis facility performance using a relative rating system for a given evaluation period. As a result of patient and other stakeholder feedback, a TEP was convened in 2015 which recommended a rating system with absolute scoring cutoffs to track changes in facility performance over time. The new system established a baseline period, in which data are collected to define measure scoring criteria and cutoff values for star categories. The 2015 TEP recommendations were implemented in the updated DFC Star Rating methodology released in October 2016.

To accommodate changes in the measures that are used to calculate the DFC Star Rating, a second TEP was convened in 2017 to provide recommendations on the addition of new and updated measures, and on re-baselining. This approach establishes a new baseline period and new scoring cutoffs. The cutoffs are selected to ensure that the new baseline proportion of facilities in each star category is the same as in the last public release so continuity would be maintained. Rebaselining was implemented for the October 2018 DFC release to accommodate the addition of new and updated measures to the DFC Star Rating. See the respective 2015 and 2017 DFC Star Rating TEP reports for a summary of the deliberations and TEP recommendations. These reports are available at <https://dialysisdata.org/content/esrd-measures>.

The Medicare DFC website displays two Star Ratings: (1) the Quality of Patient Care Star Rating (DFC Star Rating) and (2) the Survey of Patients' Experiences Star Ratings. Eleven of the DFC clinical quality measures currently reported on the Medicare DFC website are used to calculate the DFC Star Rating. Six In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) survey measures are used to calculate the separate Survey of Patients' Experiences Star Ratings. The Survey of Patients' Experiences Star Ratings were not in the scope of the 2019 TEP.

The 2019 TEP was tasked to review and provide input on options for resetting the DFC Star Rating distribution in the near future to optimize the utility of the DFC Star Rating for patients and consumers. Recent observed trends in national performance demonstrate progressive shifts in the DFC Star Rating results over the last three years. These trends may obscure underlying performance differences if many facilities become concentrated at one end of the distribution. As background, prior DFC Star Rating TEPs (referenced above) identified a strong consumer interest in the ability to follow trends in dialysis facility performance over time. In addition, TEP discussions have considered how a reporting approach that reported longitudinal facility performance trends could be reset from time to time if the DFC Star Rating system lost the ability to show meaningful differences over the range of facility performance.

This report summarizes the discussions and recommendations of the 2019 ESRD DFC Star Rating TEP meetings that included two teleconferences (May 6, 2019 and May 24, 2019) and an in-person meeting on June 6, 2019 in Baltimore, Maryland. Minutes from the pre-TEP teleconference calls are provided in Appendix C. DFC Star Rating TEP Teleconference Call #1 Minutes and Appendix D. DFC Star Rating TEP Teleconference Call #2 Minutes A Post TEP teleconference call will be held on August 21, 2019. Minutes from the post-TEP teleconference call will be provided in Appendix E. DFC Star Rating Post-TEP Teleconference Call Minutes. The complete slide presentations for the teleconference calls and in-person meeting are provided in Appendix I., Appendix J., Appendix K., and Appendix L., respectively.

3. Summary of the Technical Presentation

The TEP was provided with the following background materials: the DFC Star Rating TEP Summary Reports from 2015 and 2017, the DFC Star Rating Methodology Technical Notes, and the presentations from the October 2017 and 2018 National Provider Calls. Links to these documents are provided in Appendix F. DFC Star Rating TEP Supporting Materials List.

The majority of both pre-TEP teleconferences and the first segment of the in-person meeting were used to provide a technical presentation of methods and analyses to the TEP members. These analyses were intended to inform a discussion by the TEP members related to resetting the DFC Star Rating distribution.

The following questions framed the presentation of the technical materials and subsequent TEP discussion:

1. Is it time for the DFC Star Rating distribution to be reset?
2. How should we reset the DFC Star Rating?
3. How should CMS help DFC consumers interpret facility performance during and immediately after any transition that would be created by a resetting of the DFC Star Ratings?
4. Should re-weighting be considered for the Domain 3 measures (Total Kt/V and Hypercalcemia)?

The following technical terms were defined on the teleconference calls:

Measure Value: The original value of a facility's clinical quality measure as reported on DFC, which represents a ratio or a percentage

Measure Score: A standardized score applied to a specific measure, which has mean 0, variance 1, and takes values in the range of -2.58 to 2.58

Domain Score: A score which summarizes a facility's performance on a subset (domain) of correlated clinical quality measures. It is an average of the individual measure scores in that domain

Final Score: A continuous score calculated for each facility, which summarizes its performance on the reported clinical quality measures. It is an average of the three domain scores

Cutoff: A final score value that defines the boundary between two adjoining DFC Star Rating categories

Baseline Period: The time period (e.g. calendar year) in which data are collected for computing measure scoring criteria and defining cutoff values for the DFC Star Rating categories. These cutoffs will be used to rate facilities in future evaluation periods

Evaluation Period: The time period (e.g. calendar year) in which data are collected for the calculation of measure results and facility final scores, as reported on DFC. Final scores in the evaluation period are compared against cutoffs established in the baseline period in order to rate facilities

Trends in DFC Star Rating Distribution

Figure 1 demonstrates changes in DFC Star Rating distribution between 2015 and 2018. In October 2015, 10% of facilities were rated 1-Star. This percentage decreased to 2.8% in October 2018, with nearly two-thirds of the original 1-Star facilities increasing by two, or more, star categories. At the opposite end of the scale, 10% of facilities were rated 5-Stars in October 2015 and increased to more than 25% of facilities in October 2018. More than half of the facilities are currently 4- or 5-Stars and few facilities are 1- or 2-Stars.

Figure 1: Trends in the DFC Star Rating Distribution, October 2015 – October 2018

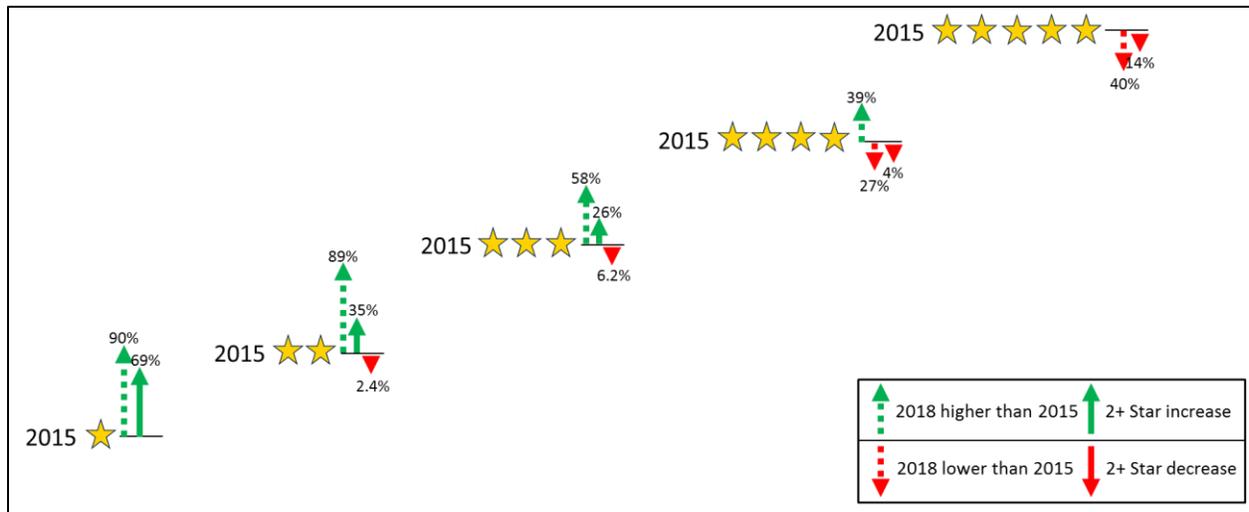


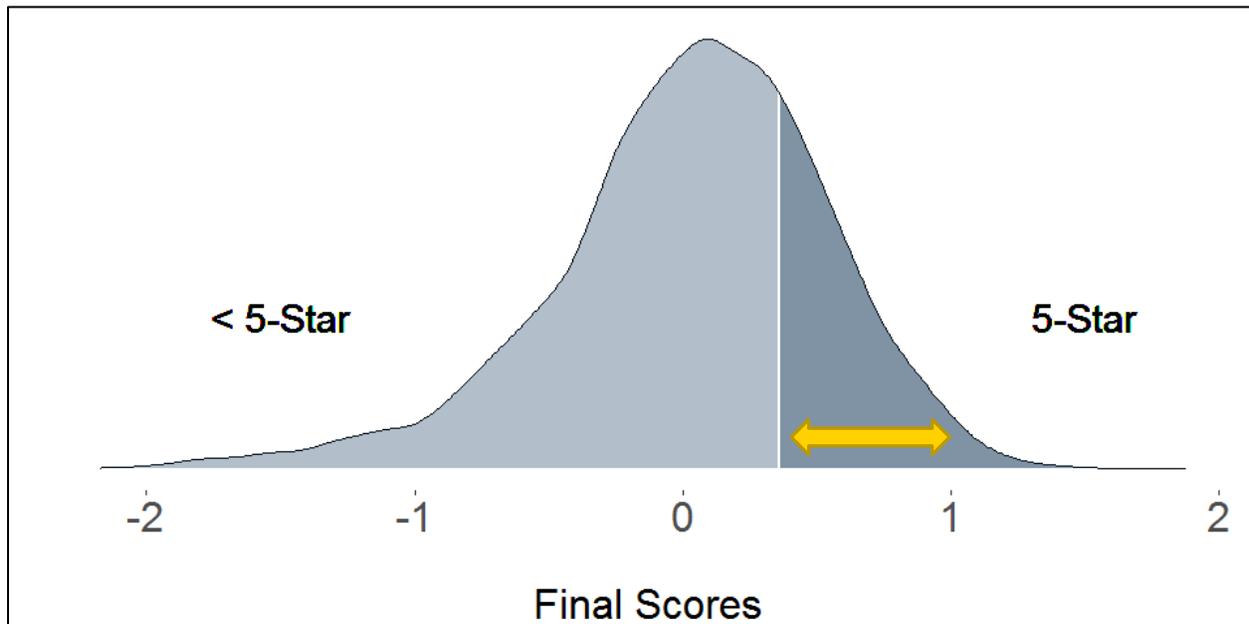
Table 1 displays the average domain scores and standard deviations by DFC release, before a new baseline was established in the October 2018 DFC release. Observed shifts in facility performance were mainly due to improvement in scores for the third measure domain, comprised of a measure of dialysis adequacy (Total Kt/V) and a measure of hypercalcemia. Improvement in the other measure domains has been gradual.

Table 1: Mean (SD) Standardized Domain Score by DFC Release

	October 2015	October 2016	April 2018
Domain 1 (SHR, SMR, STeR)	0.00 (0.69)	0.08 (0.69)	0.06 (0.70)
Domain 2 (Fistula, Catheter)	0.00 (0.85)	0.03 (0.87)	0.04 (0.86)
Domain 3 (Total Kt/V, Hypercalcemia)	0.00 (0.74)	0.36 (0.57)	0.82 (0.46)

The upward trend in DFC Star Rating reflects overall improvement in facility outcomes, while the national distribution of facility final scores has continued to demonstrate consistent variation, suggesting that heterogeneity in facility performance persists (see Figure 2 below).

Figure 2: Distribution of Facility Final Scores for the October 2018 DFC Release



Currently, some facilities have the same DFC Star Rating despite substantial variation in scores within a star category, making it difficult to distinguish performance differences between facilities. This is illustrated in Figure 2 by the relatively large spread of underlying final scores among the 5-Star facilities.

Goals and Options for Resetting the DFC Star Rating

After the 2017 TEP, CMS presented the following criteria during the October 2017 National Provider Call:

1. The DFC Star Rating distribution will be evaluated once 3 years have passed since the last reset
2. The DFC Star Ratings will be evaluated for a reset when $\leq 15\%$ of facilities are receiving 1- or 2-Stars
3. A resetting of the DFC Star Rating distribution will also include the establishment of a new baseline

With these criteria in mind, the 2019 TEP was tasked with providing input on the following three issues:

1. Is it time to reset the DFC Star Ratings?

The TEP was asked to discuss whether or not the DFC Star Ratings should be reset now to allow better differentiation of facility performance.

2. How should we reset the DFC Star Ratings?

Should CMS pre-specify the percent of facilities in each category and determine cutoffs from these proportions (e.g. 10-20-40-20-10 percent in 1-Star through 5-Star categories, respectively) or pre-

specify the cutoffs based on standard deviations from the national average score? Would the use of either Hierarchical or K-Means clustering techniques to identify the cutoff values for the five categories be a preferred alternative? Did TEP members have other suggestions for defining the cutoff thresholds associated with a reset event?

3. How should CMS help DFC consumers interpret facility performance during and immediately after any transition that would be created by a resetting of the DFC Star Rating?

Dr. Yi Li (UM-KECC) presented resetting options using pre-specified proportions or clustering approaches (see slides 20-28 in Appendix K.). His presentation included resetting the DFC Star Rating distribution to the original proportions of 10% 1-Star, 20% 2-Stars, 40% 3-Stars, 20% 4-Stars, and 10% 5-Stars (10-20-40-20-10). Under this approach the majority of facilities (66%) would experience a one star decrease in DFC Star Rating, with only five facilities experiencing a 2-star decrease.

Another option presented was to reset the DFC Star Rating using pre-set standard deviation (SD) cutoffs. A hypothetical example using cutoffs of -2, -1, +1, and +2 SD's was applied to the October 2018 DFC release data. This approach resulted in very few facilities receiving 1-Star or 5-Stars, with the majority receiving 3-Stars. It was noted the proportion of facilities in each star category could be modified by adjusting the pre-set SD values used to determine the star category cutoffs (e.g. -1, -0.5, +0.5, +1). That is, the proportion of facilities in each star category would depend on the chosen SD cutoff values.

An alternative approach presented was the application of empirical clustering techniques, which are statistical methods for grouping facilities that are more similar to each other compared other facilities. Dr. Li presented a brief introduction on two common approaches: Hierarchical and K-Means clustering. In the examples shown, if clustering methods were applied to the October 2018 release data, the distribution would result in a small proportion of facilities receiving 1- or 5-Stars. Dr. Li then presented some of the limitations of applying clustering techniques to the DFC Star Rating data (see slides 25- 27 in Appendix K.).

In response to the standard deviation method presented, one TEP Member asked if the distribution of scores would be recalculated each year to re-determine the cutoffs. Dr. Li clarified that the standard deviations would be determined using only baseline year data and would then be used as fixed cutoffs for the future years' ratings. Dr. Messina added that the original DFC Star Rating system used a method that reset the distribution every year automatically. During the 2015 DFC Star Rating TEP, the patient TEP members recommended allowing patients to see changes in performance from year to year, which resulted in the current method that uses a baseline to allow for improvements or regressions. The TEP Co-Chair stated that allowing improvement over time was important to patients at the prior TEPs. The tradeoff is that over time, if facilities "bunch-up," then a reset of the DFC Star Rating distribution may be necessary to allow users to continue to be able to differentiate higher and lower performing facilities.

Goals and Options for Reweighting the DFC Star Rating Measures or Domains

Dr. Li briefly reviewed the current DFC Star Rating methodology which applies equal weight to measures with each domain and equal weight to each domain when calculating a facility's final score (see slide 25, Appendix J.; page 9 of the DFC Star Rating Technical Notes for details on the methodology, see Appendix F. DFC Star Rating TEP Supporting Materials List).

Reweighting the measure domains may lessen the impact of any given domain and stabilize the trends in the DFC Star Rating distribution if the domain has achieved top-performance levels. The weight of measures and domains can be determined empirically or based on expert opinion, and weighting can be applied independently or simultaneously with resetting. Individual facility changes in ratings are expected with reweighting because it can reduce or eliminate the impact of measures that have very high achievement. It was noted that reweighting alone may not fix the issue of the increasing trends in facility performance.

As shown in Table 1, much of the progressive change in facility ratings has been driven by changes in the Domain 3 scores, while Domain 1 and 2 have been relatively stable across the past DFC Star Rating releases. When one domain (or measure) is found to be driving changes in star summary scores, the domain or measure with rapid achievement could be targeted as a candidate for down-weighting to reduce its impact on subsequent changes to the overall DFC Star Rating distribution. It was noted that down-weighting may only lessen, not reverse, the effects of the rapid achievement in a domain (or measure), and that resetting and reweighting are different, complementary, tools for re-calibrating the DFC Star Rating. Dr. Li then presented on the following two weighting approaches: (1) directly target domains with rapid achievement or (2) weight all domains based on the proportion of variance they explain in the data.

For the first approach, two options were presented for down-weighting Domain 3: (1) assign zero weight to Domain 3 (i.e., removal) or (2) reduce the relative weight of Domain 3 by 50%. Dr. Li presented a hypothetical analysis to show what would have happened to the change trend for the DFC Star Rating if Domain 3 had been down-weighted at baseline (see slides 13-18 in Appendix K.). For example, if the Domain 3 measures had not been included in the calculation of the DFC Star Rating since baseline, the distribution of stars in 2018 would have been close to 10% 1 star - 21% 2 stars- 39% 3-stars - 19% 4 stars - 11% 5 stars. Another example showed that a 50% reduction weight of Domain 3 might have also slowed the shift of facility star ratings, but to a lesser degree.

For the second approach, domains were weighted based on how much data variation can be explained by an individual domain. This is a data driven approach to quantify the importance of domains. Under this approach, the data directly determine the domain weights. In the analyses presented, the standardized measures explained the majority of the variation in the DFC Star Rating, while Domain 3 explained the least (see slides 56 and 57 in Appendix K.).

The TEP noted that facilities have high performance on Domain 3, and because of this high achievement, it may be more difficult differentiate between facilities. One TEP Co-Chair clarified that "variance explained," from a statistical point of view, can be thought of as how much information the particular measure provides

in discriminating performance among facilities, and gives a statistically optimal way to study the importance of measures in rating facilities. For example, the more variance a measure has, the more measure values will be different from facility to facility, which allows for better differentiation of facilities based on that given measure. The method of weighting based on variance takes into account how much information each measure provides for distinguishing facility performance. The method places more weight on the measures and domains that provide more information for distinguishing performance. In response to a TEP member question, the TEP Co-Chair and Dr. Messina clarified that not all of the variation explained by a given measure may be under the control of the facility and that some of the variation may be due to external factors (e.g., shared attribution of the measure outcome with other providers, or non-clinical factors).

One TEP Co-Chair clarified that each measure domain can be thought of as an underlying construct. For example, Domain 1 represents a concept of a type of quality outcome, and the different measures in that domain are different ways of measuring that underlying quality construct. Variance weighting for individual measures gives more weight to the measures that are better at measuring that underlying construct.

One TEP Co-Chair stated that if the variance based approach (using factor analysis) were applied at the time of reset, the results could be used to reweight the domains and measures allowing measures with the most variation to influence the distribution of ratings at the time of resetting to a new baseline.

One TEP Member stated that Domain 3 contains two measures (Total Kt/V and Hypercalcemia) that can be directly controlled by the facility. Therefore, more rapid achievement in facility performance since the baseline is expected for these measures. They stated that while facilities may have some influence on measures such as SHR, they do not exert as much direct control as on the Domain 3 measures.

Dr. Messina (UM-KECC) provided additional background, stating that reweighting alone may not be enough to correct the issue of “bunching up,” which is why it is necessary for the TEP to consider options for resetting the DFC Star Rating. It was noted that the trend towards improvement plateaued in the recent DFC release, and (1) there may not be as rapid of a trend anymore and (2) since rebaselining occurred in 2018, reweighting may not be as effective due to the high performance on the Domain 3 measures.

One TEP Co-Chair clarified that many factors can be considered, including making decisions about weighting measures when resetting. However, these methodological decisions ultimately will not change the percentage of facilities that receive each rating (though it could change which particular facilities receive each rating). Resetting refers to changing how many facilities fall into each DFC Star Rating category.

Dr. Messina clarified that removing measures can occur separately from resetting.

Follow-Up to Questions from the Teleconference Calls

RISK ADJUSTMENT

In response to TEP member questions from the TEP teleconference calls, Dr. Messina presented a slide that describes the risk adjustment used in the DFC Star Rating measures (See the slide 10 in Appendix K.). Dr. Messina re-stated there may be additional time at the end of the in-person meeting to allow TEP members to provide additional feedback, input, or recommendations on topics (such as risk adjustment and regional variation) that are beyond scope of the TEP charter/objectives. These additional items were referred to as “Parking Lot Items” throughout the in-person meeting (see the section entitled: Additional TEP Input on Topics Discussed Beyond the Scope of the TEP Charter).

REGIONAL VARIATION

Dr. Messina mentioned that questions brought up on the earlier calls about regional/geographic variation can be covered during the “parking lot” item discussion (see the section entitled: Additional TEP Input on Topics Discussed Beyond the Scope of the TEP Charter). Joel Andress, PhD (CMS) provided some additional background, stating that the current CMS policy on how to address regional variation is under review. CMS is expecting to receive a report from the Assistant Secretary for Planning and Evaluation (ASPE) that investigates risk adjustment for regional/geographic and other sociodemographic status (SDS) factors. Dr. Andress invited TEP members to provide their input on regional variation during the “parking lot” section (see the section entitled: Additional TEP Input on Topics Discussed Beyond the Scope of the TEP Charter) and the TEP’s input will be recorded and provided to the CMS team to review.

4. TEP Discussion

The TEP Co-Chairs led a discussion on the following topics:

What techniques would help users interpret the DFC Star Rating after a reset?

FOCUS OF THE DFC STAR RATING

The TEP Co-Chair stated that the original intent of the DFC Star Rating was for patients and consumers to use it as a tool similar to websites that report ratings for restaurants or hotels. The DFC Star Rating was created with the intention of informing patient choice and to help inform patient care in facilities. The TEP Co-Chair and other TEP members emphasized the importance of prioritizing patients.

PLAIN LANGUAGE COMMUNICATION ABOUT THE DFC STAR RATING

TEP members were asked for recommendations on how to best explain a future reset to patients and how to help patients interpret the DFC Star Rating. The TEP Co-Chair asked for TEP members to provide a consensus recommendation on the topic of plain language communications. TEP members provided a consensus recommendation that information from CMS explaining a reset of the DFC Star Rating should be communicated to patients and the community using plain language. Plain language communications were identified as important to ensure patients of all backgrounds understand what the reset means. Specifically, patients should be provided with a plain language explanation stating: (1) it is a reset year, (2) why a change to the DFC Star Ratings was necessary, and (3) how to interpret a reset and any changes.

TEP members recommended ensuring that any communication materials should be available for the entire dialysis community including patients, patient advocates, and provider organizations. Information about the DFC Star Rating reset should be clearly explained both on the website and at individual dialysis clinics. Several TEP members recommended that communications specifically identify (1) what measures are included in the DFC Star Rating, (2) what measures are not included in the DFC Star Rating at the time, and (3) that the DFC Star Rating measures may change in the future.

The TEP Co-Chair recommended that CMS consider reaching out to patient advocacy organizations and the TEP in the future for input on topics related to DFC Star Rating display, language, and communications.

PATIENT ENGAGEMENT AND EDUCATION

The TEP also expressed the importance of patient engagement and education when using DFC or the DFC Star Rating as a tool for patients to help them select dialysis facilities or to increase patient education about how to use the DFC Star Ratings. Several TEP members stated that they use the DFC Star Rating and DFC website as a tool when assisting patients in choosing a facility or for education. Several TEP members stated that the DFC Star Rating can be used as one tool or part of the decision-making process but should not be the only tool used when determining if a dialysis facility is the best fit for a given patient.

One TEP member specifically stated that patients need to know that the DFC Star Rating, as a tool, is valid and usable in order to have confidence in using it for themselves. Not all patients know that the DFC is intended as a tool for dialysis patients, caregivers, and their family members. Several TEP members stated it was their role (as a patients, patient advocates, providers, and leaders in patient advocacy organizations) to help communicate to patients the tools available to them, (e.g. DFC and the DFC Star Rating).

Several TEP members stated that more than half of patients “crash” into dialysis in the emergency room (i.e. have to begin emergency dialysis without any planning). For many patients, it may be difficult to process important clinical information under these circumstance because of their compromised health and overall well-being. It was also noted that many patients may not be able to use the DFC Star Rating when they initially start dialysis, and often incident dialysis patients may be assigned to a facility by their nephrologist. The TEP Co-Chair stated that with these considerations in mind, it is also important for caregivers of patients be able to use the DFC Star Rating as a decision-making tool.

Several TEP members emphasized the importance of patient education to help patients be engaged in their care. Patient engagement/education was identified as important by TEP members for information about transplant options, modality options, and to promote patients being involved in their care. For example, TEP members expressed concerns that some patients may not understand that transplant is an option. Specifically, one TEP member stated the importance of patients receiving the best dialysis treatment possible to keep patients as healthy as possible to make them better candidate for transplantation. Another TEP member stated the importance of prioritizing dignity for the individual patient in order to help patients participate and engage in their own care.

STAR RATING INTERPRETATION

TEP members stated it was important to understand how patients interpret the DFC Star Rating before setting criteria for resetting the DFC Star Rating distribution. Patients, dialysis organizations, and additional community members may have differing interpretations of what the DFC Star Rating means.

The TEP discussed how most people and patients are familiar with using websites that report star ratings for restaurants, hotels, and other services and goods. Because of this, TEP members felt that patients would likely interpret 1-Star as poor and 5-Stars as excellent. TEP member discussion further focused on different interpretations of 1-Star, 2-Stars, 3-Stars, 4-Stars, and 5-Stars and how they influence the standard of healthcare expected. Overall most TEP members agreed that 1-Star and 5-Stars should be used to clearly identify outlier facilities (i.e. very poor, unacceptable, avoided at all cost; or very good, exceptional).

Some TEP members recommended that the DFC Star Rating be based on standards of care. Dr. Andress (CMS) stated that the measures were not created to reflect standards of care. While there may be expected standards of care for a patient, it may not be clear on what the standard of care is for a facility performance measure, like mortality (e.g., “how many deaths in a facility is acceptable?”).

TEP members also noted that patients need to feel confident in the accuracy and credibility of a tool like the DFC Star Rating in order to use it. The TEP also discussed how to interpret the DFC Star Rating as relative or absolute. There was some TEP discussion on the importance of using an absolute rating to allow patients to see if facilities are improving on the measures over time, while a relative rating would show a rating based on how a given facility compares to other facilities.

Other points made by several patient TEP members were that it is important that the DFC Star Rating hold facilities accountable for the level of care provided, and several noted that if they attended a facility with a low rating, they would like to know why that facility has a low rating.

MOCKUP DISPLAY EXAMPLE

The next part of the discussion focused on examples for displaying DFC Star Rating information after a reset. Dr. Messina presented a mockup display example to show how a DFC Star Rating reset could potentially be presented at a facility level (see slide 35-37 in Appendix K.). As UM-KECC does not control or have authority over the DFC website, the mock up examples served to illustrate the concepts. In particular, the mockup illustrated how a facility's previous and current rating and percentile score could be displayed next to each other in order to display longitudinal trends. Two hypothetical examples were presented: Facility A was an example of a high performing facility. Facility A received 4-stars during the previous year, and scored a high final score in both the previous year and current year. Facility A's rating did not change during the DFC Star Rating reset. Facility B was an example of an average performing facility. Facility B experienced a drop from 4 stars to 3-stars due to the star rating reset even though their percentile score slightly increased. Additionally, it was noted a drill-down section could provide more detail on the rating and final score.

Dr. Messina asked for discussion on how much information is too much information to display. The TEP Co-Chair stated that if patients and caregivers are involved it is not possible to have "too much information" but it is important the information is displayed and organized in a useful way.

One TEP member asked if it was possible to see two years' worth of results for measures and percentiles following the period when ratings were reset. Dr. Messina stated that, while there are measure-specific drill-downs currently available on DFC, it was unclear if that level of detail would be possible to display. There was some TEP discussion on potentially having hover text or a footnote on the facility display level which could explain why a facility DFC Star Rating dropped after a reset (e.g., whether it was due to the reset or an observed decline in facility performance).

Is a reset of the DFC Star Rating needed now?

Based on the October 2018 DFC Release, about half of facilities currently receive 4- or 5-Stars (see slide 17 in Appendix I.). The current distribution shows that a majority of facilities have improved in performance over time. Originally, in the 2015 DFC Release, 3-stars was defined as "average" and 40% of facilities received 3-stars. Based on the distribution of the DFC Star Rating final scores, a majority of facilities are in the middle of the distribution (see slide 31 in Appendix K.). Because of the shift in performance that has

occurred over the past few years, a facility that used to be “average” may now be classified as “above average” based on the DFC Star Rating labels. There was some TEP discussion that this scenario may or may not reflect the typical consumer’s understanding of what “average” quality means.

The TEP Co-Chair provided background to frame the discussion on resetting. The TEP Co-Chair explained that rebaselining is conducted when there are changes to measure sets but rebaselining does not reset the distribution. When the DFC Star Rating is rebaselined, the percentage of facilities assigned to each star rating category (e.g. 3-Stars, etc.) would stay the same (to allow for longitudinal continuity), while resetting the DFC Star Rating distribution changes the percentage of facilities in each category (e.g. 3-Stars, etc.). Resetting the DFC Star Rating distribution involves resetting the bar for performance. The downside of resetting the distribution is that some facilities will receive a lower rating, despite the possibility of not worsening in terms of their absolute measure of performance. Several TEP members stated the importance of having maximum differentiation of facility performance to help patients choose facilities. Several TEP members stated that a wider spread in the proportions of facilities in each DFC Star Rating category than the current distribution would be more meaningful.

One patient TEP member specifically stated that patients may not always understand complex measures, but a star rating is more understandable. The TEP member stated that patients do not enter dialysis in optimal scenarios, but if a patient has a choice of which facility to go to, then the star rating should be a tool for patients and caregivers, since differentiation in facility performance would be important. Another patient TEP member added that, even though patients in rural areas may not have the option to switch facilities, a meaningful DFC Star Rating that differentiates facility performance might incentivize all clinics to keep focusing on improving all of the time. There was TEP discussion on whether patients would prioritize being able to continue to track star ratings over time or to see maximum differentiation among facilities. The patient TEP members reported that both are important. The TEP Co-Chair summarized that based on the TEP feedback, patients want to see trends over time, differentiation between facility performances, and that a facility provides quality care.

What distribution should the DFC Star Rating be reset to?

Setting the DFC Star Rating cutoffs determines what percentage of facilities are assigned to each star rating category (1-Star, 2-Stars, etc.) in the baseline year. After the baseline year, the distribution can shift up or down, meaning that improvement (or decline) over time can be measured. The TEP Co-Chairs presented the distribution of the October 2018 DFC Star Rating release final scores and asked TEP members to provide a recommendation on where to set the final score cutoffs for the DFC Star Rating. The TEP Co-Chairs discussed the possibility of making the distribution of stars uniform (the same number of facilities in each category) versus having the 1- and 5-Star categories represent facilities in the extremes, or any variation in between depending on the preferred distribution of final scores. The TEP Co-Chairs noted that the distribution of the final scores is fairly symmetric, suggesting that there should be equal proportions in the 1- and 5-Star categories and in the 2- and 4-Star categories.

The TEP members discussed different options for the DFC Star Rating distribution. The focus of the discussion was on what percentage of facilities should be assigned 1- and 5-Stars. TEP members expect 1-Star to be assigned to facilities that are performing very poorly and 5-Stars to be assigned to facilities that are exceptional. As a result, TEP members expected that a small percentage of facilities should be assigned 1-Star or 5-Stars. Several TEP members stated that attending a 1-Star facility should necessitate a discussion between patients and providers about the facility's quality, as some patients may consider leaving their clinic. One TEP member (a provider) noted that if one of their organization's clinics is assigned 1-Star, their organization will devote resources to address and improve the outcomes affecting its rating.

There was TEP discussion on resetting to a 10-20-40-20-10 (10% receive 1-star, 20% receive 2 –stars, 40% receive 3-stars, 20% receive 4-stars, and 10% receive 5-stars) distribution. This is the same distribution that was initially released in January 2015 (and used as the baseline distribution for the October 2015 release). The TEP discussed the advantages and disadvantages of resetting to the 10-20-40-20-10 distribution. One advantage is there is precedent in resetting to a previously implemented distribution. Some TEP members were in favor of resetting to the original distribution (10-20-40-20-10) as long as it was communicated and explained clearly to patients. This same distribution could be considered in the future. A disadvantage of resetting to 10-20-40-20-10 is that 30% of facilities would receive either 1- or 2-Stars. One TEP member was concerned that 30% was high for classifying facilities in these low star rating categories.

Some TEP members provided some alternative recommendations to consider for a potential reset, such as using 20-20-20-20-20, using a distribution with less than 30% 1- and 2-Stars combined, or using standard deviations to reset the distribution. Different levels of standard deviations (e.g. -1, -1.5, +2, and +2.5) could be considered for resetting the DFC Star Rating distribution. The TEP Co-Chair stated that the technical details can be arranged to fit the desired distribution. Several TEP members recommended using the same method and distribution for future resets of the DFC Star Rating distribution. The goal would be so that patients and dialysis organizations know what to expect in terms of the reset method and distribution that is chosen for future resets in order to establish a standard moving forward.

Discussion of Re-Weighting the DFC Star Rating Measure Domains

TEP members discussed whether highly skewed measures (i.e. Total Kt/V and Hypercalcemia) should remain in the DFC Star Rating since many facilities achieved very high performance on these measures. TEP members provided differing opinions on this topic. Some TEP members stated that if a measure is highly skewed, it may no longer provide information and should be considered for removal from the DFC Star Rating. Other TEP members expressed interest in keeping these measures in the DFC Star Rating, but down-weight them to 50% of their original weight, to continue to hold facilities accountable for these outcomes. One TEP member specifically stated that it is important to focus on improving other domains as well and referenced the importance of avoiding hospitalizations. Several provider TEP members stated that facilities had the most direct control and responsibility over the Domain 3 measures (Total Kt/V and Hypercalcemia) versus other measures like hospitalization. One provider TEP member stated that because Total Kt/V and Hypercalcemia are highly skewed measures, it would be hard for many facilities to further improve their star ratings, but it could be worth keeping those measures in the DFC Star Rating to track if facilities

experienced declines in performance on these metrics.

Other TEP members also provided their perspective on the Kt/V measure. One TEP member expressed the concern that Total Kt/V is highly skewed and this measure may have outlived its usefulness as a measure of dialysis adequacy. Several patient TEP members expressed that Kt/V is an important measure to them, not as much as a clinical value, but as a measure of quality of life and how patients feel and function. One provider TEP member shared a concern about the Hypercalcemia measure. They stated that the Hypercalcemia measure is highly skewed and also may have missing data and data accuracy issues.

While TEP members provided individual perspectives on reweighting, an official TEP recommendation was not reached during the in-person TEP meeting.

TEP Vote on Resetting the DFC Star Rating

TEP members were asked to vote on the questions below in order to provide a summary recommendation on the topic of resetting. TEP votes were anonymous in order to protect individual TEP member opinions and respect individual anonymity. All 19 TEP members voted. The results from this vote are as follows:

1. Should the Star Ratings be reset now?

Results: 16 of 19 TEP members voted yes the star ratings should be reset now.

2. Would you be willing to use 10-20-40-20-10 distribution as long as it was effectively explained?

Results: 12 of 19 voted yes they would be willing to use the 10-20-40-20-10 distribution as long as it is effectively explained to DFC users. 7 of 19 voted no to using the 10-20-40-20-10 distribution. Those that answered no to question 2, also voted on question 3 below.

3. If you answered no to #2, what is your recommended percentage for each category (1 star, 2 stars, 3 stars, 4 stars, 5 stars)? It must add up to 100%.

Results: 7 TEP members offered other star rating distribution recommendations (provided below). Three TEP members recommended a 5%-15%-60%-15%-5% distribution. (Note this would result 5% 1 star, 15% 2 stars, 60% 3 stars, 15% 4 stars, and 5% 5 stars). One TEP member recommended a 7.5%-12.5%-60%-12.5%-7.5% distribution. One TEP member recommended a 10%-10%-40%-30%-10% distribution. One TEP member recommended that the total percentage of facilities with 1 and 2 stars would be below 15%; 50% would be assigned 3 stars, and the percentage of facilities with 4 and 5 stars be below 35%. One TEP member recommended using standard deviations of 2.5 to reset.

The TEP Co-Chairs summarized the voting results. Seventeen of nineteen TEP members favored the use of a symmetrical star rating distribution for the reset. Two of nineteen TEP members favored use of an asymmetrical star rating distribution for the reset.

5. TEP Summary

Recommendations for Resetting the DFC Star Rating

1. A supermajority of TEP members (84%) agreed that the DFC Star Ratings should be reset based on the current distribution of facility DFC Star Ratings. Most TEP members agreed that the current distribution of facility DFC Star Ratings does not capture the full range of facility performance, reducing the effectiveness of the program.
2. A majority of TEP members (63%) voted to support a reset to a fixed, pre-specified distribution of 10-20-40-20-10, while a minority (21%) recommended resetting to a distribution with smaller percentages in the distribution tails (e.g. 5% 1-Star, 15% 2-Star, 60% 3-Star, 15% 4-Star, 5% 5-Star). During the pre-vote discussion, at least one TEP member stated a concern about the use of the 10-20-40-20-10 distribution, as they suggested that having 30% of facilities rated 1- or 2-star after a reset would not reflect the overall quality in some of those facilities.
 - a. As described during the TEP discussion (see page 17), setting the DFC Star Rating cutoffs determines what percentage of facilities are assigned to each star rating category (i.e., the percentage of 1-Star, 2-Star, 3-Star, 4-Star, and 5-Star facilities) in the baseline year. After the baseline year, the distribution of facilities can shift up or down based on their performance, allowing for improvement (or declines) in facility performance over time. The ability to observe improvement over time in star ratings was implemented into the star rating methodology as a result of the 2015 Star Rating TEP recommendations.
3. To allow for resetting while mitigating unintended consequences, TEP members emphasized the need for CMS to provide clear, concise educational aids and supplemental information in plain language as part of a DFC Star Rating reset, including explicit statements emphasizing (a) it is a reset year, (b) why a change to the DFC Star Ratings was necessary, and (c) how to interpret any changes.
4. Many TEP members suggested that DFC Star Rating resets should be limited in frequency and many agreed with the notion that resetting should occur using a standardized methodology in the future, to minimize the resulting confusion. Most TEP members agreed that resetting would provide net benefit to DFC users while acknowledging the potential confusion and added difficulty in interpreting the DFC Star Ratings during the reset transition period.

Additional Themes from TEP Member Comments

The comments here are intended to highlight a theme discussed or raised by multiple members that was relevant to the outcome of the TEP deliberations. Several members of the TEP commented that the DFC Star Ratings are generic and would be more useful if the ratings could be customized to reflect the individual user's "outcomes of interest." That is, users could select the quality measures most important to them and

the DFC Star Rating would be based on that. During that portion of the discussion, one member emphasized the importance of having the detailed individual measure information available in addition to the DFC Star Rating summary to allow individual users to drill down and learn more about the outcomes that are specifically relevant to their clinical situation or interests.

Outstanding Issue

Recommendations are needed on the reweighting of Domain 3, consisting of the Kt/V and Hypercalcemia measures. TEP consensus was not reached on whether Domain 3 should be weighted at 100% weight (current weight), 50% weight, or 0% weight (removal). The issue will be on a post-TEP teleconference call. See Appendix E. DFC Star Rating Post-TEP Teleconference Call Minutes for the Star Rating Post-TEP Teleconference Call Minutes.

Additional TEP Input on Topics Discussed Beyond the Scope of the TEP Charter

As referenced earlier in the report, TEP recommendations on the topic of resetting the DFC Star Rating distribution were the main focus of the TEP discussion. At the beginning of the meeting, it was stated that there may be additional time at the end of the in-person meeting to allow TEP members to provide additional feedback and input on topics (such as risk adjustment and regional variation) that are beyond scope of the TEP charter and TEP objectives. These additional topics were referred to as “Parking Lot Items” throughout the in-person meeting. Below is a summary of the discussion of these additional topics.

FUTURE AREAS OF FOCUS FOR MEASURE DEVELOPMENT

Several TEP members stated the importance of the DFC Star Rating measures being patient-driven measures and based on topics that are important to patients.

Individually, TEP members provided recommendations on potential future areas of focus on measure development and patient quality of life. Transplant measures, fluid management, patient education, and modality education (i.e. information on modalities such as home dialysis) were most commonly identified as areas of interest expressed by TEP members. Patient TEP members specifically expressed the importance of knowing how well facilities are referring patients for transplant on a comparative basis among facilities. Several TEP members also stated the importance of knowing if facilities are supportive of and appropriately utilizing home dialysis therapies.

One TEP member stated that there should be more emphasis on education about different modalities from nephrologists. Information should be made available to help establish conversations between patients and their nephrologist about modalities. One patient TEP member stated that there is good information for patients and caregivers on DFC. More is needed to promote use of these tools, such as the DFC Star Ratings, to help patients ask important questions about their care.

Individual TEP members also identified the following topics as important: anemia management, facility

communication with patients, and patient-reported outcome measures.

DFC WEBSITE DISPLAY

TEP members stated concerns about the high reading level of the DFC website content. TEP members stated that the DFC website and communication materials should be plain language accessible to a broader range of reading levels. Dr. Messina clarified that UM-KECC does not control DFC content and presentation, but provides input about measure technical content. One TEP Co-Chair stated that at the previous TEPs, the patient advocacy organizations had spoken with CMS on how to display things on DFC in an easier way for patients and the public. There are separate CMS efforts that focus on the DFC display work.

Technical terms on DFC, such as “as expected,” are used to categorize facility performance on measure like mortality and hospitalization and may not be clear to patients. The patient TEP members stated the importance of patients knowing simple summaries such as how many patients died, were admitted for bloodstream infections, or had a readmission at a dialysis facility, for example.

One patient TEP member stated an interest in seeing a customizable approach for patients to use DFC and the DFC Star Rating in the future. Using such an approach, patients could weight measures and/or measures domains according to their own preferences, assigning more weight to the measures or domains they deem most important to their decisions.

Dr. Address (CMS) expressed that it is important for CMS to do a better job in explaining the DFC Star Rating program to patients. Dr. Address stated that there will be future TEPs on the topics of adding/removing measures, or reweighting/restructuring domains.

REGIONAL VARIATION

There was TEP discussion on the topic of regional variation in the DFC Star Ratings.

One TEP member asked if UM-KECC had conducted an analysis of geographic variation. Dr. Messina (UM-KECC) stated that UM-KECC had presented a poster at ASN in 2018 examining facility distance and patient selection of facility. Based on the analysis, it appears that a majority of patients have the option to choose a higher rated dialysis facility within 10 miles from their current facility.

It was stated that previous analyses suggested some level of regional variation in ratings but it may be unclear how much of that is region specific factors or due to disparities in care. It was acknowledged there are sociodemographic factors that may affect regional variations in care.

Some TEP members stated concerns about access to healthy foods to establish or maintain a healthy diet, dental care, healthcare, and vascular surgeons as being challenges in some regional areas.

One patient TEP member asked if the website could display DFC Star Ratings by region or rural vs urban location, for example. Dr. Messina stated that the topic of an individualized star rating (by using customizable search settings to score facilities) has been brought up in the past. Dr. Messina stated that redesign of DFC capabilities was beyond the scope of this TEP. Dr. Andress (CMS) asked if a separate PDF report that provides results by different characteristics (for example, geographic region, urban or rural location, etc.) could be developed in order to be responsive. This would be in place of a customizable DFC website search. The TEP patient member responded that anything to be responsive could be helpful.

The TEP discussion referenced the ASPE study and its findings on sociodemographic factors and quality measures. Dr. Andress explained that, as part of measure development and testing, analyses are performed that look at the impact of sociodemographic risk adjustment on the standardized outcome measures.

PALLIATIVE DIALYSIS

A provider TEP member asked for patient TEP member opinions on whether palliative dialysis should be provided as a treatment option. One patient TEP member agreed that patients should be given an option to have palliative dialysis as opposed to being required to undergo three treatments a week.

The TEP Co-Chair stated that it is important for providers to discuss patient life aspirations with patients and caregivers in order to best inform a patient treatment plan. The TEP Co-Chair said that understanding a patient's motivation or aspirations helps inform how to best treat that patient.

One patient TEP member stated that the frequency of treatment may be an individualized, shared decision for patients and providers. Some patients may be seeking palliative care or some patients may be regaining residual kidney function. Both circumstances may affect how often a patient is seeking dialysis treatment.

OTHER DISCUSSION TOPICS

One patient TEP member asked if the DFC Star Rating system could be increased to be a 10 star rating system in order to allow for more measurement of longitudinal growth and to show more differentiation. Another TEP member recommended that half-star increments should be considered for the DFC Star Rating system. Other TEP members raised issues about interpretability of the DFC Star Rating.

A few TEP members stated concerns about the length and lack of follow-up for the In-Center Hemodialysis (ICH) CAHPS survey. One patient TEP member expressed the desire to have a follow-up after completing the survey. The TEP member stated that patients are more likely to complete the survey if there was feedback (i.e. to show that the results are being used to improve care). A provider TEP member stated that it is difficult for the survey to be actionable because facilities only receive aggregated results.

One patient TEP member stated they want to see facilities held accountable for what they are responsible for. The patient TEP member stated that, in terms of bloodstream infections, some of the responsibility is on the patient to maintain a clean vascular access site. The patient TEP member also stated the importance

of patient dental care/cleanliness as part of patient health. They stated the importance of restoring dignity to the individual patient as part of the patient's care.

Patient TEP members also identified measures important to patients including avoiding bloodstream infections, mortality, and hospitalization.

6. Public Comment Period

After the TEP discussions concluded, time was set aside for a public comment period. Several members of the public were in attendance to observe the in-person TEP meeting discussions. The public comment period is intended as an opportunity for members of the public (who are not serving on the TEP) to provide their comment so that it is recorded in the TEP summary documentation. One member of the public, Kathy Lester from Kidney Care Partners (KCP), provided a public comment. The comment is transcribed below:

“Well, thank you, my name is Kathy Lester, and I am counsel to the Kidney Care Partners, which is an organization of providers, manufactures, patient advocates, physicians, nurses, and other professionals in the dialysis field and first, I just want to thank each and every one of you for the time that you put into sitting at this TEP, to CMS for having this TEP, making it a public TEP and having this dialogue. As you know KCP has been very involved in quality from the first step of trying to create a value-based purchasing program and working with Congress and CMS to succeed in getting that done and now in looking at 5-star and trying to refine it so that it does meet the patients’ needs. And I know how much time it takes, and I appreciate you all doing this. Very much appreciate this dialogue today and the important issue of resetting the rates. I think we would agree what many of the patients said, even just now, about the need to really rethink some of the measures in this program. Many of these measures are topped out. And while they are important to report and make sure everyone has access to that information, they may not be appropriate for a Star Rating or even in the QIP program. So very much look forward to continue dialogue on that, and getting to those meaningful measures that people want to use to select their facilities. We also really appreciate the points that were made by several patients, and our members feel this as well, that stars have certain connotations that are very, very difficult to break in peoples’ minds. So very encouraged by the dialogue of how do we take patients at where they are? And where those perceptions are, and make sure that we are getting that right information to them. And one of the things that we have thought about and very much look forward to further conversations with CMS and maybe even all of you at a future TEP, is really returning – as Joel said, these programs to what they are meant to be. So the QIP is a value based purchasing program. And DFC, with or without stars, is about giving patients information about quality measures and really having that information out there to access. So we’ve thought and put some time into a way that could be done. So you don’t have competing programs that conflict with each other, same measures with different specifications in both programs, but really divide them out in a way, and I think that’s another thing that I would encourage a future TEP or discussion around because it would create that consistency and reduce some of the confusion that we’ve seen. But again very much appreciate your time, and dedication, and passion around this issue. And look forward to further conversations.”

After the public comment, Paul Conway stated the following:

“Hi, I would like to thank Kathy for her comments and I would also like to thank you for the letter you sent to me and to Dr. Sugar which I will ask it to be entered into the public record that was sent before the TEP. Unfortunately, neither one of us had a chance to read the entire document, but I would like that to be put in the public record for the TEP.”

7. Appendices

- Appendix A. List of TEP Members and Contractor Staff
- Appendix B. Technical Expert Panel Charter
- Appendix C. DFC Star Rating TEP Teleconference Call #1 Minutes
- Appendix D. DFC Star Rating TEP Teleconference Call #2 Minutes
- Appendix E. DFC Star Rating Post-TEP Teleconference Call Minutes
- Appendix F. DFC Star Rating TEP Supporting Materials List
- Appendix G. TEP Composition Form
- Appendix H. Public Comment Letter from Kidney Care Partners (KCP)
- Appendix I. Teleconference Call 1 Presentation
- Appendix J. Teleconference Call 2 Presentation
- Appendix K. In-Person Meeting Presentation
- Appendix L. Post-TEP Teleconference Call Presentation

Appendix A. List of TEP Members and Contractor Staff

The following individuals were selected to serve on the TEP:

Name and Credentials	Organizational Affiliation, City, State	Conflicts of Interest Disclosed
Paul T. Conway, BA <i>TEP Co-Chair</i>	Board of Directors Member; Chair of Public Policy and Global Affairs, American Association of Kidney Patients, Falls Church, VA	None
Catherine A. Sugar, PhD, MS <i>TEP Co-Chair</i>	Professor, Departments of Biostatistics, Statistics & Psychiatry, University of California, Los Angeles, CA Director, Semel Institute Statistics Core, University of California, Los Angeles, Los Angeles, CA	None
Mark Andaya, MS, RN	Director of Quality Assurance and Performance Improvement, The Rogosin Institute, New York, NY	None
Andrew Conkling, BS	Vice President, Dialysis Patient Citizens, Arab, AL	None
Lorien Dalrymple, MD, MPH	Vice President of Epidemiology and Research, Fresenius Medical Care North America, Waltham, MA Volunteer Clinical Faculty Associate Professor, Dept. of Medicine, Division of Nephrology, University of California, Davis, Sacramento, CA	Employed by and share-options in Fresenius Medical Care; Member of the Kidney Care Quality Alliance Steering Committee and has participated in research related to quality measures
Sharon Dickson, RN, MSN	Regional Quality Manager, Fresenius Kidney Care, Lancaster, SC	None

Name and Credentials	Organizational Affiliation, City, State	Conflicts of Interest Disclosed
Dawn Edwards	Patient Advocate, NxStage, New York, NY Health Ambassador, The Rogosin Institute, New York, NY	None
Derek Forfang	Patient Advocate/Chair, Kidney Patient Advisory Council, Forum of ESRD Networks, San Pablo, CA Patient Advocate/Chair, Kidney Advocacy Committee, National Kidney Foundation, San Pablo, CA	Chair of the Public Policy Committee for the National Kidney Foundation
Monica Fox	Outreach Associate/Patient Advocate, National Kidney Foundation of Illinois, Flossmoor, IL	None
Lonnie Green	Patient Advocate/Subject Matter Expert, IPRO ESRD South Atlantic Network #6, Kennesaw, GA Patient Advocate/Subject Matter Expert, National Patient and Family Engagement – Learning Action Network, Kennesaw, GA	None
Mark Johnson	Patient Advocate/Network Patient Representative, Heartland Patient Advisory Council, ESRD Network 12, Atlantic, IA Patient Advocate/Subject Matter Expert, National Patient and Family Engagement – Learning Action Network, Atlantic, IA	None
Mark Joseph, MD	Pediatric Nephrologist, Pediatric Kidney Disease and Hypertension Centers, Phoenix, AZ Medical Director, Phoenix Pediatric Dialysis Center, Phoenix, AZ	None
Richard Knight, MBA	President, American Association of Kidney Patients, Bowie, MD Adjunct College Instructor, Bowie State University, Bowie, MD	None

Name and Credentials	Organizational Affiliation, City, State	Conflicts of Interest Disclosed
Mahesh Krishnan, MD, MPH, MBA	Group Vice President of Research and Development, DaVita Kidney Care, Washington D.C.	Employee of DaVita
Michael “Jack” Lennon, MBA	Executive Director, Improving Renal Outcomes Collaborative, Cincinnati, OH	None
Nicole Stankus, MD, MSc	Medical Director, DaVita Stony Island Dialysis Center Associate Professor, Department of Medicine Section of Nephrology, University of Chicago, Chicago, IL	Medical Director of the DaVita Stony Island Dialysis Center and a member of the DaVita Physician Council
Caprice Vanderkolk, RN, MS, BC-NE	Nurse Manager, Adult and Pediatric Dialysis Program, University of Minnesota/Fairview Hospitals, Minneapolis, MN	Secretary of the National Renal Administrators Association
Curtis Warfield, BS, MS	Patient Reviewer, National Kidney Foundation, Indianapolis, IN Senior Quality Analyst, State of Indiana, Indianapolis, IN	None
David “Dave” White	Patient Advocate/Board of Director Member, American Association of Kidney Patients, Hillcrest Heights, MD Healthcare Consultant, Quality Insights Renal Network 5, Hillcrest Heights, MD	Family member stock in Amgen, CareDX, and Proteon Therapeutics. Grants with Quality Insights Renal Network 5. Serves on the Board for Network 5 and the University of North Carolina Kidney Center

The following individuals are contractor staff at the University of Michigan Kidney Epidemiology and Cost Center:

Name and Credentials	Organizational Affiliation, City, State	Conflicts of Interest Disclosed
Joseph Messana, MD	Director and Professor of Nephrology/Internal Medicine, University of Michigan, Kidney Epidemiology and Cost Center	None
Yi Li, PhD	Professor of Biostatistics, University of Michigan, Kidney Epidemiology and Cost Center	None
Claudia Dahlerus, PhD, MA	Principal Research Scientist, University of Michigan, Kidney Epidemiology and Cost Center	None
Richard Hirth, PhD	Professor of Health Management and Policy, University of Michigan, Kidney Epidemiology and Cost Center	None
Peisong Han, PhD	Professor of Biostatistics, Department of Biostatistics, University of Michigan, Kidney Epidemiology and Cost Center	None
Stephen Salerno, MS	Graduate Student Research Assistant, University of Michigan, Kidney Epidemiology and Cost Center	None
Wolf Gremel, MS	Senior Research Analyst, University of Michigan, Kidney Epidemiology and Cost Center	None
Jingya Gao, MS	Research Analyst, University of Michigan, Kidney Epidemiology and Cost Center	None
Jennifer Sardone, BA	Project Senior Manager, University of Michigan, Kidney Epidemiology and Cost Center	None
Jordan Affholter, BA	Associate Project Manager, University of Michigan, Kidney Epidemiology and Cost Center	None
Brandon Frye, BA	Project Support Analyst, University of Michigan, Kidney Epidemiology and Cost Center	None
Karen Wisniewski, MPH	Lead Research Analyst, University of Michigan, Kidney Epidemiology and Cost Center	None
Lan Tong, MPH	Lead Research Analyst, University of Michigan, Kidney Epidemiology and Cost Center	None
Casey Parrotte, PMP	Lead Project Manager, University of Michigan, Kidney Epidemiology and Cost Center	None

Appendix B. Technical Expert Panel Charter

Project Title:

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

TEP Nomination Period:

March 6, 2019 - April 5, 2019

Project Overview:

The Centers for Medicare & Medicaid Services (CMS) has contracted with the University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) to act as quality measure developer and DFC technical content support contractor, under the *Kidney Disease Quality Measure Development, Maintenance, and Support* contract. The contract number is 75FCMC18D0041, task order number 75FCMC18F0001. As part of this contract, UM-KECC convenes technical expert panels (TEPs) to provide valued consumer and provider input for both quality measure development and the Dialysis Facility Compare (DFC) Quality of Patient Care Star Rating (DFC Star Rating). This TEP Charter has been developed in response to CMS' request for TEP recommendations related to options for resetting the DFC Star Rating baseline distribution. We are seeking interest from individuals with relevant experience, expertise, and a variety of perspectives to serve on this TEP, including patients.

A DFC Star Rating TEP was first convened in 2015 to review the original star rating methodology and presentation of the DFC Star Ratings on the DFC website. The 2015 TEP provided several recommendations that were implemented in the updated DFC Star Rating methodology released in October 2016. A second TEP was convened in 2017. TEP members provided recommendations on candidate measures proposed for inclusion in the DFC Star Ratings. See the respective 2015 and 2017 DFC Star Rating TEP reports for a summary of the deliberations and TEP recommendations. These are available at:

<https://dialysisdata.org/content/esrd-measures>.

CMS developed the DFC Star Rating to help healthcare consumers (including patients and caregivers) understand CMS quality measures and more easily identify differences in overall quality when selecting dialysis facilities, as part of CMS' broader initiative for all of the Medicare Compare sites to make quality information more accessible to patients, caregivers, providers and policymakers.

The Medicare DFC website displays two star ratings: (1) the Quality of Patient Care Star Rating (DFC Star Rating) and (2) the Survey of Patients' Experiences Star Rating. Eleven of the DFC quality measures currently reported on the Medicare DFC website are used to calculate the Quality of Patient Care Star Rating (DFC Star Rating). Six In-center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) Survey Measures are used to calculate the separate Survey of Patients' Experiences Star Ratings. The Survey of Patients' Experiences Star Ratings will not be covered as a discussion topic during this TEP.

This TEP will be expected to review and provide input on options for resetting the DFC Star Rating. This is in anticipation of the need to recalibrate the DFC Star Rating distribution in the near future to optimize the utility of the DFC Star Ratings for patients and other consumers. Recent observed trends in national performance demonstrate progressive shifts in the DFC Star Rating results over the last three years. This may obscure underlying performance differences if many facilities become concentrated at one end of the distribution. As background, prior DFC Star Rating TEPs (referenced above) identified a strong consumer interest in the ability to follow trends in dialysis facility performance over time. In addition, TEP discussions have considered how a reporting approach, that reported longitudinal facility performance trends, could be reset from time to time if the star rating system lost the ability to show meaningful differences over the range of facility performance.

The TEP will be expected to represent a diversity of perspectives and backgrounds. Members will be selected based on their personal experience as patients, caregivers and providers, or based on methodological expertise. Given that the audience for the DFC Star Ratings is primarily patients, the TEP will have ample representation from patients and patient advocates.

We anticipate that the in-person meeting will take place over one day. There will also be additional pre- and follow-up teleconference calls. TEP members are expected to attend all meetings.

Project Objectives:

The University of Michigan Kidney Epidemiology and Cost Center (UM-KECC), through its contract with the Centers for Medicare and Medicaid Services (CMS), will convene a technical expert panel to obtain recommendations on options for resetting DFC Star Ratings distribution. Input from the TEP will inform the development of a methodology for resetting the DFC Star Ratings distribution. The final methodology developed is intended to ensure that the DFC Star Ratings to continue to be informative by reflecting meaningful performance differences among facilities.

TEP Objectives:

The 2019 Star Rating TEP will:

1. Provide recommendations on options for resetting the DFC Star Ratings. The TEP recommendations will be used to inform the development of a methodology for resetting the DFC Star Ratings distribution. The final methodology developed is intended to allow the DFC Star Ratings to continue to reflect meaningful performance differences among facilities.

The task listed above is the only discussion topic for this TEP.

Scope of Responsibilities:

The role of each TEP member is to provide advisory input to UM-KECC.

Role of UM-KECC: As the CMS measure development contractor, UM-KECC has a responsibility to support the development and implementation of ESRD quality measures for public reporting, and the development and implementation of the DFC Star Rating methodology. The UM-KECC moderators will work with the TEP chair(s) to ensure the TEP meeting discussions are focused. During discussions, UM-KECC moderators may 1) advise the TEP and chair(s) on the needs and requirements of the CMS contract and the timeline, and 2) provide specific guidance and criteria that must be met with respect to CMS requirements.

Role of TEP chair(s): Prior to the in-person TEP meeting, one or two TEP members are designated as the chair(s) by UM-KECC and CMS. The TEP chair(s) are responsible, in partnership with the moderator, for directing the TEP to meet the objectives of the TEP.

Duties and Role of TEP members: As defined by CMS in the Measure Management System Blueprint, TEPs are advisory to the measure contractor. In this advisory role, the primary duty of the TEP is to review the TEP supporting materials, and provide recommendations to UM-KECC regarding the resetting of the DFC Star Ratings Distribution.

In May and June 2019, TEP members will be expected to attend pre-TEP conference calls as necessary; and attend a one-day in-person meeting that will take place in June 2019 (date to be finalized at a later date) in Baltimore, MD; and attend additional follow-up teleconference meeting and provide follow-up written feedback and comments as needed (via e-mail).

The TEP will review, edit and adopt a final charter at the first teleconference. The first teleconference will focus on the overall tasks and goals/objectives of the TEP.

During the in-person meeting, the TEP will review the considerations around resetting the DFC Star Ratings distribution and provide recommendations on how to reset the baseline distribution. The key deliverable of the TEP in-person meeting includes a summary report documenting the discussions and proposed recommendations that are made during the in-person meeting.

At the end of the in-person meeting the TEP chair(s) and TEP members will present proposed recommendations. Subsequent to the in-person meeting, the TEP chair(s) will have additional contact with UM-KECC moderators to work through further discussion of the proposed recommendations. After the in-person meeting, approximately between June 2019 and August 2019, TEP members will be asked to review and provide input on a summary report of the TEP meeting discussions and other necessary supporting documents.

Guiding Principles:

Potential TEP members must be aware that:

- Participation on the Technical Expert Panel is voluntary
- Discussion will be recorded during the in-person meeting and tele-conferences
- Proceedings of the in-person meeting will be summarized in a report that is disclosed to the general public
- Patient TEP participants can keep their names confidential, if they elect to do so. Patient TEP participants may elect to remain anonymous in all TEP proceedings. They should notify UM-KECC if they choose to have their names omitted from the TEP roster, in-person meetings, and all meeting minutes.
- If a TEP member has chosen to disclose private, personal data, that material and those communications are not covered by patient-provider confidentiality
- All questions about confidentiality will be answered by the TEP organizers
- All potential TEP members must disclose any current and past activities that may pose a potential conflict of interest for performing the tasks required of the TEP
- All potential TEP members must commit to the expected time frame outlined for the TEP
- All issues included in the TEP summary report will be voted on by the TEP members
- The TEP summary report will include the results of TEP votes taken for specific recommendations
- Numerical voting results and written opinions of the TEP members will be included, if requested

Estimated Number and Frequency of Meetings:

- TEP members should expect to attend one or two 2-hour teleconference calls prior to the in-person meeting
- The in-person meeting will be held for one-day in June 2019 in Baltimore, MD. The exact meeting date will be finalized after the call for nominations period
- After the in-person meeting, additional conference calls (2-5 calls) may be needed

Date Approved by TEP: TBD

TEP Membership: TBD

Appendix C. DFC Star Rating TEP Teleconference Call #1 Minutes

Kidney Disease Quality Measure Development, Maintenance, and Support Project End Stage Renal Disease (ESRD) Dialysis Facility Compare (DFC) Star Rating Technical Expert Panel (TEP) Teleconference Call #1 Minutes

May 6, 2019, 3:00 pm – 5:00 pm EDT

TEP Members*	UM-KECC	CMS
Paul Conway, BA	Yi Li, PhD	Jesse Roach, MD
Catherine Sugar, PhD	Joseph Messana, MD	
Mark Andaya, MS, RN	Claudia Dahlerus, PhD	
Andrew Conkling, BS	Richard Hirth, PhD	
Lorien Dalrymple, MD, MPH	Lan Tong, MPH	
Sharon Dickson, RN, MSN	Brandon Frye, BA	
Dawn Edwards	Jordan Affholter, BA†	
Derek Forfang	Stephen Salerno, MS	
Monica Fox	Wolf Gremel, MS	
Lonnie Green		
Mark Johnson		
Mark Joseph, MD		
Richard Knight, MBA		
Mahesh Krishnan, MD, MPH, MBA		
Michael "Jack" Lennon, MBA		
Nicole Stankus, MD, MSc		
Caprice Vanderkolk, RN, MS, BC-NE		
Curtis Warfield, BS, MS		
David White		

**All TEP members were in attendance; † Jordan Affholter (UM-KECC) is the contact person for the TEP members
UM-KECC = University of Michigan Kidney Epidemiology and Cost Center; CMS = Centers for Medicare & Medicaid Services*

INTRODUCTIONS

Jordan Affholter (UM-KECC) opened the TEP call, noting it was open to the public, would be recorded, and that the last five minutes were set aside for public comments.

Dr. Joseph Messana (UM-KECC) introduced himself Joe Messana and referred to a list of the of the internal workgroup members related to the DFC Star Rating at the University Of Michigan Kidney Epidemiology and Cost Center (UM-KECC). He then welcomed the Star Rating TEP members and thanked them for their willingness to participate. Dr. Messana noted that the administrative contact person is Jordan Affholter and referenced that his email is included in the presentation should any questions arise from TEP members. He then introduced the CMS leads for DFC and ESRD Measures as Golden Horton, MS and Jesse Roach, MD and asked that the CMS leads introduce themselves and provide any opening comments to the TEP.

The CMS project leads introduced themselves Dr. Jesse Roach (ESRD Quality Measures Lead) and Ms. Golden Horton (Dialysis Facility Compare Lead).

Dr. Messina then introduced Mr. Paul Conway and Dr. Catherine Sugar as the TEP Co-chairs and stated that they have both served in this role for the two prior TEP meetings in 2015 and 2017.

Dr. Messina thanked the TEP Co-Chairs and conducted the TEP member ordered roll call. TEP members' names and affiliation are listed in Appendix A. List of TEP Members and Contractor Staff. It was confirmed that the TEP members have had the opportunity to review the charter for this TEP and stated that their voluntary participation is evidence of agreement with and willingness to abide by the scope of the TEP charter. Next, the CMS Measures Blueprint was reviewed noting the role of TEP members, and UM-KECC, in the TEP process. The TEP is being asked to give their best opinions and KECC will do its best to create an accurate and objective record of those opinions to inform the public discussion about any decision that might come out of these TEP proceedings.

TEP PRESENTATION AND DISCUSSION

Dr. Messina provided brief background on the prior two technical expert panels in 2015 and 2017, stating at the 2015 TEP, there was a strong preference from the patient group for absolute grading of facilities rather than relative grading. The DFC Star Rating methodology was modified to accommodate that preference. The 2017 TEP discussed the addition or replacement of measures and the impact of adding new measures to the DFC Star Ratings. There was also preliminary discussion of rebaselining and resetting. The complete slide presentation is provided in Appendix I..

A figure was presented showing the shift in facility DFC Star Rating categories since 2015. There is an overall trend of improvement for those facilities with a longitudinal history. This upward movement has resulted in "bunching up," leading to the discussion of resetting the DFC Star Rating in the form of three questions:

1. Is it time to reset the DFC Star Rating?
2. How should we reset the DFC Star Ratings?
3. How do we help DFC consumers interpret facility performance during and immediately after any transition that would be created by a resetting of the DFC Star Ratings?

Technical terms were defined before moving into the technical portion of the presentation. Additional background was provided on the history of the DFC Star Rating and the notions of rebaselining and resetting. The DFC Star Rating was first released in 2015 and was implemented using a relative rating system. The system has the following 5 components:

1. Standardization of Measures
2. Domain Grouping
3. Domain Score Computation
4. Calculation of Continuous Final Scores

5. Grouping of Final Scores into Five Categories using Rankings

Dr. Yi Li (Professor of Biostatistics) stated that a TEP was convened in 2015 to evaluate the DFC Star Rating methodology. The TEP recommended an absolute cutoff rating system to track changes in performance over time. The new system established a baseline period, in which data are collected to define measure scoring criteria and cutoff values for star categories. The absolute system captures improvement in performance, but there are two challenges. First, a majority of the facilities may end up with 4- and 5-Stars. As a result, it would be difficult to distinguish the facilities performance. Another challenge is that with new measures or updated measures the final scores may not be comparable across years.

Dr. Li stated that the rating system needs to be modified to accommodate new measures or updated measures. To accommodate changes in measures, the 2017 TEP recommended rebaselining. This approach establishes a new baseline period and new scoring cutoffs. The cutoffs are selected to retain the same proportion of facilities in each star category as in the last public release so continuity would be maintained. The fundamental difference between rebaselining and resetting is that rebaselining will maintain the same proportions of star categories as in the last release and resetting will not do so.

Dr. Li further presented a figure which shows the change in the proportion of facilities in each star category over time. In October 2015, 10% of all the facilities were rated 1-Star. In 2016, the percentage dropped to 4.9%, in April 2018 the percentage dropped to 2.9%, and in October 2018 to 2.8%. On the other hand, in October 2015 only 10% of all the facilities reached the 5-Star status, whereas in 2016 the percentage increased to 16.8%. In April 2018 and October 2018, more than 25% of facilities had reached the 5-Star status. More than half of the facilities are currently receiving 4- or 5-Star status.

Dr. Li confirmed that there has been an upward shift in the proportion of facilities receiving 4- or 5-Stars. This does reflect the improvement of the final scores. But, on the other hand, there is almost the same variation in final scores over the years. Even with more facilities receiving higher ratings, heterogeneity in performance still exists. The shift was mainly due to improvement in Domain 3, which consists of Total Kt/V and Hypercalcemia. The improvement in the other domains was more gradual.

One TEP Co-Chair stated that, while the improvement in facility performance over the course of three TEPs is encouraging, there is now a situation in which facilities have the same rating despite the fact that there is substantial variation in facility scores within a star rating category. If all facilities have four or five stars, then it may be challenging for patients to see differences in facility performance. The TEP Co-Chair asked the question of whether or not the DFC Star Rating has reached the point where it is warranted to reset its distribution to differentiate performance, while still maintaining the ability for patients to observe change over time and recognize the positivity of the increases in facility performance thus far.

One TEP Member asked how patients would understand a reset, particularly for a facility that would potentially experience a decrease in rating, and how socio-economic differences would be accounted for.

Dr. Messina stated that the question of how patients would understand a reset is a primary point of discussion for the TEP members at the in-person meeting, particularly if the TEP recommends resetting and

has a reasonable plan for doing so. The TEP is asked to provide input on what information should be made available to patients on DFC in order to facilitate a better understanding of the DFC Star Rating reset.

One TEP Co-Chair stated that it would be useful to present the scores in each domain that were typical of a 5-Star facility previously and the scores in each domain that are typical of a 5-Star facility after a reset. This information would be important for patients to understand both relative and absolute changes in performance trends and to understand the implications of the reset on current quality measure standards.

Dr. Messina asked if there were any questions from the TEP members at this point in the teleconference and stated that there would be plenty of opportunity for TEP members to ask clarifying questions on the background and technical material presented both on the calls and at the in-person meeting.

One TEP Member asked, when considering rebaselining, if this must be done within the rules of the pre-defined measure domains, or if re-allocation could be performed within domains.

Dr. Messina stated that this presents both a technical and a policy question. Historically, when the DFC Star Rating was first developed, one of the criteria that [UM-KECC] was asked to follow was to utilize all the measures presented on the DFC. Over time, the measure set on DFC and/or the DFC Star Rating have changed. The question was asked as to whether the TEP would be able to discuss which measures have achieved their maximum incentives at this point and whether reweighting individual measures was on the table. Arbitrarily moving measures to different domains is a technical question for Dr. Li.

Dr. Li stated that, given an established measure set, factor analysis will be conducted to determine the structure of the domains and whether the domains will change as a result of updates to the measures.

Dr. Messina clarified that the assignment of measures to domains was done using factor analysis (a statistical method). Going outside of that technical approach and assigning domains based on face-validity might be outside the charge of the current TEP, but alternative weighting is something that will be addressed and left to the TEP to decide whether it is worth considering.

One TEP Co-Chair clarified that certain measures have been driving improvements in performance and it would be worth considering if these measures have reached their “ceiling.” Removing these measures could be viewed as a rebaselining, however, this would not address resetting the distribution. Even if measure removal and rebaselining is deemed appropriate, this may not be enough to fix the issue of “bunching up.” They asked if UM-KECC had performed any calculations to determine the impact of removing certain measures from the DFC Star Rating.

Dr. Li confirmed that these calculations had been carried out and would be discussed on the next call and agreed that simply rebaselining would not be sufficient to address the issue of “bunching up.”

One TEP Member asked if there was any requirement to use measures present in the Quality Incentive Program (QIP), as most of the measures in the DFC Star Rating are QIP measures.

Dr. Messina clarified that DFC and QIP are separate programs, although CMS has made efforts to harmonize the two programs in terms of the measures or measures categories. There are no requirements to utilize QIP measures. The DFC Star Rating are limited to using measures that are available on DFC. That is the only program the TEP has the ability to advise on. It is beyond the scope of this TEP to develop new measures or domains for DFC, and the DFC Star Rating is limited to only the measures currently on DFC.

One TEP Member asked if the domains carried equal weight in all historic DFC releases and if results would be presented from calculations as a result of reweighting domains.

Dr. Li clarified that the domains have historically carried equal weight and stated that results would be presented for down-weighting the third domain and other strategies for reweighting the domains.

Dr. Messina stated that the information presented during the remainder of this teleconference was meant to motivate future discussion at the next teleconference and the in-person meeting.

Dr. Messina stated that after the 2017 TEP, CMS formulated the following policy. This was informed, in part, by the 2017 TEP discussions and was presented during CMS' October 2017 National Provider Call:

1. The DFC Star Rating distribution will be evaluated once 3 years have passed since the last reset
2. The DFC Star Rating would be evaluated for a reset when $\leq 15\%$ of facilities receive 1- or 2-Stars
3. A resetting of the DFC Star Rating distribution will also include the establishment of a new baseline

Dr. Messina stated that updates to the scoring cutoffs and Star Rating distribution follow the criteria:

1. When the ability to differentiate facility-level performance is reduced
2. Resetting defines new baseline scoring cutoffs for facilities to be rated
3. Proportions of facilities in each Star Rating category are reset, creating a new distribution

CLOSING REMARKS

Dr. Messina stated that recommendations for resetting will be a primary focus of the TEP and will be covered in greater detail on the next teleconference call. To improve usability of DFC Star Rating during the reset transition, development of tools to assist DFC users' interpretation should be considered. He asked the TEP to consider the following questions in preparation for the in-person meeting:

1. What additional information or display options tools can help users interpret star ratings in the transition after resetting?
2. Please consider potential DFC display options or other tools that could assist patients with interpretation during and after DFC Star Ratings resetting

Jordan Affholter stated that the second pre-TEP teleconference call would be scheduled for May 24, 2019 from 3:00-5:00 pm, EST, and the in-person TEP meeting would be June 6, 2019 from 8:30 AM-4:00 pm, EST.

Dr. Messina thanked the TEP members for their participation and stated that UM-KECC would begin soliciting public comments early and that TEP members could sign off after any initial public comments, but the call line would remain open for the designated time of the call to allow for additional public comment.

PUBLIC COMMENTS

No public comments were received during this teleconference call.

Appendix D. DFC Star Rating TEP Teleconference Call #2 Minutes

Kidney Disease Quality Measure Development, Maintenance, and Support Project End Stage Renal Disease (ESRD) Dialysis Facility Compare (DFC) Star Rating Technical Expert Panel (TEP) Teleconference Call #2 Minutes

May 24, 2019, 3:00 pm – 5:00 pm EDT

TEP Members*	UM-KECC	CMS
Paul Conway, BA	Yi Li, PhD	Jesse Roach, MD
Catherine Sugar, PhD	Joseph Messana, MD	
Mark Andaya, MS, RN	Claudia Dahlerus, PhD	
Andrew Conkling, BS	Richard Hirth, PhD	
Lorien Dalrymple, MD, MPH	Lan Tong, MPH	
Sharon Dickson, RN, MSN	Brandon Frye, BA	
Dawn Edwards	Jordan Affholter, BA†	
Derek Forfang	Stephen Salerno, MS	
Monica Fox	Wolf Gremel, MS	
Lonnie Green		
Mark Johnson		
Mark Joseph, MD		
Richard Knight, MBA		
Mahesh Krishnan, MD, MPH, MBA		
Michael "Jack" Lennon, MBA		
Nicole Stankus, MD, MSc		
Caprice Vanderkolk, RN, MS, BC-NE		
Curtis Warfield, BS, MS		
David White		

**All TEP members were in attendance; † Jordan Affholter (UM-KECC) is the contact person for the TEP members
UM-KECC = University of Michigan Kidney Epidemiology and Cost Center; CMS = Centers for Medicare & Medicaid Services*

INTRODUCTIONS

Jordan Affholter (UM-KECC) stated the TEP call was open to the public, being recorded, and that the last five minutes were set aside for public comments.

Dr. Joseph Messana (UM-KECC) welcomed everyone to the call and thanked the TEP members. He then stated that the main topic for the call would be the shift in the DFC Star Rating distribution, but in response to the TEP, UM-KECC added information on measure reweighting. He then conducted an ordered roll call.

DFC STAR RATING TEP PRESENTATION AND DISCUSSION

Dr. Messana presented three questions for the TEP to consider throughout this TEP process: (1) is it time to reset the Star Ratings? (2) When it is time to reset, how should we do it? (3) How do we help DFC consumers interpret facility performance on star ratings during and immediately after the transition?

Dr. Messina then provided an introduction to rebaselining and resetting. After the 2015 TEP, the DFC Star Rating was changed to allow DFC users to follow changes in facility performance from year to year (see Appendix F. DFC Star Rating TEP Supporting Materials List for the DFC Star Rating Technical Notes link). The individual measures scores, domain scores, and summary score for each facility are calculated for each year and a DFC Star Rating is calculated based on score cutoff thresholds from the original baseline year, which was calendar year 2014 or the data from the October 2015 DFC release. This was a recommendation of the original 2015 DFC Star Rating TEP (see Appendix F. DFC Star Rating TEP Supporting Materials List for the 2015 DFC Star Rating Summary Report link). For years with major changes to the quality measure definitions or set of measures included in the DFC Star Rating, UM-KECC has to determine if the addition of different measures influenced the facility star ratings. If the DFC Star Rating is influenced, then it will be necessary to “rebaseline,” so that DFC users can more easily interpret the star rating categories assigned to facilities. This allows DFC users to continue to compare year to year changes in overall facility scores despite the change in measure definitions. Resetting is needed if a large percentage of facilities are rated 4- or 5-Stars. Resetting determines a new distribution of ratings to report the full range of performance on measures included in the DFC Star Rating, with the goal of providing more information to DFC users. Dr. Messina then introduced Dr. Yi Li (UM-KECC), a professor of biostatistics.

Dr. Li reported that since 2015, there has been an upward shift in the proportion of facilities receiving 4- or 5-Stars. The DFC Star Ratings were rebaselined in April 2018 due to changes in the measure set. Since rebaselining, the rate of increase of 4- and 5-Stars has slowed. 54% of facilities currently receive 4- or 5-Stars. Dr. Li then presented the options for how to reset the DFC Star Rating distribution:

1. Pre-specify the percent of facilities in each category and determine cutoffs from these proportions
2. Pre-specify the cutoffs based on standard deviations from the average national score

Dr. Li explained that a fixed approach was used in the first DFC Star Rating release. Clustering methods are an alternative approach, by which facilities that are more similar to each other are grouped:

1. Hierarchical: Begin grouping facilities two-at-a-time, then combine groups, based on how close their final scores are until there are 5 groups
2. K-Means: Create 5 groups of facilities by minimizing the difference from the average score in each group and maximizing the difference from the average scores in other groups

Dr. Li stated that UM-KECC would be seeking recommendations from TEP members on other resetting options. He then presented some limitations of clustering (see Appendix J. for the full slide presentation). He stated clustering is most appropriate for grouping facilities based on measures with natural gaps. He then asked if any TEP member had questions at this point in the presentation.

One TEP Co-Chair asked about the practical impact of using the fixed distribution versus clustering options.

Dr. Messina stated that in adopting the decision to set absolute scoring thresholds allowed DFC users to follow longitudinal trends in facility performance. The downside is that statistically average facilities are

now falling into above average categories and DFC users are losing information. The practical impact of a reset would be to return to a distribution that includes more information about facility performance. There are differences in facility performance within the higher categories, and resetting would allow for better differentiation in facility performance. Regardless of using the fixed distribution versus the clustering options, the net result would be a reset to a distribution that is similar to the October 2015 DFC release.

The TEP Co-Chairs clarified that the difference between using a fixed distribution versus clustering options for resetting the DFC Star Rating is how to set the cutoffs and then opened the discussion to the TEP.

One TEP Member asked how a reset is received at the facility level. The TEP member suggested that historic ratings could be displayed over time to help explain the impact of a DFC Star Rating reset.

Dr. Messina clarified that facility improvement has continued over the past few years, prompting the need for a TEP discussion on resetting the DFC Star Rating. Continued improvement creates a situation where it is difficult to differentiate between extraordinary facilities and those not performing as well. The October 2018 DFC release featured rebaselining due to changes to the measure set. Resetting has not yet occurred.

One TEP Member asked if rebaselining or resetting accounts for regional or geographic variation.

Dr. Messina clarified that, per CMS, the DFC Star Rating is meant to provide a national comparison.

One TEP Member asked how often clinics are evaluated. The TEP member further asked about what factors are accounted for when comparing facilities, in particular variability in vascular access and Kt/V.

Dr. Messina clarified that a majority of the clinical measures on DFC are used in the calculation of the DFC Star Rating, including many intermediate measures. Factor analysis is used to group highly correlated measures into domains of care. Final scores are calculated by averaging these scores from these domains. Factor analysis is a data-driven approach that shows how related the measures are to each other.

One TEP Co-Chair clarified that part of the question being asked was whether or not differences in individual patient characteristics are adjusted for in the calculation of the DFC Star Rating.

One TEP Member asked if, additionally, socio-economic differences were considered in the DFC Star Rating.

Dr. Messina clarified that the DFC Star Rating is a balanced summary of the clinical measures reported on DFC. Many individual measures are extensively risk-adjusted for characteristics like socio-demographics and incident/prevalent comorbidities. Kt/V and Hypercalcemia are not risk adjusted, but these measures have reserve status with the National Quality Forum and most facilities have high achievement on them.

One TEP Member asked how frequently facilities were rated.

Dr. Messina stated that the DFC Star Rating is refreshed annually in October, except in extenuating circumstances such as a natural disaster, which affects facilities. There is a preview period in July for facilities to review these data. Some measures are updated quarterly, while others are updated annually.

One TEP Member asked if the increase in facility performance was due to the DFC Star Rating, or if there were other factors incentivizing increased quality at the facility level.

Dr. Messina clarified that there is a correlation between improved performance and the implementation of the DFC Star Rating, but that definitive statements about causality could not be drawn. There is no way of assigning credit to any one factor, and there may be several intrinsic factors for improving patient care.

The TEP Member asked, given the above discussion, who the DFC Star Rating should be speaking to.

One TEP Co-Chair stated that the intent of the DFC Star Rating was to provide a tool to dialysis patients to help make informed decisions about where they receive care.

One TEP Member asked where cutoffs would be established using the standard deviation approach, if there is a precedent for resetting the other Compare site star ratings and if so, what were consumer reactions.

Dr. Messina clarified that he was not familiar with all the Compare sites, but the Nursing Home Compare Star Ratings were reset due to rapid dichotomization of stars. He was not aware of how this was received.

One TEP Member asked how socio-demographic or psycho-social factors that vary regionally were being addressed in the DFC Star Rating and how facilities with disparate populations were being accounted for.

Dr. Messina asked that discussion of regional or socio-demographic variation be tabled for the in-person meeting. These points had been previously adjudicated in the conception of the DFC Star Rating and the purpose of this TEP is not to re-examine previous policy decisions. UM-KECC would provide material on regional variation and risk-adjustment for the in-person TEP discussion.

Dr. Li stated that during the in-person TEP, he will also present potential DFC displays and other tools to assist patient interpretation during or after a reset. The TEP will also discuss options for weighting domains. The goal of grouping measures in domains is to avoid domination of the DFC Star Rating by a single measure. The third domain is driving the shift in facility performance. Reweighting domains may lessen the impact of one domain and stabilize the change in distribution when some domains reach top-performance levels:

Option 1: Weight based on the proportion of variance in the data explained by the domains:

- a. Weight domains based on proportion of variance explained for the data
- b. Weight the individual measures based on their contributions to each domain

Option 2: Down-weight Domain 3:

- a. Down-weight Domain 3, containing the two measures with historically highest achievement
- b. Consider removal of the Domain 3 measures (Total Kt/V and Hypercalcemia)

Dr. Li presented a table which showed that weighting the domains adjusts the influence of measures on the final rating calculated for each facility. He stated that weight of measures and/or domains can be done empirically or based on expert opinion and can be applied independently or simultaneously with resetting. Individual facility changes in ratings are expected with re-weighting. Weighting can reduce or eliminate the impact of measures that have very high achievement, allowing facilities to concentrate on other domains.

One TEP Co-Chair asked if the marginal distributions of DFC Star Rating for the table presenting the impact of re-weighting were fixed by design. They asked if results could be provided not fixing the margins, in order to show the impact of reweighting on the longitudinal progression of the ratings.

Dr. Li clarified that the marginal distributions were fixed by design to illustrate the concept of re-weighting and how it could impact individual facilities' ratings. He stated that the longitudinal results of which would be discussed at the in-person TEP meeting.

Dr. Richard Hirth (UM-KECC) clarified that the table shows which facilities received which rating and how that would differ under re-weighting. There are two takeaways: (1) only one facility had a 2-Star difference in rating, and (2) facilities doing equally well in all domains would not experience a different rating under, facilities performing better in the first two domains would experience a higher rating, and facilities that would experience a lower rating were the facilities whose performances were driven by Domain 3.

One TEP Member asked if the observed influence of Kt/V and Hypercalcemia is due to the performance being compared to standards established in 2014 and if this same level of influence would still be observed since these measures now have high achievement. They also asked if high-performing measures should be included in the DFC Star Ratings and what the future implications are of resetting the DFC Star Rating.

Dr. Messina clarified that this is why the TEP has been asked to consider re-weighting or removing Domain 3 from the DFC Star Rating. Resetting and reweighting are different, complementary, tools for re-calibrating the DFC Star Rating. The TEP will be asked to consider a combination of resetting and reweighting.

One TEP Co-Chair stated that measures with high performance do not provide information to differentiate between facilities. Such measures should be removed from the DFC Star Rating. However, if there is a measure in which most facilities perform well, but a subset of facilities do not, this measure still provides information for discriminating performance. Considering both resetting and reweighting is necessary.

Dr. Li stated that one option for reweighting is to weight domains based on the proportion of variance explained. Since there is less variation in Kt/V and Hypercalcemia, these measures contribution would be

less than other measure than have greater variability in performance.

The TEP Co-Chair stated if a measure has poor performance for all facilities, then there is no differentiation but there is information to be gained. The DFC Star Rating needs to convey both absolute and relative information. Weighting by variance would maximize the ability to differentiate facility performance.

One TEP Member asked if there was any indication of what the reset baseline year would be.

Dr. Messina stated that it would not be possible to answer this question until after the TEP deliberations and after CMS has had the opportunity to discuss the TEP recommendations and potential policy decisions.

One TEP Member asked about weighting considerations with respect to regional performance differences.

One TEP Member asked what the data delay is between the period of collection and the annual DFC refresh.

Dr. Messina stated that DFC Star Ratings are released for public use in October, eight to nine months after the December 31st close of the evaluation period in the prior year for that DFC release. Data are near-complete in the spring of the following year. All scores are made available to facilities in July to facilitate questions or concerns about specific data issues. Calculations are finalized and made public in October.

One TEP Member asked if there would be opportunity to discuss the influence of individual measures on the DFC Star Rating, specifically in the context of measures that are difficult to collect and report.

Dr. Messina clarified that the TEP has been tasked with considering reduction in weight or exclusions of domains in the DFC Star Ratings. It is beyond the charge of the TEP to adjudicate the content of DFC. There would be an opportunity to discuss other issues at the end of the in-person meeting.

CLOSING REMARKS

Dr. Messina stated that UM-KECC would begin soliciting public comments early and that TEP members could sign off after any initial public comments. He then thanked the TEP members for their participation on the teleconference and stated that he was looking forward to the in-person discussion.

PUBLIC COMMENTS

No public comments were received during this teleconference call.

Appendix E. DFC Star Rating Post-TEP Teleconference Call Minutes

Kidney Disease Quality Measure Development, Maintenance, and Support Project End Stage Renal Disease (ESRD) Dialysis Facility Compare (DFC) Star Rating Technical Expert Panel (TEP) Post TEP Teleconference Call Minutes

August 21, 2019, 1:00 pm – 3:00 pm ET

TEP Members	UM-KECC	CMS
Paul Conway, BA	Yi Li, PhD	Golden Horton, MS
Catherine Sugar, PhD	Joseph Messana, MD	
Mark Andaya, MS, RN	Claudia Dahlerus, PhD	
Andrew Conkling, BS	Peisong Han, PhD	
Lorien Dalrymple, MD, MPH	Jennifer Sardone, PMP	
Sharon Dickson, RN, MSN	Lan Tong, MPH	
Dawn Edwards	Jingya Gao, MS	
Derek Forfang	Jordan Affholter, BA	
Lonnie Green		
Mark Johnson		
Richard Knight, MBA		
Mahesh Krishnan, MD, MPH, MBA		
Michael “Jack” Lennon, MBA		
Nicole Stankus, MD, MSc		
Caprice Vanderkolk, RN, MS, BC-NE		
Curtis Warfield, BS, MS		
David White		
	Unable to Attend	
Unable to Attend	Richard Hirth, PhD	Unable to Attend
Mark Joseph, MD	Karen Wisniewski, MPH	Joel Andress, PhD
Monica Fox	Wolf Gremel, MS	Jesse Roach, MD

† Jordan Affholter (UM-KECC) is the contact person for the TEP members

UM-KECC: University of Michigan Kidney Epidemiology and Cost Center; CMS: Centers for Medicare & Medicaid Services

INTRODUCTIONS

Jordan Affholter (UM-KECC) opened the technical expert panel (TEP) call, noting it was open to the public, would be recorded, and that the last five minutes were set aside for public comments.

Dr. Joseph Messana (UM-KECC) thanked TEP members for joining the call. Dr. Messana explained that the focus of the teleconference call discussion would be on the options for weighting Domain 3 of the Dialysis Facility Compare (DFC) Star Rating. While TEP members provided individual perspectives on reweighting Domain 3, an official TEP recommendation was not reached during the in-person TEP meeting. The post-TEP call was convened in order to allow for further TEP discussion and to receive a final TEP recommendation on this topic. Dr. Messana conducted the TEP member ordered roll call.

TEP PRESENTATION

Dr. Messina provided a brief summary of the previous TEP discussion on weighting and a presentation of the weighting options for Domain 3 in the DFC Star Rating. The post-TEP Call presentation slides are provided in Appendix L.

Dr. Messina stated that TEP members will be provided with a voting form for weighting Domain 3 in the DFC Star Rating calculation. He stated that the vote would be private. The results will be anonymized and only UM-KECC will know how specific TEP members voted. Domain 3 contains the Kt/V and Hypercalcemia measures and the four options to vote on are:

- Option 1 - Reweight Domain 3 at 0% (Remove Domain 3 from the DFC Star Rating Calculation)
- Option 2 - Reweight Domain 3 at 50% (Down-weight Domain 3 at 50% of its current weight)
- Option 3 - Weight Domain 3 at 100% (Maintain Domain 3 at its current weight)
- Option 4 - None of the above (if TEP members choose this option, then they will provide a recommended percentage for weighting Domain 3 in the comment box).

Dr. Messina handed the meeting over to the TEP co-chairs to lead a discussion of the weighting options. UM-KECC and the TEP co-chair clarified that the purpose of this call is for the TEP to provide recommendations on how to weight Domain 3. One TEP co-chair stated that there are two potential reasons for reweighting Domain 3: (1) to control how much importance is assigned to Domain 3 in the DFC Star Ratings, and (2) to limit the ability for Domain 3 to drive changes in the DFC Star Ratings. In response to the TEP Co-Chair, Dr. Yi Li (UM-KECC) explained that the Kt/V and Hypercalcemia measures are highly skewed and have experienced high achievement. The Domain 3 measures have relatively small variation, especially in comparison to the Domain 1 and 2 measures. The TEP co-chair further stated that if the ability of facilities to improve their Domain 3 performance is maxed out, it is less likely the measures in Domain 3 will drive improvement going forward in the DFC Star Ratings.

TEP DISCUSSION AND TECHNICAL CLARIFICATIONS

One TEP member asked what would happen if the measures scores in Domain 3 start to decline in the future. Dr. Messina stated that, even if a measure is not included in the DFC Star Ratings, they may continue to be monitored (i.e., they will remain on DFC).

In response to comparisons made to how measures are handled in the Quality Incentive Program (QIP), Dr. Messina reminded TEP members that while the TEP members could make general recommendations about other programs, the QIP is a separate CMS program from DFC and decisions are made by a separate group (at CMS). Dr. Messina stated this current TEP was convened to provide recommendations only on the DFC Star Ratings.

One TEP member asked if down-weighting Domain 3 would result in the ability to add other measures such as the transplant measures to the DFC Star Rating. Dr. Messina clarified that CMS announced during the most recent National Provider Call that there are no plans to add new measures to the DFC Star Ratings at this time. Down-weighting Domain 3 in the context of this TEP recommendation would result in redistributing weight to Domains 1 and 2.

In response to the questions about the current weighting approach, one TEP co-chair explained that all domains are weighted equally, and within each domain individual measures are weighted equally. For example, Total Kt/V and Hypercalcemia both have a 50% contribution to a facility's Domain 3 score. Assigning less weight to Domain 3 would result in greater weight to be placed on Domains 1 and 2.

UM-KECC and the TEP co-chairs clarified that if the Domain 3 measures are recommended for removal from the DFC Star Ratings that does not necessarily result in the Domain 3 measures being removed from public reporting on the Dialysis Facility Compare (DFC) website.

Dr. Messina clarified that Domain 3 has driven the upward shift in the DFC Star Ratings from 2015 to 2018. One TEP co-chair explained that a future re-setting of the DFC Star Rating distribution would shift the star distribution back towards the center. Dr. Li added that down-weighting Domain 3 would not address the issue of the DFC Star Rating shifting towards a higher proportion of 4- and 5-stars, but may assist in preventing future rapid shifts. Down-weighting Domain 3 may assist in limiting how much influence the Domain 3 measures have on the calculation of a facility's final score.

One TEP member asked what year of data would be used to reset the baseline. Dr. Messina (UM-KECC) stated that the baseline year for a future reset has not yet been determined. The final considerations of the baseline year is a CMS policy decision. Dr. Messina stated that the baseline year is more likely to be based on more recent data than older data. Dr. Messina explained that the most recent preliminary data shows very little change in performance on the Domain 3 measures. The rapid shift in facility performance on the Domain 3 measures (that occurred from 2015-2018) does not appear to be continuing.

The TEP co-chair confirmed that if Domain 3 is down-weighted by 50%, that would result in Domains 1, 2, and 3 accounting for 40%, 40% and 20%, respectively, in the final score of the Star Rating calculation. In response to a question whether Domains 1 and 2 could be assigned different weights, Dr. Messina clarified that TEP is only being asked to provide a recommendation for Domain 3's weighting, but TEP members could provide additional comments/recommendations in the comment box of the voting form.

TEP DISCUSSION AND RECOMMENDATIONS

Several TEP members recommended that the Domain 3 measures should be down-weighted. Reasons they provided included: (1) allowing for differentiation among facility performance, (2) limiting the amount of influence the Domain 3 measures have, (3) keeping the measure in the system for patients, and (4) allowing for greater focus on the measures in Domains 1 and 2 (which include important measures such as hospitalization, mortality; and vascular access).

Several TEP members emphasized the importance that the DFC site be useful for patients. They stated that down-weighting the Domain 3 measures does not reduce the value an individual patient would place on the Kt/V or Hypercalcemia measures, and down-weighting would only change the way the DFC Star Ratings would be calculated. Several TEP members supported the notion that, regardless of their weight, the Kt/V and Hypercalcemia measures should remain on the DFC site.

One provider TEP member referenced a previous TEP discussion, where patients interpreted Kt/V as related to quality of life. The TEP members suggested this may be one reason to keep it in the DFC Star Rating. The TEP member further stated the literature does not provide much support for the Kt/V measure improving patient outcomes. The TEP member stated that the Hypercalcemia measure is more informative for future events than the Kt/V measure.

One patient TEP member stated that the Domain 3 measures (Kt/V and Hypercalcemia) should not be down-weighted and referenced that these measures are very important to patients. The TEP member stated that the Kt/V and Hypercalcemia values are the important values in relation to patient longevity and evaluating treatment effectiveness. The TEP co-chair supported the statement that the Kt/V and Hypercalcemia measures are important to patients. The TEP co-chair clarified that the focus of this TEP call is how to weight these measures in the DFC Star Rating system in order to distinguish between facility performance in the DFC Star Ratings.

One provider TEP member stated that the Domain 1 measures (SHR, SMR, SRR, and STeR) carry more importance than the Domain 3 measures. The TEP member stated their opinion that the Kt/V measure is not very stable and that facilities may take repeated measurements in order to obtain a good Kt/V value. The TEP member stated that almost all dialysis clinics are prescribing enough dialysis. Kt/V may be considered overly important based on old research studies. The TEP member stated that Kt/V is easy to influence so the majority of facilities can influence the value. If a facility does not score well on Kt/V, it may be due to patient populations. Another provider TEP member explained that facilities score very high on both Kt/V and Hypercalcemia measures. The TEP member explained if a facility does not score well on the Domain 3 measures, it is often due to data reporting issues resulting in missing values.

Dr. Messina suggested that Kt/V and Hypercalcemia may also be functioning as a proxy for overall facility processes of care. He suggested that, in his clinical experience, both measures rely on facility monitoring and care team response. Dr. Messina closed the TEP discussion.

CLOSING REMARKS and TEP Voting Instructions

UM-KECC explained that TEP members will be sent a voting form after the teleconference call. TEP members will be asked to provide a recommendation on the weighting options for Domain 3 of the DFC Star Rating. The TEP was asked to vote on one of the following options:

- Option 1 - Reweight Domain 3 at 0% (Remove Domain 3 from the DFC Star Rating Calculation),
- Option 2 - Reweight Domain 3 at 50% (Down-weight Domain 3 at 50% of its current weight),
- Option 3 - Weight Domain 3 at 100% (Maintain Domain 3 at its current weight),
- Option 4 - None of the above (if TEP members choose this option, then they will provide a recommended percentage for weighting Domain 3 in the comment box).

TEP members were asked to email back their vote to Jordan Affholter (UM-KECC) by Friday August 23, 2019. Next, UM-KECC will produce the post-TEP call minutes for the TEP members to review for accuracy. A summary of the TEP weighting options vote would be provided in the post-TEP call minutes document. Once the post-TEP call minutes are finalized, the document will be included in the full TEP summary report.

The final version of TEP summary report will be provided to CMS. CMS will make final decisions on updates to the DFC Star Rating system and when those changes will take place.

Dr. Messina thanked the TEP members for their participation on the call. The call was closed after the public comment period.

PUBLIC COMMENTS

No public comments were received during this teleconference call.

APPENDIX – Summarized Voting Results from the Weighting Options TEP Vote

TEP members were asked to vote on the Domain 3 weighting options. TEP votes were anonymous in order to protect individual TEP member opinions and respect individual anonymity. 19 of 19 TEP members voted. The results from this vote are as follows:

- 0 TEP members (0%) voted for Option 1 - Reweight Domain 3 at 0% (Remove Domain 3 from the DFC Star Rating Calculation)
- 15 TEP members (79%) voted for Option 2 - Reweight Domain 3 at 50% (Down-weight Domain 3 at 50% of its current weight)
- 3 TEP members (16%) voted for Option 3 Option 3 - Weight Domain 3 at 100% (Maintain Domain 3 at its current weight)
- 1 TEP member (5%) voted for Option 4 - None of the above (if TEP members choose this option, then they will provide a recommended percentage for weighting Domain 3 in the comment box).
 - One TEP member voted to reweight Domain 3 to 25%. The TEP member recommended to eventually reweight the individual measures within Domain 3 so that the Kt/V measure would account for 25% of Domain 3's total score and Hypercalcemia would account for 75% of Domain 3's total score.

Appendix F. DFC Star Rating TEP Supporting Materials List

1. Technical Notes on the Updated DFC Star Rating Methodology for the October 2018 Release

This document describes the updated DFC Star Rating methodology as of the October 2018 DFC Star Rating Release. This methodology was informed by the 2017 TEP recommendations and is the methodology that is currently being used for the DFC Star Ratings. The link is as follows:

https://dialysisdata.org/sites/default/files/content/Methodology/Updated_DFC_Star_Rating_Methodology_for_October_2018_Release.pdf

2. DFC Star Rating TEP Summary Report from the 2017 TEP

This document provides a detailed summary of the discussions and recommendations from the 2017 DFC Star Rating Technical Expert Panel. Note that a high level summary of the discussion and recommendations is provided on pages 36-38 of this document. The link is as follows:

https://dialysisdata.org/sites/default/files/content/ESRD_Measures/ESRD_DFC_Star_Ratings_TEP_Summary_Report_2017.pdf

3. DFC Star Rating TEP Summary Report from the 2015 TEP

This document provides a detailed summary of the discussions and recommendations from the 2015 DFC Star Rating Technical Expert Panel. Note that a high level summary of the discussion and recommendations is provided on pages 34 through 37 of this document. The link is as follows:

https://dialysisdata.org/sites/default/files/content/ESRD_Measures/ESRD_DFC_Star_Rating_TEP_Summary_Report_2015.pdf

4. October 2018 CMS National Provider Call Presentation

This document provides the presentation materials for the October 2018 CMS National Provider Call, where CMS presented the 2018 DFC release measures and methodology updates. The link is as follows:

<https://www.cms.gov/Medicare/End-Stage-Renal-Disease/ESRDGeneralInformation/Downloads/October-2018-Dialysis-Facility-Compare-National-Provider-Call-Slide-Deck.pdf>

5. October 2017 CMS National Provider Call Presentation

This document provides the presentation materials for the October 2017 CMS National Provider Call, where CMS presented guidelines for resetting the DFC Star Rating distribution. The link is as follows:

<https://www.cms.gov/Medicare/End-Stage-Renal-Disease/ESRDGeneralInformation/Downloads/October-25-Dialysis-Facility-Compare-National-Provider-Call-Slide-Deck.pdf>

Transcripts and Q&A's for the public CMS National Provider Calls are provided at the following link:

<https://www.cms.gov/Medicare/End-Stage-Renal-Disease/ESRDGeneralInformation/index.html>

Appendix G. TEP Composition Form

The Technical Expert Panel Composition Form is provided on the next several pages.

TECHNICAL EXPERT PANEL COMPOSITION (MEMBERSHIP) LIST

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

Dates:

- Two pre-TEP conference calls will be held on May 6 and May 24, 2019.
- The in-person TEP meeting will be held on June 6, 2019 from 8:30am - 4:00pm EDT in Baltimore, MD

Project Overview:

The Centers for Medicare & Medicaid Services (CMS) has contracted with The University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) to act as quality measure developer and DFC technical content support contractor, under the Kidney Disease Quality Measure Development, Maintenance, and Support contract. The contract number is 75FCMC18D0041, task order number 75FCMC18F0001. As part of this contract, UM-KECC convenes technical expert panels (TEPs) to provide valued consumer and provider input for both quality measure development and the Dialysis Facility Compare (DFC) Quality of Patient Care Star Rating (DFC Star Rating).

These individuals were selected and have agreed to serve on the DFC Quality of Patient Care Star Rating TEP for this project.

Name, Credentials, and Professional Role	Organizational Affiliation, City, State	Consumer Perspective	Clinical Content	Performance Measurement	Coding and Informatics	Conflict of Interest Disclosure
Paul T. Conway, BA <i>TEP co-chair</i>		X				None
Board of Directors Member; Chair of Public Policy and Global Affairs	American Association of Kidney Patients (AAKP) Falls Church, VA					

Name, Credentials, and Professional Role	Organizational Affiliation, City, State	Consumer Perspective	Clinical Content	Performance Measurement	Coding and Informatics	Conflict of Interest Disclosure
Catherine A. Sugar, PhD, MS <i>TEP co-chair</i> Professor Director	Departments of Biostatistics, Statistics & Psychiatry University of California, Los Angeles Semel Institute Statistics Core, University of California, Los Angeles Los Angeles, CA			X		None
Mark Andaya, MS, RN Director of Quality Assurance and Performance Improvement	The Rogosin Institute New York, NY		X	X		None
Andrew Conkling, BS Vice President	Dialysis Patient Citizens (DPC) Arab, AL	X				None
Lorien Dalrymple, MD, MPH Vice President of Epidemiology and Research Volunteer Clinical Faculty Associate Professor	Fresenius Medical Care North America (FMCNA) Waltham, MA Dept. of Medicine, Division of Nephrology, University of California, Davis Davis, CA		X	X		Disclosure: Served as Member of the Kidney Care Quality Alliance (KCQA) Steering Committee and has participated in research related to quality measures.

Name, Credentials, and Professional Role	Organizational Affiliation, City, State	Consumer Perspective	Clinical Content	Performance Measurement	Coding and Informatics	Conflict of Interest Disclosure
Sharon Dickson, RN, MSN Regional Quality Manager	Fresenius Kidney Care Lancaster, SC		X	X		None
Dawn Edwards Patient Advocate Health Ambassador	NxStage New York, NY The Rogosin Institute New York, NY	X				None
Derek Forfang Patient Advocate/Chair Patient Advocate/Chair	Kidney Patient Advisory Council, Forum of ESRD Networks San Pablo, CA Kidney Advocacy Committee, National Kidney Foundation San Pablo, CA	X		X		None
Monica Fox Outreach Associate/Patient Advocate	National Kidney Foundation of Illinois Flossmoor, IL	X				None

Name, Credentials, and Professional Role	Organizational Affiliation, City, State	Consumer Perspective	Clinical Content	Performance Measurement	Coding and Informatics	Conflict of Interest Disclosure
<p>Lonnie Green Patient Advocate/Subject Matter Expert</p> <p>Patient Advocate/Subject Matter Expert</p>	<p>IPRO ESRD South Atlantic Network #6 Kennesaw, GA</p> <p>National Patient and Family Engagement – Learning Action Network (NPFE-LAN) Kennesaw, GA</p>	X			X	None
<p>Mark Johnson Patient Advocate/Network Patient Representative</p> <p>Patient Advocate/Subject Matter Expert</p>	<p>Heartland Patient Advisory Council, ESRD Network 12 Atlantic, IA</p> <p>National Patient and Family Engagement – Learning Action Network (NPFE-LAN) Atlantic, IA</p>	X				None
<p>Mark Joseph, MD Pediatric Nephrologist</p> <p>Medical Director</p>	<p>Pediatric Kidney Disease and Hypertension Centers Phoenix, Arizona</p> <p>Phoenix Pediatric Dialysis Center Phoenix, Arizona</p>		X	X		None

Name, Credentials, and Professional Role	Organizational Affiliation, City, State	Consumer Perspective	Clinical Content	Performance Measurement	Coding and Informatics	Conflict of Interest Disclosure
Richard Knight, MBA President Adjunct College Instructor	American Association of Kidney Patients (AAKP) Bowie, MD Bowie State University Bowie, MD	X				None
Mahesh Krishnan, MD, MPH, MBA Group Vice President of Research and Development	DaVita Kidney Care Washington D.C.		X	X		Disclosure: Employee of DaVita.
Michael "Jack" Lennon, MBA Executive Director	Improving Renal Outcomes Collaborative Cincinnati, OH	X		X		None
Nicole Stankus, MD, MSc Medical Director Associate Professor	DaVita Stony Island Dialysis Center Department of Medicine Section of Nephrology, University of Chicago Chicago, IL		X	X		Disclosure: Medical Director of the DaVita Stony Island Dialysis Center and a member of the DaVita Physician Council.
Caprice Vanderkolk, RN, MS, BC-NE Nurse Manager	Adult and Pediatric Dialysis Program, University of Minnesota/Fairview Hospitals Minneapolis, MN		X	X		None

Name, Credentials, and Professional Role	Organizational Affiliation, City, State	Consumer Perspective	Clinical Content	Performance Measurement	Coding and Informatics	Conflict of Interest Disclosure
<p>Curtis Warfield, BS, MS Patient Reviewer</p> <p>Senior Quality Analyst</p>	<p>National Kidney Foundation Indianapolis, IN</p> <p>State of Indiana Indianapolis, IN</p>	X		X		None
<p>David “Dave” White Patient Advocate/Board of Director Member</p> <p>Healthcare Consultant</p>	<p>American Association of Kidney Patients (AAKP) Hillcrest Heights, MD</p> <p>Quality Insights Renal Network 5 Hillcrest Heights, MD</p>	X				Disclosure: Family members own stock in Amgen, CareDX, and Proteon Therapeutics.

Appendix H. Public Comment Letter from Kidney Care Partners (KCP)

The Public Comment Letter from Kidney Care Partners (KCP) is provided on the next several pages.



June 5, 2019

Dr. Catherine Sugar
Co-Chair
DFC Star Rating TEP
Associate Professor In-Residence
Department of Biostatistic
UCLA Fielding School of Public Health
Los Angeles, CA 90095-1772

Mr. Paul Conway
Co-Chair
DFC Star Rating TEP
Chair of Policy and Global Affairs;
Immediate Past President
American Association of Kidney Patients
14440 Bruce B. Downs Blvd.
Tampa, FL 33613

Re: Dialysis Facility Compare (DFC) Star Rating Technical Expert Panel (TEP)

Kidney Care Partners (KCP) appreciates the continued attention and monitoring of the DFC Star Rating program. In anticipation of the in-person TEP on June 6, KCP wanted to share our previous recommendations related to topped out measures and better aligning the DFC Star Rating program with the statutorily mandated Quality Incentive Program (QIP). Specifically, we recommend that CMS first remove topped out measures in Domain 3, which the contractor has indicated are driving the increasing number of 4 and 5 stars, before resetting the cut points. We hope that the TEP will consider these comments during its upcoming meeting.

KCP is an alliance of members of the kidney care community that includes patient advocates, dialysis care professionals, providers, and manufacturers organized to advance policies that improve the quality of care for individuals with both CKD and irreversible kidney failure, known as ESRD.

I. KCP believes It Is Important to Remove Topped Out before Resetting the Star Rating Cut Points.

As we have noted in the past, KCP continues to believe that the star rating program must meet two basic principles. First, the star ratings should accurately reflect the quality of care provided by dialysis facilities and not be distorted to meet a pre-determined number of facilities in particular star levels. Second, topped out measures should be removed from quality programs; even though such measures could still be reported publicly, they should not be incorporated in the QIP or five star programs.

Given these principles, we do not believe it is appropriate to re-set the cut points for the DFC star ratings to address the increase in facilities receiving 4 or 5 stars. As the data presented to the TEP during the most recent call indicates, Domain 3 (which is comprised of the Kt/V and hypercalcemia measures) is driving the star rating shift not the cut points. Before any discussion about re-setting occurs, we would hope that CMS would either

reduce the weight of Domain 3 in comparison to the other domains or retire the domain entirely based upon its QIP measure factors for evaluating measures.

This approach would be consistent with the factors 1 and 2 in the CMS rulemaking:

- Factor 1. Measure performance among the majority of ESRD facilities is so high and unvarying that clinically meaningful distinctions in improvement or performance can no longer be made (for example, the measure is topped-out).
- Factor 2. Performance or improvement on a measure does not result in better or the intended patient outcomes.

Therefore, we ask that the TEP pause on re-setting the cut points and instead focus on reweighting or retiring the Kt/V and hypercalcemia for the star rating.

II. KCP supports focusing the ESRD QIP on Meaningful Measures and Recommends Streamlining the ESRD QIP and Dialysis Facility Compare (DFC) to Reduce Administrative Burden and Improve Transparency for Patients, Caregivers, and Consumer

We also ask that the TEP continue its efforts to make sure that the star ratings serve the needs of patients and caregivers. To that end, KCP has recommended in previous comment letters about the ESRD QIP that CMS align the QIP and star ratings programs by affirming that the QIP is a pay-for-performance (P4P)/value-based purchasing (VBP) program, which was the intent of the Congress when it established the program, while the DFC is meant to be a public reporting program. Attaching star ratings to DFC confuses consumers and the separate roles of each program. While both are quality accountability programs, DFC is best described as a quality assurance program.

To address this problem and, most importantly, to empower patients and provide them with reliable tools they can use to make decisions about their health care, KCP recommends that CMS separate the programs clearly by using different measures in each program, using the star ratings based on the ESRD QIP penalty distribution, and improving the functionality of the DFC website.

The ESRD QIP would include a parsimonious set of measures consistent with the recommendations in Attachment B. The public reporting certificates required by the statute should be returned to the previous format that includes meaningful information, not just the number that provides patients with no specific information on the measures. If CMS continues to promote star ratings, the stars should be incorporated into the ESRD QIP certificates and be set using the QIP penalty distribution. MedPAC also has supported

eliminating the star ratings on DFC.¹ All measures should be valid, reliable, feasible, and be NQF-endorsed.

The DFC would be a public reporting, quality assurance, program. This return to its purpose would in no way diminish the program; rather it would allow DFC to achieve its intended purpose. Public reporting is considered by NQF and others to be an accountability program, so measures publicly reported should not be viewed as “second class.” Measures that are important, but not in the QIP, could be publically visible in DFC. To improve the patient experience, the DFC website should be upgraded in a way that allows patients and caregivers to understand the site and use it more often. Specifically, we recommend that CMS:

- Allow patients to compare facilities using multiple measures at the same time, consistent with the recommendations of the TEP, rather than the current approach of being able to compare facilities using only one measure. This capability is standard in many online tools.
- Establish a true mobile experience for patients that allows them to use their mobile devices to access the system.
- Engage with all in the kidney care community to encourage its use among stakeholders.

While the DFC would report on measures that are not in the QIP, it could list the QIP measures – using the same specifications, benchmarks, and results. KCP has been discussing with CMS staff that the specifications for the “same” measures, but different programs, do not align, leading we believe to anomalies in penalties and star ratings. Having specifications for the “same” measure that differ based on the program is confusing and unduly burdensome. This would allow patients and caregivers to compare all the measures in one easy place and eliminate the confusing inconsistencies among the programs. Because star ratings are more aligned with the Total Performance Score requirements of the QIP, they should be used for the QIP TPS, while the DFC should provide a more detailed and comprehensive assessment of facilities that can be accessed in a manner that allows users to tailor the results to their individual needs.

Once the purpose of the two programs is clearly delineated, the measures used in each program should be refined. First, all measures in the programs should be valid,

¹MedPAC, Comment Letter to CMS on ESRD PPS CY 2017 (July 2016) (“In our August 15, 2014 comment letter to your predecessor, the Commission questioned why CMS believed it necessary to develop a second quality system for dialysis facilities. We also raised concerns that beneficiaries and their families might be confused if a facility’s star and QIP scores diverge, which could occur because the measurement systems use different methods and measures to calculate a facility’s performance score. The Commission believes the ESRD quality measurement process needs greater simplicity and clarity. Moving to one quality measurement system that is based on a reasonable number of outcomes-based performance measures would be easier to understand for beneficiaries and their families and would reduce administrative costs for providers and CMS.”)

Dr. Sugar
Mr. Conway
June 5, 2019
Page 4 of 14

reliable, and feasible and meet the scientific acceptability criterion for measure endorsement used by the NQF.

Based upon the measures currently included in the two programs and measures under current development, KCP recommends that CMS use the measures in Attachment B for the ESRD QIP, with an important caveat. Specifically, if there is no measure that has been endorsed in the domain or the measure currently being used has been rejected by the NQF, CMS (working with KCP and the kidney care community) should prioritize addressing the problems with the existing measure and refine it or develop a new measure that would meet the NQF criteria and submit the measure to NQF for endorsement. Once endorsed, it would be added to the ESRD QIP or the DFC.

Even with this bifurcated approach, CMS should not simply create more and more measures. In brief, each program should contain a parsimonious set of measures that distinguish performance among facilities, that measure facility actions rather than actions of other providers, and that matter to patients.

III. Conclusion

KCP looks forward to continuing to work with the TEP and CMS as efforts to refine the star ratings continue. Please contact Kathy Lester at klester@lesterhealthlaw.com or 202-534-1773 if you have questions or would like to discuss the specifics of these recommendation in detail.

Sincerely,



Allen Nissenson
Chairman
Kidney Care Partners

cc: Dr. Kate Goodrich
Dr. Michelle Schreiber
Dr. Jesse Roach

Appendix A: Measure Recommendations

Please note for a full discussion about recommendations regarding the specifications of these measures, please KCP's August 2018 ESRD QIP Proposed Rule Comment Letter.

A. QIP Measure Recommendations

In making these recommendations, the KCP spent several months with a cross-sectional work group of our members. All voices of the community were represented – patients, facilities, physicians, nurses, technicians, manufacturers, and suppliers. This group carefully reviewed the reports from the various CMS quality and measure development technical expert panels (TEPs), comments from non-KCP members, recommendations from MedPAC, and the CMS Meaningful Measures Initiative, as well as policies and measures used in other Medicare P4P/VBP programs. As a result of this work, KCP recommends that CMS use the following measures in the ESRD QIP.

- **Standardized hospitalization rate measure**
 - The current ratio measure should be abandoned. A true risk-standardized rate measure should be developed. CMS can start with the current numerator and denominator, and build a valid risk model from there. CMS should eliminate the manipulation of the current ratio, which merely applies a multiplication factor to convert the ratio into a rate.
 - CMS should target the measure to admissions that are within the control of dialysis facilities, focusing on “avoidable” hospitalizations—i.e., avoidable because the measure focuses on reasons for admissions that can be stopped with appropriate medical intervention by the facility. There is no reason to hold dialysis facilities responsible for hospitalizations out of their control, when other providers have more targeted measures. While an all-cause measure may make more sense in a hospital or broader health care setting that treats patients for multiple conditions, dialysis facilities provide a single service – dialysis treatments – and should be held accountable for what they can control. CMS has been testing a similar measure for skilled nursing facilities through its innovation center.²
 - Assessments of standardized ratio measures of hospitalization (as well as mortality and readmission) have demonstrated that such standardized measures are highly imprecise. For example, the standardized hospitalization ratio is estimated so imprecisely that nearly three quarters (74.7 percent) of facilities have confidence intervals that span from the top to the bottom quintiles of overall performance. Put simply, the imprecision makes it

²See <https://innovation.cms.gov/initiatives/rahnfr-phase-two/>.

impossible to determine if an individual facility is among the best or the worst performing facilities. Such consideration would apply equally to standardized 'rates' (the currently reported metrics) which are derived as scaled up versions of their corresponding standardized ratio. A better approach would be to simply develop an actual risk-standardized rate rather than try to convert the existing ratio to a rate.

- KCP continues to recommend development of true risk standardized rates (not the CMS "conversion factor" rates). As we have noted in the preceding bullet, the ratios are highly imprecise and make it impossible to distinguishing quality among facilities. Penalizing facilities based on scores that do not have meaningful differences, as we have just described, is inappropriate.
- **Standardized readmissions rate measure**
 - Like the hospitalization measure, the current readmissions ratio measure should be abandoned and a true risk-standardized rate should be developed, as noted above.
 - Again, as previously noted, CMS should target the measure to re-admissions that are within the control of dialysis facilities and focus the measure on "avoidable" readmissions—*i.e.*, avoidable because the measures focus on reasons for readmissions that can be stopped with appropriate medical intervention by the facility.
 - The concerns with the confidence interval noted above apply here as well.
- **Catheter > 90 Days Clinical Measure**
 - The current catheter > 90 days measure should be maintained as is, but the VAT topic would be eliminated.
 - Clinical consensus is that one of the most important factors in dialysis patient outcomes is the removal of a catheter after 90 days. While the placement of a fistula often is preferred, it is not the medically appropriate choice for all patients, including a fistula and/or graft measure only dilutes the impact of the removal of catheter measure in the TPS. Adopting it alone would appropriately emphasize the importance of removing catheters.
- **Bloodstream infection measures**
 - While KCP supports having a bloodstream infection measure, it needs to meet the scientifically acceptable measure development criteria.

- The two current measures, NHSN Dialysis Event Reporting Measure and Infection Monitoring: National Healthcare Safety Network (NHSN) Bloodstream Infection in Hemodialysis Patients Clinical Measure, should be revised to include a single, valid and reliable BSI outcomes measure.
- As discussed in greater detail in KCP's 2016 and 2017 comment letters and articulated by several members of previous TEPs, the current outcome measure is not valid and has produced errant results. Retaining it provides patients and caregivers with inaccurate information that may lead to medical decisions that are contrary to their goals. The NHSN BSI Measure is inappropriate as a clinical measure because it is not valid, as shown by the measure developer, CDC's et al. own research, and CMS's own data. It is also unreliable for facilities with small census populations. CMS has stated that its review shows that as many as 60-80 percent of dialysis events may be under-reported with the NHSN BSI measure. We have heard during TEP meetings that this amount now might be slightly lower, but even at half this value, it still remains unacceptably high. The high under-reporting rate associated with this measure demonstrates that the measure is simply not a valid measure. A lack of validity means that we cannot be certain that the measure results in scientifically acceptable findings. Making sure that measures are valid in the context of public reporting and value-based purchasing is essential to the success of these programs. Providers are being incentivized to change their behavior to improve the results of the measure. If the measure is not valid, these changes may not be appropriate to implement. In addition, if the measure is not producing valid findings, it does not help patients who are trying to use measures to make informed decisions about their care.
- The Dialysis Event Reporting Measure specifications now incorporate the reporting of several subjectively interpreted signs of infection (*e.g.*, swelling, redness). This expansion of the reporting protocol is highly subjective, burdensome, and does not contribute to the measure's underlying premise—to identify BSIs verified by positive blood cultures. These modifications will not serve the purpose of reducing BSI events.
- CMS should make the development of a valid and reliable measure that meets the NQF endorsement criteria a top priority for its work.
- **Patient Experience of Care: In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) Survey Clinical Measure**
 - This measure should be revised along the lines KCP has previously recommended and as outlined in Appendix A.

- A home dialysis CAHPS and a pediatric CAHPS survey should be established as well.
- With respect to changes in the ICH CAHPS specifications, KCP seeks clarification on the proposed elimination under Additional Information: “Missing data are not included in the calculations. Only data from a ‘completed’ survey are used in the calculations.” If CMS means that the elimination of the completed survey requirement permits the use of data wherein only global ratings questions are answered or only all questions pertaining to a composite are answered, KCP opposes this change. Similarly, if CMS is proposing imputation of missingness, this may be problematic, particularly with how ‘skip questions’ in the ICH CAHPS are handled. KCP recognizes the potential to increase sample size in this manner, but KCP has on multiple occasions proposed administering the instrument by domains in a manner that both reduces burden and maintains the scientific integrity of the testing. An approach that merely increases response rate by accepting answers from incomplete surveys calls into question validity and introduces cherry-picking of questions and domains. Again, KCP supports burden reduction and increased sample sizes, but not at the expense of scientific acceptability.
- **Anemia management measure**
 - The current two measures in this domain Standardized Transfusion Ratio (STrR) Clinical Measure and Anemia Management Reporting Measure should be replaced with a Hgb < 10 g/dL measure. While it will be necessary to develop updated specifications, exclusions, and business rules, CMS has developed a similar measure several years ago that would be an appropriate starting point. We are aware such a measure was not endorsed by NQF, but believe NQF’s updated evidence algorithm provides a path for its consideration anew.
 - A lower hemoglobin measure is preferable as an outcome measure to a reporting measure. Most importantly, this measure is actionable by physicians and will have a direct and positive impact on an issue of critical importance to patients.
 - It also is a better measure than the STrR because facilities and physicians have access to patient hemoglobin data in the facility, whereas they do not have access to STrR data. ; moreover, it is actionable by physicians and will have a direct a positive impact on an issue of critical import to patients. Additionally, we have identified a significant validity issue with the STrR data since the ICD-9 to ICD-10 conversion. KCP has historically been concerned about under-counting and has documented that different coding practices for

transfusions leads to under-reporting (Appendix C). Put simply the STTrR's validity is in further question due to increased under-reporting by hospitals after the switch to ICD-10.

- Overall, we have found that for the STTrR measure, 545 of 4,541 of hospitals (12.0 percent) had an estimated reduction in transfusion coding >80 percent after the ICD-10 conversion, and 979 of 4,541 hospitals (21.6 percent) had an estimated reduction in transfusion coding >50 percent. As the technical appendix documents, such reductions occur for both non-critical access and critical access hospitals and are geographically widespread.
- While there is currently a downward trend in transfusion utilization in the United States, it defies logic that such a significant proportion of hospitals would reduce their transfusions by 80 percent, or even 50 percent after the conversion to ICD-10. Rather, we believe the original concern regarding under-reporting has been exacerbated. **Because there is no requirement that the ICD-10 procedure or value codes be used for a facility to be paid, valid transfusion claims that include only revenue codes will be missed by the STTrR.** With the switch to ICD-10 codes, we hypothesize that even more hospitals are using only revenue codes, and no accompanying ICD-10 procedure or value codes, which are required for the STTrR. Dialysis facility performance that may appear to have drastically improved on the STTrR (fewer transfusions), may in fact solely be due to hospitals not including the ICD-10 codes specified by the measure. Conversely, facilities associated with hospitals that use ICD-10 and revenue codes appear to perform poorly.
- Further to this point, the largest hospital by volume with a >80 percent apparent reduction in transfusion was a facility in the Northeast. In the last year before ICD-10-PCS and the first year after ICD-10-PCS, a blood transfusion occurred during 10.0 percent and 0.1 percent of hospitalizations, respectively. A dialysis facility (or facilities) associated with this hospital will show a significant improvement in the StrR due to the ICD-10 implementation and change in the hospital's reporting practices.
- In summary, the STTrR's validity is in question as well due to the under-reporting by hospitals after the switch to ICD-10. A review of the claims suggests that a substantial percentage of hospitals simply stopped including ICD-10 procedure codes for blood transfusions during hospitalizations, making it now impossible to determine if a transfusion has occurred.

- **Serum Phosphorus**

- KCP supports maintaining the serum phosphorus measure as part the QIP and eliminating the hypercalcemia measure (as described below). Physicians rely upon the serum phosphorus level to make clinical decisions.
- We understand that the Agency must comply with the Protecting Access to Medicare Act (PAMA). To this end, the serum phosphorus measure is a more appropriate measure to meet the statutory requirement than the hypercalcemia measure given that serum phosphorus but not serum calcium is impacted by oral only medications.

- **Transplant measure**

- KCP agrees that it is important to have a transplant measure in the ESRD QIP. However, the two current measures – Percentage of Prevalent Patients Waitlisted (PPPW) and Standardized First Kidney Transplant Waitlist Ratio for Incident Dialysis patients (SWR) – are not appropriate because NQF has recommended against endorsement. In addition, facilities do not have control over how the transplant waitlists work, as KCP has commented in the past, so the measures are not actionable.
- Regarding the specifications for the SWR, we note that during the NQF Renal Standing Committee’s consideration of the SWR in June 2018, the Committee discussed whether a patient with a previous transplant was excluded. CMS responded in the affirmative. Our impression is that this satisfied the NQF Committee. The specifications proposed for the QIP, however, eliminate this exclusion. We request justification for the modification of this exclusion.
- CMS should prioritize developing an appropriate transplant measure that is actionable by dialysis facilities. A measure that recognizes what is actionable by facilities would better support the Meaningful Measures Initiative priority area of increased focus on effective communication and coordination. The problem is not with facility assessment and evaluation, but with the criteria hospitals set for the waitlists. We recognize the need to avoid a “check-box measure,” but believe that a transplant measure must be actionable.

B. DFC Measure Recommendations

- **KCQA UFR Measure**

- KCP continues to believe that fluid management is an important quality area, which is why it funded the KCQA to undertake such measure development. The KCP members identified addressing fluid management as the highest priority from KCP’s *Strategic Blueprint for Kidney Care Quality*. We commend CMS for indicating it is using KCQA’s NQF-endorsed measure, 2701:

Avoidance of Utilization of High Ultrafiltration Rate (≥ 13 ml/kg/hour); but that indications also means that CMS should use the specifications for this measure that NQF-endorsed and not modify them.

- KCP requests justification as to why the Additional Information item, “A facility is excluded from a reporting month if its certification date falls on or after the first day of the reporting month (the scenario can only occur once during January 2019-June 2019)” has been struck. We recognize the cases are likely rare, but when they do occur, those facilities should be excluded (and the dates altered to reflect future payment years).
- This measure should be part of the DFC, but not the QIP.
- **KCQA Medication Reconciliation (MedRec) Measure**
 - KCP supports the KCQA MedRec [Measure](#), as evidenced by our prioritizing its development using community resources. However, as noted above, we believe the ESRD QIP should include a parsimonious set of measures that can be relied upon over time to provide an overarching assessment of facility performance. More specific outcomes measures should reside in the DFC.
 - With respect to the specifications, rather than strike the definitional elements of “medication reconciliation,” we recommend the specifications restore the endorsed verbiage as “Additional Information/Definition” to ensure standardized reconciliation. Additionally, page 148 of the Proposed Rule notes the measure is calculated using administrative claims; this should be deleted, as claims are not required for the measure. Finally, page 150 states that the measure “is endorsed by NQF as #2988.” Given the specification changes, it is more accurate to state “the specifications are based on NQF #2988.”
- **NHSN Healthcare Personnel Influenza Vaccination Reporting Measure**
 - KCP continues to believe that influenza vaccination of healthcare personnel is an important public health concept and has supported including NHSN Healthcare Personnel Influenza Vaccination as a reporting measure, but the performance period needs to be aligned with the CDC’s guidelines and the NQF’s standard specifications for influenza immunization measures. Specifically, both define the acceptable immunization period as “October 1 or whenever the vaccine became available.” Vaccine shipments typically begin in August, and the measure should be specified to allow for this fact. The measure also lacks the ability for facilities to batch submit. Thus, as currently specified the measure should be eliminated from the QIP under Factor 3 because it does not align with clinical practice.

- Because this area is important, but not a critical driver of key patient outcomes, it is more appropriate that the measure be in DFC.
- **Kt/V Dialysis Adequacy Comprehensive Clinical Measure**
 - While dialysis adequacy is a core metric of facility performance, there is little gap in performance, so under Factor 1 it should not be included in the QIP. However, it remains an important measure to patients and should be included in DFC, but merely as a reported measure and not one that attributes to penalties or star ratings.
 - Unfortunately, the current pooled measure masks performance for home dialysis and pediatric patients. CMS indicated the purpose of creating the pooled measure was to address the problem that most facilities that care for pediatric patients do not meet the minimum sample size for their pediatric population. If the measure is eliminated from the QIP and included in DFC, the individual measures for adequacy should be what is reported and accessible to patients and caregivers. What is paramount is that patients have access to information that is personally meaningful to them. Pooling the adequacy measure serves none of the patients.
- **Vascular Access Type (VAT) Measure Topic – Arteriovenous Fistula (AVF) Clinical Measure/Standardized Fistula Measure**
 - As noted already, reduction in catheters drives better patient outcomes more than the placement of a fistula, so under Factor 5, the VAT Topic and AVF measure should not be included in the QIP. However, understanding performance on this measure in a public way is important and it should be included in DFC.
 - We recommend that the specifications be edited to explicitly state that the patient must be on maintenance HD using an AVF “without a dialysis catheter present” to emphasize importance of removing long-term catheters. We also note that the denominator should use a “patient-months” construction (as do the numerator and measure description).
- **Clinical Depression Screening and Follow-Up Reporting Measure**
 - Clinical Depression Screening does not drive a core outcome for patients, but is important more generally to the population. Inclusion in the QIP dilutes the TPS and make it more difficult for the QIP to drive improvement. However, this measure should be used in the DFC and publicly available to patients and caregivers.
- **Standardized Mortality Rate measure**

- Like the hospitalization measure, the current morality ratio measure should be modified to be a true risk-standardized rate, as noted above.

- **Patient Reported Outcome Measure**

- KCP supports further development of a measure in this domain.

C. Measures That Should Not Be Used in QIP or DFC

- **Pain Assessment and Follow-Up Reporting Measure**

- KCP agrees with the CMS proposal to eliminate this measure from the ESRD QIP because “measure performance among the majority of ESRD facilities is so high and unvarying that meaningful distinctions in improvements or performance can no longer be made.”³ If distinctions among facilities cannot be made by a measure, it is not appropriate or useful to patients to include the measure on DFC as well.

- **Hypercalcemia Clinical Measure**

- KCP has consistently raised concerns with the use of the hypercalcemia measure. NQF has concluded that the hypercalcemia measure is topped out and placed the measure in Reserve Status because of high facility performance and minimal room for improvement. Similarly, the Measure Applications Partnership (MAP) did not support the measure in its 2016 report. Thus, the hypercalcemia measure also should be eliminated under Factor 1.

- In previous rulemaking, the preamble indicated that despite these facts, CMS felt bound to maintain the measure because the statute requires including measures specific to oral-only drugs.⁴ It has stated that hypercalcemia is the only measure of which we are aware that meets the statutory requirements in PAMA for an NQF-endorsed quality measure of conditions treated with oral-only medications. The measure focused on the administration of oral Sensipar® (cinacalcet), which with the development and launch of the IV Parsabiv® (etelcalcetide), is no longer an oral-only drug. Because there are no longer any oral-only calcimimetics, the hypercalcemia measure is no longer required by the statute and thus the rationale for maintaining this topped out measure is no longer relevant.

- **Emergency Department Utilization**

³83 *Fed. Reg.* at 34338.

⁴42 U.S.C. § 1395rr(h)(2)(e).

- This measure should not be included in either the QIP or DFC because it has failed to be endorsed by the NQF. The measure was rejected for low and/or insufficient validity and/or reliability by NQF Methods Panel. A measure that is not reliable or valid should not be used because its results cannot be trusted to be accurate. It would seem clearly to come within Factor 2 because “[p]erformance or improvement on a measure does not result in better or the intended patient outcomes,” since the measure is not accurately measuring performance. In addition, as noted above, the fact that NQF did not endorse the measure cannot be circumvented by referencing the authority the Congress provided is no measure has been endorsed by NQF in a particular domain.

Appendix I. Teleconference Call 1 Presentation

The Teleconference Call 1 Presentation is provided on the next several pages.



Dialysis Facility Compare (DFC) Star Rating TEP

TELECONFERENCE CALL #1

MAY 6, 2019 · 3:00 – 5:00 PM, EST

TEP Teleconference # 1 Call Agenda

1. TEP Member Introductions (15 minutes)
2. TEP Overview and Charter Approval (10 minutes)
3. Background on the DFC Star Ratings (25 minutes)
4. Presentation and Discussion of the Shift in the Star Rating Distribution (45 minutes)
5. Public Comment (5 minutes)

UM-KECC Star Rating Team

- Yi Li, PhD
- Joseph Messana, MD
- Claudia Dahlerus, PhD
- Richard Hirth, PhD
- Peisong Han, PhD
- Casey Parrotte, PMP
- Jennifer Sardone, PMP
- Lan Tong, MPH
- Karen Wisniewski, MPH
- Wolf Gremel, MS
- Jingya Gao, MS
- Stephen Salerno, MS
- Brandon Frye, BA
- **Contact Person:**
Jordan Affholter, BA
affjorda@med.umich.edu

CMS Representatives

Centers for Medicare & Medicaid Services, Division of Quality Measurement:

- Joel Andress, PhD · ESRD Measures Development Lead
- Jesse Roach, MD · Nephrologist, Clinical Subject Matter Expert
- Golden Horton, MS · Dialysis Facility Compare Lead

TEP Members

TEP Co-Chairs:

- Paul Conway, BA

Board of Directors Member, Chair of Public Policy and Global Affairs · *American Association of Kidney Patients*

- Catherine Sugar, PhD

Professor, Departments of Biostatistics, Statistics & Psychiatry · *University of California, Los Angeles*
Director · *Semel Institute Statistics Core, University of California, Los Angeles*

TEP Members:

- Mark Andaya, MS, RN

Director of Quality Assurance and Performance Improvement · *The Rogosin Institute*

- Andrew Conkling, BS

Vice President · *Dialysis Patient Citizens*

TEP Members

- Lorien Dalrymple, MD, MPH

Vice President of Epidemiology and Research · *Fresenius Medical Care North America*
Volunteer Clinical Faculty Associate Professor · *University of California, Davis*

- Sharon Dickson, RN, MSN

Regional Quality Manager · *Fresenius Kidney Care*

- Dawn Edwards

Patient Advocate · *NxStage*, Health Ambassador · *The Rogosin Institute*

- Derek Forfang

Patient Advocate/Chair · *Forum of ESRD Networks*, *National Kidney Foundation*

- Monica Fox

Outreach Associate/Patient Advocate · *National Kidney Foundation of Illinois*

TEP Members

- Lonnie Green

Patient Advocate/Subject Matter Expert · *IPRO ESRD Network 6, National Patient and Family Engagement – Learning Action Network (NPFE-LAN)*

- Mark Johnson

Patient Advocate/Subject Matter Expert · *ESRD Network 12, National Patient and Family Engagement – Learning Action Network (NPFE-LAN)*

- Mark Joseph, MD

Pediatric Nephrologist · *Pediatric Kidney Disease and Hypertension Centers*
Medical Director · *Phoenix Pediatric Dialysis Center*

- Richard Knight, MBA

President · *American Association of Kidney Patients*, Adjunct College Instructor · *Bowie State University*

- Mahesh Krishnan, MD, MPH, MBA

Group Vice President of Research and Development · *DaVita Kidney Care*

TEP Members

- Michael “Jack” Lennon, MBA

Executive Director · *Improving Renal Outcomes Collaborative*

- Nicole Stankus, MD, MSc

Associate Professor of Medicine / Nephrology · *University of Chicago*
Medical Director · *DaVita Stony Island Dialysis Center*

- Caprice Vanderkolk, RN, MS, BC-NE

Manager, Adult and Pediatric Dialysis Program · *University of Minnesota/Fairview Hospitals*

- Curtis Warfield, BS, MS

Patient Reviewer · *National Kidney Foundation*, Senior Quality Analyst · *State of Indiana*

- David White

Patient Advocate/Board of Directors Member · *American Association of Kidney Patients*
Healthcare Consultant · *Quality Insights Renal Network 5*

TEP Objectives

Provide recommendations on options for resetting the DFC Star Ratings

- The TEP recommendations will be used to inform the development of a methodology for resetting the DFC Star Ratings
- The final methodology developed is intended to allow the DFC Star Ratings to continue to reflect meaningful performance differences among facilities

Questions

ANY TEP QUESTIONS ABOUT THE OBJECTIVES?

Star Rating Overview

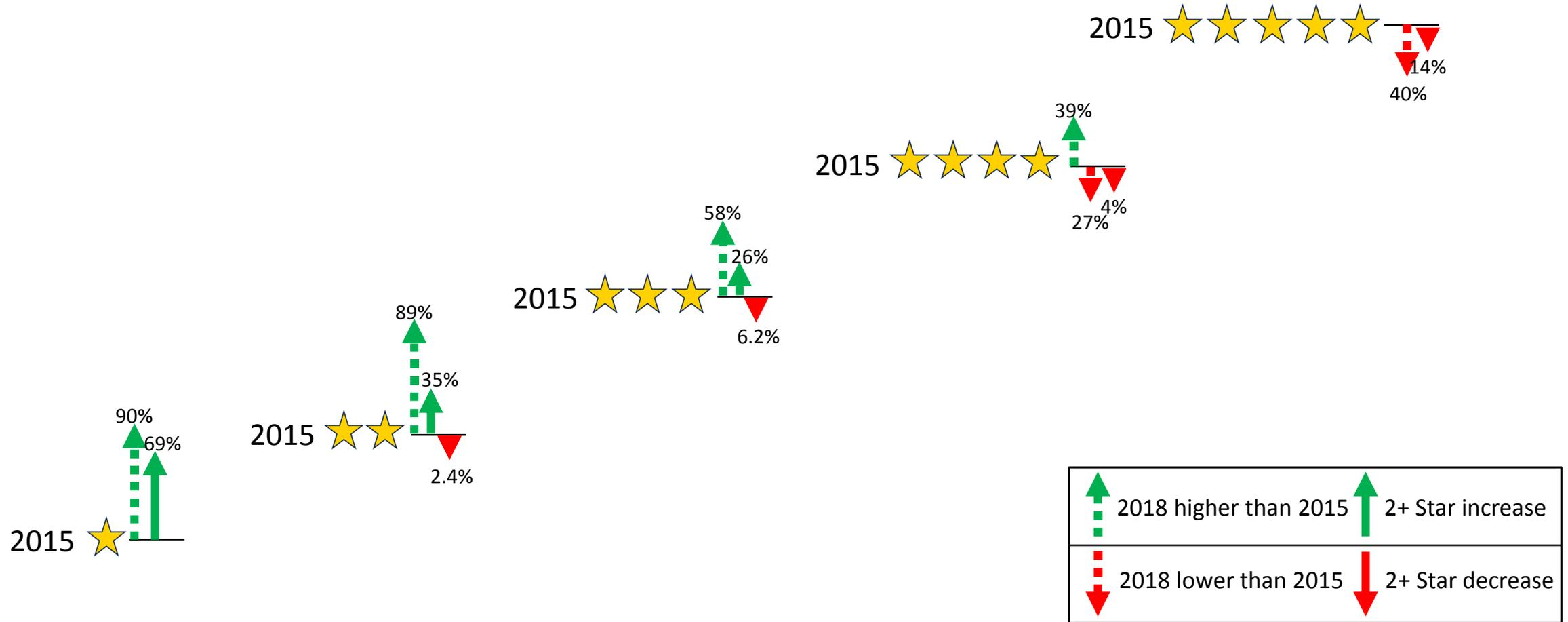
The Centers for Medicare and Medicaid Services (CMS), through a contract with UM-KECC, developed the Dialysis Facility Compare (DFC) Quality of Patient Care Star Rating System:

- To rate the quality of care provided by dialysis facilities
- To provide patients, families, and caregivers information to easily compare dialysis facilities

Since the first DFC Star Rating was released in January 2015, KECC facilitated two TEPs, both of which provided recommendations regarding revisions to the DFC Star Rating methodology

The Shift in Star Ratings

Change from 2015 to 2018



Key Concepts and Terminology

Measure Score: A standardized score applied to a specific measure, which has mean 0, variance 1, and takes values in the range of -2.58 to 2.58

Measure Value: The original value of a facility's clinical quality measure as reported on DFC, which represents a ratio or a percentage

Domain Score: A score which summarizes a facility's performance on a group of correlated clinical quality measures (domain). It is an average of the individual measure scores in that group

Final Score: A continuous score calculated for each facility, which summarizes its performance on the reported clinical quality measures. It is an average of the three domain scores

Cutoff: A value of the final score that defines the boundary between two adjoining Star Rating categories

Key Concepts and Terminology, continued

Rebaselining: Establishing a new baseline year and rescoring measures

- Defines new baseline scoring cutoffs for facilities to be rated
- The cutoffs let Star Rating distribution remain unchanged from the past release to allow for continuity over time

Resetting: Establish new scoring cutoffs and Star Rating distribution

- Defines new baseline scoring cutoffs for facilities to be rated
- The cutoffs reset Star Rating distribution to improve the ability to differentiate facility performance

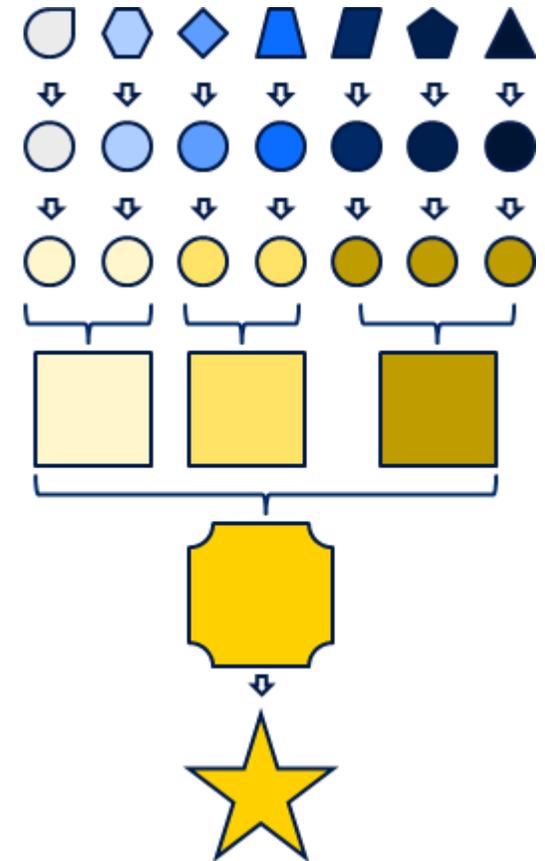
Baseline Period: The time period (e.g., calendar year) in which data are collected for computing measure scoring criteria and defining cutoff values for Star Rating categories. The cutoffs will be used to rate facilities in future evaluation periods

Evaluation Period: The time period (e.g., calendar year) in which data are collected for calculation of measure results and facility Star Rating scores, reported on DFC. Final scores in the evaluation period are compared against cutoffs established in the baseline period in order to rate facilities

Original DFC Star Ratings

1. Measure values were standardized to measure scores
2. Factor analysis on measure scores identified 3 measure domains
3. Measure scores were averaged into domain scores
4. Domain scores were averaged into a final score
5. Final scores were grouped into five categories based on a facility's performance relative to all other facilities in a given period:

10% 1-Star, 20% 2-Star, 40% 3-Star, 20% 4-Star, 10% 5-Star



TEP I - Recommendation

Establish a Baseline to Account for Changes in Facility Performance Over Time:

- **Baseline Period:** Time period, typically a calendar year, in which data are collected to define measure scoring criteria and cutoff values for Star Rating categories.
- **Evaluation Period:** The time period, typically a calendar year, in which data are collected for calculation of measure results and facility Star Rating scores, reported on DFC.

Implication of using a baseline: improvement in absolute measure values may translate to improvement in final scores. This may result in bunching of the DFC Star Rating distribution over time

Star Rating Distribution by DFC Release



TEP II - Recommendations

The 2017 TEP was asked to provide recommendations on the inclusion of new and updated measures, and criteria for updating the star rating in the future

Update Measure Set for October 2018 Release:

- Inclusion of new measures: Pediatric PD Kt/V and SRR
- Update current measure definitions: SMR, SHR, STrR, Fistula, and Catheter

Maintain Longitudinal Continuity of the DFC Star Rating Distribution:

- *Rebaseline* the DFC Star Ratings by establishing new baseline scoring cutoffs but maintaining the proportion of facilities in each Star Rating category

Rebaselining DFC Star Rating Cutoffs

April 2018 DFC Release Distribution with Previous Measure Set

Standardized Mortality Ratio (SMR)

Standardized Hospitalization Ratio (SHR)

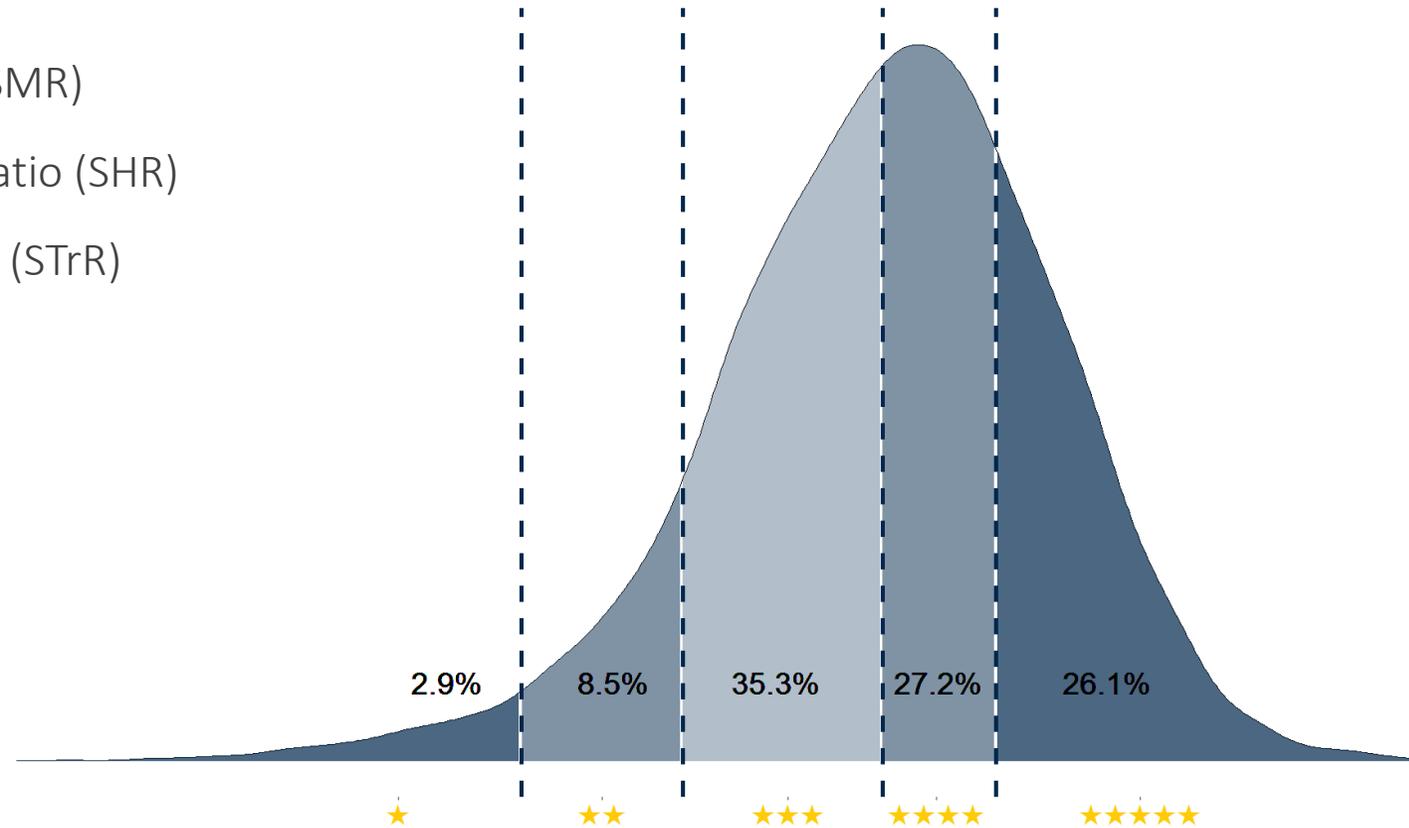
Standardized Transfusion Ratio (STrR)

% Fistula (Fistula)

% Catheter (Catheter)

Hypercalcemia

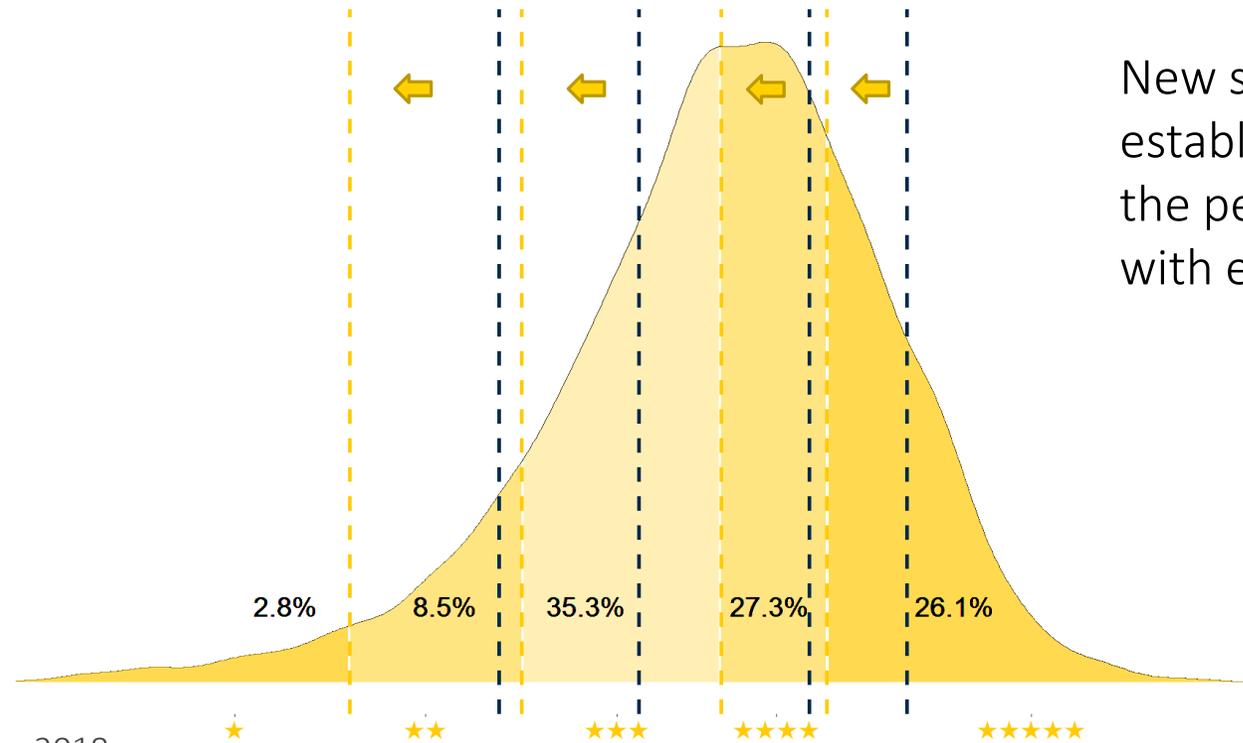
Total Kt/V



Rebaselining DFC Star Rating Cutoffs

April 2018 **Baseline** Distribution with **Updated** Measure Set

- Standardized Mortality Ratio (SMR)*
- Standardized Hospitalization Ratio (SHR)*
- Standardized Transfusion Ratio (STrR)*
- Standardized Readmissions Ratio (SRR)†
- Standardized Fistula Rate (SFR)*
- Long-Term Catheter Rate (LTCR)*
- Hypercalcemia*
- Total Kt/V‡



New score cutoffs are established to maintain the percent of facilities with each Star Rating

*Updated in October 2018; †Added in October 2018;
‡Component Added in October 2018

Rebaselining

Establishing a new baseline year and rescoring measures:

- **Criteria:** Necessary when new measures are added or current measures are updated/removed
- Defines new baseline scoring cutoffs for facilities to be rated
- Cutoffs allow proportions of facilities in each Star Rating category to **remain unchanged** compared to the previous release
- Carried out for the October 2018 DFC Release

Shift in the Star Rating Distribution

PRESENTATION AND DISCUSSION

Final Score Distribution Shift

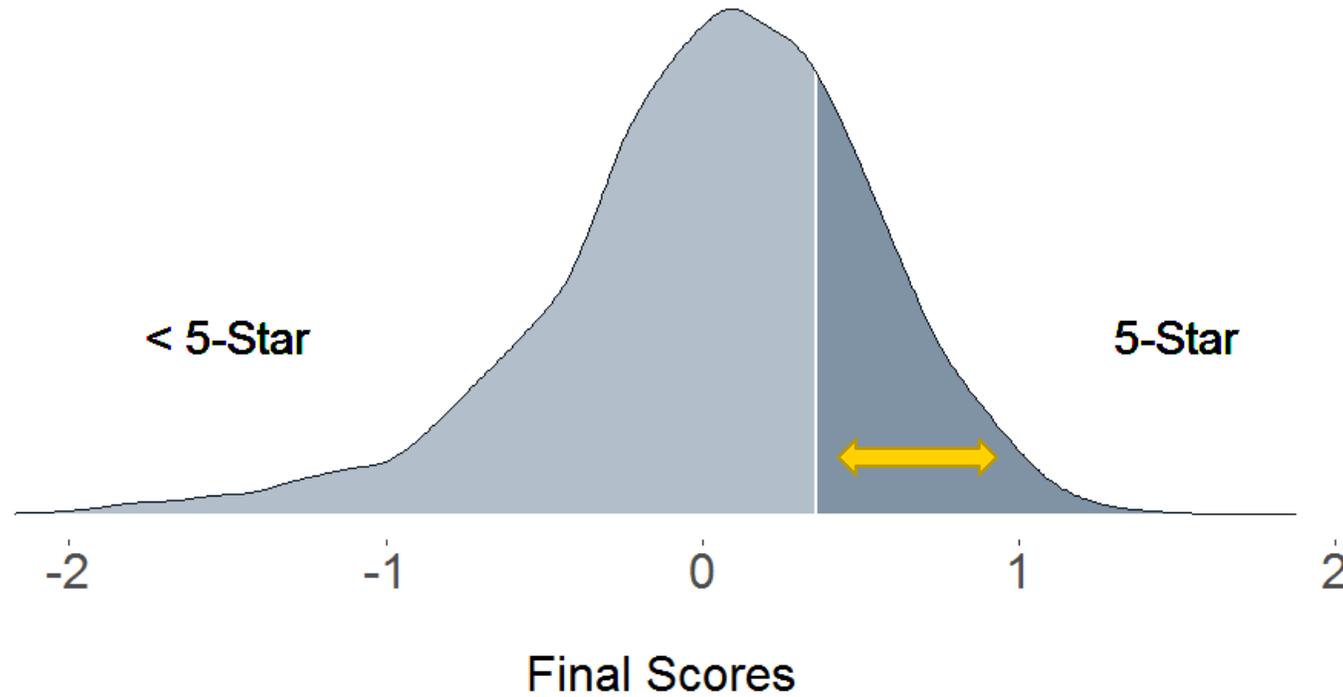
- From October 2015 to April 2018, there has been an upward shift in the proportion of facilities receiving a 4- or 5-Star Rating (77.9% increase)
- The increasing upward shift has slowed since rebaselining

	Min.	Q1	Med.	Mean	Q3	Max.
Oct. 2015 (Baseline)	-2.35	-0.31	0.05	0.01	0.36	1.77
Oct. 2016	-2.06	-0.12	0.20	0.17	0.49	1.80
Apr. 2018	-2.03	0.03	0.34	0.31	0.62	1.73
Oct. 2018*	-2.17	-0.26	0.07	0.03	0.38	1.88

**Rebaselined with April 2018 as baseline*

Variation within 5-Star Facilities

October 2018 DFC Release



What is Driving this Shift?

Mean (SD) domain scores by release year before rebaselining:

	Oct. 2015	Oct. 2016	Apr. 2018
Domain 1 (SHR, SMR, STrR)	0.00 (0.69)	0.08 (0.69)	0.06 (0.70)
Domain 2 (Fistula, Catheter)	0.00 (0.85)	0.03 (0.87)	0.04 (0.86)
Domain 3 (Total Kt/V, Hypercalcemia)	0.00 (0.74)	0.36 (0.57)	0.82 (0.46)

- Mean domain scores increased the most for Domain 3
- Results from sensitivity analyses (not shown) were consistent when applying the same set of measure definitions to all measures across all releases

Questions

ANY TEP QUESTIONS ON THE CONTENT PRESENTED?

Star Rating Reset Background

- After the 2017 TEP, CMS formulated the following policy informed, in part, by the TEP summary
- Presented during CMS' October 2017 National Provider Call:
 - The DFC Star Rating distribution will be evaluated once 3 years have passed since the last reset
 - The DFC Star Ratings will be evaluated for a reset when $\leq 15\%$ of facilities are receiving 1- or 2-Stars
 - A resetting of the DFC Star Rating distribution will also include the establishment of a new baseline

Resetting

Update scoring cutoffs and Star Rating distribution:

- **Criteria:** When the Star Rating's ability to differentiate facility-level performance is reduced (e.g. compression of the Star Ratings due to progressive shifts in facility performance)
- Defines new baseline scoring cutoffs for facilities to be rated
- Proportions of facilities in each Star Rating category are **reset, creating a new Star Rating distribution**

Recommendations for resetting will be a primary focus of the TEP and will be covered in greater detail on the next teleconference call

How to Interpret Star Ratings after Resetting?

- To improve usability of DFC Star Ratings during the reset transition, development of tools to assist DFC users' interpretation should be considered
- In-person meeting discussion:
 - What additional information or display options tools can help users interpret star ratings in the transition after resetting?
 - UM-KECC will develop some concepts to help stimulate brainstorming discussion
 - Prior to the In Person TEP, please consider potential DFC display options or other tools that could assist patients with interpretation during and after DFC Star Ratings resetting

Upcoming TEP Meetings

- Teleconference Call #2: May 24, 2019 from 3:00 – 5:00 PM, EST
- In-Person TEP Meeting: June 6, 2019 from 8:30 AM – 4:00 PM, EST

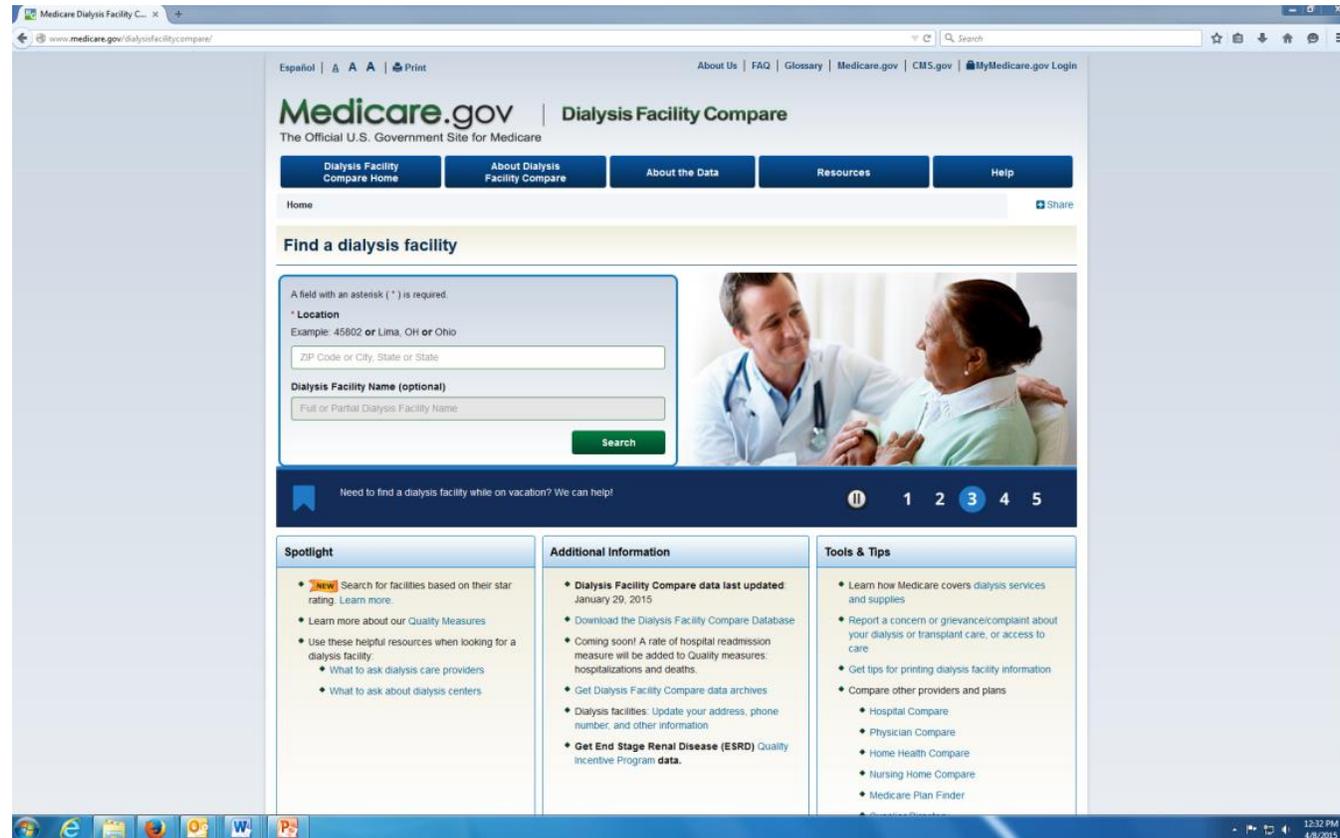
Public Comment Period

4:55 – 5:00 PM, EST

End of Teleconference Call #1

Appendix

Dialysis Facility Compare Website



www.medicare.gov/dialysisfacilitycompare/

DFC Star Ratings Display

Dialysis facility profile

[Back to Results](#)

General information

Survey of patients' experiences

Quality of patient care

UNIV OF MI DIALYSIS CLINICS - LIVONIA

19900 HAGGERTY RD STE 106
LIVONIA, MI 48152
(734) 432-7870

Quality of patient care star rating: ★★★★★

Survey of patients' experiences star ratings: ★★★★★●

[Learn more about star ratings for dialysis centers](#)

Distance: 4.1 miles

[Add to my Favorites](#)
[Map and Directions](#)

Dialysis center information

[Learn why these characteristics and services are important.](#)

- [Shifts Starting After 5PM:](#) Yes
- [In-Center Hemodialysis:](#) Yes
- [Number of Hemodialysis Stations:](#) 16
- [Peritoneal Dialysis:](#) Yes
- [Home Hemodialysis Training:](#) No
- [Type of Ownership:](#) Non-Profit
- Corporate Name: UNIVERSITY OF MICHIGAN
- Center's Initial Date of Medicare Certification or Recertification: 07/01/1999

Star Rating History

01-2014	Development of DFC Star Ratings began
06-2014	DFC Star Ratings were announced
07-2014	First DFC Star Ratings preview period
01-2015	First DFC Star Ratings release
04-2015	First technical expert panel
07-2016	Preview period for updated methodology refresh
10-2016	DFC Star Ratings refreshed with updated methodology
02-2017	Second technical expert panel
07-2018	Preview period for second update refresh
10-2018	DFC Star Ratings refreshed with second update
06-2019	Third technical expert panel

Star Rating Distribution by DFC Release

DFC Release	1-Star	2-Stars	3-Stars	4-Stars	5-Stars
October 2015 (Baseline)	579 (10.0%)	1,163 (20.0%)	2,332 (40.0%)	1,168 (20.0%)	584 (10.0%)
October 2016	300 (4.9%)	873 (14.4%)	2,386 (39.4%)	1,483 (24.5%)	1,019 (16.8%)
April 2018	180 (2.9%)	525 (8.5%)	2,188 (35.3%)	1,689 (27.2%)	1,621 (26.1%)
April 2018 (Baseline)	178 (2.8%)	534 (8.5%)	2,220 (35.3%)	1,713 (27.3%)	1,641 (26.1%)
October 2018	190 (2.9%)	536 (8.2%)	2,311 (35.2%)	1,810 (27.5%)	1,724 (26.2%)
April 2019 (Hypothetical)	169 (2.6%)	497 (7.6%)	2,342 (35.7%)	1,776 (27.1%)	1,770 (27.0%)

Measure Distribution by DFC Release

Mean (SD) clinical quality measure values by release year before rebaselining:

	Oct. 2015 (n = 5,872)	Oct. 2016 (n = 6,061)	Apr. 2018 (n = 6,203)	Relative Change 2015 - 2018
Domain 1: SHR ^{*,§}	1.00 (0.31)	1.01 (0.32)	1.02 (0.32)	2.00%
Domain 1: SMR ^{*,§}	1.02 (0.29)	0.99 (0.28)	0.98 (0.26)	-3.92%
Domain 1: STrR ^{*,§}	1.00 (0.57)	0.94 (0.55)	0.94 (0.51)	-6.00%
Domain 2: Fistula	64.43 (11.16)	66.67 (11.23)	66.85 (11.13)	3.76%
Domain 2: Catheter [*]	10.22 (6.59)	11.09 (7.00)	11.10 (6.90)	8.61%
Domain 3: Total Kt/V	89.95 (8.98)	91.79 (6.68)	95.34 (4.68)	5.99%
Domain 3: Hypercalcemia [*]	2.22 (2.62)	1.43 (1.52)	0.74 (1.01)	-66.67%

[§]Risk-adjusted to Oct. 2015 release standards in the Oct. 2016 and Apr. 2018 releases; ^{*}Lower values indicate better performance

Hypercalcemia Definition Changes

Measure Definition	2012*	2013	2014	2015	2016
1) Patient-Months with Missing Calcium Excluded from Denominator & Numerator					
Pt-months w/ Hypercalcemia (Num.)	58,414	103,400	105,538	73,464	38,795
Eligible Pt-Months (Denom.)	2,056,337	4,249,846	4,574,120	4,824,216	5,211,293
<i>Avg. Uncorrected Calcium > 10.2 mg/dL</i>	2.8%	2.4%	2.3%	1.5%	0.7%
2) Patient-Months with Missing Calcium Included in Denominator					
Pt-months w/ Hypercalcemia (Num.)	58,414	103,400	105,538	73,464	38,795
Eligible Pt-Months (Denom.)	2,304,039	4,703,748	4,885,182	5,043,512	5,302,261
<i>Avg. Uncorrected Calcium > 10.2mg/dL</i>	2.5%	2.2%	2.2%	1.5%	0.7%
3) Patient-Months with Missing Calcium Included in Denominator & Numerator					
Pt-months w/ Hypercalcemia (Num.)	255,141	479,099	365,076	262,278	129,763
Eligible Pt-Months (Denom.)	2,304,039	4,703,748	4,885,182	5,043,512	5,302,261
<i>Avg. Uncorrected Calcium > 10.2mg/dL</i>	11.1%	10.2%	7.5%	5.2%	2.4%
<i>Missing</i>	8.3%	7.9%	5.3%	3.7%	1.7%

* Includes data from July to December 2012 only

Rebaselining: October 2018 Release

- The October 2018 DFC release used the April 2018 Star Rating distribution to establish new scoring cutoffs in order to maintain longitudinal continuity
- The October 2018 release used the new measure specifications applied to the April 2018 release data to establish a new set of final score cutoffs
- The cutoffs were calculated reproduce the Star Rating distribution previously achieved for the April 2018 release using the prior measures and methodology
- The April 2018 release served as an evaluation period for the old measure set and as a baseline period for the updated measure set

Appendix J. Teleconference Call 2 Presentation

The Teleconference Call 2 Presentation is provided on the next several pages.



Dialysis Facility Compare (DFC) Star Rating TEP

TELECONFERENCE CALL #2

MAY 24, 2019 · 3:00 – 5:00 PM, EDT

TEP Teleconference # 2 Call Agenda

1. Review of Star Rating Shift
2. TEP Discussion on Resetting the Star Ratings
3. Potential Options for Weighting the Star Ratings
4. Public Comment (4:55 – 5:00 EDT)

UM-KECC Star Rating Team

- Yi Li, PhD
- Joseph Messana, MD
- Claudia Dahlerus, PhD
- Richard Hirth, PhD
- Peisong Han, PhD
- Casey Parrotte, PMP
- Jennifer Sardone, PMP
- Lan Tong, MPH
- Karen Wisniewski, MPH
- Wolf Gremel, MS
- Jingya Gao, MS
- Stephen Salerno, MS
- Brandon Frye, BA
- **Contact Person:**
Jordan Affholter, BA
affjorda@med.umich.edu

CMS Representatives

Centers for Medicare & Medicaid Services, Division of Quality Measurement:

- Joel Andress, PhD · Measure Development Subject Matter Expert
- Jesse Roach, MD · Nephrologist, ESRD Measures Development Lead
- Golden Horton, MS · Dialysis Facility Compare Lead

TEP Members

TEP Co-Chairs:

- Paul Conway, BA

Board of Directors Member, Chair of Public Policy and Global Affairs · *American Association of Kidney Patients*

- Catherine Sugar, PhD

Professor, Departments of Biostatistics, Statistics & Psychiatry · *University of California, Los Angeles*
Director · *Semel Institute Statistics Core, University of California, Los Angeles*

TEP Members:

- Mark Andaya, MS, RN

Director of Quality Assurance and Performance Improvement · *The Rogosin Institute*

- Andrew Conkling, BS

Vice President · *Dialysis Patient Citizens*

TEP Members

- Lorien Dalrymple, MD, MPH

Vice President of Epidemiology and Research · *Fresenius Medical Care North America*
Volunteer Clinical Faculty Associate Professor · *University of California, Davis*

- Sharon Dickson, RN, MSN

Regional Quality Manager · *Fresenius Kidney Care*

- Dawn Edwards

Patient Advocate · *NxStage*, Health Ambassador · *The Rogosin Institute*

- Derek Forfang

Patient Advocate/Chair · *Forum of ESRD Networks*, *National Kidney Foundation*

- Monica Fox

Outreach Associate/Patient Advocate · *National Kidney Foundation of Illinois*

TEP Members

- Lonnie Green

Patient Advocate/Subject Matter Expert · *IPRO ESRD Network 6, National Patient and Family Engagement – Learning Action Network (NPFE-LAN)*

- Mark Johnson

Patient Advocate/Subject Matter Expert · *ESRD Network 12, National Patient and Family Engagement – Learning Action Network (NPFE-LAN)*

- Mark Joseph, MD

Pediatric Nephrologist · *Pediatric Kidney Disease and Hypertension Centers*
Medical Director · *Phoenix Pediatric Dialysis Center*

- Richard Knight, MBA

President · *American Association of Kidney Patients*, Adjunct College Instructor · *Bowie State University*

- Mahesh Krishnan, MD, MPH, MBA

Group Vice President of Research and Development · *DaVita Kidney Care*

TEP Members

- Michael “Jack” Lennon, MBA

Executive Director · *Improving Renal Outcomes Collaborative*

- Nicole Stankus, MD, MSc

Associate Professor of Medicine / Nephrology · *University of Chicago*
Medical Director · *DaVita Stony Island Dialysis Center*

- Caprice Vanderkolk, RN, MS, BC-NE

Manager, Adult and Pediatric Dialysis Program · *University of Minnesota/Fairview Hospitals*

- Curtis Warfield, BS, MS

Patient Reviewer · *National Kidney Foundation*, Senior Quality Analyst · *State of Indiana*

- David White

Patient Advocate/Board of Directors Member · *American Association of Kidney Patients*
Healthcare Consultant · *Quality Insights Renal Network 5*

For TEP Discussion

- Is it time to reset the Star Ratings?
- How should we do it?
- How do we help DFC consumers interpret facility performance on star ratings during and immediately after the transition?

Key Concepts

Rebaselining: Rescoring of measures when establishing a new baseline year

- Defines new baseline scoring cutoffs for facilities to be rated
- Star Rating proportions remain unchanged to allow for continuity over time

Resetting: Update scoring cutoffs and Star Rating distribution

- Defines new baseline scoring cutoffs for facilities to be rated
- Star Rating proportions are reset to improve the ability to differentiate facility performance

Star Rating Shift

RECAP OF CALL #1 AND DISCUSSION

Star Rating Distribution by DFC Release

- There has been an upward shift in the proportion of facilities receiving a 4- or 5-Star Rating
- The rate of increase in the proportion of 4- and 5-Star facilities has slowed since re-baselining



Resetting the DFC Star Ratings

POTENTIAL OPTIONS

Resetting the Star Rating Distribution

- When the Star Rating's ability to differentiate facility-level performance is reduced (e.g. bunching of the Star Ratings due to shifts in facility performance)

Potential Options for Resetting:

1. Reset using (a) set proportions for the star categories or (b) based on standard deviations from the mean of the final score
2. Reset using empirical clustering methods such as K-means or hierarchical clustering
3. Other potential approaches from the TEP?

Option 1: Reset to a Pre-Specified Distribution

Based on Fixed Proportions:

Pre-specify the percent of facilities in each category and determine the cutoffs based on these proportions (e.g. 10%-20%-40%-20%-10%)

Based on Standard Deviations:

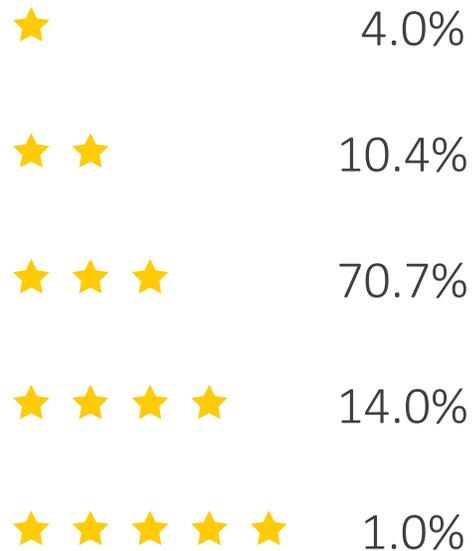
The final score cutoffs are determined based on “how far” facilities’ scores are (in standard deviations) from the average national scores

E.g. $< -2 = 1\text{-Star}$, $-2 \text{ to } -1 = 2\text{-Stars}$, $-1 \text{ to } 1 = 3\text{-Stars}$, $1 \text{ to } 2 = 4\text{-Stars}$, $> 2 = 5\text{-Stars}$

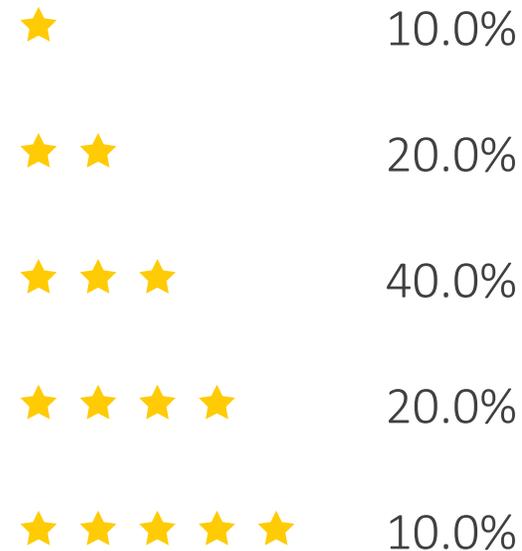
Option 1: Reset to a Pre-Specified Distribution

Example using the October 2018 release:

Stars Based on Standard Deviations:



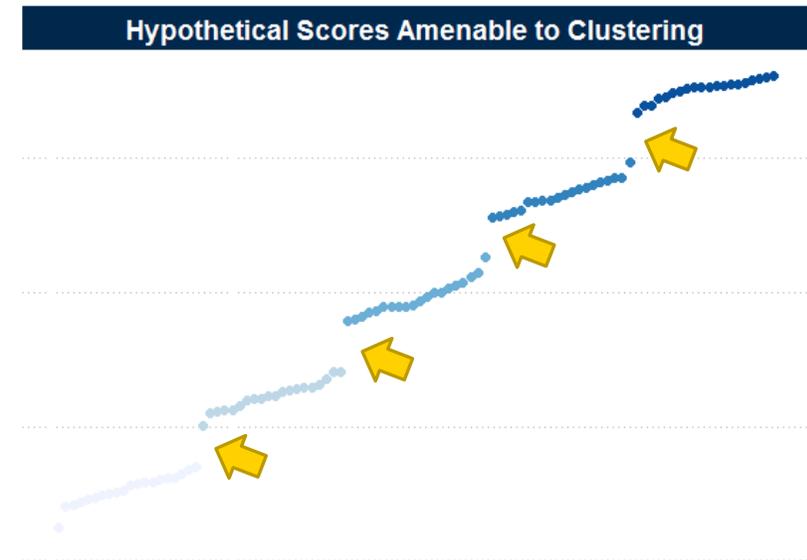
Stars Based on Fixed Proportions:



Option 2: Reset with Clustering Methods

What is clustering?

- Group facilities that are more similar to each other compared to facilities in other groups
- Clustering is most appropriate for grouping measures that have **natural gaps** across scores



Option 2: Reset with Clustering Methods

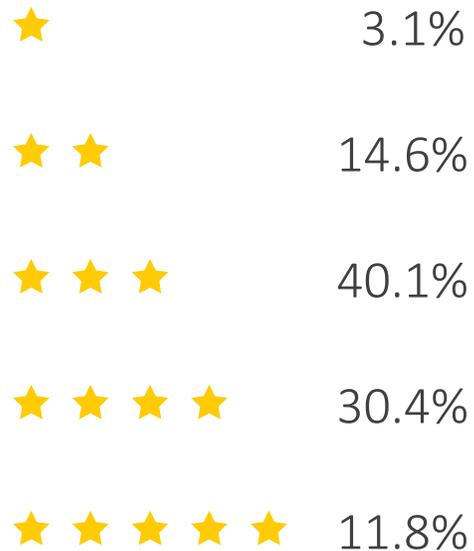
Types of clustering methods considered:

- **Hierarchical:** Begin grouping facilities two-at-a-time, then combine groups, based on how close their final scores are until there are 5 groups
- **K-Means:** Create 5 groups of facilities by minimizing the difference from the average score in each group and maximizing the difference from the average scores in other groups

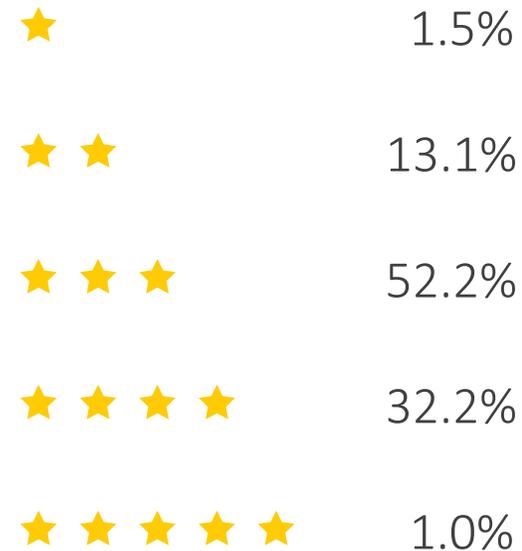
Option 2: Reset with Clustering Methods

Example using the October 2018 release:

Stars Based on Hierarchical Clustering:



Stars Based on K-Means Clustering:



Option 2: Reset with Clustering Methods

Some limitations of clustering methods:

Hierarchical Clustering:

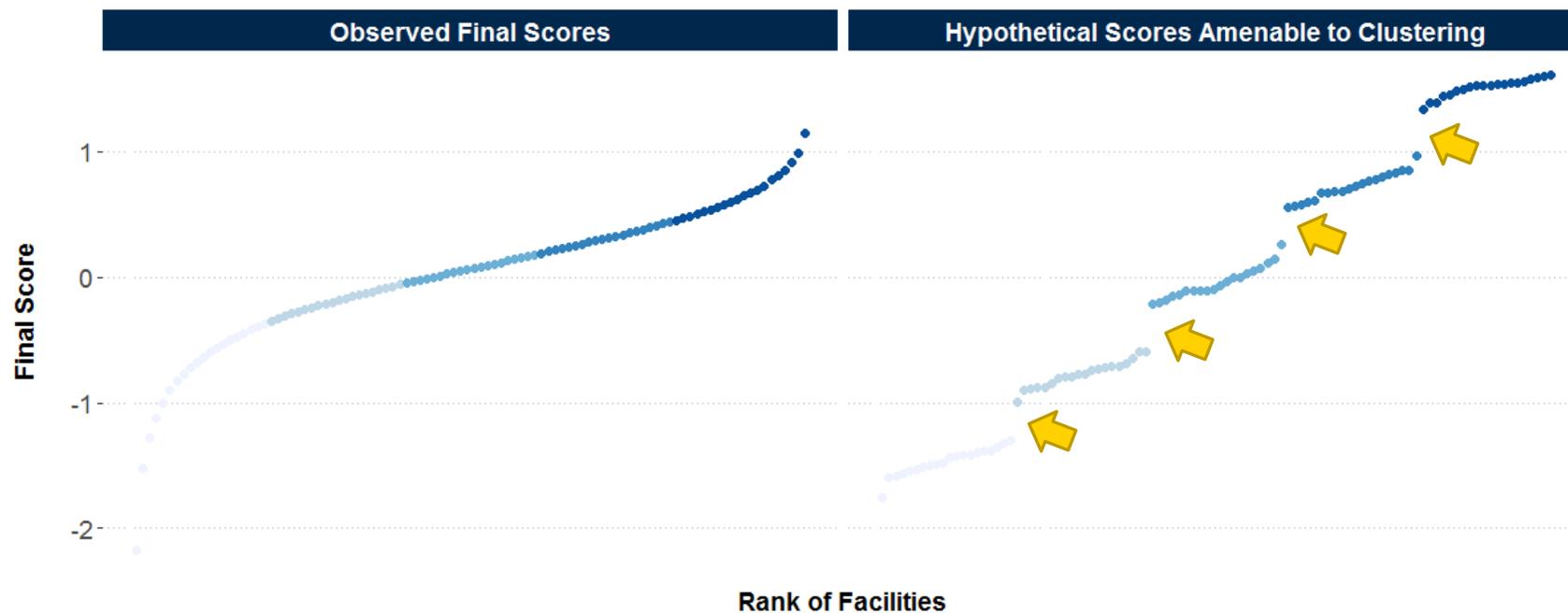
- Asymmetrical distribution of facilities compared to the fixed proportion methodology
- Groups are sensitive to which year of data are used and sensitive to outlying facilities
- Facilities that are much above or below average will likely form their own categories

K-Means Clustering:

- Very small proportions of 1- and 5-star categories, respectively.
- Method is sensitive to the choice of initial cluster centers as it optimizes cluster centers

Option 2: Reset with Clustering Methods

- Clustering is most appropriate for grouping facilities based on measures with **natural gaps**
- Clustering algorithms are shown to be unstable when categorizing continuous measures



Questions

ANY TEP QUESTIONS ON THE CONTENT PRESENTED?

In-Person TEP Meeting Discussion Topics

Options for Resetting the Star Ratings

- Discuss the options to reset to a pre-specified distribution, clustering, or other potential approaches.
- Discuss potential DFC display options or other tools that could assist patients with interpretation during and after DFC Star Ratings resetting

Re-Weighting the Star Rating Domains

- Weight domains by proportion of variance explained by each domain
- Reweight only Domain 3 in the Star Ratings

Reweighting the Star Rating Domains

POTENTIAL OPTIONS

Background

- DFC Star Ratings use factor analysis to group quality measures into domains based on relatedness
- The goal is to avoid domination of the Star Ratings by a single measure
- Factor analysis on the current measure set identifies three domains
- Individual measures within domains are currently averaged with equal weight
- The three domains are equally weighted when calculating a final score

What is Driving the Star Rating Shift?

Mean (SD) domain scores by release year before rebaselining:

	Oct. 2015	Oct. 2016	Apr. 2018
Domain 1 (SHR, SMR, STrR)	0.00 (0.69)	0.08 (0.69)	0.06 (0.70)
Domain 2 (Fistula, Catheter)	0.00 (0.85)	0.03 (0.87)	0.04 (0.86)
Domain 3 (Total Kt/V, Hypercalcemia)	0.00 (0.74)	0.36 (0.57)	0.82 (0.46)

- Mean domain scores increased the most for Domain 3
- May need to consider reweighting domains to lessen the impact of one domain on shifting and to stabilize the change of distribution or when some domains have reached the top-performance level

Re-weighting Options

Option 1: Weight based on Proportion of Variance Explained

- Weight domains based on proportion of variance explained for the data
- Weight the individual measures based on their contributions to each domain

Option 2: Downweight Domain 3

- Downweight Domain 3, containing the two measures with historically highest achievement
- Consider removal of Domain 3 measures (Total Kt/V and Hypercalcemia)

Example: Equal Weight vs. 50% Weight

Star Rating Agreement for Equal Weight vs. 50% Weight of Domain 3 (Total Kt/V, Hypercalcemia)

Equal Weight	Downweight Domain 3					Total
	1	2	3	4	5	
1	559	98	0	0	0	657
2	98	1,010	206	0	0	1,314
3	0	206	2,247	175	1	2,629
4	0	0	176	1,061	77	1,314
5	0	0	0	78	579	657
Total	657	1,314	2,629	1,314	657	6,571

Weighting of Measures or Domains

- Weighting adjusts influence of measures:
 - Weight measures and/or domains empirically or based on expert opinion
 - Can be applied independently or along with resetting
 - Individual facility changes in star ratings are expected with reweighting
 - Weighting can reduce or eliminate the impact of measures that have very high achievement, allowing facilities to concentrate on other Star Rating domains.
- At the in-person TEP meeting, we will solicit TEP feedback on reweighting the Star Ratings

Upcoming In-Person TEP Meeting

JUNE 6, 2019 FROM 8:30 AM – 4:00 PM, EDT

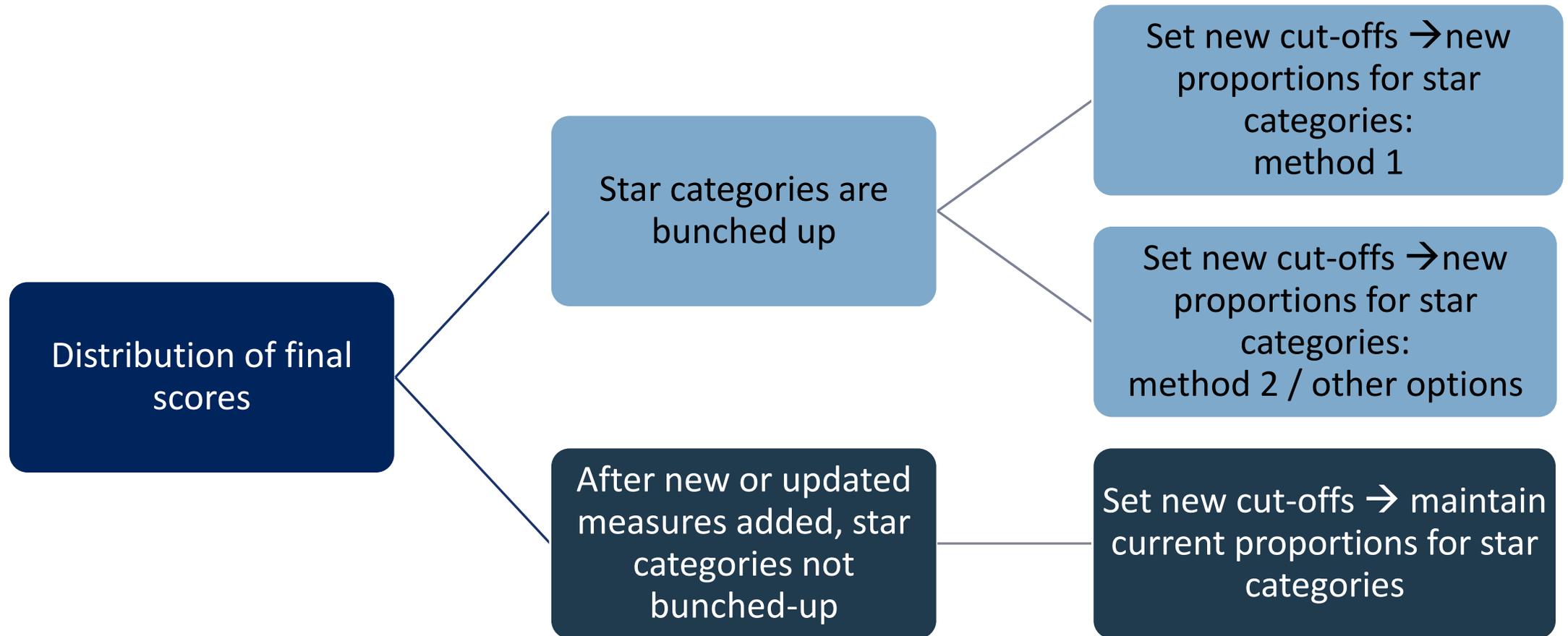
Public Comment Period

4:55 – 5:00 PM, EST

End of Teleconference Call #2

Appendix

Rebaselining star rating cutoffs for final scores



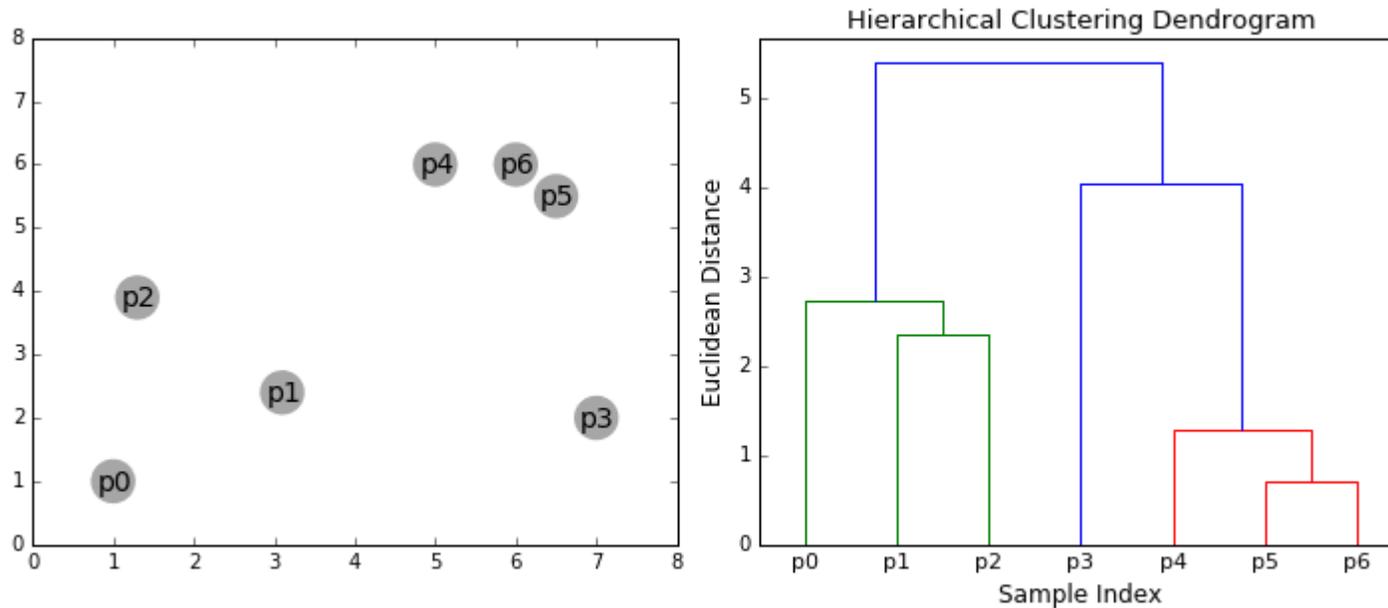
Option 2: Reset with Clustering Methods

Hierarchical Clustering (Ward's Method):

- Clusters are formed iteratively. Specifically, this method pairs individual facilities, then groups of facilities, until the desired number of groups is achieved
- This type of clustering does not produce one unique set of five clusters, but rather a series of partitions until five groupings are created
- Clustering criteria minimizes the total within-cluster variance (error sums of squares) between facilities' final scores while maximizing the between-cluster variance
- The distance used to implement Ward's method is Euclidian

Option 2: Reset with Clustering Methods

Hierarchical Clustering (Ward's Method):



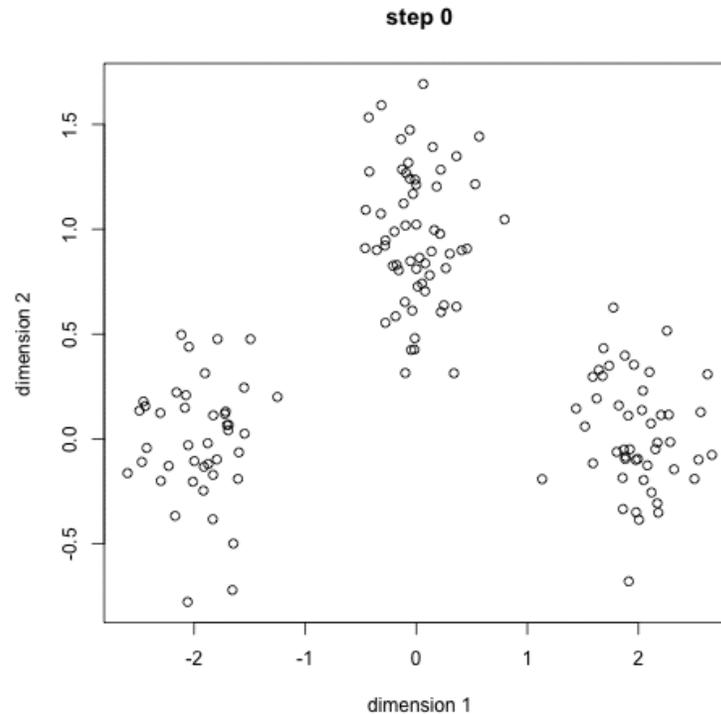
Option 2: Reset with Clustering Methods

Centroid-Based Clustering (K-Means):

- K-means clustering is currently used in the calculation of the Hospital Compare Star Rating
- The number of clusters must be pre-specified, as well as the initial centers (seed) of the clusters which are often chosen randomly, in this case from the final scores
- We allowed the initial cluster centers to be chosen by the SAS default options
- The algorithm then iteratively assigns facilities to clusters and re-calculate cluster centers until the difference in final scores within a category are minimized

Option 2: Reset with Clustering Methods

Centroid-Based Clustering (K-Means):



Appendix K. In-Person Meeting Presentation

The In-person Meeting Presentation is provided on the next several pages.



Dialysis Facility Compare (DFC) Star Rating TEP

IN-PERSON MEETING

JUNE 6, 2019

8:30 AM – 4:00 PM EDT

In Person TEP Agenda

8:30 – 9:00	Registration
9:00 – 9:15	TEP Introductions and Disclosure of Conflicts of Interest (TEP Chairs, TEP Members, CMS Representatives, UM-KECC Facilitators)
9:15 – 9:30	Overview of Objectives and Agenda
9:30 – 10:30	Weighting Domains and/or Measures; Brief Recap of Resetting Examples
10:30 – 10:45	Break
10:45 – 12:00	TEP Discussion - Is it time to Reset? What Method?
12:00 – 12:45	Lunch
12:45 – 2:30	Continuation of Resetting Discussion; How to Assist DFC Users during Transition?
2:30 – 2:45	Break
2:45 – 3:45	TEP Discussion and Recommendations
3:45 – 4:00	Public Comment Period

UM-KECC Star Rating Team

- Yi Li, PhD
- Joseph Messana, MD
- Claudia Dahlerus, PhD
- Richard Hirth, PhD
- Peisong Han, PhD
- Casey Parrotte, PMP
- Jennifer Sardone, BA
- Lan Tong, MPH
- Karen Wisniewski, MPH
- Wolf Gremel, MS
- Jingya Gao, MS
- Stephen Salerno, MS
- Brandon Frye, BA
- **Contact Person:**
Jordan Affholter, BA
affjorda@med.umich.edu

CMS Representatives

Centers for Medicare & Medicaid Services, Division of Quality Measurement:

- Joel Andress, PhD · Measure Development Subject Matter Expert
- Jesse Roach, MD · Nephrologist, ESRD Measures Development Lead
- Golden Horton, MS · Dialysis Facility Compare Lead

TEP Member Introductions

AND CONFLICT OF INTEREST DISCLOSURES

TEP Members

TEP Co-Chairs:

- Paul Conway, BA

Board of Directors Member, Chair of Public Policy and Global Affairs · *American Association of Kidney Patients*

- Catherine Sugar, PhD

Professor, Departments of Biostatistics, Statistics & Psychiatry · *University of California, Los Angeles*
Director · *Semel Institute Statistics Core, University of California, Los Angeles*

TEP Members:

- Mark Andaya, MS, RN

Director of Quality Assurance and Performance Improvement · *The Rogosin Institute*

- Andrew Conkling, BS

Vice President · *Dialysis Patient Citizens*

TEP Members

- Lorien Dalrymple, MD, MPH

Vice President of Epidemiology and Research · *Fresenius Medical Care North America*
Volunteer Clinical Faculty Associate Professor · *University of California, Davis*

- Sharon Dickson, RN, MSN

Regional Quality Manager · *Fresenius Kidney Care*

- Dawn Edwards

Patient Advocate · *NxStage*, Health Ambassador · *The Rogosin Institute*

- Derek Forfang

Patient Advocate/Chair · *Forum of ESRD Networks*, *National Kidney Foundation*

- Monica Fox

Outreach Associate/Patient Advocate · *National Kidney Foundation of Illinois*

TEP Members

- Lonnie Green

Patient Advocate/Subject Matter Expert · *IPRO ESRD Network 6, National Patient and Family Engagement – Learning Action Network (NPFE-LAN)*

- Mark Johnson

Patient Advocate/Subject Matter Expert · *ESRD Network 12, National Patient and Family Engagement – Learning Action Network (NPFE-LAN)*

- Mark Joseph, MD

Pediatric Nephrologist · *Pediatric Kidney Disease and Hypertension Centers*
Medical Director · *Phoenix Pediatric Dialysis Center*

- Richard Knight, MBA

President · *American Association of Kidney Patients*, Adjunct College Instructor · *Bowie State University*

- Mahesh Krishnan, MD, MPH, MBA

Group Vice President of Research and Development · *DaVita Kidney Care*

TEP Members

- Michael “Jack” Lennon, MBA

Executive Director · *Improving Renal Outcomes Collaborative*

- Nicole Stankus, MD, MSc

Associate Professor of Medicine / Nephrology · *University of Chicago*
Medical Director · *DaVita Stony Island Dialysis Center*

- Caprice Vanderkolk, RN, MS, BC-NE

Manager, Adult and Pediatric Dialysis Program · *University of Minnesota/Fairview Hospitals*

- Curtis Warfield, BS, MS

Patient Reviewer · *National Kidney Foundation*, Senior Quality Analyst · *State of Indiana*

- David White

Patient Advocate/Board of Directors Member · *American Association of Kidney Patients*
Healthcare Consultant · *Quality Insights Renal Network 5*

Risk Adjustment for Standardized Measures

	SMR	SHR	STrR	SRR	SFR
Age	✓	✓	✓	✓	✓
Sex	✓	✓		✓	✓
Race	✓				
Ethnicity	✓				
Diabetes as Cause of ESRD	✓	✓	✓	✓	✓
Duration of ESRD	✓	✓	✓	✓	✓
Nursing Home Status in Previous Year	✓	✓	✓		✓
BMI at Incidence	✓	✓	✓	✓	✓
Calendar Year	✓	✓	✓		
Comorbidities at Incidence	✓	✓	✓	✓	✓
Prevalent Comorbidities	✓	✓		✓	✓
Inability to Ambulate/Transfer					✓
Nephrologist Care Prior to ESRD					✓

Discussion of Weighting

What is Driving the Star Rating Shift?

Mean (SD) domain scores by release year

	Oct. 2015	Oct. 2016	Apr. 2018
Domain 1 (SHR, SMR, STrR)	0.00 (0.69)	0.08 (0.69)	0.06 (0.70)
Domain 2 (Fistula, Catheter)	0.00 (0.85)	0.03 (0.87)	0.04 (0.86)
Domain 3 (Total Kt/V, Hypercalcemia)	0.00 (0.74)	0.36 (0.57)	0.82 (0.46)

- Mean domain scores increased the most for Domain 3
- May need to consider reweighting domains to lessen the impact of one domain on shifting and to stabilize the change of distribution or when some domains have reached the top-performance level

Weighting Approaches

Directly Target Domains with Rapid Achievement

- Down-weight Domain 3, which contains the two measures with historically highest achievement

Weight based on Proportion of Variance Explained

- Weight domains based on proportion of variance they explain in the individual measures
- Weight the individual measures based on their contributions to each domain

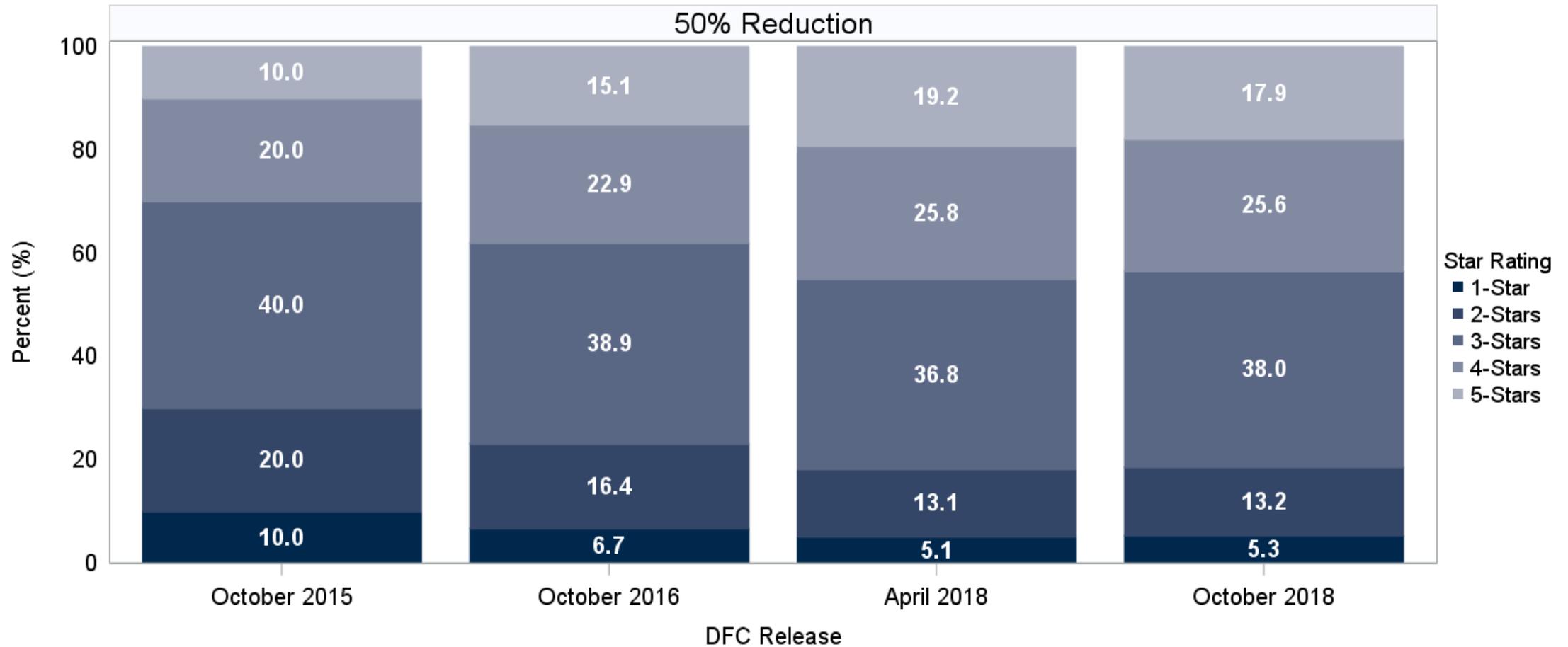
Down-Weight/Remove Domain 3 (Kt/V and Hypercalcemia)

- Down-weight Domain 3 which contains high achieving measures: Total Kt/V and Hypercalcemia
 1. 0% weight (removal) of Domain 3
 2. 50% of current Domain 3 weight

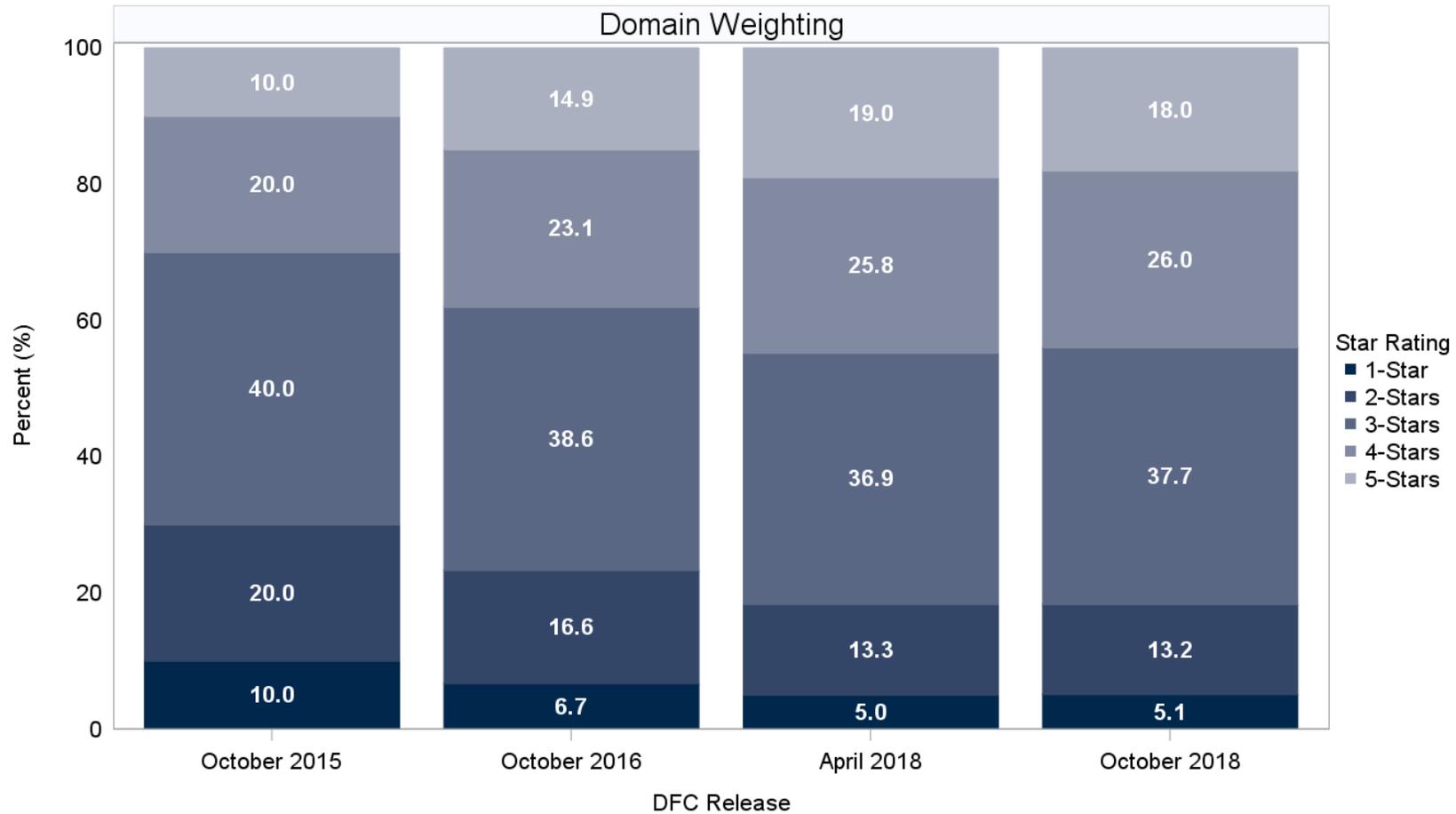
Hypothetical Star Rating Trends Across DFC Releases: Removal of Domain 3



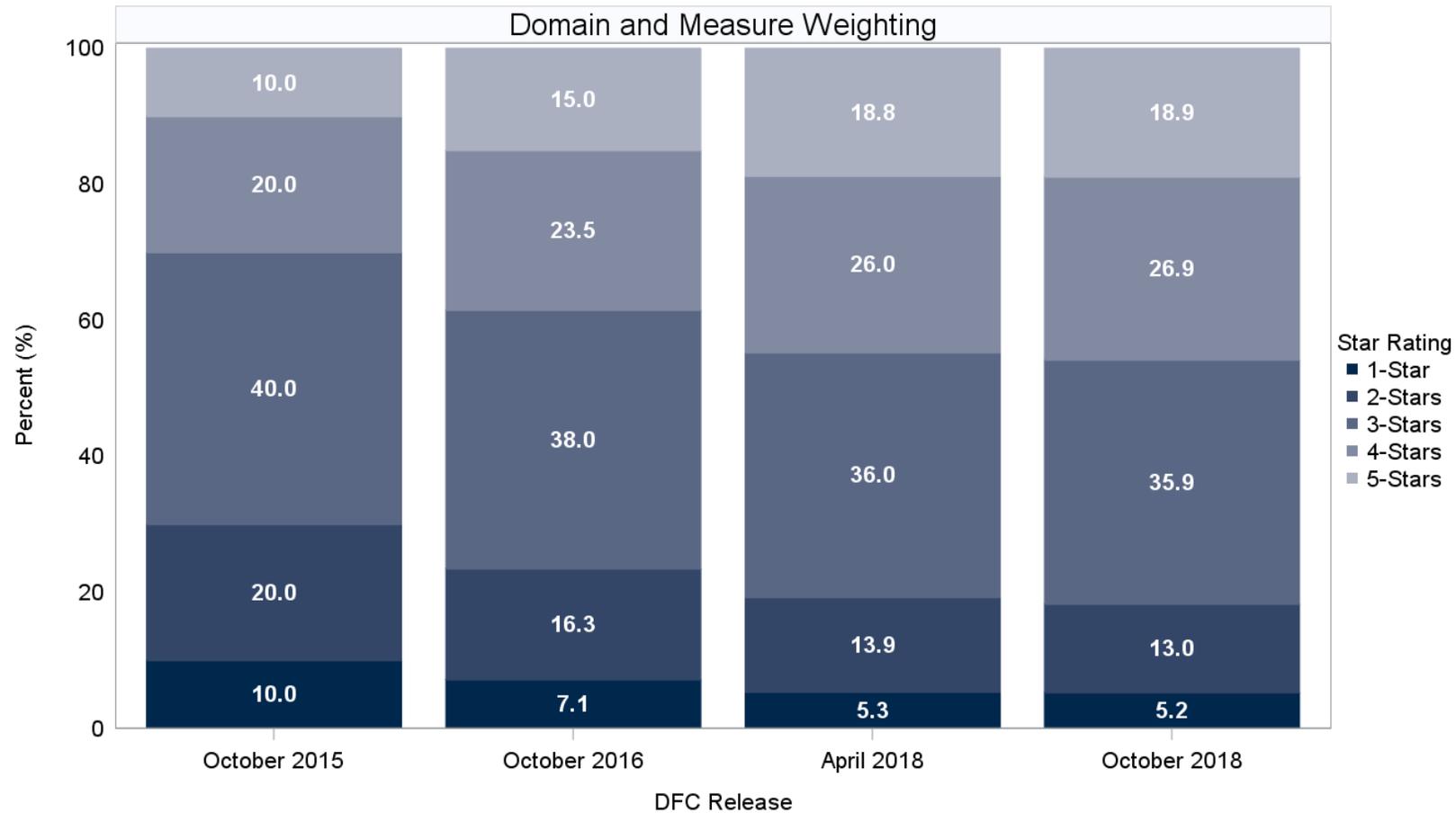
Hypothetical Star Rating Trends Across DFC Releases: 50% Weight of Domain 3



Hypothetical Star Rating Trends Across DFC Releases: Domain Weighting



Hypothetical Star Rating Trends Across DFC Releases: Measure + Domain Weighting



Discussion of Weighting Examples

TEP INPUT AND QUESTIONS

Star Rating Reset

RECAP AND TEP DISCUSSION

Resetting

Update scoring cutoffs and Star Rating distribution:

- **Reset** when the Star Rating may not differentiate facility-level performance (e.g. compression of the Star Ratings due to progressive shifts in facility performance)
- Defines new baseline scoring cutoffs for facilities to be rated
- Proportions of facilities in each Star Rating category are **reset to new numbers**

Reset to a Pre-Specified Distribution

Based on Fixed Proportions:

Directly choose the percent of facilities in each category and determine the cutoffs based on these proportions (e.g. the current 10-20-40-20-10 rule)

Based on Standard Deviations:

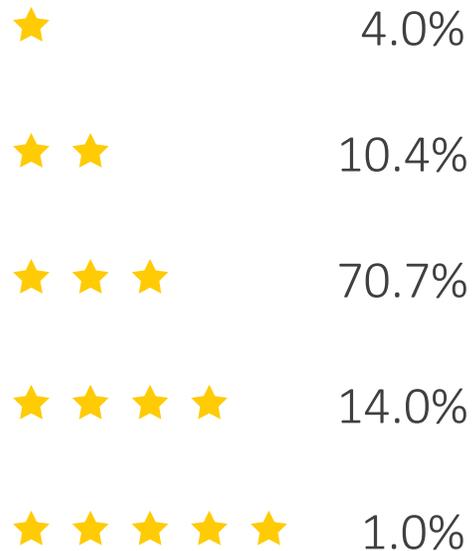
The Star Rating cutoffs are determined based on “how far” the facilities scores are (in terms of standard deviations) from the average national scores

E.g. $< -2 = 1\text{-Star}$, $-2 \text{ to } -1 = 2\text{-Stars}$, $-1 \text{ to } 1 = 3\text{-Stars}$, $1 \text{ to } 2 = 4\text{-Stars}$, $> 2 = 5\text{-Stars}$

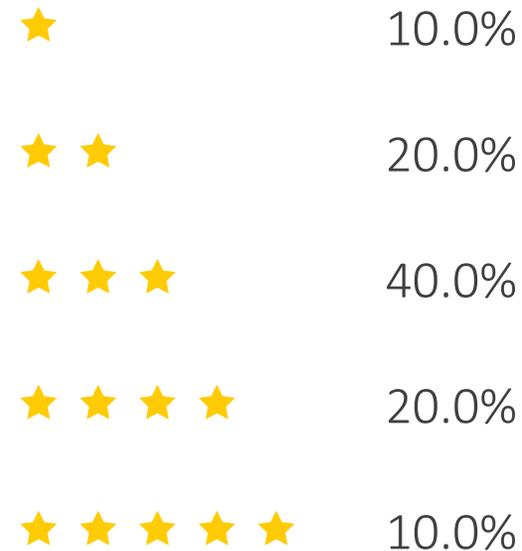
Reset to a Pre-Specified Distribution

Example using the October 2018 release:

Stars Based on Standard Deviations:



Stars Based on Fixed Proportions:



Reset to Fixed Proportions

October 2018 DFC Star Rating Distribution vs. Resetting to Fixed Proportions (10-20-40-20-10):

Calculation	1-Star	2-Stars	3-Stars	4-Stars	5-Stars
Reported on DFC (2018)	190 (2.89)	536 (8.16)	2,311 (35.17)	1,810 (27.55)	1,724 (26.24)
Reset Distribution	657 (10.00)	1,314 (20.00)	2,629 (40.00)	1,314 (20.00)	657 (10.00)

- 4,332 facilities (66%) would have experienced a 1-star decrease in rating if the Star Rating distribution were to have been reset to 10-20-40-20-10 in October 2018
- Only 5 facilities (0.07%) would have experienced a 2-star shift. The remaining 2,234 facilities (34%) would have received the same rating after the reset

Reset with Clustering Methods

What is clustering?

- Statistical method for grouping facilities that are more similar to each other compared to facilities in other groups

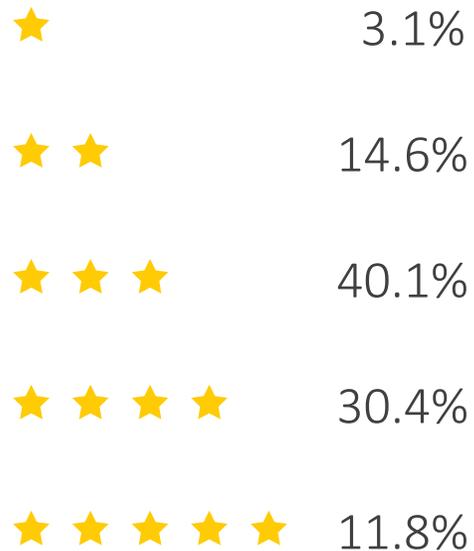
Types of clustering methods considered:

- **Hierarchical:** Begin grouping facilities two-at-a-time, then combine groups, based on how close their final scores are until there are 5 groups
- **K-Means:** Create 5 groups of facilities by minimizing the difference from the average score in each group and maximizing the difference from the average scores in other groups

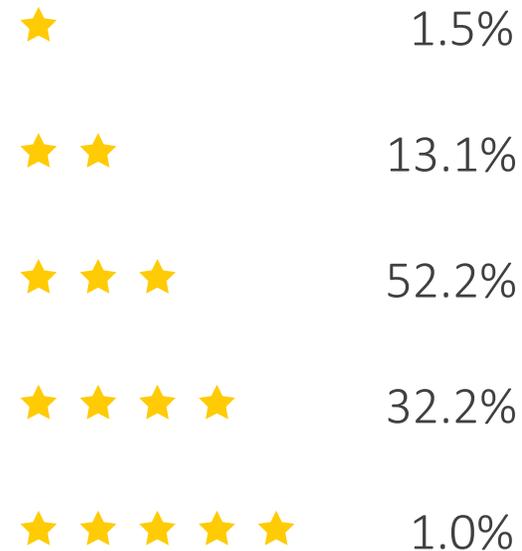
Reset with Clustering Methods

using the data for the October 2018 release

Stars Based on Hierarchical Clustering:



Stars Based on K-Means Clustering:



Reset with Clustering Methods

Some limitations of clustering methods to inform DFC Star Rating reset:

Hierarchical Clustering:

- Asymmetrical distribution of facilities compared to the fixed proportion methodology
- Groups are sensitive to which year of data are used and sensitive to outlying facilities
- Facilities that are much above or below average will likely form their own categories

K-Means Clustering:

- Very small proportions of 1- and 5-star categories, respectively
- Obtained clusters do not necessarily represent separable types, but rather an efficient grouping of the data

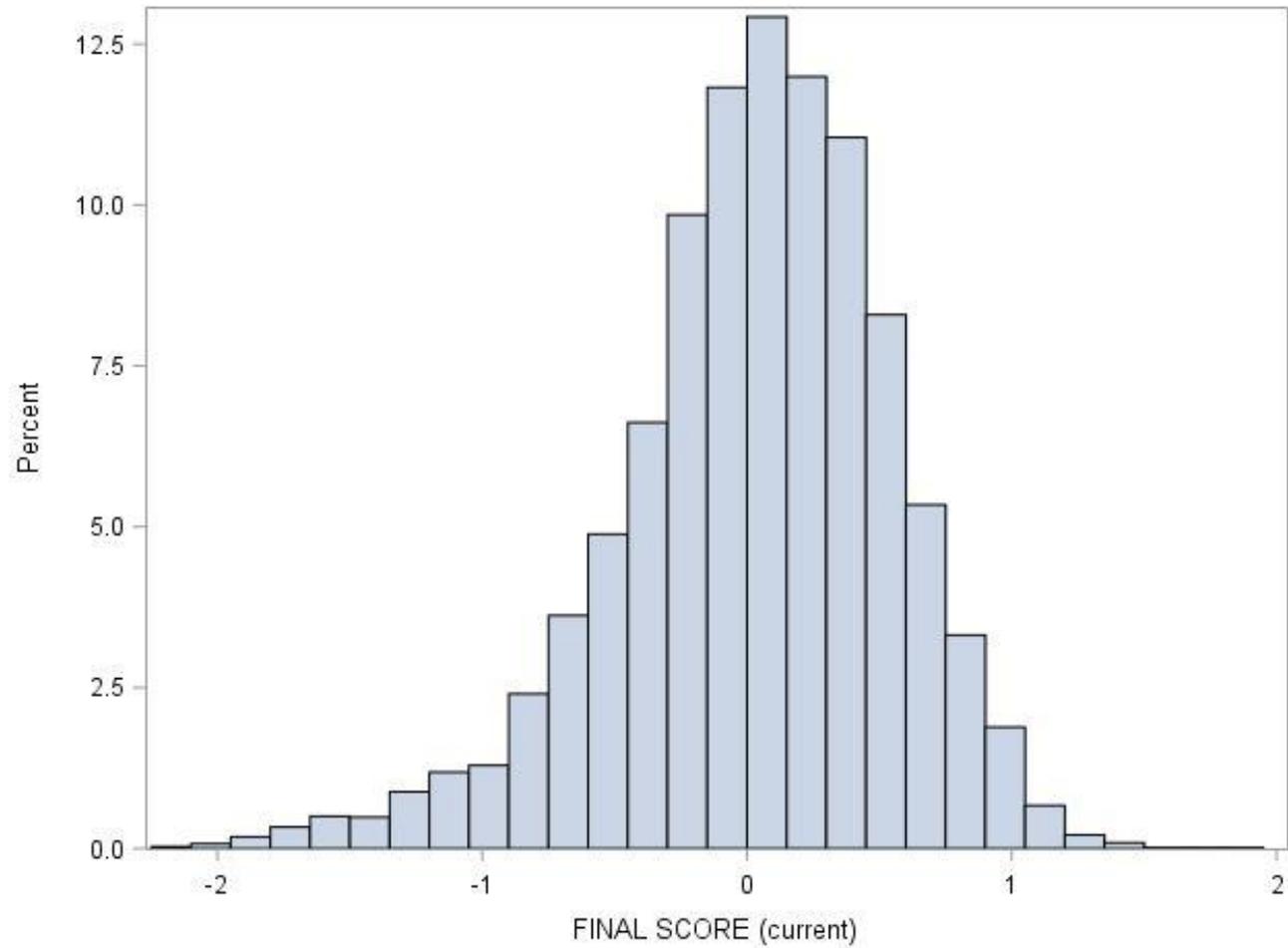
Main Topics to Discuss

- Is it time to reset the Star Ratings?
- How should we do it?
- How do we help DFC consumers interpret facility performance on Star Ratings during and immediately after the transition?

TEP Discussion on Resetting

Lunch

12:00 – 12:45 PM (EDT)



Questions

1. Should the Star Ratings be reset now?
2. Would you be willing to use 10-20-40-20-10 distribution as long as it was effectively explained?
3. If you answered no to #2, what is your recommended percentage for each category (1 star, 2 stars, 3 stars, 4 stars, 5 stars)? It must add up to 100%.

TEP Discussion on Resetting

Interpreting Star Ratings on DFC after a Reset

- After star ratings are re-set what tools on DFC can help consumers interpret facility performance on star ratings during and immediately after the transition?
- What type of information and displays would consumers find useful?
- We developed a few mock-ups to spark discussion

Mock-Up Example

General information

Survey of patients' experiences

Quality of patient care

EXAMPLE FACILITY

FACILITY ADDRESS

PHONE NUMBER

DISTANCE

[Add to my Favorites](#)
[Map and Directions](#)

Quality of patient care

- Get more information about the quality of patient care measures
- Get more information about how the quality of patient care measure data are collected and reported

Quality of patient care rating, facility percentile score and facility continuous final score

Hover Text: Medicare assigns the star ratings based on data that facilities report to Medicare. More stars mean that dialysis centers perform better on the quality of care measures shown on the website. Patient survey results are not included Quality of Patient Care Star Rating.

Hover Text: The percentage of dialysis facilities that your facility exceeded based on its performance ranking. The range of the score is 0 -100.

[View More Details](#)

Hover Text: This is a numeric representation of your facility's final score from which your facility's star rating is based on. The range of the score is 0 -100.

	<u>Quality of patient care rating</u>	EXAMPLE FACILITY Facility percentile score <i>Higher percentages are better</i>	EXAMPLE FACILITY Facility continuous final score <i>Higher scores are better</i>
Current Year	★●●●●	8%	6
Previous Year	★★●●●	7%	5

Mock-Up Examples

Facility A	<u>Quality of patient care rating</u>	Facility percentile score <i>Higher percentages are better</i>	Facility continuous final score <i>Higher scores are better</i>
Current Year	★★★★●	67%	71
Previous Year	★★★★●	66%	70

Facility B	<u>Quality of patient care rating</u>	Facility percentile score <i>Higher percentages are better</i>	Facility continuous final score <i>Higher scores are better</i>
Current Year	★★★★●●	52%	52
Previous Year	★★★★●	50%	51

Mock-Up Example – Drilldown Details

Example Percentile Cutoffs

Star Rating Category	Percentile Cutoffs in Current Year	Percentile Cutoffs in Previous Year
1 star	0%-8%	0%-5%
2 stars	9% - 18%	6% - 15%
3 stars	19% - 53%	16% - 45%
4 stars	54%-84%	46%-75%
5 stars	85%-100%	76%-100%

Example Facility Continuous Final Score Cutoffs

Star Rating Category	Facility Cutoffs in Current Year	Facility Cutoffs in Previous Year
1 star	0 - 6	0 - 4
2 stars	7 - 17	5 - 15
3 stars	18 - 53	16 - 50
4 stars	54 -80	51 -75
5 stars	81 -100	76 -100

Wrap-Up and Next Steps

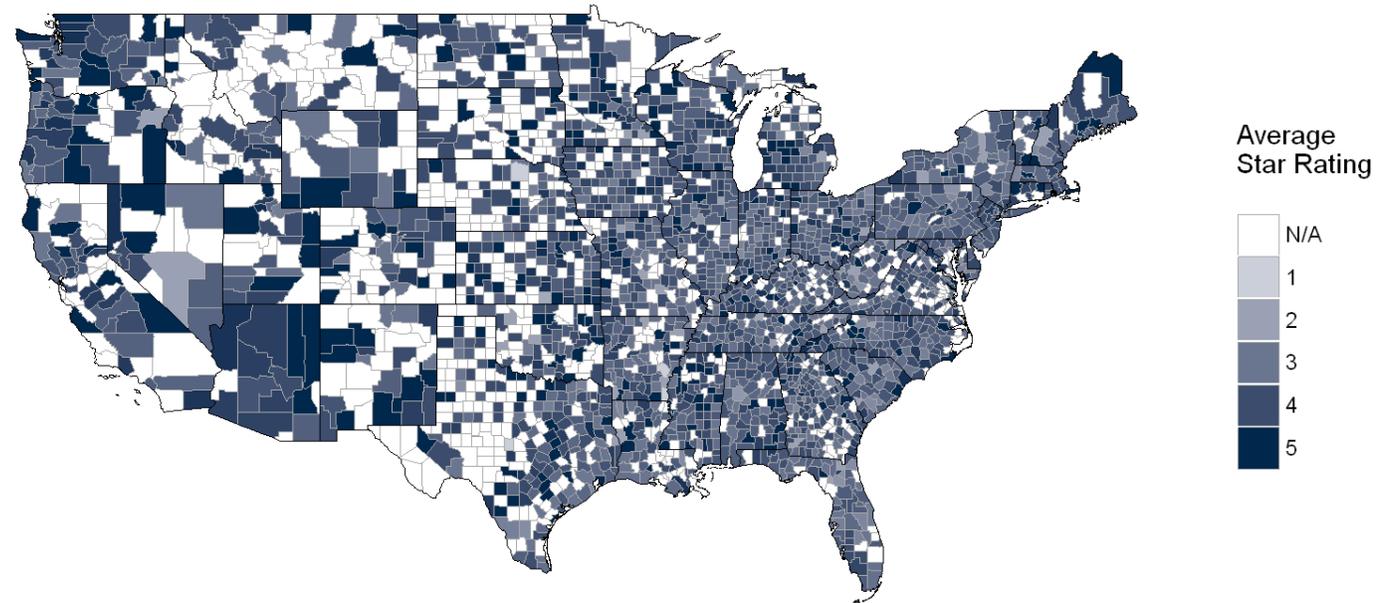
SUMMARY OF RECOMMENDATIONS AND DISCUSSION

Public Comment Period

3:45 – 4:00 PM, EDT

Appendix: Teleconference II Questions

Geographic Co-location of Dialysis Facilities



Geographic Co-location of Dialysis Facilities

The map displays the county-level average Star Rating as of the April 2018 DFC release. No clear pattern for the distribution of facility performance is apparent.

The histogram displays the distribution of the minimum distance a patient has to travel beyond their closest facility to the nearest higher rated facility:

- 65.3% of facilities are < 10 miles from a higher rated facility as of the April 2018 DFC release
- This proportion is up from 59.9% in October 2016 release and 54.4% in the October 2015
- As of April 2018, 41.8% of patients go to the facility closest to them, compared to 39.8% in October 2016 and 39.6% in October 2015

Exclusions for Vascular Access Measures

Exclusions that are implicit in the denominator definition include:

- Pediatric patients (<18 years old)
- Patients on Peritoneal Dialysis
- Patient-months for less than a complete reporting month at the same facility

Exclusions are applied to the denominator for patients with a catheter that have limited life expectancy:

- Patients under hospice care in the current reporting month
- Patients with metastatic cancer in the past 12 months
- Patients with end stage liver disease in the past 12 months
- Patients with coma or anoxic brain injury in the past 12 months

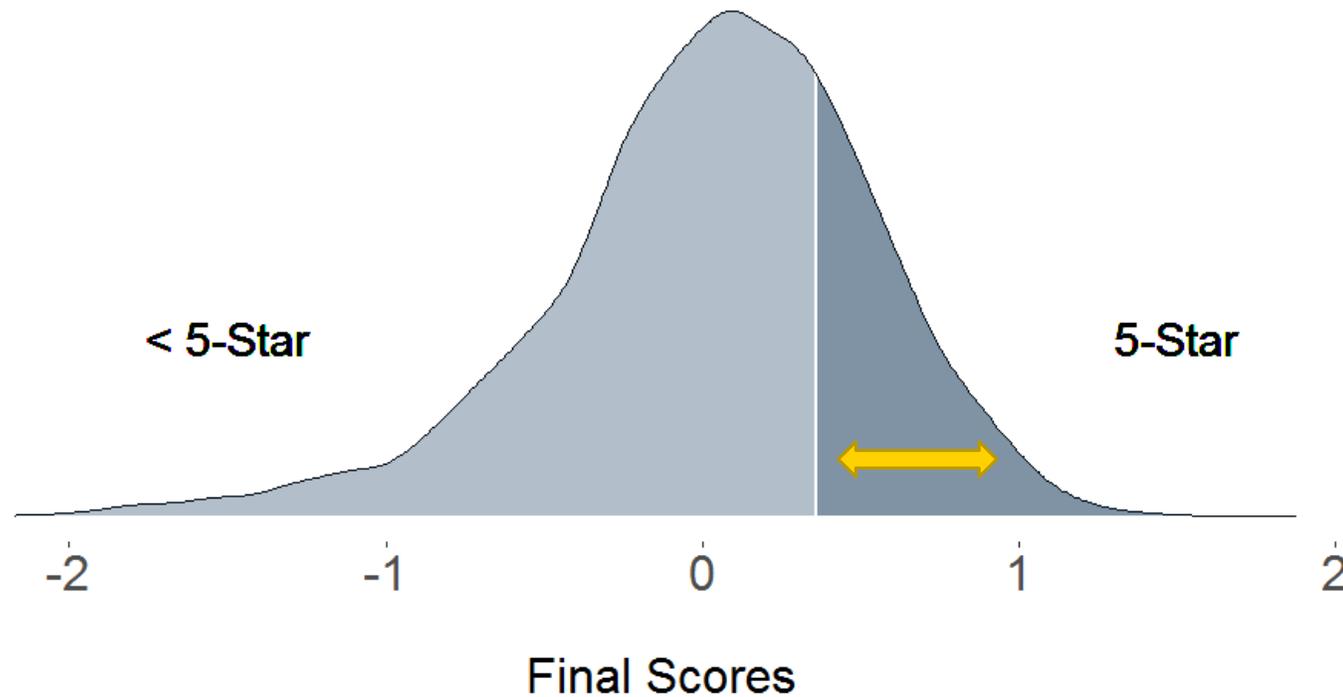
Appendix: Rebaselining and Resetting

Star Rating Distribution Shift



Variation of Final Scores within 5-Star

October 2018 Release



Considerable variation in facility performance beyond the 5-Star cutoff

Key Concepts

Rebaseline: Establish a new baseline year and rescore measures

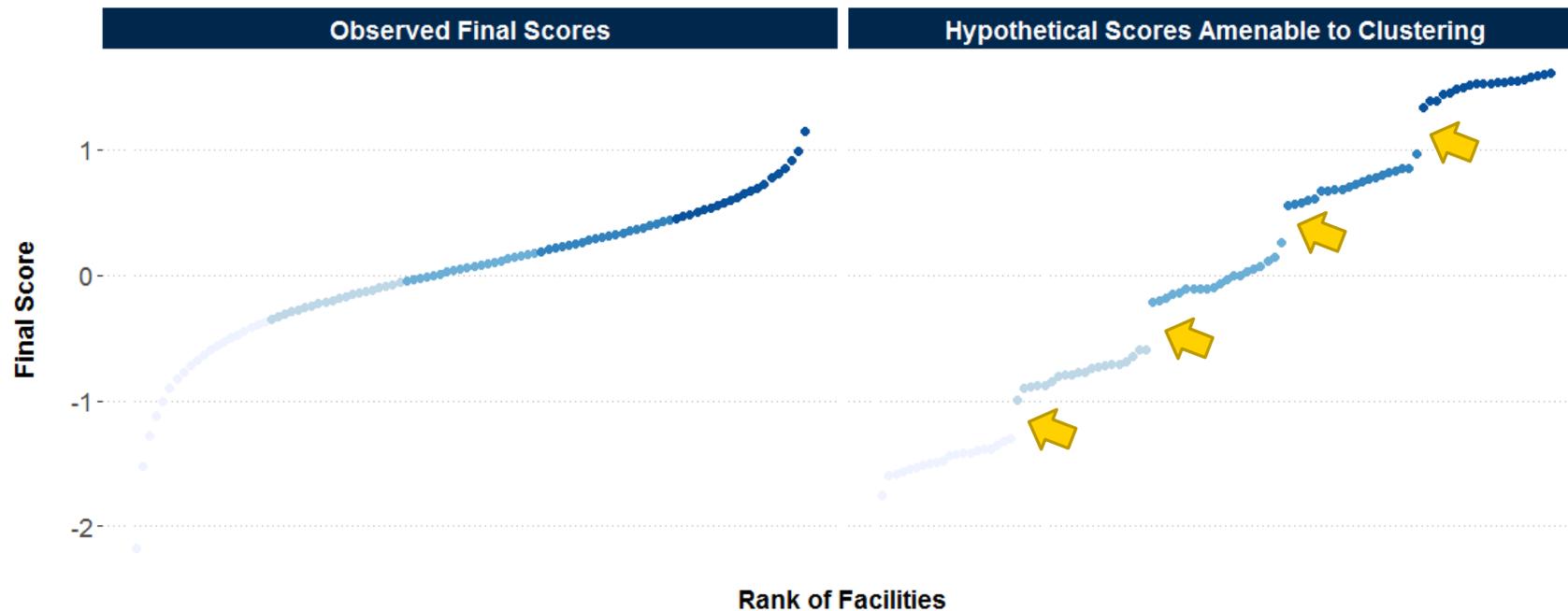
- Defines new baseline scoring cutoffs for facilities to be rated
- Star Rating proportions **remain unchanged compared to the last public release**

Reset: Update scoring cutoffs and Star Rating distribution

- Defines new baseline scoring cutoffs for facilities to be rated
- Star Rating proportions are **reset to improve the ability to differentiate facility performance**

Resetting with Clustering Methods

- Clustering is most appropriate for grouping facilities based on measures with **natural gaps**
- Hierarchical clustering is more unstable when categorizing continuous measures



Clustering

- Though methods such as k-means tend to be much more stable than hierarchical clustering, the results cannot necessarily be interpreted as “clusters,” as they do not necessarily represent separable “types” but rather the most “efficient” grouping of the data

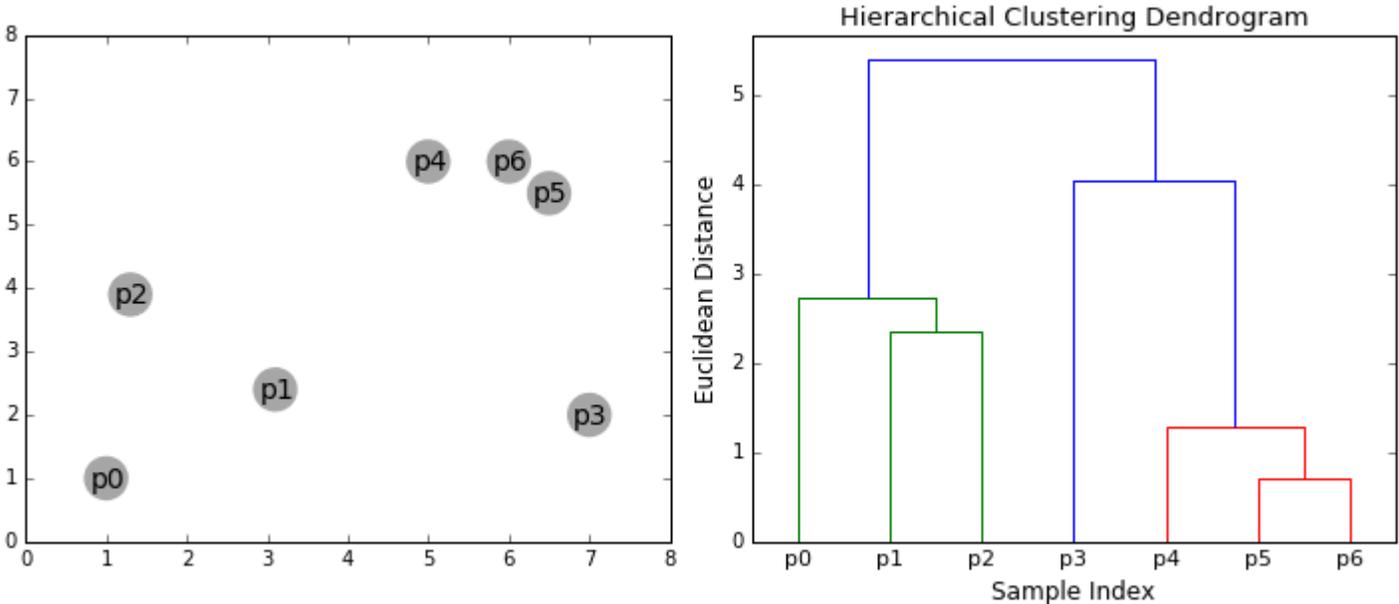
Reset with Clustering Methods

Hierarchical Clustering (Ward's Method):

- Clusters are formed iteratively. Specifically, this method pairs individual facilities, then groups of facilities, until the desired number of groups is achieved
- This type of clustering does not produce one unique set of five clusters, but rather a series of partitions until five groupings are created
- Clustering criteria minimizes the total within-cluster variance (error sums of squares) between facilities' final scores while maximizing the between-cluster variance
- The distance used to implement Ward's method is Euclidian

Example 2: Reset with Clustering Methods

Hierarchical Clustering (Ward's Method):



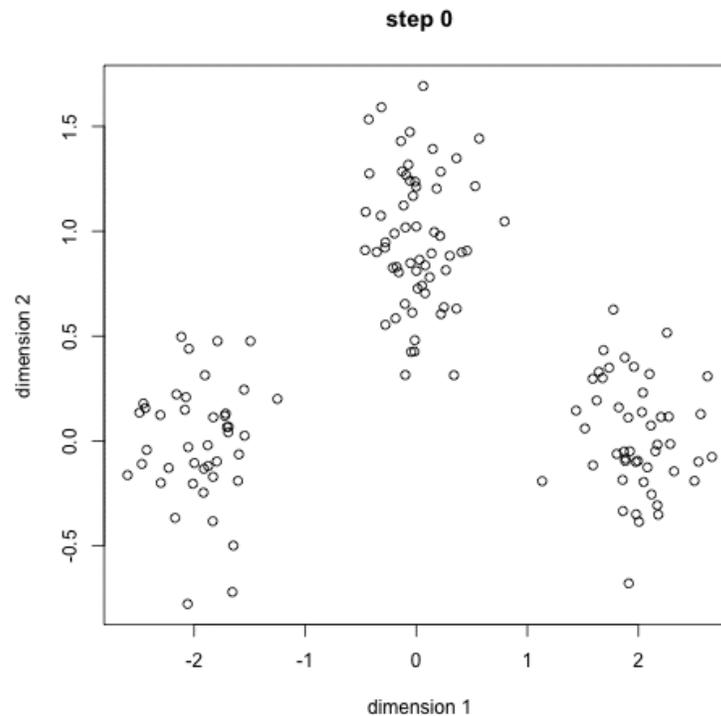
Reset with Clustering Methods

Centroid-Based Clustering (K-Means):

- The number of clusters must be pre-specified, as well as the initial centers (seed) of the clusters which are often chosen randomly, in this case from the final scores
- We allowed the initial cluster centers to be chosen by the SAS default options
- The algorithm then iteratively assigns facilities to clusters and re-calculate cluster centers until the difference in final scores within a category are minimized

Reset with Clustering Methods

Centroid-Based Clustering (K-Means):



Appendix: Weighting

Weighting of Measures or Domains

- Individual measures within domains are currently averaged with equal weight
- Domains are equally weighted when calculating final scores
- Weighting adjusts influence of measures:
 - Weight measures and/or domains empirically or based on expert opinion
 - Can be considered in conjunction with resetting

Rotated Factor Pattern from Factor Analysis

Measure	October 2015 Baseline			April 2018 Baseline		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
	0.78	0.65	0.34	0.92	0.72	0.63
SHR	0.55	0.13	0.06	0.63	0.10	0.10
SMR	0.36	0.12	0.11	0.33	0.01	0.11
SRR	-	-	-	0.51	0.05	0.03
STrR	0.53	0.10	0.01	0.32	0.07	0.03
Fistula	0.15	0.54	0.13	0.09	0.57	0.20
Catheter	0.15	0.55	0.11	0.08	0.56	0.17
Total Kt/V	0.15	0.09	0.39	0.17	0.18	0.52
Hypercalcemia	-0.01	0.07	0.37	0.04	0.19	0.52

Values can be interpreted as the correlation between the individual measures and each factor

Weight by Variance Explained

Calculated Weights Based on Factor Analysis Results

Measure	October 2015 Baseline			April 2018 Baseline		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
	0.44	0.37	0.19	0.41	0.32	0.27
SHR	0.38	0	0	0.35	0	0
SMR	0.25	0	0	0.18	0	0
SRR	-	-	-	0.29	0	0
STrR	0.37	0	0	0.18	0	0
SFR	0	0.50	0	0	0.50	0
LTCR	0	0.50	0	0	0.50	0
Total Kt/V	0	0	0.51	0	0	0.50
Hypercalcemia	0	0	0.49	0	0	0.50

Total Kt/V Quartiles by DFC Release

	100% Max.	75% Q3	50% Med.	25% Q1	0% Min.
Oct. 15	100.00	93.81	90.76	86.83	0.00
Oct. 15*	100.00	94.01	91.06	87.00	0.00
Oct. 16	100.00	96.00	93.10	89.49	12.44
Oct. 16*	100.00	96.15	93.37	89.51	0.00
Apr. 18	100.00	97.93	96.27	93.76	18.32
Oct. 18	100.00	98.08	96.62	94.34	0.00

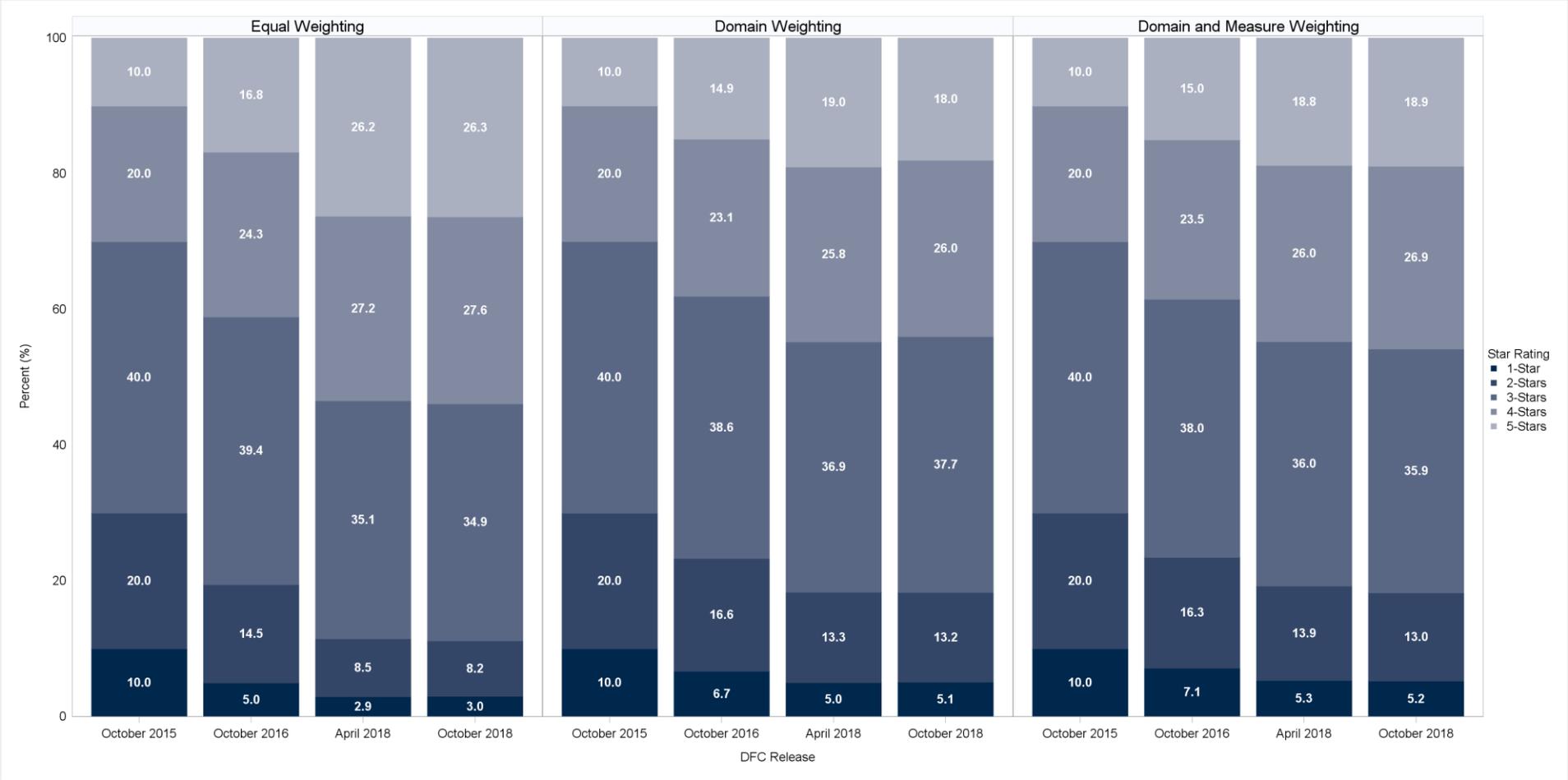
* Recalculated Using Current Measure Definitions

Hypercalcemia Quartiles by DFC Release

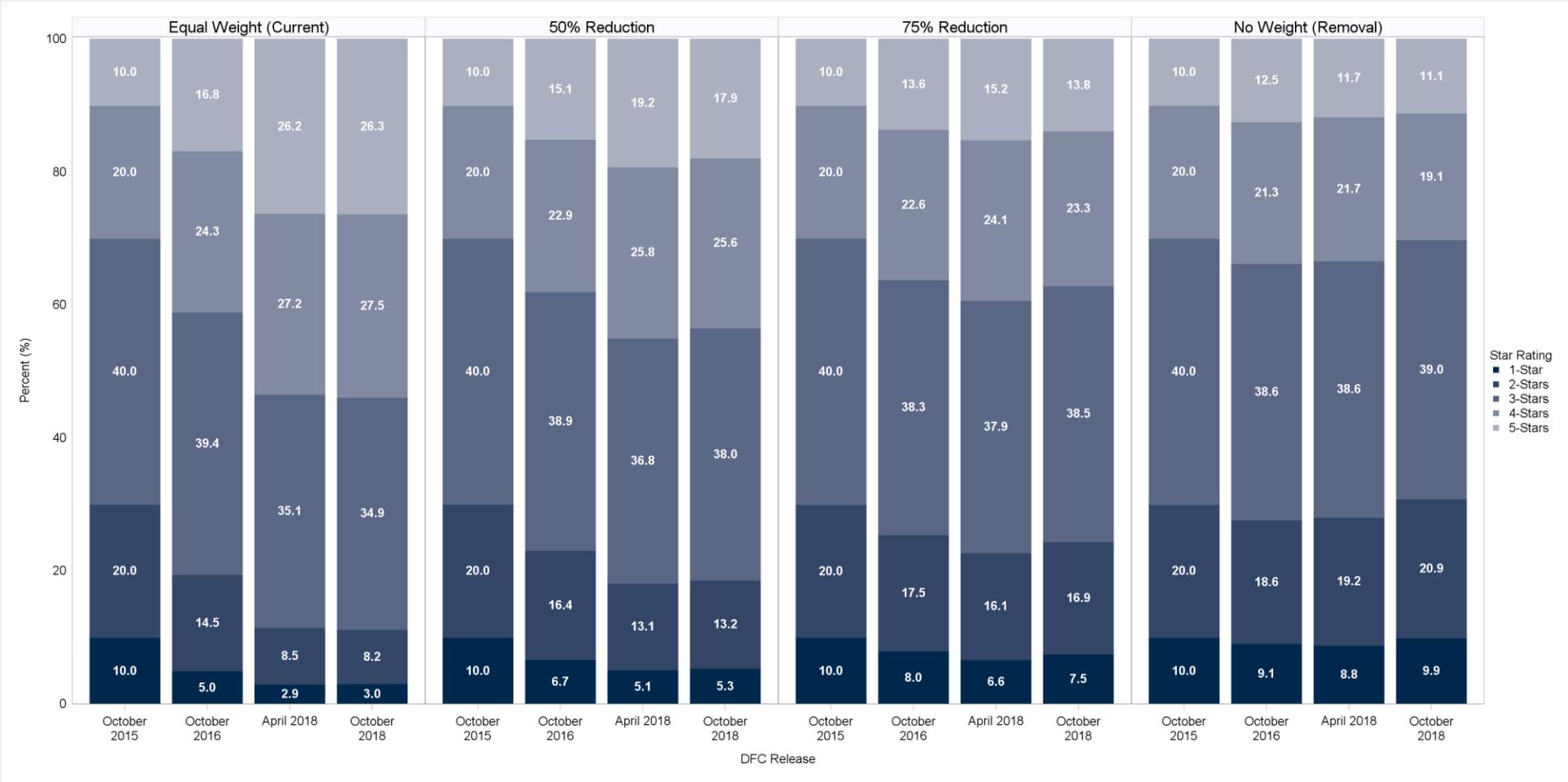
	100% Max.	75% Q3	50% Med.	25% Q1	0% Min.
Oct. 15	56.32	3.13	1.37	0.47	0.00
Oct. 15*	40.19	4.22	2.21	1.00	0.00
Oct. 16	35.16	1.94	1.03	0.43	0.00
Oct. 16*	35.79	3.02	1.93	1.12	0.00
Apr. 18	21.88	1.04	0.44	0.05	0.00
Oct. 18	100.00	2.76	1.43	0.52	0.00

* Recalculated Using Current Measure Definitions

Historic Star Ratings Weighting by Variance



Historic Star Ratings Weighting Domain 3



Appendix L. Post-TEP Teleconference Call Presentation

The Post-TEP Teleconference Presentation is provided on the next several pages.



Dialysis Facility Compare (DFC) Star Rating TEP

POST TEP TELECONFERENCE CALL

AUGUST 21, 2019

1:00PM-3:00PM EDT

Post TEP Teleconference # 1 Call Agenda

1. TEP Discussion on Weighting Options for Domain 3 (115 minutes)
2. Public Comment (5 minutes)

UM-KECC Star Rating Team

- Yi Li, PhD
- Joseph Messana, MD
- Claudia Dahlerus, PhD
- Richard Hirth, PhD
- Peisong Han, PhD
- Casey Parrotte, PMP
- Jennifer Sardone, BA
- Lan Tong, MPH
- Karen Wisniewski, MPH
- Wolf Gremel, MS
- Jingya Gao, MS
- Stephen Salerno, MS
- Brandon Frye, BA
- **Contact Person:**
Jordan Affholter, BA
affjorda@med.umich.edu

CMS Representatives

Centers for Medicare & Medicaid Services, Division of Quality Measurement:

- Joel Andress, PhD · Measure Development Subject Matter Expert
- Jesse Roach, MD · Nephrologist, ESRD Measures Development Lead
- Golden Horton, MS · Dialysis Facility Compare Lead

TEP Members

TEP Co-Chairs:

- Paul Conway, BA

Board of Directors Member, Chair of Public Policy and Global Affairs · *American Association of Kidney Patients*

- Catherine Sugar, PhD

Professor, Departments of Biostatistics, Statistics & Psychiatry · *University of California, Los Angeles*
Director · *Semel Institute Statistics Core, University of California, Los Angeles*

TEP Members:

- Mark Andaya, MS, RN

Director of Quality Assurance and Performance Improvement · *The Rogosin Institute*

- Andrew Conkling, BS

Vice President · *Dialysis Patient Citizens*

TEP Members

- Lorien Dalrymple, MD, MPH

Vice President of Epidemiology and Research · *Fresenius Medical Care North America*
Volunteer Clinical Faculty Associate Professor · *University of California, Davis*

- Sharon Dickson, RN, MSN

Regional Quality Manager · *Fresenius Kidney Care*

- Dawn Edwards

Patient Advocate · *NxStage*, Health Ambassador · *The Rogosin Institute*

- Derek Forfang

Patient Advocate/Chair · *Forum of ESRD Networks*, *National Kidney Foundation*

- Monica Fox

Outreach Associate/Patient Advocate · *National Kidney Foundation of Illinois*

TEP Members

- Lonnie Green

Patient Advocate/Subject Matter Expert · *IPRO ESRD Network 6, National Patient and Family Engagement – Learning Action Network (NPFE-LAN)*

- Mark Johnson

Patient Advocate/Subject Matter Expert · *ESRD Network 12, National Patient and Family Engagement – Learning Action Network (NPFE-LAN)*

- Mark Joseph, MD

Pediatric Nephrologist · *Pediatric Kidney Disease and Hypertension Centers*
Medical Director · *Phoenix Pediatric Dialysis Center*

- Richard Knight, MBA

President · *American Association of Kidney Patients*, Adjunct College Instructor · *Bowie State University*

- Mahesh Krishnan, MD, MPH, MBA

Group Vice President of Research and Development · *DaVita Kidney Care*

TEP Members

- Michael “Jack” Lennon, MBA

Executive Director · *Improving Renal Outcomes Collaborative*

- Nicole Stankus, MD, MSc

Associate Professor of Medicine / Nephrology · *University of Chicago*
Medical Director · *DaVita Stony Island Dialysis Center*

- Caprice Vanderkolk, RN, MS, BC-NE

Manager, Adult and Pediatric Dialysis Program · *University of Minnesota/Fairview Hospitals*

- Curtis Warfield, BS, MS

Patient Reviewer · *National Kidney Foundation*, Senior Quality Analyst · *State of Indiana*

- David White

Patient Advocate/Board of Directors Member · *American Association of Kidney Patients*
Healthcare Consultant · *Quality Insights Renal Network 5*

Domain 3 Weighting Options

Weighting of Measures or Domains

- Individual measures within domains are currently averaged with equal weights
- Domains are equally weighted when calculating final scores
- Weighting adjusts influence of measures and can be done empirically or based on expert opinions

What is Driving the Star Rating Shift?

Mean (SD) domain scores by release year

	Oct. 2015	Oct. 2016	Apr. 2018
Domain 1 (SHR, SMR, STrR)	0.00 (0.69)	0.08 (0.69)	0.06 (0.70)
Domain 2 (Fistula, Catheter)	0.00 (0.85)	0.03 (0.87)	0.04 (0.86)
Domain 3 (Total Kt/V, Hypercalcemia)	0.00 (0.74)	0.36 (0.57)	0.82 (0.46)

- Domain 3 contains two measures which have achieved historically the highest performance
- Mean domain scores increased the most for Domain 3
- Down-weighting could typically be applied if maximal performance is reached in a specific domain
- Consider down-weighting Domain 3 to lessen its impact on the shift in Star Rating and to stabilize changes in the distribution

Domain 3 (Kt/V and Hypercalcemia) Weighting Options

- During the previous TEP Discussions, the following options were presented for weighting Domain 3 which contains high achieving measures: Total Kt/V and Hypercalcemia
 1. 0% weight (removal of Domain 3 from the DFC Star Rating Calculation)
 2. Down-weight Domain 3 at 50% of its current weight
 3. Maintain Domain 3 its current weight (100%)

Recap of TEP Discussion on the Weighting Options

- Some TEP members stated that if a measure is highly skewed, it may no longer provide useful information and should be considered for removal from the DFC Star Ratings.
- Other TEP members expressed interest in keeping the Domain 3 measures in the DFC Star Ratings, but down-weight them to 50% of their original weight, to continue to hold facilities accountable for these outcomes.
- Several patient TEP members said Kt/V is an important measure, not as much as a clinical value, but as a measure of quality of life and how patients feel and function.
- An official TEP recommendation was not reached during the in-person TEP meeting.

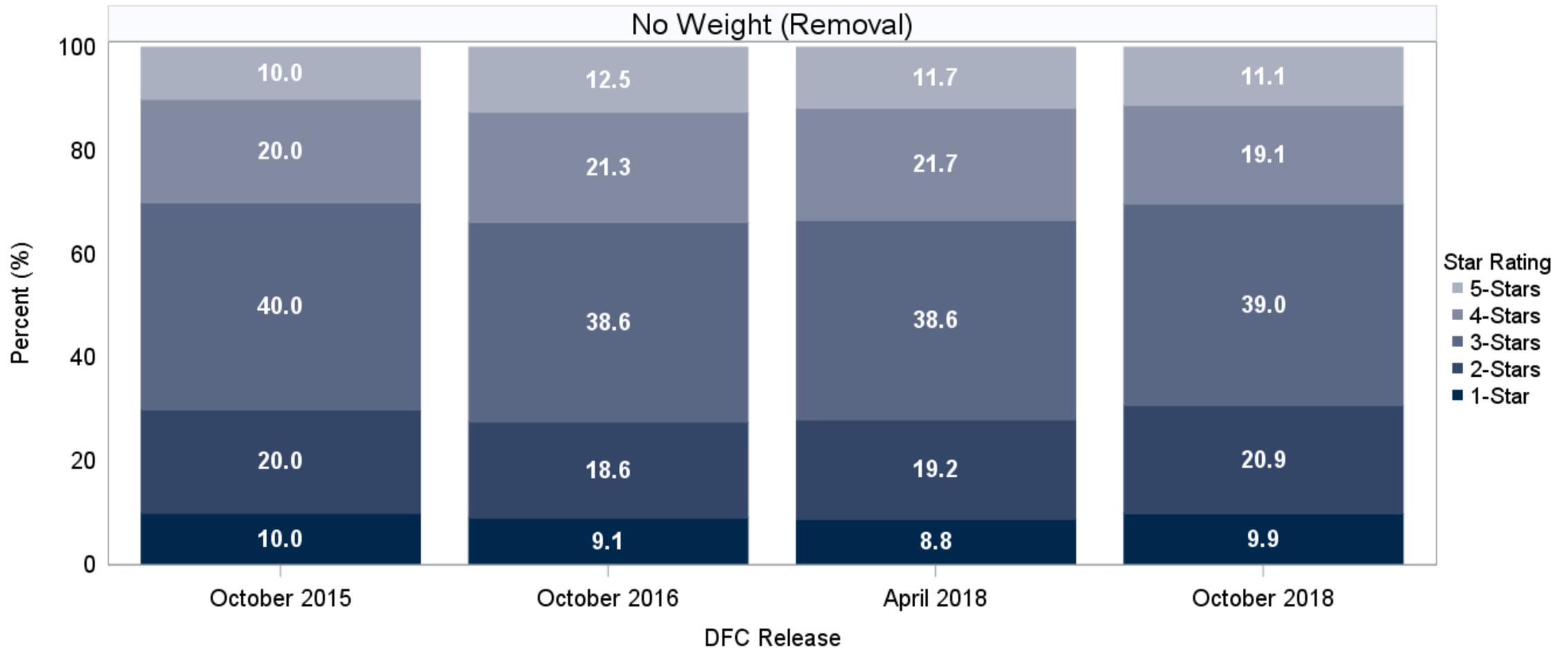
TEP Discussion on the Weighting Options

Public Comment Period

2:55 – 3:00 PM, EDT

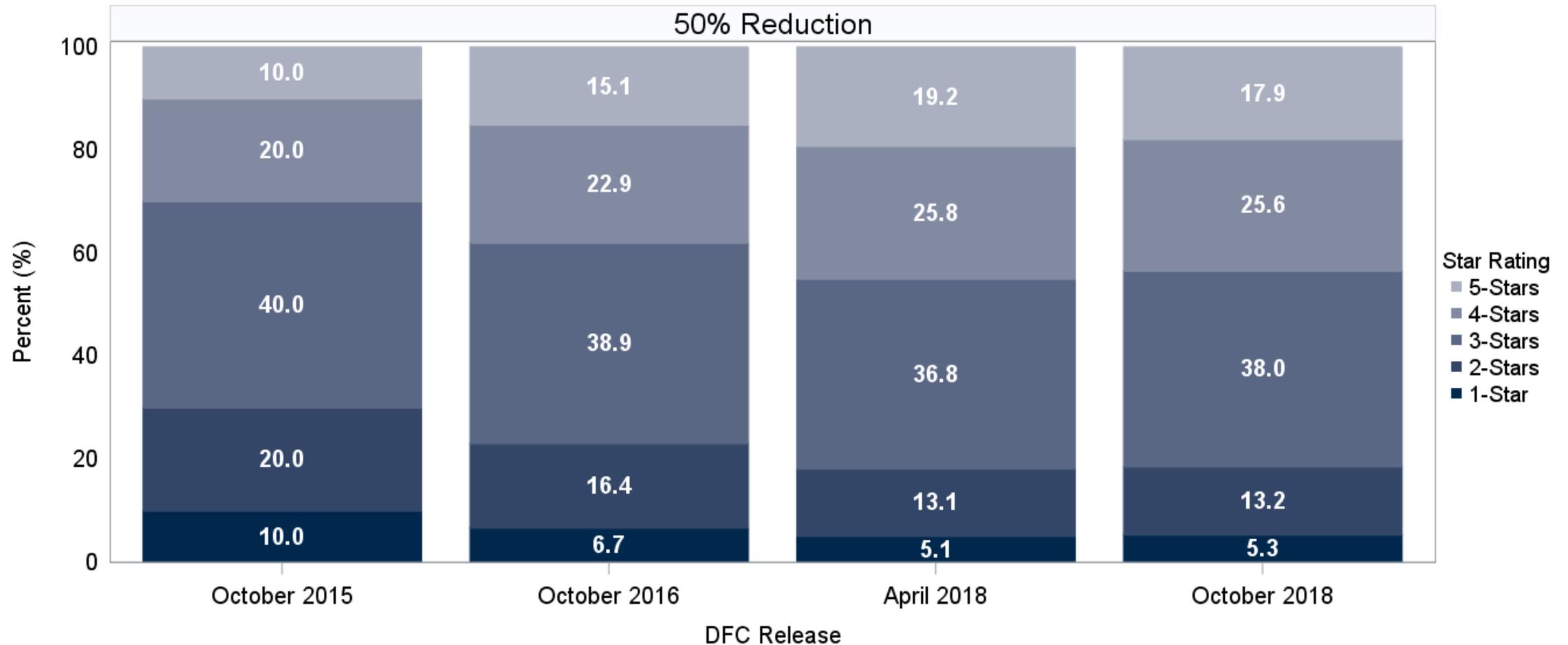
Appendix: Weighting

Hypothetical Star Rating Trends Across DFC Releases: 0% to Domain 3



Hypothetical Star Rating Trends

Across DFC Releases: 50% of its current weight on Domain 3



Total Kt/V Quartiles by DFC Release

	100% Max.	75% Q3	50% Med.	25% Q1	0% Min.
Oct. 15	100.00	93.81	90.76	86.83	0.00
Oct. 15*	100.00	94.01	91.06	87.00	0.00
Oct. 16	100.00	96.00	93.10	89.49	12.44
Oct. 16*	100.00	96.15	93.37	89.51	0.00
Apr. 18	100.00	97.93	96.27	93.76	18.32
Oct. 18	100.00	98.08	96.62	94.34	0.00

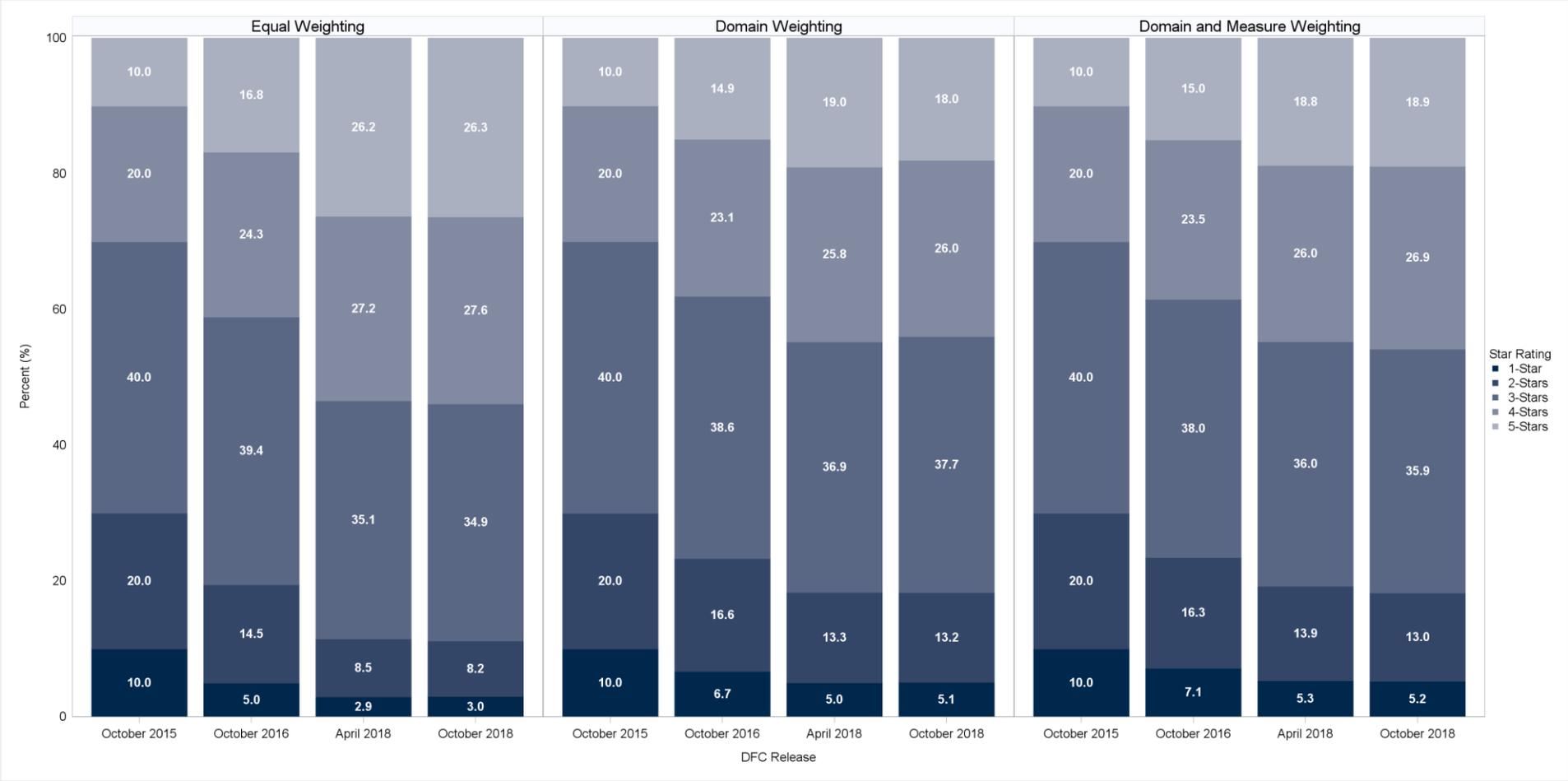
* Recalculated Using Current Measure Definitions

Hypercalcemia Quartiles by DFC Release

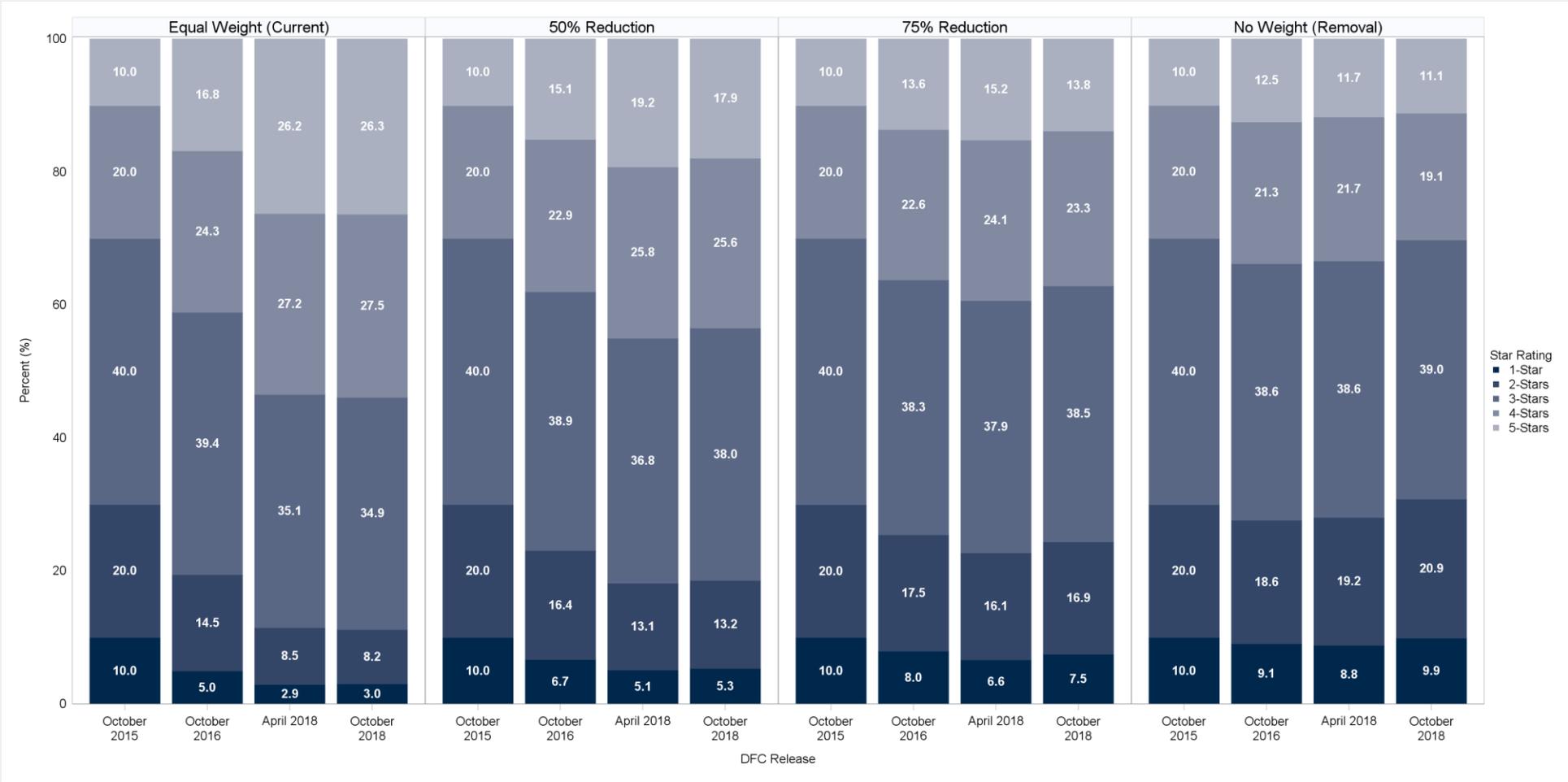
	100% Max.	75% Q3	50% Med.	25% Q1	0% Min.
Oct. 15	56.32	3.13	1.37	0.47	0.00
Oct. 15*	40.19	4.22	2.21	1.00	0.00
Oct. 16	35.16	1.94	1.03	0.43	0.00
Oct. 16*	35.79	3.02	1.93	1.12	0.00
Apr. 18	21.88	1.04	0.44	0.05	0.00
Oct. 18	100.00	2.76	1.43	0.52	0.00

** Recalculated Using Current Measure Definitions*

Historic Star Ratings Weighting by Variance



Historic Star Ratings Weighting Domain 3



Reset to Fixed Proportions

October 2018 DFC Star Rating Distribution vs. Resetting to Fixed Proportions (10-20-40-20-10):

Calculation	1-Star	2-Stars	3-Stars	4-Stars	5-Stars
Reported on DFC (2018)	190 (2.89)	536 (8.16)	2,311 (35.17)	1,810 (27.55)	1,724 (26.24)
Reset Distribution	657 (10.00)	1,314 (20.00)	2,629 (40.00)	1,314 (20.00)	657 (10.00)

- 4,332 facilities (66%) would have experienced a 1-star decrease in rating if the Star Rating distribution were to have been reset to 10-20-40-20-10 in October 2018
- Only 5 facilities (0.07%) would have experienced a 2-star shift. The remaining 2,234 facilities (34%) would have received the same rating after the reset

Star Rating Distribution Shift

