



**ESRD Quality Measure Development, Maintenance, and Support Project**  
**End Stage Renal Disease Dialysis Facility Compare (DFC) Star Ratings**  
**Star Rating Post-TEP Conference Call #1 Minutes**  
 August 21, 2015 11:00am – 1:00pm (ET)

TEP Members	UM-KECC	CMS
Paul Conway	Yi Li	Joel Andress
J. Richard Landis	Joseph Messana	Elena Balovlenkov
Franklin Maddux	Richard Hirth	Sofia Martinez
Allen Nissenson	Claudia Dahlerus	
Jane Pendergast	Chris Harvey	
John Reynolds	ZeZhi (Zac) Zhang	
Tonya Saffer	Natalie Scholz	
Chris Sarfaty	Cindy Liao	
Catherine Sugar	Jennifer Sardone	
Sumi Sun	Jordan Affholter	
David White		

**Introductions**

UM-KECC welcomed everyone to the Post-TEP Star Rating conference call, and thanked the TEP members for their time. A brief overview of the agenda was presented in order to describe what was going to be discussed on the conference call including presentations on possible methods for measure score imputation and applying thresholds. The presentations were followed by a period for TEP discussions. UM-KECC acknowledged that the TEP made several recommendations for revisions to the star rating methodology, but explained that the investigation of the imputation and thresholding methods were issues that were able to be investigated in a relatively short period of time. The call was open to the public and fifteen minutes were set aside from 11:55am to 12:10pm for public comments.

The members of the UM-KECC team were briefly introduced followed by brief statements from CMS. CMS informed TEP members that the final report for the Star Rating TEP has been posted online for public review. CMS will also be arranging a public call that will present the results and recommendations from the Star Rating TEP to the public and make the public aware of some of the updates. CMS noted they will be reaching out to the Public Reporting/Patient and Consumer Understanding workgroup to put together a focus group to help move forward with the consumer and patient recommendations.

TEP members briefly introduced themselves.

**Approaches to Imputation of Missing Measure Values**

The first part of the meeting presented analysis results assessing a different approach to the imputation of missing measure values. The reason for imputation is that if there is not enough information to score facilities due to high missingness of measure values, then there will be uncertainty in the ratings. Because of this,

imputation is used so that a domain with one individual measure missing can still be scored and allow for the calculation of a domain score.

UM-KECC began by noting the current imputation approach involves replacing missing measures with the national average of that measure. In response to TEP recommendations, UM-KECC investigated the implementation of an alternative method of imputing missing measure values, specifically the nearest neighbor approach that was suggested from the TEP methodology workgroup.

Based on analyses presented using the nearest neighbor approach, UM-KECC proposed that instead of using the national average, a local average is used based on an optimal number of similar facilities (that have share similar characteristics), or what is called “nearest neighbors.” UM-KECC described this approach in more detail.

UM-KECC presented the optimal number of nearest neighbors (or “k”) for all of the measures in the star rating. UM-KECC presented an analysis looking at the changes in star rating from using the old imputation method to using new imputation method. The analyses showed that 97% of facilities achieve the same star rating using both methods. Overall, the method demonstrated statistical improvements over the current method using the national average. UM-KECC presented the option to apply the new method based on nearest neighbors or keep the current (national average method).

### **Discussion about Imputation Methods**

One TEP member asked about the term “neighbor”. In response, the methodology workgroup co-chair clarified that nearest neighbor is a statistical term, and does not necessarily mean the actual facilities closest geographically to a facility with missing values.

One TEP member asked about the symmetry in the changes by star rating category using the imputation method. UM-KECC explained that because the star rating categories are determined by fixed cut-offs, the number of facilities that move up (or down) corresponds to an equal number of facilities moving down (or up) a category. Another TEP member stated the nearest neighbor approach strengthens statistical confidence. The TEP recommended that “location,” which is one of the facility level characteristics used when determining nearest neighbors, should be defined instead as rural versus non-rural since that is the actual characteristic that is being measured.

### **Two Approaches to Threshold Based Scoring for Individual Measures**

UM-KECC presented analyses on two approaches to threshold scoring. UM-KECC’s evaluation of thresholds was in response to feedback received from the patient and consumer workgroup at the in-person TEP meeting in April. The patient and consumer workgroup stated they strongly prefer use of an absolute threshold for scoring measures over relative rankings when possible as this could improve the interpretability of the star ratings. With absolute thresholds, facilities would have a target for which to aim. UM-KECC gave an overview of basic assumptions that must be met in determining an absolute threshold. UM-KECC also provided an overview of an alternative approach to scoring that they evaluated using the achievement threshold adopted in the ESRD Quality Incentive Program (QIP). Both approaches were assessed as way to score a subset of the current DFC measures used in the star rating.

In their first set of analyses to evaluate identification of an absolute threshold, UM-KECC examined whether there was a point where higher no longer indicated better performance on a measure, explaining how they plotted measure values against performance outcomes. For example, the identification of an absolute threshold could be established if there was a clear point where there was a reduction in a statistical correlation with better

outcomes, at higher levels of Kt/V. UM-KECC explained that they did not observe these in their analyses, thereby lacking empirical evidence on which to base an absolute threshold. UM-KECC next explained that as an alternative they investigated threshold scoring methods used for the ESRD QIP which uses relative achievement thresholds as part of their method for determining achievement scores. UM-KECC explained they evaluated using the lower threshold of achievement use in the QIP program, which is the 15<sup>th</sup> percentile of national performance of a measure. They presented analyses showing the impact on final scores using the 15<sup>th</sup> percentile, and scoring facilities whereby they would receive full or partial credit depending on whether they met/exceeded the measure threshold (receive 100 points), or fell below the threshold (receive less than 100 points).

After presenting the results UM-KECC highlighted advantages of using thresholds for individual measure scoring, such as easier interpretability by patients and other consumers because patients will know that the star ratings depend on facilities meeting these targets. UM-KECC also identified several disadvantages in terms of statistical tradeoffs. One included the potential negative impact on small facilities that have a greater chance of going below the threshold (because of random error) versus larger facilities. UM-KECC also recommended that a more in-depth evaluation of two scoring methods by a CMS contractor from the patient/consumer understanding perspective would be important. An evaluation would provide information on which approach is more easily understood and usable by patients and other consumers.

The presentation on the threshold methods concluded with a summary of three options for TEP members to consider: threshold all the non-standardized DFC measures (option 1), or just threshold a subset of the highly skewed measures (option 2). Hypercalcemia was labelled as potentially fitting this description at the TEP along with Kt/V, and potentially the catheter measure; or (option 3) retain the current method for measure scoring. On top of all these options, additional information could be added to the website about the typical profile of the facility by star rating category. This would give the average value in each star rating category for each measure. These average values can be used as targets for facilities and can give patients a general idea of the makeup of the different star rating facilities.

## **Discussion and Feedback**

A key issue brought up by a few TEP members during the call was about the current use of fixed percentiles (10-20-40-20-10) to assign the star ratings. Several TEP members stated their understanding was that use of scoring thresholds would also involve changing the current 10-20-40-20-10 fixed cut-offs used for the star rating categories. For example, a TEP member stated that using a threshold to score a measure will not make much of a difference if the final determination of star rating categories is based on the fixed cut-offs (fixed percentiles). CMS responded they would bring the concern of the TEP to CMS senior leadership and ask if the use of the fixed percentiles (10-20-40-20-10) were a policy decision that could be changed. CMS stated they would report back to the TEP on this question.

One TEP member recommended for UM-KECC to look at a dynamic or adjusted threshold model for smaller size facilities.

Several TEP members stated their concerns about the lack clinically significant/empirically determined cutoffs for a threshold. UM-KECC concurred based on their analysis results which did not yield an empirically based threshold for the skewed measures, where most facilities are scoring very high. One TEP member stated if the threshold does not have external validation and is highly sensitive depending on what threshold is used, then it should not be a practical update to the methodology using this method.

With respect to the fixed percentiles (current method), one TEP member stated a change in the fixed percentile cutoffs for the star ratings may be necessary because of the impact of the several skewed measures. However, one TEP member replied that there may be difficulty in establishing objective cut-offs for star rating categories to replace the 10-20-40-20-10 percentile cut-offs.

One TEP member stated concerns that determining rating categories based on actual (versus relative) facility performance may result, for example, in a similar pattern as the historical trend toward grade inflation. They explained that with absolute cut-offs, over time all of the one-star facilities may move into the two-star category, so star categories may need to be redefined. The TEP member suggested a second option which is presenting both the relative distribution based on fixed percentiles, and an absolute distribution.

One TEP member asked about any updates or additions to the measure set included in the star rating, specifically measures of patient reported outcomes, and patient satisfaction. CMS stated they are working on measure development and meeting soon with stakeholders to discuss the topic of measure development in these areas.

Several TEP members stated that the primary audience of the DFC star ratings is patients and consumers and therefore needs to be understandable to this primary audience. The current methodology (i.e., using fixed-percentiles for star rating categories, based on relative performance) is confusing to consumers. Several Public Reporting/Patient and Consumer Understanding workgroup TEP members reinforced the need for CMS to examine several of the other workgroup recommendations and for both workgroups to meet jointly moving forward in order for all members to have the respective methodological and patient/consumer perspectives of one another. It was also clarified that UM-KECC is responsible for development and updates to the methodology, but some of the other TEP recommendations for changes to DFC are the responsibility of other CMS contractors that work on the website and consumer testing.

### **Public Comments**

There were no public comments.

### **Closing Remarks**

UM-KECC thanked everyone for participating in the teleconference call, and for providing their questions, suggestions, and comments. All TEP members were encouraged to email additional comments and suggestions to UM-KECC.

CMS thanked everyone for participating in the Post-TEP teleconference call. CMS stated they will be in communication with the TEP to let them know when the next TEP call is and of any decisions that can be reported, and the information for that public call will be sent out soon.