

PUBLIC COMMENT SUMMARY REPORT

Project Title: End Stage Renal Disease (ESRD) Dialysis Facility Compare (DFC) Measures

Dates: The Call for Public Comment ran from October 25, 2018 to December 31, 2017. The Public Comment Summary was made available on May 22, 2018.

Project Overview: The Centers for Medicare & Medicaid Services (CMS) has contracted with The University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) to review the methodology developed to produce the DFC Star Ratings. The contract name is the ESRD Quality Measure Development, Maintenance, and Support contract. The contract number is HHSM-500-2013-130171.

Project Objectives: In October 2017, CMS requested Public Comment on the inclusion of additional measures to Dialysis Facility Compare (DFC) measures in order to:

- Increase transparency in the process and selection criteria
- Allow for increased input from the community on candidate measures
- Increase opportunity for the inclusion of externally developed measures on DFC

Comments were requested on the following measures:

- Percentage of Prevalent Patients Waitlisted (PPPW)
- Standardized Waitlist Ratio (SWR)

Information About the Comments Received:

Public comments were solicited by email and through a National Provider Call (teleconference). Five responses were received on this topic.

Stakeholder Comments—General and Measure-Specific

Percentage of Prevalent Patients Waitlisted (PPPW)

One commenter expressed interest in additional exclusions for patients under 75 who were not eligible for a transplant for clinical reasons, as well as an exclusion for patients who chose not to pursue a transplant.

Response: *We wish to re-emphasize that some indirect adjustment or exclusion for comorbidity is included in the current measure specifications in the form of exclusions for nursing home admission and age 75 years or greater, as well as adjustment for age. There were several arguments in favor of not adjusting further for comorbidity that informed the decision about the current measure specifications. Based on the literature, most patients on dialysis stand to benefit from transplantation and in fact, it is difficult to identify any subgroups that do not do better with transplantation as compared with remaining on dialysis. As such, adjustment or exclusion of patients with certain comorbidities risks disadvantaging their access to the benefits of transplantation. Furthermore, certain comorbid conditions (such as the presence of an active infection) may reflect poor care delivered at dialysis facilities and therefore may not necessarily be appropriate to for adjustment. The issue of comorbidities was debated substantially as part of*

the Access to Transplantation TEP and there was no consensus about whether to adjust for specific comorbidities and/or for which comorbidities in particular.

We acknowledge the importance of patient autonomy to make decisions about transplantation. However, it is important that patients make informed decisions about their health. Many patients may refuse transplantation out of fears and anxieties that could be allayed with proper education and support about the benefits of transplantation, which can be provided by dialysis facilities. In this manner, dialysis facilities can have a substantial influence on decision-making by patients.

One commenter asked that the effect of the regional/geographic variance in transplantation policy on dialysis facility performance on this measure be considered prior to implementation.

Response: *Based on available data, we are unable to determine which transplant center did (or would) evaluate a given patient, except of course for patients who are already wait listed. It is not possible to estimate, by center, the fraction wait listed (of those evaluated).*

We examined regional variation by using a transplant center rate adjustment based on historical waitlist data weighted by zip code. However, we ultimately decided against including transplant center adjustment for the following reasons:

- 1. The transplant center rate adjustment is not statistically significant in the model, and is unstable dependent on how a small percent of missing values are handled.*
- 2. The C-Index for both the model with and without this adjustment is 0.72, suggesting no improvement in discrimination with inclusion of the effect.*
- 3. The IUR decreases from 0.82 to 0.79 after adding the SWTR effect to the PPPW model, suggesting a decline in reliability of the measure with inclusion of the effect.*

One commenter believes the PPPW is an intermediate outcome measure (not process).

Response: *Thank you for the comment. CMS and UM-KECC have consulted with the National Quality Forum, which has classified this measure as a process measure.*

Standardized First Kidney Transplant Waitlist Ratio for Incident Dialysis Patients (SWR)

One commenter expressed concern with using comorbidity data from the 2728 for adjustment for incident comorbidities.

Response: *Comorbidities reported on this form have been found to be useful predictors of mortality, suggesting that the most salient comorbidities are reported (Wolfe et al., 2000); Ashby et al., 1998; Roys et al., 1999). The comorbidities from the CMS Form 2728 included in the SWR model were chosen based on their association with first year mortality. Additionally, we believe that it is reasonable to expect dialysis facilities to have an awareness of patient comorbidities at incidence. When dialysis facilities receive an intake call, they receive an extract of the patient's chart, which includes current conditions/comorbidities. Facilities would review that chart before accepting a patient. Dialysis facilities also attest to the accuracy of the information reported on the 2728 prior to submission.*

One commenter requested more information regarding the identification of statistically meaningful differences (specifically, reporting p-values and a detailed description of measure scores). The commenter noted that they are unable to assess the statistical significance of these findings without p-values. Additionally, they note that with large sample sizes, even statistically significant differences in performance may not be clinically meaningful, and a detailed description of measure scores (e.g., distribution by quartile, mean, median, standard deviation, outliers) should be provided.

Response: Regarding the first point, the label “better than expected” denotes statistical significance (i.e., significantly better than the national average, covariate adjusted; $p < 0.05$). Analogously, “worse than expected” also reflects the facility performing significantly ($p < 0.05$) worse than the covariate adjusted national average.

To your second comment regarding clinical importance, although the sample is indeed very large at the national level, this has little impact on statistical power at the facility level. For instance, if a given facility has 500 patients, standard errors computed for that facility will be heavily influenced by the $n=500$, with the size of the overall population (e.g., 500,000) playing a minor role. From this perspective, there is little chance of clinically unimportant finding being detected as statistically significant at the facility-level.

To your third point, we would be pleased to provide additional descriptive data, in order to demonstrate the meaningful variability in the SWR distribution across facilities.

One commenter noted that they prefer normalized rates or year-over-year improvement in rates instead of a standardized ratio.

Response: *The measure has been specified and calculated as a standardized ratio, but could be expressed as a standardized rate (as stated in the MIFs for each measure).*

Further, note that most regression analyses (of binary or count responses) in the clinical and epidemiologic literature are based on ratios. For example, when logistic regression is employed, odds ratios are typically reported. When Poisson regression is carried out, rate ratios are reported. In the context of survival analysis, hazard ratios are the measure used most often. To summarize, ratio measures are well accepted in the published literature.

General

One commenter recommended an exclusion for patients with active malignant cancer within five the past five years.

Response: *Consideration of a history of malignancy about transplant candidacy is complex, and depending on the nature of the malignancy wait times prior to active listing can vary from none to 5 years or more. As such, it would be difficult to make a specific exclusion for this, not to mention the difficulty in capturing the information in claims. Although a history of malignancy may bear on a waitlisting decision for an individual patient, these measures are at the facility level and it is highly unlikely that facilities would be systematically disadvantaged by the lack of an exclusion for these types of patients.*

One commenter made a recommendation regarding measure implementation, suggesting that the facility-level data be compiled and if the resulting measures appear to be reliable after exposure to interested parties, that they be reported on Dialysis Facility Compare but not included in the calculation of star ratings. They urge CMS to continue to prioritize development of a valid and broadly acceptable

transplant access measure, perhaps informed by experiences in the ESRD Seamless Care Organization program.

Response: *Thank you for your comments.*

One commenter was concerned that the proposed measures will increase the burden of additional data collection on dialysis facilities

Response: *The proposed waitlist measures are calculated with existing data sources. No new data collection is proposed or required for the reporting of these measures.*

A number of commenters were concerned about the degree to which performance on these measures is under the control of the dialysis facility. Specifically, commenters are concerned that these measures are limited in terms of actionability by the dialysis center, as the ultimate decision on waitlist status is made by the transplant center and the patient.

Response: *Waitlisting for transplantation is the culmination of a variety of preceding activities. These include (but are not limited to) education of patients about the transplant option, referral of patients to a transplant center for evaluation, completion of the evaluation process and optimizing the health of the patient while on dialysis. These efforts depend heavily and in many cases, primarily, on dialysis facilities. Although some aspects of the waitlisting process may not entirely depend on facilities, such as the actual waitlisting decision by transplant centers, or a patient's choice about the transplantation option, these can also be nevertheless influenced by the dialysis facility. For example, through strong communication with transplant centers and advocacy for patients by dialysis facilities, as well as proper education, encouragement and support of patients during their decision-making about the transplantation option. The waitlisting measures were therefore proposed in the spirit of shared accountability, with the recognition that success requires substantial effort by dialysis facilities. In this respect, the measures represent an explicit acknowledgment of the tremendous contribution dialysis facilities can be and are already making towards access to transplantation, to the benefit of the patients under their care.*

One commenter noted that neither of the transplantation access metrics are NQF-endorsed.

Response: *The SWR and PPPW will be submitted to NQF for endorsement in early April.*

One commenter believes age as the only sociodemographic risk variable is insufficient, suggesting additional adjustment for geography and other biological and socioeconomic factors that may influence waitlisting rates. The commenter did state that they do not support adjustment for waitlisting based on economic factors or by race or ethnicity.

Response: *We agree that financial and other social issues can pose substantial barriers to waitlisting for patients. However, they do not take away from the fact that many patients with these issues will still stand to benefit substantially from transplantation as compared with remaining on dialysis. As such, it is expected that dialysis facilities will work with transplant centers, advocate for patients and assist them in overcoming barriers to waitlisting to the extent possible. We also recognize that even with the best efforts, not all dialysis patients will ultimately be suitable candidates for waitlisting. Thresholds for the measures are assessed at the facility level. Examination of facility level measures essentially allows comparison of an individual facility's performance to a consensus standard, empirically set by the achievement of dialysis*

facilities across the nation. Through comparison with the performance of other facilities, these measures may help individual dialysis facilities identify opportunities for improvement in their waitlisting rates.

Regarding geography, we examined this extensively and ultimately decided against including adjustment for the following reasons:

1. The transplant center rate adjustment is not statistically significant in the model, and is unstable dependent on how a small percent of missing values are handled.
2. The C-Index (a measure of goodness of fit) for both the model with and without this adjustment is 0.72, suggesting no improvement in discrimination with inclusion of the effect.
3. The IUR decreases from 0.82 to 0.79 after adding the SWTR effect to the model, suggesting a decline in reliability of the measure with inclusion of the effect. The inter-unit reliability (IUR) measures the proportion of the total variation of a measure that is attributable to the between-facility variation, the true signal reflecting the differences across facilities

One commenter agreed that the access to transplantation measures should not apply to persons with a limited life expectancy, so they were pleased to see the exclusion for hospice patients.

Response: Thank you for your comment.

One commenter express concern that the risk model testing yielded an overall C-statistic of 0.72 for the PPPW and 0.67 for the SWR, raising concerns that the models will not adequately discriminate performance. They feel that a C-statistic of 0.8 is a more appropriate indicator of a model's goodness of fit, predictive ability, and validity to represent meaningful differences among facilities.

Response: We believe that the C-statistics of 0.72 (PPPW) and 0.67 (SWR) are considered to be a good fit based on recent literature, and note that they are similar in magnitude to other current NQF endorsed quality measures implemented by CMS. As we refine the risk model in the future, we will work to improve the model's ability to discriminate performance between facilities.

One commenter expressed concerns about reliability for small facilities, and requested that IUR be reported by facility size.

Response: Given the established effect of sample size on IUR calculations, it is expected that large facilities will have higher IUR values and small facilities will have lower IUR values for any given measure. Using the empirical null method, facilities are flagged if they have outcomes that are extreme when compared to the variation in outcomes for other facilities of a similar size. That is, smaller facilities have to have more extreme outcomes compared to other smaller facilities to be flagged. Note that for each measure respectively, facilities with fewer than 11 eligible patients or 2 expected events (SWR) or 11 eligible patients (PPPW) are not included in the respective measure calculations.

One commenter had questions about the identification of meaningful differences reported in the testing form. They note that with large sample sizes, even statistically significant differences in performance may not be clinically meaningful. A detailed description of measure scores, such as distribution by quartile, mean, median, standard deviation, outliers, should be provided to allow stakeholders to assess the measure. Therefore, before CMS adopts the PPPW or SWR measures, it should provide these data to allow for a thorough review of the measures' performance.

Response: We would be pleased to provide additional descriptive data, in order to demonstrate the meaningful variability in the distribution across facilities. We will incorporate such data in future iterations of the testing materials.

Regarding clinical importance, note that the large size of the population at the national level has little impact on facility-level statistical power. For instance, if a given facility has 500 patients, standard errors computed for that facility will be heavily influenced by the $n=500$, with the size of the overall population (e.g., 500,000) playing a minor role. From this perspective, there is little chance of clinically unimportant findings being detected as statistically significant at the facility-level.

Overall Analysis of the Comments and Recommendations

CMS and UM-KECC appreciate the time dedicated to reviewing and providing comments on the proposed candidate measures for DFC. The two transplant waitlist measures (SWR and PPPW) will be made available to facilities in the Dry Run area of the QDFC preview report, beginning with the July 2018 preview period. Public reporting of these measures on DFC is anticipated in October 2019.

Public Comment Verbatim Report

Date Posted	Text of Comments	Name, Credentials, and Organization of Commenter	Type of Organization	Recommendations/Actions Taken
May 22, 2018	See appendix	<p>Ralph Atkinson, MD Chair, Forum of ESRD Networks Medical Advisory Council (MAC)</p> <p>Donald Molony, MD, President, Forum of ESRD Networks</p>	Provider Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.
May 22, 2018	See appendix	Frank Maddux, M.D., Chairman, Kidney Care Partners (KCP)	Patient Advocacy Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.
May 22, 2018	See appendix	Kevin Longino, CEO, National Kidney Foundation	Patient Advocacy Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.
May 22, 2018	See appendix	Nancy Pierce RN BSN CNN Dialysis Director St. Peters Hospital	Individual	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.
May 22, 2018	See appendix	Jackson Williams, Director of Regulatory Affairs, Dialysis Patient Citizens	Patient Advocacy Organization	We thank you for your feedback. Stakeholder comments will be reviewed by measure developers and taken under consideration. Responses to comment themes are provided above.

President

Donald A. Molony, MD
Houston TX

President-Elect & MAC Chair

Ralph Atkinson III, MD
Nashville, TN

Secretary

Chris Brown
Cranbury, NJ

Treasurer

Stephanie Hutchison, MBA
Seattle, WA

Past-President

John Wagner, MD, MBA
Brooklyn, NY

EDAC Chair

Chris Brown
Cranbury, NJ

EDAC Vice-Chair

Susan Caponi, MBA, RN, BSN
Lake Success, NY

KPAC Chair

Derek Forfang
San Pablo, CA

KPAC Vice-Chair

Joseph Karan
Land O Lakes, FL

Members-At-Large

Natasha Avery, DrPH, MSW, CHES
Ridgeland, MS

Kam Kalantar-Zadeh, MD, MPH, PhD
Orange, CA

Kelly M. Mayo, MS
Tampa, FL

Stephen Pastan, MD
Atlanta, GA

Jeff Perlmutter, MD
Rockville, MD

Katrina Russell, RN, CNN
Seattle, WA

Ad Hoc Members

Louis Diamond
MB, ChB, FACP, FCP(SA)
Washington DC

Andrew Howard, MD, FACP
Alexandria, VA

Forum Coordinator

Dee LeDuc
Birchwood, WI

December 29, 2017

University of Michigan - KECC
1415 Washington Heights
Ann Arbor, MI 48109-2029

RE: Forum Medical Advisory Council Response to UM-KEC Proposed Transplant Metrics

To Whom It May Concern,

The Forum of ESRD Networks Medical Advisory Council (MAC) appreciates the opportunity to comment on the UM-KECC proposed ESRD Access to Kidney Transplantation Measure Development. MAC members are the Medical Review Board Chairs of the 18 ESRD Networks. We represent a broad range of Nephrology experience in both dialysis and renal transplantation. We agree that renal transplantation is the preferred way to treat ESRD in eligible patients and are keenly interested in ways to improve outcomes in our patients eligible for transplantation.

The two proposed metrics, Percentage of Prevalent Patients Waitlisted (PPPW) and Standardized First Kidney Transplant Waitlist Ratio for Incident Dialysis Patients (SWR), were released for comment on October 25, 2017. Unfortunately, the MAC only became aware of them 10 days ago. While we are submitting comments to meet the deadline of December 31, 2017 we respectfully reserve the right to provide supplemental comments at a later date. We additionally request an opportunity to work with UM-KECC to assure the Forum and its members are made aware of future comment periods that relate to the ESRD community of patients we serve.

The ESRD Access to Kidney Transplantation TEP convened on April 20 and 21, 2015 to consider measure development related to patient access to kidney transplantation. One of the criterion for recommended measures was that it be “feasible without creating undue burden for dialysis facilities” (https://dialysisdata.org/sites/default/files/content/ESRD_Measures/Access_To_Kidney_Transplantation_TEP_Summary_Report.pdf , pg4). Dialysis facilities under the CMS Conditions for Coverage are increasingly saddled with data collection and reporting that is intended to improve patient outcomes. CROWNWEB usability and data accuracy has been a particular focus of the Forum and the MAC, and these ongoing issues could affect the accuracy of the proposed measures. It is the opinion of the MAC that the proposed measures will increase the burden of additional data collection on dialysis facilities

The TEP additionally considered “the degree to which performance on a measure is under the control of the dialysis facility”

(https://dialysisdata.org/sites/default/files/content/ESRD_Measures/Access_To_Kidney_Transplantation_TEP_Summary_Report.pdf , pg 8). The MAC commends the TEP for recognizing these realities in their deliberations. Unfortunately, metrics indexed to rates of waitlisting rather than rates of referral for transplant explicitly holds dialysis facilities accountable for a quality metric outside of the dialysis unit’s direct influence, much less control. Decisions regarding which patients are or are not selected for waitlisting are determined by a transplant center, not a dialysis facility. Furthermore, different transplant centers utilize different medical, surgical, financial, and psychosocial criteria to determine qualification for listing. Since none of those criteria are (or should be) either controlled nor even influenced by referring dialysis facilities, it follows that dialysis facilities should not be held responsible for how these criteria are applied to patients referred from the dialysis facility. Instead, the MAC recommends a metric which assess rates of patient referral for transplantation. Initiated referrals, unlike rates of waitlisting, are a discrete action which is under the control of the attending physicians in a dialysis facility.

In sum, these proposed measures will not foster what will truly improve access to kidney transplantation, that of a productive collaboration between transplant centers and the facilities they serve.

Sincerely,



Ralph Atkinson III, MD
Chair, Medical Advisory Council (MAC)



Donald A. Molony, MD
President, National Forum of ESRD Networks



December 12, 2017

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

Dear Dr. Goodrich,

On behalf of Kidney Care Partners (KCP) and its members, I am writing to share comments on the changes that were presented on October 25 for the methodology and new measures for Star Ratings of dialysis facilities. We appreciate the Agency's efforts to work with the kidney care community to revise the Star Rating methodology and opportunity to comment on measures under consideration for Dialysis Facility Compare (DFC)/ESRD Five Star Rating Program (ESRD Five Star). As you know, the appropriate implementation of ESRD Five Star is a top priority for the members of KCP. It is critically important to create a system that is accurate, transparent, and easy for patients, family members/caregivers, and other consumers to understand.

In this letter, we provide comments on the candidate measures proposed during the October 25 call, as well as comments about the suggestions related to how stars are determined under the program.

I. Candidate Measures

KCP recognizes the fundamental importance of improving transplantation rates for patients with ESRD, but does not support the attribution of successful/unsuccessful waitlisting to dialysis facilities. As we have noted in previous letters, KCP believes that while a referral to a transplant center, initiation of the waitlist evaluation process, or completion of the waitlist evaluation process may be appropriate facility-level measures that could be used in ESRD quality programs, the Percentage of Prevalent Patients Waitlisted (PPPW) and Standardized First Kidney Transplant Waitlist Ratio for Incident Dialysis Patients (SWR) are not.

Waitlisting per se is a decision made by the transplant center and is beyond a dialysis facility's locus of control. We further recommend CMS explore a care coordination measure with mutual facility-transplant center responsibilities. Lastly,

we note that a completion of the waitlist process measure and a waitlisting measure should be developed for transplant centers. Transplantation is a multi-party process: To optimally drive improvement, measurement of all key parties should be done.

A. Comments Relevant to both the PPPW and SWR Measures.

Several of KCP's concerns apply to both the PPPW and SWR measures.

1. ***PPPW and SWR: NQF endorsement.*** KCP notes that neither of the transplantation access metrics are NQF-endorsed, a general pre-requisite for KCP to support inclusion of a measure in any accountability program.
2. ***PPPW and SWR: Facility attribution.*** As just noted, KCP strongly objects to attributing successful/unsuccessful placement on a transplant waitlist to dialysis facilities. The transplant center decides whether a patient is placed on a waitlist, not the dialysis facility. One KCP member who is a transplant recipient noted there were many obstacles and delays in the evaluation process with multiple parties that had nothing to do with the dialysis facility—e.g., his private pay insurance changed the locations where he could be evaluated for transplant eligibility on multiple occasions, repeatedly interrupting the process mid-stream. Penalizing a facility each month through the PPPW and SWR for these or other delays is inappropriate. Again, KCP emphasizes our commitment to improving transplantation access, but we believe other measures within the facility's appropriate sphere of control should be pursued.
3. ***PPPW and SWR: Age as the only sociodemographic risk variable.*** KCP strongly believes age as the only sociodemographic risk variable is insufficient. We believe other biological and demographic variables are important, and not accounting for them is a significant threat to the validity of both measures.

Geography, for instance, should be examined, since regional variation in transplantation access is significant. Waitlist times differ regionally, which will ultimately change the percentage of patients on the waitlist and impact performance measure scores. That is, facilities in a region with long wait times will "look" better than those in a region with shorter wait times where patients come off the list more rapidly—even if both are referring at the same rate.

Additionally, criteria indicating a patient is "not eligible" for transplantation can differ by location—one center might require evidence of an absence of chronic osteomyelitis, infection, heart failure, etc., while another may apply them differently or have additional/different criteria. The degree to which

- these biological factors influence waitlist placement must be accounted for in any model for the measure to be a valid representation of waitlisting. Moreover, transplant centers assess a myriad of demographic factors—e.g., family support, ability to adhere to medication regimens, capacity for follow-up, insurance-related issues, etc. Given transplant centers consider these types of sociodemographic factors, any waitlisting measure risk model should adjust for them. Of note, like the Access to Kidney Transplantation TEP, KCP does not support adjustment for waitlisting based on economic factors or by race or ethnicity.
4. **PPPW and SWR: Hospice exclusion.** We note that an exclusion for patients admitted to hospice during the month of evaluation has now been incorporated into both measures. KCP agrees that the transplantation access measures should not apply to persons with a limited life expectancy and so is pleased to see this revision.
 5. **PPPW and SWR: Risk model fit.** We note that risk model testing yielded an overall C-statistic of 0.72 for the PPPW and 0.67 for the SWR, raising concerns that the models will not adequately discriminate performance. Smaller units, in particular, might look worse than their actual performance. We reiterate our long-held position that a minimum C-statistic of 0.8 is a more appropriate indicator of a model's goodness of fit, predictive ability, and validity to represent meaningful differences among facilities.
 6. **PPPW and SWR: Stratification of reliability results by facility size.** CMS has provided no stratification of reliability scores by facility size for either measure; we are thus unable to discern how widely reliability varies across the spectrum of facility sizes. We are concerned that the reliability for small facilities might be substantially lower than the overall IURs, as has been the case, for instance, with other CMS standardized ratio measures. This is of particular concern with the SWR, for which empiric testing has yielded an overall IUR of only 0.6—interpreted as “moderate” reliability by statistical convention.¹ To illustrate our concern, the *Standardized Transfusion Ratio for Dialysis Facilities* (STrR) measure (NQF 2979) was also found to have an overall IUR of 0.60; however, the IUR was only 0.3 (“poor” reliability) for small facilities (defined by CMS as ≤ 46 patients for the STrR). Without evidence to the contrary, KCP is thus concerned that SWR reliability is similarly lower for small facilities, effectively rendering the metric meaningless for use in performance measurement in this group of providers. KCP believes it is incumbent on CMS to demonstrate reliability for all facilities by providing data by facility size.

¹ Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.

7. ***PPPW and SWR: Meaningful Difference.*** We note that with large sample sizes, as here, even statistically significant differences in performance may not be clinically meaningful. A detailed description of measure scores, such as distribution by quartile, mean, median, standard deviation, outliers, should be provided to allow stakeholders to assess the measure. Therefore, before CMS adopts the PPPW or SWR measures, it should provide these data to allow for a thorough review of the measures' performance.

B. Comments Relevant to Only PPPW Measure

Process vs. intermediate outcome measure. The CMS Measure Information Form identifies the PPPW as a process measure. KCP believes the PPPW is an intermediate outcome measure and recommends the form indicate such.

C. Comments Relevant to Only the SWR Measure

1. ***Incident comorbidities incorporated into risk model.*** We note that eleven incident comorbidities—heart disease, inability to ambulate, inability to transfer, COPD, malignant neoplasm/cancer, PVD, CVD, alcohol dependence, drug dependence, amputation, and needs assistance with daily activities—have been incorporated into the SWR risk model. All are collected through the CMS Form 2728. As we have noted before, we continue to be concerned about the validity of the 2728 as a data source and urge CMS to work with the community to assess this matter.
2. ***Meaningful differences in performance.*** Testing results presented in the documents released for review indicate that the SWR can distinguish differences in performance in approximately 16% of facilities (i.e., 8.7% of facilities were classified as “better than expected” and 6.9% as “worse than expected”); these results are interpreted in the documents as demonstration of “practical and statistically significant differences in performance across facilities based on their proportion of patients placed on the transplant waitlist.” We first note, however, that we are unable to assess the statistical significance of these findings as p-values are not provided. Additionally, we note that with large sample sizes, as here, even statistically significant differences in performance may not be clinically meaningful. Per NQF measure testing guidance,² a detailed description of measure scores (e.g., distribution by quartile, mean, median, standard deviation, outliers) should ideally be provided to allow for assessment of this endorsement criterion.

²National Quality Forum. Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties. January 2011. Available at: http://www.qualityforum.org/Measuring_Performance/Improving_NQF_Process/Measure_Testing_Task_Force_Final_Report.aspx.

We urge CMS to provide these data to facilitate transparency and to allow for a thorough review of the measure's performance.

3. **Rate vs. ratio.** Notwithstanding our many concerns regarding attribution and risk adjustment of this measure, consistent with our comments on other standardized ratio measures (e.g., SHR, SMR), KCP prefers normalized rates or year-over-year improvement in rates instead of a standardized ratio. We believe comprehension, transparency, and utility to all stakeholders is superior with a scientifically valid *rate* methodology.

D. Conclusion

In sum and for the reasons stated above, we do not believe that the PPPW and SWR measures should be added to DFC/Five Star.

II. Awarding Star Ratings

A. Re-baselining Target Percentages

Based on the October 25 presentation, we understand that CMS will consider a re-baselining when the percentage of facilities receiving 1 or 5 Stars falls below 15 percent. KCP asked Discern Health to model the impact of this policy. The current percentage of facilities at 1 or 2 Stars is 19 percent. That fact coupled with the improvement in many of the measures used to calculate the Star Rating would result in a likely re-baselining during the next year (or, if not, the following year).

Unfortunately, it is not possible to be completely certain because the guidance around the re-baselining timeline remains ambiguous. It is possible that once the 15 percent of 1 or 2 Stars is met, re-baselining would occur simultaneously. Alternatively, it could occur the year after the 15 percent threshold was met. In either scenario, the forced distribution CMS uses would negatively impact Star Ratings at numerous facilities, because it would double the number of facilities with 1 or 2 Stars with no concurrent drop in quality. Such a significant shift in Star Ratings is misleading to patients who mistakenly believe that their dialysis facility quality has dropped. We also remain concerned that no other Five Star program requires such re-baselining or a predetermined distribution of stars. Therefore, we once again strongly recommend that CMS eliminate the overly burdensome nature of having two different public reporting systems and use the statutorily mandated ESRD Quality Incentive Program (QIP) and its methodology for public reporting. Stars could be assigned based on the payment reduction tiers.

B. Updating the Star Ratings for 2018

KCP appreciates that CMS has recognized that the lack of nursing home data has resulted in the DFC measures being inaccurate. However, with CMS updating certain measures January 1, 2018, we believe that the stars awarded for 2017 will not provide an accurate representation of facility quality and will inappropriately mislead patients and consumers.

Further, we are concerned that the lack of updates to the Kt/V and hypercalcemia measure data, and the decision to impute current performance based on prior years' data. This may result Star Ratings that do not accurately portray facility performance. Discern Health analyzed two full years of DFC data for the Kt/V and hypercalcemia measures (data released in June 2016 and June 2017). For each measure, Discern analyzed the year-to-year correlation of the measures by facility (for both the raw score and the z-score used in the Star Ratings). Both Kt/V and hypercalcemia demonstrated significant year-to-year volatility across facilities. Performance in one year for any individual facility is not a strong predictor of that facility's performance in the following year. For Kt/V, the year-to-year correlation coefficient for measure performance is .27. This suggests that there is significant year-to-year variation in performance by facility. More than half of facilities saw a year-to-year change of ± 0.5 or greater in their z-score. The results are similar for the hypercalcemia, where the year-to-year correlation coefficient for measure performance is .44 and more than half of facilities saw a year-to-year change of ± 0.35 or more in their z-score. Even small differences in z-scores can signify importance differences in performance, and have a material impact on Star Ratings.

Given the year-to-year volatility in measure results, we are concerned that using last year's measure performance to impute scores for current Star Ratings may lead to many facilities receiving Star Ratings that do not convey their current performance. Therefore, we ask again, that CMS eliminate this problem by using the ESRD QIP scores as the basis for assigning star ratings.

III. Conclusion

Once again, we want to thank you and your team for addressing some of the concerns we have raised in previous letters. We reiterate our commitment to working with you to resolve the outstanding issues that will allow the Star Rating program to achieve the Agency's goal and be a useful tool for patients, caregivers, and consumers. Please do not hesitate to contact Kathy Lester at klester@lesterhealthlaw.com or (202) 534-1773 if you have questions or would like to discuss these recommendations.

Sincerely,



Frank Maddux, M.D.
Chairman
Kidney Care Partners

cc: Jean Moody-Williams, RN, MPP, Deputy Director, Center for Clinical Standards and Quality

Elena K. Balovlenkov, MS, RN, CHN, Technical Lead, Dialysis Facility Compare, Division of Quality Measurement, Centers for Medicare & Medicaid Services

Joel Andress, PhD, End-Stage Renal Disease Measures Development Lead, Division of Quality Measurement, Centers for Medicare & Medicaid Services

Appendix A: KCP Members

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



National
Kidney
Foundation®

30 E. 33rd Street
New York, NY 10016

Tel 212.889.2210
Fax 212.689.9261
www.kidney.org

December 31, 2017

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

Submitted via email: dialysisdata@umich.edu

Dear Dr. Goodrich,

The National Kidney Foundation appreciates the opportunity to comment on updates to the Dialysis Facility Compare website and the Star Ratings program as presented on the October 25, 2017 webinar. The National Kidney Foundation is.....

I. Proposed Measures for DFC:

a. Percentage of Prevalent Patients Waitlisted (PPPW)

The National Kidney Foundation believes the PPPW is a meaningful measure for patients, but request that changes be made to the measure before it is used on DFC. Some patients under age 75 may not be eligible for transplantation due to other clinical reasons. In addition, in some cases even the most informed and educated patient may ultimately choose not to pursue a transplant. Limited, but additional exclusions to account for these circumstances should be evaluated.

Ultimately, the decision on whether a patient is listed for a transplant is made by the transplant center that evaluated the patient (and the patient's desire for a transplant). These are complex decisions that include many factors and vary by transplant center and geographic region, which would make nationwide comparisons of waitlist percentages difficult to interpret. This is of concern as the stated intent of DFC is to help patients and their family members understand the quality of care delivered in dialysis facilities to help them make decisions about which facility to choose. DFC uses national comparisons to determine performance on individual measures. Given the geographical and center variation in transplant center criteria for adding patients to the waitlist this may not provide patients with the most accurate picture of how well a dialysis facility is doing in assisting patients to remain on the transplant waitlist. The effect of this

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

Tel 212.889.2210
Fax 212.689.9261
www.kidney.org

variance in transplantation policy on dialysis facility performance on this measure should be considered prior to implementation.

b. Standardized First Kidney Transplant Waitlist Ratio for Incident Dialysis Patients (SWR)

The National Kidney Foundation does not support including this measure on DFC. NKF appreciates the intent of this measure to ensure that patients are waitlisted as early as possible after starting dialysis, if they were not already waitlisted. However, we are concerned this measure is limited in terms of actionability by the dialysis center as the ultimate decision on waitlist status is made by the transplant center and the patient. Dialysis facilities have a role in educating patients about transplant and supporting their active listing. However, incident dialysis patients, who were not listed before starting dialysis, may be more complex and have comorbidities that make them ineligible for the waitlist during the first year. While it is the responsibility of the dialysis facility to work to improve the health and functional status of dialysis patients during the first year, much of the final decision, regarding acceptance to a transplant list, is beyond their control. In addition, dialysis units involved in education and care coordination in the transition of advanced CKD to ESRD would not be recognized for pre-emptively having patients on the waitlist.

II. Public Reporting and Patient Engagement

While the National Kidney Foundation supports the intent of DFC to help patients and families make informed decisions about where to receive care, we remain concerned that having two quality rating programs causes significant confusion. Additionally, having some measures contribute to the star ratings and other measures that do not is an added layer of complexity that is not readily transparent. Consumers should reasonably be able to rely on the visual of star ratings to determine whether the facility provides meaningful, high quality care. Currently, there are so many nuances around what the star ratings include and don't include, how they are assigned, and the differences between the stars, DFC, and the performance certificate score used for the QIP that it is unlikely that most patients find this public reporting helpful. At worst these ratings could be used for marketing and private payer contracts that could mislead patients and limit their options of where to receive dialysis.

The National Kidney Foundation appreciates the considerable changes that CMS has made to the DFC website and star ratings in response to patient input, including recommendations by the National Kidney Foundation. Specifically, we appreciate that a navigation tab has been added to the website that includes the results from patient experience surveys using the in-center hemodialysis CAHPS®, providing additional information about what other patients think about the quality of care delivered in their dialysis facility. We also appreciate this tab clarifies that this information is not included in the star ratings.

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

Tel 212.889.2210
Fax 212.689.9261
www.kidney.org

We recognize that CMS is using the best tools at its current disposal to try and provide consumers with clear information about the quality of care delivered in dialysis facilities. However, there still are many gaps in information for what the star ratings and DFC provide and confusion persists. The National Kidney Foundation continues to believe a substantial redesign of DFC and the star ratings is necessary to achieve the intent of helping patients make decisions about where to receive care and how to determine the quality of care delivered in local dialysis facilities. We offer the following recommendations.

1. DFC should be a more personalized tool that allows users to filter results according to their preferred criteria. This concept was supported by many of the patients who participated in the 2015 ESRD Star Ratings technical expert panel.
2. CMS should either align the ratings with the payment reduction thresholds in the QIP to avoid confusion of having two quality reporting programs that include the same measures, but are scored differently, or completely differentiate the star ratings by allowing them to be consumer driven. Consumer driven ratings could be structured in such a way that DFC users anonymously rate facilities using categories that are most meaningful to patients (i.e., how safely patients feel the care is delivered, how clean the facility is, how engaged patients feel they are about treatment decisions, etc.).
3. A blend of these two approaches could also generate a creative and easily understandable approach where a CMS quality rating is displayed as well as a direct consumer rating – this approach would be similar to movie review websites where there is a critic rating and an audience rating. However, even in a blended approach the CMS rating should still be aligned with the QIP measures.

Again, we appreciate that CMS has convened and incorporated the feedback of several TEPs and focus groups that included patients to improve DFC and the star ratings. However, we believe substantial improvements including, website redesign are necessary to empower patients to make decisions about where to receive care. We are also concerned that as currently constructed the star ratings could be used to mislead patients and restrict access to care. The National Kidney Foundation would be happy to further discuss our recommendations and opportunities to improve quality reporting programs to better assist patients in making informed decisions about where to receive treatment.

Sincerely,
Kevin Longino

Kevin Longino
CEO
Kidney Transplant Recipient

Greetings!

As a dialysis nurse and facility director with 43 years of nephrology experience, I have a few comments on the new measures:

1. I think it's wonderful to have transplants looked at--great job!
2. In determining the facility percentages and standardized ratio, I think credit should be given to living related donor transplants no matter when they occur as opposed to happening in the first year. Patients may have some medical issues (stabilization of auto-immune disease, needed cardiac surgery, weight loss needed, etc.) that need to be cleared before the living transplant will happen, and donors may not come forward until later on. We see this happening a lot. We frequently have living donor transplants after the first year. We've sent patients to different transplant centers for work-ups because they have been turned down by one center after the initial work-up. They were accepted by another center, but the whole process takes time to happen. These patients had living donors waiting, but it took a while to happen.. After 6 years, one of our obese patients lost enough weight to be a transplant candidate, and a relative come forward to donate. This process needs encouragement nation-wide!
3. I also think an exclusion should be made for patients with active malignant cancer within 5 years. There is no transplant center which will accept a patient with this diagnosis.

Thanks for allowing me to give input!

Nancy Pierce RN BSN CNN
Dialysis Director
St. Peters Hospital
2475 Broadway
Helena, MT 59601
406-444-2281

December 21, 2017

Elena Balovlenkov, R.N.
Technical Lead for Dialysis Facility Compare
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Mail Stop S3-02-01
Baltimore, MD 21224

Re: Addition of New Measures to Dialysis Facility Compare and Other DFC Matters

Dear Ms. Balovlenkov:

In the October National Provider Call it was announced that CMS is considering the addition of two new measures, Percentage of Prevalent Patients Waitlisted (PPPW) and Standardized First Kidney Transplant Waitlist Ratio for Incident Dialysis Patients (SWR), to Dialysis Facility Compare in 2019.

We share the concerns expressed by others that while referral to a transplant center, initiation of the waitlist evaluation process, and completion of the waitlist evaluation process are activities for which a dialysis facility should be held accountable, the decision whether to list a patient for transplant is made by a transplant center and this outcome is not fully within the dialysis facility's control. As such, we cannot at this time endorse the use of these measures in the Quality Incentive Program or for inclusion in the star rating calculation.

However, we do believe there is value to transparency on as important a matter as access to transplants, and we don't think reporting can be put off indefinitely until the "perfect" transplant measure is developed. We suggest that the facility-level data be compiled and if the resulting measures appear to be reliable after exposure to interested parties, that they be reported on Dialysis Facility Compare but not included in the calculation of star ratings. This will permit scrutiny of facilities' track records and identification of any outlier facilities; and empower patients to initiate an informed dialogue with their health care professionals if they have concerns about delays in their own journey to transplants. We urge CMS to continue to prioritize

development of a valid and broadly acceptable transplant access measure, perhaps informed by experiences in the ESRD Seamless Care Organization program.

We wanted to also take this opportunity to convey our thoughts on rebaselining of the star ratings. As we have said before, we agree that the distribution must not skew too much in favor of four- and five-star facilities, as we believe that such plaudits should indicate above-average and excellent performance. But we also know that one- and two-star ratings carry a stigma among consumers and believe they should be reserved for truly underperforming clinics.

Now that DFC has posted star ratings for some length of time, patients may be monitoring them with an eye toward longitudinal performance (particularly of their “home” clinic) rather than the cross-sectional performance that the tool is designed to illustrate. As such, we think rebaselining should take place as infrequently as possible so as to accurately convey improvements and declines in performance at a given clinic, and avoid unnecessarily alarming patients in this era of steadily improving ESRD outcomes. The recent announcement of a review of CMS’ quality measurement portfolios to emphasize “meaningful measures” suggests to us the likelihood that there soon could be a new collection of measures used to comprise the DFC star ratings. If so, we think that rebaselining should wait until the new measure set is finalized to avoid multiple rebaselining. We also hope that a rollout of a new measure set will await adoption of techniques to account for social risk factors, especially if the revamped measure array emphasizes so-called “big dot” outcome measures.

Since you are in the rating business we imagine that you have been following news coverage of the recent controversies surrounding Rottentomatoes.com, the website that aggregates star ratings of movies. The volume of debate over this site, its methodologies, and reporting procedures is surprising given the comparatively frivolous nature of film reviews and their subjects. Nevertheless, these controversies illustrate the high stakes that are perceived by businesses and consumers, who use these ratings to decide how to spend three-hour blocks of their time. The stakes are high because negative reviews divert consumers from spending time and money on certain movies, and low ratings are understood as meaning that a movie is a waste of time. We continue to think it imposes a high cognitive burden on consumers to be asked to apply different meanings to star ratings on DFC than they apply to other goods and services, and indeed to different services rated on Medicare.gov.

We think that CMS executes a more intuitive star scale in its ratings of Medicare Advantage plans and Part D Prescription Drug Plans. In 2017, all 364 of combined MA-PD plans rated received ratings of 2 ½ stars or above, and of 55 PDPs rated, all scored 2 ½ stars or above. We understand that the Center for Medicare’s policy has been to terminate the contracts of 1- and 2-star plans. We think that this is in line with consumer expectations; that is, that as a trusted source of health insurance, beneficiaries and caregivers perceive the Medicare program as avoiding entities that can’t muster the 2 ½ star rating considered the borderline dividing watchable from unwatchable films. We think consumers expect that a two-star dialysis facility would be one operating under some type of corrective action plan, and a one-star facility would be one that is terminated. We urge you to confer with your colleagues in CM to understand their reasoning and make an effort to assign ratings more uniformly across programs.

Finally, in response to Administrator Verma's measures announcement at the Health Care Payment Learning and Action Summit, and subsequent briefing by your CCSQ colleagues, our Board of Directors held an informal discussion of quality measurement priorities. I must report that the concept of limiting quality programs to "big dot" measures does not seem to have intuitive appeal to patients. But more in line with CMS priorities, our Board members continue to place value on measures of preventable harm and patient-reported outcomes. As you know, the recent TEP identified patient safety as a potential subject of a PROM. We encourage further exploration in these areas.

We look forward to engaging CMS in dialogue as adjustments for social risk factors and the meaningful measurement framework are implemented. We sincerely hope that in the near future we will feel comfortable urging patients and their caregivers to utilize DFC when choosing a clinic.

Respectfully submitted,

Jackson Williams
Director of Regulatory Affairs

cc: Joel Andress, Ph.D.
Kate Goodrich, M.D.