



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 system 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