



October 30, 2014

Marilyn Tavenner  
Administrator  
Centers for Medicare and Medicaid Services  
200 Independence Avenue, SW  
Washington, DC 20201

Dear Administrator Tavenner,

I am writing on behalf of Kidney Care Partners (KCP) in follow up to our ongoing discussions regarding the launch of ESRD Five Star. We appreciate your willingness to collaborate with the dialysis community to ensure the accuracy of reporting and to support high quality dialysis care. As we have noted previously, KCP supports transparent quality programs that provide accurate and useful information to patients and consumers. However, we have been deeply troubled by, among other things, the methodology of ESRD Five Star that distorts actual performance quality and may not support effective consumer choice of the provider that best meets their needs. We also remain concerned that the CMS staff and the contractor during national calls discussing ESRD Five Star continue to indicate that no changes will be made to the program prior to January, despite your letter to the contrary.

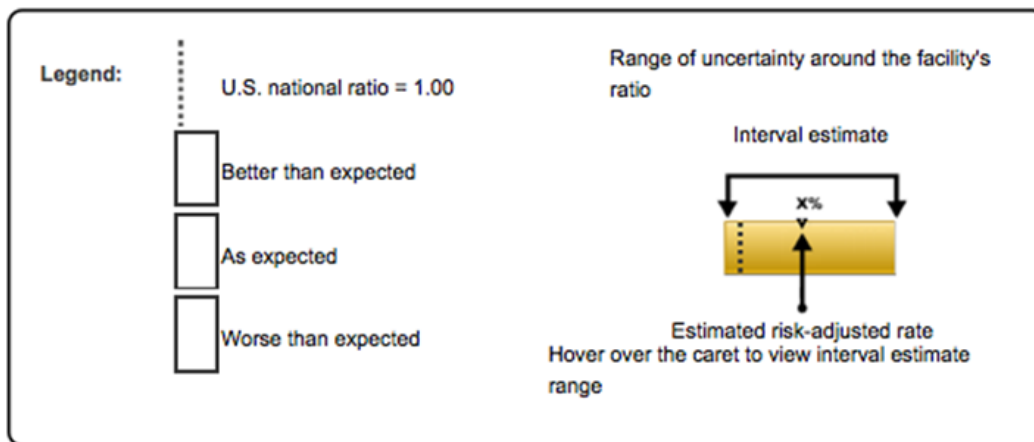
KCP is encouraged that the CMS leadership team has agreed to examine our concerns in an effort to address them before CMS finalizes ESRD Five Star. Since our most recent meeting with you, KCP has developed a set of scoring methodology recommendations that could be implemented prior to the desired January 2015 launch. These recommendations do not resolve all of the issues that we have identified; however, we offer them as a way to improve the methodology while allowing the Agency to maintain its current timeline for launch. KCP is continuing to work on ways to resolve our ongoing serious concerns that cannot be addressed in a comprehensive manner before January 2015. These concerns include data integrity issues and the need for a clear definition of quality. However, adopting scoring changes that more accurately reflect the actual quality performance of facilities would be a step in the right direction.

## **I. Short-Term Recommendations**

KCP has worked to develop a consensus set of recommendations that CMS would have the ability to implement and still meet its January 2015 timeframe. These four methodological recommendations seek to bring statistical validity to the methodology and to a small, but important degree, address some of the critical flaws of the proposed methodology.

### A. Use the Categorical Results for the Standardized Ratio Measures

Despite our concerns about the methodology and data for the standardized ratio measures, we understand that CMS is planning to use these measures as part of ESRD Five Star. If the Agency does include these measures, KCP recommends that CMS use the categorical results that are publicly reported via Dialysis Facility Compare (DFC), rather than the numerical values, for the standardized ratio measures. As noted on the legend to the standardized ratio measures on DFC, there is a significant amount of statistical uncertainty inherent in the measures, due to the way they have been designed. DFC explains this uncertainty to consumers with the following legend.



The yellow area shows the “range of uncertainty around the facility’s ratio.” This uncertainty means that if a facility has a ratio of 8, the actual ratio could be between 4 and 12. DFC compensates for this uncertainty by categorizing facilities into one of three categories: better than expected, as expected, or worse than expected.

In ESRD Five Star, CMS proposes to use the numerical value as the basis for scoring and ignore that uncertainty. A better approach would be to use the categorical results from DFC, assign a numerical score based on those categories, and then add the points to assign a domain score. In other words, a facility that is better than expected would receive 2 points; one that is as expected would receive 1 point; and one that is worse than expected would receive 0 points.

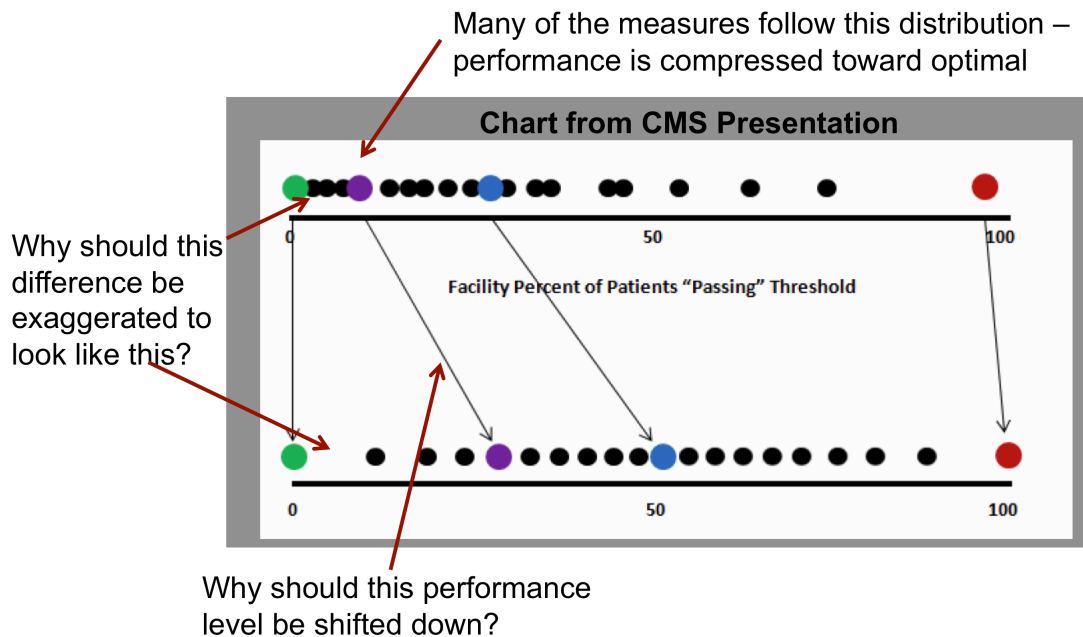
This recommendation is superior to the proposed use of the numerical value because it recognizes the statistical uncertainty of the standardized measures that CMS, itself, has consistently acknowledged in presenting the data on DFC. Because CMS already has these data from DFC, it is a straightforward effort to assign the values and add them to create the domain score prior to the January 2015 launch.

Based upon our experience with the data, implementing this recommendation would not require a significant amount of time.

## B. Use a Z-Score Model

Second, KCP recommends that CMS use a z-score model to score measures, instead of the proposed percentile model. While we understand that the contractor and staff continue to state that the ESRD Five Star methodology does not create a forced bell curve, the reality is that by selecting a percentile method to standardize the measures, the results have been significantly distorted.

The chart below comes from the initial CMS National Provider Call slides that described the ESRD Five Star methodology. The top line represents the actual performance of facilities, while the bottom line shows how CMS has applied the percentile methodology.



The application of the percentile methodology distorts the difference among facilities, in some cases taking small differences and magnifying them, in other cases shrinking significant performance differences.

Unlike the percentile methodology, a z-score methodology preserves the relative magnitude of performance differences across facilities. As the Agency's own data show, the performance distribution among facilities is not symmetrical. It becomes symmetrical only when applying a methodology that distorts the actual performance distribution.

Z-scores would allow CMS to describe performance by expressing it as a number of standard deviations from the mean—a concept that is more readily understandable. To apply this methodology, CMS would calculate a z-score for each facility for each measure. Then, it would average the scores within each domain. Finally, it would average the domains for an overall score. Using the most recent data available from CMS, the application of the z-score methodology and the use of categorical values for the standardized ratio measures shows that the distribution of performance is not a normal bell curve, but rather an asymmetrical curve.

We strongly urge CMS to use z-scores to preserve the magnitude of relative performance, rather than force the ranks as the percentile methodology does. Again, the application of the z-score methodology can be performed with existing data in a relatively straightforward manner in a short timeframe.

### C. Define Star Assignment by Setting Performance Criteria

Third, KCP recommends that CMS define star assignment by setting performance criteria and not percentage targets. As proposed, ESRD Five Star relies on an arbitrary determination of the number of facilities that can be assigned to each star category. Further, as we discuss in the following section, no accounting of inherent statistical measurement error has been applied. Under the CMS approach, 10 percent of facilities will always receive five stars and 70 percent of facilities will always receive 3 or fewer stars. Because the distribution of performance, as described above, is actually asymmetrical, the distribution is not a normal bell curve. Given this fact, the 10-20-40-20-10 percentage distribution of star assignments cannot be justified.

A more statistically valid and appropriate approach would be to use boundaries that are symmetrical and related to the DFC performance categories. The chart below shows the recommended cut points. The number of facilities in each star category would be determined by the performance on all of the measures across all domains.

	Base Criteria	“Plain English” Description
★★★★★	All three domain z-scores $\geq .25$	“The facility’s performance in <u>each quality domain</u> is better than expected.”
★★★★	Overall z-score $\geq .25$ (but does not meet 5-star criteria)	“The facility’s <u>overall performance</u> is better than expected.”
★★★	Overall z-score $\geq -.25$ and $< .25$	“The facility’s overall performance is as expected.”
★★	Overall z-score $\leq -.25$ (but does not meet 1-star criteria)	“The facility’s <u>overall performance</u> is worse than expected.”
★	All three domain z-scores $< -.25$	“The facility’s performance in <u>each quality domain</u> is worse than expected.”


Using performance criteria avoids the use of an arbitrary percentage distribution of stars. It provides symmetrical meaning to the stars and more closely reflects facility performance. This approach would provide more meaningful data to beneficiaries and consumers than the proposed ESRD Five Star assignment method.

**D. Adjust the Star Category Cut Points for Statistical Confidence**

Finally, KCP recommends that CMS adjust the star category boundaries to provide statistical confidence in the final stars. As we have described during previous meetings, there are several issues related to the data. While providing statistical confidence for the final stars does not address all of these concerns, it is a step in the right direction.

Based upon our analysis, using a 3 standard error confidence interval would provide almost 100 percent certainty that a facility is assigned to the right category. In contrast, a 2 standard error confidence interval would provide 95 percent certainty, while a 1 standard error confidence interval would provide only about 67 percent certainty. The chart below shows how the adjustment would be applied.

Z-Score Before Adjustment		"Net Z-Score" After Adjustment for Statistical Confidence
.25	Subtract 3 Standard Errors	0.222745
-.25	Reduces uncertainty that performance is below cut point	-0.27726



This recommendation compensates for the statistical uncertainty around boundaries for the star cut points.

**E. Impact of Recommendations**

Because the short-term adjustments would come closer to aligning with actual performance, the number of facilities in each category would shift, but the four short-term recommendations would not result in all facilities being defined as 4 or 5 stars. Thus, in adopting these recommendations, the Agency would not be shifting all facilities into the 4 or 5 star categories. The chart below shows the percentage of facilities in the different star categories if the KCP recommendations were implemented in total.

	Criteria After Adjustment for Statistical Confidence	Number of Facilities	Percentage of Facilities
★★★★★	All three domain z-scores $\geq .25 - 3SE$	1319	21%
★★★★	Overall z-score $\geq .222$ (but does not meet 5-star criteria)	1272	20%
★★★	Overall z-score $\geq -.277$ and $< .222$	2174	35%
★★	Overall z-score $\leq -.277$ (but does not meet 1-star criteria)	1311	21%
★	All three domain z-scores $< -.25 - 3SE$	160	3%

## II. Conclusion

We appreciate your commitment to working with us to improve ESRD Five Star before the January 2015 launch. As noted, KCP believes that it will take more than a few months to address all of the community's concern. However, we do believe that some short-term adjustments would improve the program in important ways and allow us to continue working with CMS to address the more challenging problems during the coming year. In sum, we strongly urge you to adopt the recommendations outlined in this letter. But even if CMS were to make each of these four recommendations, ESRD Five Star would be significantly improved, but not all of the problems would be resolved. We would like to continue working with you on these issues as well.

Thank you again for your willingness to engage with us on this important program. Kathy Lester will be in touch to schedule a meeting with your team to answer any questions about the recommendations and discuss their implementation.

Sincerely,



Edward R. Jones, M.D.  
 Chairman  
 Kidney Care Partners

cc: Patrick Conway, M.D., Deputy Administrator for Innovation & Quality, CMS Chief Medical Officer, Director of the Center for Clinical Standards and Quality  
 Kate Goodrich, M.D., Director of the Quality Measurement and Health Assessment Group  
 Elena Balovlenkov, R.N., Technical Lead for Dialysis Facility Compare  
 Joel Andress, Ph.D., Center for Quality Measurement in the Health Assessment Group