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

Summary Report

In-Person Meeting, Baltimore, MD

February 21, 2017

End-Stage Renal Disease Star Ratings Technical Expert Panel Summary

The Centers for Medicare & Medicaid Services (CMS) has contracted with The University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) to convene a Technical Expert Panel (TEP), including patients and experts with relevant experience, to obtain recommendations on potential quality measures to include in the Dialysis Facility Compare (DFC) Star Ratings.

TEP Objectives

The objectives of the End-Stage Renal Disease (ESRD) DFC Star Ratings TEP are described in the TEP charter provided (Appendix A) to TEP members prior to the in-person meeting. The TEP was tasked with using existing data and the experiences and expertise of its members to formulate recommendations for UM-KECC regarding the star rating statistical methodology, consideration of quality measures for retirement, future implementation, or updates, and the presentation of the star ratings on the DFC website.

TEP In-Person Meeting

The in-person ESRD TEP was convened in Baltimore, MD, on February 21, 2017.

The TEP consisted of individuals from the following areas of expertise or experience:

- Consumer, patient, and family perspectives on treatment;
- Biostatistics and statistics methodology;
- Clinical treatment of ESRD;
- Dialysis organization operation;
- Performance measurement and quality improvement.

The TEP was tasked with discussing the following topic areas:

- Measures to be added into the star ratings;
- Measure updates;
- Reporting patient-reported outcomes in the star ratings;
- Re-baselining to set the star rating thresholds.

The following individuals participated in this TEP:

Name and Credentials	Organizational Affiliation, City, State	Conflicts of Interest Declared
Paul T. Conway, BA *TEP co-chair* *President* *Board Member* *Board Member*	 American Association of Kidney Patients (AAKP) Mid-Atlantic Renal Coalition (MARC) Polycystic Kidney Disease Foundation (PKDF) Falls Church, VA	None
Catherine A. Sugar, PhD, MS *TEP co-chair* *Director* *Professor*	 Semel Institute Statistics Core, University of California, Los Angeles Los Angeles, CA Departments of Biostatistics, Statistics & Psychiatry University of California, Los Angeles, Los Angeles, CA	None
Lorien Dalrymple, MD, MPH *Vice President of Epidemiology and Research* *Volunteer Clinical Faculty, Associate Professor*	Fresenius Medical Care North America (FMCNA) Department of Medicine, Division of Nephrology, University of California, Davis	Employed by Fresenius Medical Care NA and member of the KCQA Steering Committee. Fresenius Medical Care (FMC) is a member of Kidney Care Partners (KCP), and the Chief Medical Officer for Fresenius Medical Care is the chair of KCP.
Amanda Grandinetti, MPH *Senior Specialist, Performance Measures and Analysis* *Kidney Action Committee Member*	American Academy of Dermatology National Kidney Foundation	None

Name and Credentials	Organizational Affiliation, City, State	Conflicts of Interest Declared
Mark Joseph, MD *Nephrologist* *Clinical Assistant Professor* *Co-Chair*	Pediatric Kidney Disease and Hypertension Centers, Phoenix, Arizona Department of Pediatrics, University of Arizona, Tucson, Arizona Public Policy Committee, American Society of Pediatric Nephrology	Member of the Horizon Pharmaceutical speaker's bureau for nephropathic cystinosis
Richard Knight, MBA *Vice President/Chair of Public Policy* *Board Member*	American Association of Kidney Patients (AAKP), New Carrollton, MD Mid-Atlantic Renal Coalition (MARC)	None
Jewell Kyle, RN, BSN, CNN *RN Charge Nurse/Staff Educator*	Chattanooga Kidney Centers, LLC	None
J. Richard Landis, PhD, MS *Professor of Biostatistics* *Professor of Statistics*	University of Pennsylvania Perelman School of Medicine Philadelphia, PA University of Pennsylvania (Wharton School), Philadelphia, PA	Currently working on two hemodialysis clinical trials.
Allen R. Nissenson, MD, FACP *Chief Medical Officer* *Emeritus Professor of Medicine*	DaVita HealthCare Partners El Segundo, California David Geffen School of Medicine at University of California, Los Angeles Los Angeles, CA	Full-time employee, DaVita, Inc. Co-chair of the Kidney Care Quality Alliance (KCQA)
Chris Sarfaty, MSW, LICSW *Patient-Centered Collaborative Care Coach* *Licensed Independent Clinical Social Worker*	Coaching for Health Professionals Whately, MA Baystate Franklin Medical Center	None

Name and Credentials	Organizational Affiliation, City, State	Conflicts of Interest Declared
Nicole Stankus, MD, MSc *Medical Director* *Associate Professor of Medicine*	DaVita Stony Island Dialysis Center The University of Chicago Chicago, IL	Member of the DaVita Physician Council
Sumi Sun, MPH *VP of Analytics and Quality Strategy*	Satellite Healthcare, San Jose, CA	None
David M. White *Board of Directors Member* *Acting Chair* *Chair on the Patient Advisory Committee and Medical Review Board member*	American Association of Kidney Patients (AAKP) Kidney Health Initiative's Patent and Family Partnership Council (PFPC) Mid-Atlantic Renal Coalition (MARC) Hillcrest Heights, MD	None

*One TEP member withdrew and did not participate in the TEP teleconferences, in-person meetings, and voting.

Contractor Staff University of Michigan Kidney Epidemiology and Cost Center (UM KECC)
Yi Li, PhD, *Professor of Biostatistics/Principal Investigator*
Joseph Messana, MD, *Swartz Collegiate Professor of Nephrology, University of Michigan Health System and Interim Director, UM-KECC*
Richard Hirth, PhD, *Professor of Health Management and Policy*
Claudia Dahlerus, PhD, MA, *Principal Research Scientist*
Ji Zhu, PhD, *Professor of Statistics*
Stephen Salerno, BS, *Graduate Student Research Assistant*
Karen Wisniewski, MPH, *Lead Research Analyst*
Natalie Scholz, MPH, *Lead Research Analyst*
Cindy Liao, MS, MPH, *Research Analyst*
Casey Parrotte, PMP, *Project Manager/ Research Analyst*
Jordan Affholter, BA, *Research Analyst*

1. Introduction

This report summarizes the discussions and recommendations of the ESRD DFC Star Ratings TEP convened on February 21, 2017, in Baltimore, MD, as well as the three pre-TEP teleconference calls conducted on February 2, 13, and 14, 2017.

2. Overview

The TEP was tasked with consideration of quality measures for retirement, future implementation, or update, and the presentation of the star ratings on the DFC website. In the course of preliminary discussions held on two pre-TEP teleconferences, the following topics were the main focus of the in-person meeting:

- Measures to be added into the star ratings;
- Measure updates;
- Integration of patient-reported outcomes into the star ratings;
- Re-baselining to update star rating thresholds.

3. Preliminary Activities

3.1 Environmental Scan and Literature Review

Prior to the in-person TEP meeting, UM-KECC presented the TEP members with the National Quality Forum (NQF) Measure Information Form (MIF) background materials related to the measures considered for inclusion in the star rating methodology. The TEP was also provided the DFC Star Ratings TEP Summary Report from 2015, the Star Rating Methodology Technical Notes, and the presentation from the October 5, 2016, and the August 31, 2016 National Provider Calls.

3.2 TEP Charter

The DFC Star Ratings TEP Charter was publicly posted with the nomination materials, and was distributed to the TEP members for review. The TEP Charter is included as Appendix A.

3.3 Pre-TEP Teleconference Calls

Three 90-minute preliminary teleconference calls preceded the in-person TEP meeting. These calls were held on February 2, 13, and 14, 2017.

The first part of the first pre-TEP teleconference call focused on introduction of the TEP members, the role of the TEP, the TEP Charter, and TEP objectives. The pre-TEP teleconference calls served as an introduction to (1) the candidate measures proposed for inclusion, (2) how to report patient-reported

outcomes in the star rating, and (3) how to re-baseline the star ratings. The pre-TEP teleconference minutes were provided to the TEP before the in-person meeting on February 21, 2017.

The pre-TEP teleconference minutes and public comments are included as Appendices C, D, and E.

4. TEP Meeting

4.1 Introductions

Yi Li, PhD (Professor of Biostatistics/Principal Investigator, University of Michigan) welcomed everyone to the in-person TEP meeting, and thanked the TEP members for their participation on behalf of the University of Michigan team for this project funded by the Centers for Medicare & Medicaid Services (CMS). He emphasized that the team had taken the TEP input very seriously.

After the opening statement, Dr. Li asked the University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) team to introduce themselves. Joseph Messana, MD (Clinical Nephrologist, University of Michigan), Claudia Dahlerus, PhD (Principal Research Scientist at UM-KECC), Richard Hirth, PhD (Health Economist and Professor of Health Management and Policy, University of Michigan), Ji Zhu, PhD (Professor of Statistics, University of Michigan), Stephen Salerno (Biostatistics Graduate Student, University of Michigan), and Jordan Affholter (Research Analyst, University of Michigan) introduced themselves to the TEP. Dr. Li also identified the other UM-KECC workgroup members who were present: Karen Wisniewski (Lead Research Analyst), Natalie Scholz (Lead Research Analyst), Cindy Liao (Research Analyst), and Casey Parrotte (Project Manager/ Research Analyst).

Dr. Li, Dr. Messana, Dr. Hirth, and Dr. Dahlerus are the TEP co-facilitators for the DFC Star Ratings TEP.

Joel Andress, PhD (CMS) introduced himself as the Project Contracting Officer's Representative (COR) for CMS's development of the star rating methodology. Dr. Andress thanked all of the TEP members for volunteering their time to this project. Elena Balovlenkov, RN, introduced herself as the Dialysis Facility Compare (DFC) lead for Public Reporting and thanked the TEP members for their participation. Dr. Jesse Roach introduced himself as a nephrologist and Medical Officer working at CMS.

The following ICH CAHPS (In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems) representatives were in attendance or participated over the phone: Scott Scheffler, MS (RTI), Judy Lynch (RTI), Julia Zucco, PhD (CMS), Elizabeth Goldstein, PhD (CMS), and Celia Eicheldinger, MS (RTI).

The following CDC representatives participated: Shunte Moon, PhD, dialysis epidemiologist and representative from the Centers for Disease Control and Prevention (CDC) was in attendance. Priti Patel, MD, Medical Officer and Dialysis Activity Leader at the Centers for Disease Control and Prevention (CDC) participated via teleconference.

CMS' public reporting contractors NORC and Ketchum were in attendance. Abigail Fredenburg from Ketchum introduced herself. Rebecca Catterson from NORC was also present.

The TEP Members introduced themselves and disclosed their conflicts of interest. The TEP member affiliations and conflicts of interests are documented in the TEP In-Person Meeting section.

Dr. Li (UM-KECC) reviewed the TEP member role and responsibilities and the three TEP objective items from the TEP charter, which include: recommendations to CMS about measure additions, approach to including and reporting patient-reported outcomes in star ratings, and when and how to re-baseline to update the star rating thresholds. The in-person meeting agenda was also reviewed and it was noted there would be a public comment period at the end of the in-person meeting.

4.2 Preliminary Voting Results

Dr. Claudia Dahlerus (UM-KECC) explained that she was going to present the results of the TEP member preliminary voting on the candidate measures that took place after the second pre-TEP teleconference call (voting form was circulated after the second teleconference and TEP members were asked to vote on each measure using the 5-point Likert scale). The TEP members were asked to rate the measures for potential inclusion in the star ratings (see Appendix G for the preliminary voting form). For the Fistula, Catheter, SMR, SHR, and STrR measures, TEP members were asked if the updated version of the measure should replace the current version of the measure used in the DFC Star Ratings. For the Pediatric PD Kt/V, NHSN SIR, ICH CAHPS, and SRR measures, the TEP members were asked to rate the measures for potential inclusion in the star ratings. The results of the preliminary voting were used to help inform and prioritize agenda items for the in-person TEP meeting discussion.

Before reviewing the preliminary voting results, Dr. Dahlerus stated that 13 out of the 14 TEP members completed the preliminary voting form. One TEP member only rated the ICH CAHPS measure and abstained from rating the other measures, therefore 13 participants rated the ICH CAHPS measure, while 12 TEP members rated all of the measures. Dr. Dahlerus presented the preliminary measure voting results which are displayed in the table below.

Table 1: Preliminary Measure Voting Form Results

Measure	% Strongly Agree/Agree	% Neither Agree nor Disagree	%Strongly Disagree/Disagree
Standardized Fistula Rate	92%	8%	0%
Long-term Catheter Rate	92%	8%	0%
Standardized Mortality Ratio (SMR)	84%	8%	8%
Standardized Hospitalization Ratio (SHR)	92%	0%	8%
Standardized Transfusion Ratio (STrR)	75%	8%	17%
Pediatric Peritoneal Dialysis Adequacy: Achievement of Target Kt/V (Pediatric PD Kt/V)	58%	42%	0%
The National Healthcare Safety Network Bloodstream Infection measure (NHSN SIR)	50%	8%	42%
In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)	70%	15%	15%
Standardized Readmission Ratio (SRR)	50%	17%	33%

Dr. Dahlerus stated that the Standardized Fistula Rate, Long-term Catheter Rate, Standardized Mortality Ratio (SMR), Standardized Hospitalization Ratio (SHR), Standardized Transfusion Ratio (STrR) measures are the updated candidate measures that were updated as part of the endorsement process with the National Quality Forum (NQF). Dr. Dahlerus stated that the voting results demonstrate overall consensus

on adding the updated candidate measures (Fistula, Catheter, SMR, SHR, STrR), and the Pediatric PD Kt/V candidate measure to the star rating October 2018 release.

Based on the voting results and discussions during the pre-TEP teleconferences, the NHSN SIR, SRR, and ICH CAHPS measures were identified as topics for further TEP discussion at the in-person meeting.

Although the ICH CAHPS measure did receive a majority vote in favor of adding this measure to the star ratings, it was added as a discussion point because there were outstanding questions from the second pre-TEP conference call about which form of ICH CAHPS measures should be used (i.e., whether to use the top-box, which reflects the percentage of patients that selected "Always" or "Yes", or, to use the linear mean score based on all the data). Dr. Dahlerus clarified that the ICH CAHPS group (CMS and RTI) would present an overview of scoring options for the ICH CAHPS measures and take questions from the TEP members.

Dr. Dahlerus stated that after the New Candidate Measure discussion, TEP members will be asked to take another vote on the SRR and NHSN SIR in order to determine whether clear consensus could be reached for those measures. In addition, the TEP will also be asked to vote on the scoring options for the ICH CAHPS measure.

4.3 Candidate Measures Discussion

4.3a National Healthcare Safety Network Bloodstream Infection measure (NHSN SIR) Discussion

Dr. Dahlerus (UM-KECC) explained that the TEP co-chairs will lead the discussion on the NHSN SIR measure and the CDC will be able to answer TEP questions regarding the NHSN SIR measure.

One of the TEP co-chairs stated that monitoring bloodstream infections (BSIs) is an issue that is very important to patients. The TEP co-chair stated that facility staff have guidelines they must follow for preventing BSIs. The TEP co-chair explained that every person who approaches a dialysis patient may be a potential source of infection to the patient, stressing the importance of infection prevention. The TEP co-chair stated that the discussion during the pre-TEP teleconference call mainly focused on the technical rigor of the measure. They reinforced that it should not be forgotten that infection prevention is very important to patients.

One TEP member agreed that the BSI issue is important. The TEP member stated that the discussion around validity is even more important because of this topic's significance (infection prevention and safety). The TEP member stated that because of this, it is important to have accurate information so that patients are not misled with incorrect data. The TEP co-chair agreed that technical rigor is imperative.

To address the underreporting issues surrounding the infection measure, Dr. Shunte Moon (CDC) stated that the CDC is working with CMS to increase validation of the data reported to NHSN. Dr. Moon stated that the opinion of the CDC is that the measure should not be dismissed due to underreporting. He stated that the studies referenced in the Quality Incentive Program (QIP) Final Rules from PY2016 – PY2020 were done prior to mandating of NHSN SIR reporting for QIP. Dr. Moon stated the CDC conducted a study validating 2015 data in Georgia, which had a larger sample size than previous validation studies. In response to concerns about the finding of previous studies Dr. Moon referenced one CDC validation study involving 30 dialysis facilities in Georgia, which showed that 29% underreporting of bloodstream infections occurred. A second validation study conducted in Colorado

showed that 24% underreporting occurred. As a result of these unpublished studies, Dr. Moon stated that the CDC is working on determining where underreporting occurs and is also working to determine which facilities have zero infections. Dr. Moon also referenced the graph displayed (during an earlier teleconference) that showed a large number of facilities that reported zero BSIs. In response to this concern, he stated the CDC believes many of those facilities may truly have zero bloodstream infections given the small facility size. Dr. Moon acknowledged that the CDC is aware that underreporting does happen, but the large range reported of missing data may not be accurate. Dr. Moon stated that CDC is working with CMS and other stakeholders to determine when underreporting occurs. Dr. Moon stated that the measure should not be dismissed, because it is a valuable measure to patients, and is used for prevention. For example, Dr. Moon stated that the CDC has observed lower infection rates among facilities that use infection prevention interventions. Dr. Moon stated the CDC is working on additional studies that have larger sample sizes to improve validation of the measure.

After this overview, Dr. Priti Patel (CDC) provided additional background. She stated that the NHSN SIR is an important outcome and the CDC believes this measure is valid and provides important information. Dr. Patel stated that the CDC has concerns with the small sample-size studies that suggest the NHSN SIR measure is invalid due to underreporting (i.e., the studies cited in the ESRD QIP rule and referenced earlier in this discussion). The 2012 and 2014 studies (referenced in the QIP rules) that cite high amounts of missing data were conducted before NHSN mandatory reporting, and relied on data from one electronic medical system, but they do not reflect the validity of the current reported data. Dr. Patel explained that the current reported data is different than the data reported in the earlier 2012 or 2014 validation studies. Dr. Patel provided further background on the more recent and larger Georgia and Colorado validation studies. For example, the Georgia study targeted facilities with high numbers of catheters (where one might expect more vascular access infections) but reported either no BSIs or a low number of BSIs. Dr. Patel stated that the underreporting rate across the facilities was about 29%, noting this is smaller than the underreporting numbers referenced in the 2012 and 2014 validation studies included in the ESRD QIP rule. For example, the Colorado study observed a 24% underreporting rate. Dr. Patel explained that it is reasonable to conclude that the data quality is higher now than at the time the earlier validation studies (referenced in the QIP rules) were conducted as evidence of underreporting.

In reference to facility size and reliability, Dr. Patel stated the CDC and CMS are investigating an adjustment to account for different levels of reliability in facilities with fewer patients. For example, a facility with 60 patients and zero BSIs may have a different reliability than a facility with 800 patients and zero BSIs. The measure does not currently have a reliability adjustment, but that may be considered through future measure maintenance.

The TEP co-chairs thanked Dr. Patel (CDC) for providing the background information and opened up the discussion to the TEP members.

One TEP member asked a clarifying question about the CDC Colorado and Georgia validation studies. The TEP member asked if the validation rates differed by facility. Both CDC representatives stated they did not know if the CDC has looked at the different validation rates by facility. Dr. Moon stated that most studies found that most underreporting was due to misunderstanding the reporting protocol and that CDC is educating facilities on reporting guidelines. Dr. Moon stated that as the CDC continues their education outreach more facilities will better understand how to report the SIR measure.

One TEP member stated that the NHSN SIR measure is important and is a work in progress. The TEP member stated concerns about how the previous 2012 and 2014 validation studies show 60%-80% underreporting, and the more recent Georgia and Colorado validation studies show (about) 30% underreporting. The TEP member expressed concerns about a measure that has a 30% underreporting

rate. The TEP member stated that while there is a lot of education being done for the NHSN SIR measure, they do not believe the measure is currently ready for use in the star ratings. The TEP member stated it is unlikely that many of the facilities that report zero infections actually have zero infections. The TEP member stated that often infections can occur in the facility or at the patient's home, for example, a large number of infections may be related to how catheters are cared for at a patient's home. TEP member did recognize some facilities may have zero infections, but it would likely be a small percentage. The TEP member stated it would be beneficial to have an infection measure, but that this measure is not ready at this time.

Dr. Patel (CDC) responded to the TEP member's statements saying that a majority of the time catheters are being accessed in the dialysis facility, and did not agree that infections are frequently caused by the home environment. Dr. Patel stated that there is both underreporting and overreporting, and stated that the SIR measure is important and should not be dismissed because the measure is not perfect.

One of the TEP co-chairs then clarified that the more important question is whether the SIR measure orders ("ranks") facilities correctly by performance (i.e., measure scores). For example, if all of the facilities underreport by 20%, but are in the correct performance order, then the measure could still be useful for the star rating, even if the measure is not entirely accurate. The TEP co-chair asked if the CDC believes that even with the underreporting that the facilities are ordered correctly to allow for relative comparisons.

The CDC's Dr. Patel stated "we don't 100% know" if the facilities are ordered correctly by performance. She went on to state that underreporting may be balanced by overreporting, but the CDC does not know to what extent, and they currently cannot determine if the level of underreporting or overreporting is consistent.

One TEP member clarified their earlier statement about the number of infections that may occur at home. The TEP member agreed with Dr. Patel that the majority of, but not all, BSIs occur in the facility, hence the facility may not know about those infections happening at home. The TEP member stated that the level of underreporting across facilities is unknown. The TEP member clarified that if there is a minor difference in underreporting then it may not matter, but if there is a large difference in the percentage of underreporting across facilities, then that may be an issue. The TEP member described a hypothetical to illustrate their point. They said it is possible that diligent facilities (good at reporting all the time) are also the most likely to have high infection reporting rates, therefore, in terms of public reporting on the measure they would be considered as very poor performing facilities. The TEP member stated this is the opposite of the intention of the measure. The TEP member further stated that if there are consistent standards about drawing and reporting blood cultures, then facilities could be fairly compared against each other. But absent these formal standards (for drawing and reporting blood cultures) they stated that the best way to have a low SIR is to never draw blood cultures (resulting in an unintended consequence of the current measure).

Dr. Patel (CDC) responded that CDC did not agree, and stated that every facility should be diligent about drawing blood cultures and providing the best patient care.

In response to the discussion about blood cultures, another TEP member stated that if a blood culture is drawn and is shown positive for infection, and it persists after 21 days, then the persisting infection is counted as a new event. The TEP member stated that their facility treats many chronic infections. Based on how the measure is defined, an infection persisting after 21 days could be considered a new (second) infection, and then, their facility could get penalized twice for the same infection. The TEP member asked if anyone has looked at catheter rates in facilities that have recorded zero percent infections as

one might expect facilities with zero infections would likely have a low percentage of catheters being used as well. Dr. Patel (CDC) first clarified that it is a very small percentage of patients that have more than one reported infection event (in response to the example about a persistent infection lasting more than 21 days). Dr. Patel then addressed the TEP member's question about catheter rates and clarified that the CDC has not looked at that information yet, but will try to follow-up and report that information.

The discussion continued with other comments from TEP members. One TEP member stated that over the course of a dialysis session there are an estimated 30 times when hand hygiene is needed, therefore, there is a high probability for an infection event to happen even if a patient has the most attentive dialysis technician. The TEP member stated that they have received very good care at the facility but acknowledged facility staff still have a very difficult job in infection prevention. The TEP member felt it is important for the SIR measure to be included in the star rating because that would encourage facility staff to further increase attentiveness to bloodstream infection prevention practices if they know that their facility will be measured on this clinical outcome. The TEP member further recognized that many patients are uncomfortable asking facility staff to wash hands or change gloves and therefore sometimes the best treatment practices do not happen despite the facility's efforts. Because of this, they felt that it is important to increase the data available for reporting on bloodstream infections. The TEP member stated that despite underreporting, the NHSN SIR is an important measure and stressed the need for increased emphasis on reporting, in order to get more valid data. TEP member stated that from the perspective of patient-centered outcomes, it makes sense to include this measure and continue to make improvements to the measure. The TEP member stated that while they wanted facilities to be compared fairly, they also wanted to make sure patients receive the best care. The TEP member stated that there may be a short-term productivity cost effect for implementing this measure, but over the long-term, overall patient care would improve.

One TEP member asked if there is a separate comparison (measurement) for pediatric and adult patients for the NHSN SIR measure. Dr. Patel (CDC) responded there is not a separate comparison for pediatric and adult patients right now but the CDC is working with pediatric providers to create reports in NHSN where pediatric facilities could be compared against pediatric facilities. CDC is still investigating this topic. The TEP member also stated that the pediatric community cannot support this measure at this time because pediatric specific comorbidities and conditions (associated with higher infection risk) are not included. The TEP member also stated the home environment plays a large role in infections for pediatric patients and especially among very young pediatric patients (i.e., where they may pull at the catheter and expose their access site).

The TEP discussion returned to the topic of underreporting and questions about facilities with zero infections reported. One TEP member stated they were interested in how underreporting is detected. They stated some facilities may have zero BSIs because they are a vigilant facility, but as in the example described earlier (by another TEP member) other facilities may have zero BSIs because they are not reporting properly. The TEP member asked if random sampling has been done in order to better understand the pattern of underreporting, for example, if blood cultures were collected for a certain week (for a random set of facilities) to estimate whether infections are being detected correctly.

Dr. Moon (CDC) responded that CDC performs validation by comparing the electronic medical record system results to results reported in NHSN. The CDC has seen overreporting on items such as more than one positive blood culture within 21 days. The CDC is working with the facilities to provide education on the correct way to report.

One TEP co-chair stated that there may be underreporting because tests of blood cultures are not being conducted. The TEP co-chair stated that it may be necessary to go to a facility for a week and check the blood cultures and compare versus what was recorded in the record. The TEP member asked if there was data where blood culture samples were taken to estimate underreporting.

Dr. Patel (CDC) stated that infections are relatively infrequent occurrences and most facilities may have about one bloodstream infection in a year. The CDC does not recommend blood culture surveillance because of other considerations such as false positives. The CDC validates infection events by reviewing labs, paper medical records at the facility, hospital records, patient absences at the facility, hospitalizations, and other means to determine validation.

One TEP member stated that a large number of infections occur because of poor (less skilled) dialysis technicians. The TEP member stated that it can be very time intensive for facilities to request blood cultures or hospitalization records. Dr. Patel (CDC) agreed that it is onerous to determine if a patient was admitted to the hospital for bloodstream infection, and, that it may not be easy for facilities to get that information. Dr. Patel stated that a majority of infections are bloodstream infections, and that there are multiple ways patients can get infections (at the facility) including through vascular access, contaminated hands or contaminated water, and therefore the CDC believes the measure is capturing infections that are under the control of facility practices.

At this point in the TEP discussion, Dr. Andress (CMS) summarized two competing issues that emerged from the discussion. Issue one is the agreement that the topic of patient infection is important and it merits consideration in terms of promoting quality improvement. Issue two is about the accuracy of the data. Because of the cost involved to determine that, it is unlikely we will know that information in the near future. Therefore, Dr. Andress stated, CMS is looking for the TEP's recommendation on moving forward with the current NHSN measure.

The discussion continued. One TEP member stated that from a nurse's perspective, patients may have bloodstream infections for various reasons (e.g., homelessness, drug use, or missing treatment), and that for these reasons the infection may not be the fault of the facility. The TEP member also asked how CROWNWeb relates to the statistics and reporting of infection. Dr. Andress (CMS) responded that CROWNWeb is not collecting infection data at this time and therefore does not provide this information. Dr. Andress further stated the best way to capture unreported infections is through medical chart review. The TEP member asked why reporting of infection could not be implemented into CROWNWeb as reviewing charts is time consuming and inefficient. They recommended that CROWNWeb be used for reporting of infections.

Another TEP member stated that TEP members agree that BSIs are important and clarified that the NHSN BSI measure is both publically reported on DFC and used in the QIP program. The TEP member further stated it is the responsibility of the TEP to determine if the NHSN SIR measure also belongs in the star ratings. The TEP member also acknowledged that the CDC is working to gain more information and that the amount of underreporting is unknown at this time.

Another TEP member agreed (with the previous statement), noting the NHSN SIR is already on DFC for patients to view and therefore they already have access to information about facility infection rates. The TEP member went on to explain that because of the poor validity (underreporting) the current SIR measure was not ready to be used in the star ratings. The TEP member stated that the dialysis community is looking to control the number of measures (which has increased), therefore getting the measure correct is more important than quickly adding another measure to the star rating.

The TEP co-chair recalled the intent of the star ratings, referring to the (former) Secretary of Health and Human Services Kathleen Sebellius, (former) CMS Administrator Marilyn Tavenner and Dr. Patrick Conway (CMS Chief Medical Officer). They set the expectations that the star rating program would be for patients to easily discern and compare quality among facilities. Since then, there have been three years of messaging to patients that DFC is the place for them to go to determine the best place for care for themselves or for others. The measures chosen should be measures that are the most important to patients and that should be communicated to patients. The technical rigor is the responsibility of the agency that develops the measures. The TEP co-chair clarified that bloodstream infections are a fundamental concern to patients and is something they want to know about. The TEP co-chair stated that every time a patient sees a new person (facility staff) or sees something that is not addressed, that can be a threat to a patient's life. The star rating program therefore must address the question of whether the facility works diligently to protect the patient's life. The TEP co-chair then closed the discussion on the NHSN SIR measure and asked that the TEP move to the second topic of the ICH CAHPS measure.

4.3b In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) Discussion

Scott Scheffler from the Research Triangle Institute (RTI) provided an overview of the ICH CAHPS survey, including an explanation of how CAHPS surveys are used in star ratings for other compare sites (such as Home Health Compare). Scott Scheffler emphasized that the information about ICH CAHPS scoring presented by CMS and RTI during this meeting is not meant to be presumptive or suggestive, but to help facilitate TEP discussion of the ICH CAHPS and reporting options.

One of the TEP co-chairs clarified two items regarding the approach for including the ICH CAHPS in the star ratings. One point was to note that what is being presented here by RTI uses a linear mean based on a continuous measure. This is different from the "top-box" results that are reported on DFC and were presented on the pre-TEP calls. The TEP co-chair stated the UM-KECC team only has access to the top-box scores, and not the linear scores. As a result, the preliminary analyses presented to the TEP on the pre-TEP call are distinct from the methodology that was being presented at the in-person meeting by RTI.

The second point of clarification had to do with cluster analysis utilized by RTI for the other star rating programs. The TEP co-chair explained that instead of deciding ahead of time what percentage of your data (e.g., facilities) will fall into several categories, clustering allows you to look for natural breaks in the data. The TEP co-chair stated they were unclear about the specific clustering methodology used by RTI (for the other star rating programs).

Scott Scheffler (RTI) responded by stating that the data for top-box is different than the linearized means data. He agreed with the TEP co-chair about the clustering methodology differences and was not sure how data were combined for the other CAHPS programs (e.g., Home Health) but he stated that this is something they are currently investigating with the other RTI teams.

The TEP co-chair asked why the linearized data has not been used on DFC or released to UM-KECC and if it would be possible to use it in the future. Dr. Andress (CMS) responded by saying that the decision to use top-box data was driven by the need to consider how the data are presented on the website for consumers' use versus how one incorporates this into a complex methodology for the star ratings. He explained that Hospital Compare uses a split approach of presenting the data; that is, the data is presented using the top-box approach on the public reporting site but the linearized data are used to calculate the hospital star rating. CMS stated that this is an option to explore for the DFC Star Ratings.

The TEP co-chair asked CMS to clarify whether using linearized scores for the measure is an option, or is the methodology restricted to only top-box data because that is what is currently reported. The TEP co-chair also asked if the linearization score data would be made available to UM-KECC. CMS stated that they believe the data will be made available to UM-KECC. CMS asked RTI to verify.

RTI stated that they believe they are capable of doing the work to produce the linearized data for UM-KECC in 6-9 months, however, further discussion with CMS would be needed. Dr. Andress (CMS) agreed and further clarified that UM-KECC will have access to the data but did not at the time of the meeting. In addition, CMS noted that they are asking TEP members to provide feedback as to whether this is something they want to pursue (using linearized scores), or if they would prefer to use the top-box scores as it may be more interpretable by consumers.

The TEP co-chair responded that knowing the percentage of subjects who responded "strongly agree or agree" is very interpretable, but many TEP members on the pre-TEP calls expressed some confusion about combining all the individual items to create a single score. Furthermore, the TEP co-chair expressed concern that using the top-box scores will result in losing much information about the distribution of the scores. The TEP co-chair recommended using the full set of data for all the item responses to create the linearized scores. They also emphasized that if items are combined the result will no longer have the single item percentage interpretation. The TEP co-chair asked the patient representatives on the TEP to weigh in on the interpretability of the results and the path forward for incorporating the measure into the star ratings.

The other TEP co-chair asked CMS to clarify the intent of the discussion of this measure. CMS responded by stating that the current star ratings currently lack any direct information about how patients experience care. CMS believes that this is an important issue because there are two purposes to the provision of dialysis care: extending life and quality of life. CMS further explained that CMS has heard from patients that they would like to see information provided directly by patients, which led to the inclusion of the ICH CAHPS on DFC. While it is not part of the star ratings at this time, patients can get patient experience information (from DFC) and start the conversation with the facility as to what the scores mean.

A further point was made about the importance of patient experience of care information. A TEP member stated that validity means different things to different groups (clinicians, statisticians, patients). For patients, validity means, "Do I care about this?" This TEP member felt that if there are no patient-reported outcome measures in the star ratings then many patients are not going to care and will not use the DFC website to inform their decisions about care. The TEP member further emphasized that the star rating is created for patients, so it is critically more important that it be included than for the measure to be "correct". They felt the measure experts can work on getting the technical components right as time progresses.

Another TEP member agreed that this is a good argument for having a separate rating for patient experience and a separate rating for the clinical measures, in order to make the patient experience information stand out more. They asked RTI how they handled instances where the surveys are administered multiple times during the year. The TEP member also asked if there are any benchmarks for the response rate that indicate that it would be an accurate representation of the whole population.

Judy Lynch (RTI) responded by saying that Home Health CAHPS is administered on a monthly basis and the results are analyzed based on 12 months of data. ICH CAHPS is administered twice a year, but for public reporting they use the two most recent survey administrations. RTI further explained the rules are not yet developed for a minimum number of completed surveys for ICH CAHPS for the star ratings,

but there are rules in place for other surveys. For Home Health CAHPS (HH CAHPS) star ratings, the minimum number of completed surveys is 40 over four quarters of 12 months of data.

Another TEP member asked RTI to clarify if they are resampling from the same population for the Home Health (HH CAHPS) and Hospital CAHPS (H CAHPS). RTI responded that for HH CAHPS there is a rule that patients cannot be sampled more than once. For ICH CAHPS, the patient population is very small so some of the same patients are re-sampled in multiple survey periods. Elizabeth Goldstein (CMS) added that for H CAHPS if a patient has multiple hospitalizations within a given month they cannot be sampled more than once.

A TEP member asked RTI/CMS, what is the median number of completed surveys in a 12-month period for H CAHPS, HH CAHPS, or ICH CAHPS? The RTI team agreed to make that information available after the meeting.

A TEP member then asked the statisticians on the TEP if there is an issue with the ranking of facilities when there is a lack of statistical precision when facilities do not meet the goal of having 200 completed surveys in a 12-month period (which they noted may be very rare).

The TEP co-chair responded that one would get an unbiased estimate of the typical patient response even with a small number of surveys, provided that the completed surveys are representative of the patient mix. For example, one does not need as many as 200 surveys to get a reasonably accurate result. However, the problem is the facilities that have an extremely low number of completed surveys. In these cases it would not be appropriate to calculate the measure for those facilities without the minimum number of surveys.

Another TEP member noted that it is important to realize that adult facilities have a lot of older patients, and that the prevalence of dementia for patients over 80 years old is 30%. There is concern about the validity of the responses from this population. The TEP member then asked if there have been any attempts to examine differences between those that complete the surveys versus those that do not.

Scott Scheffler (RTI) noted that there is a common misconception that the completed surveys are done by younger and healthier patients, however, they have examined the demographics of those who respond to the survey and have found that not to be the case. RTI has also been tracking survey burden, and they do not believe that they are seeing any evidence of survey burden. RTI consistently finds that a patient that responded in the past tends to respond to the survey again.

RTI further added they have examined response rates by comorbidity conditions (e.g., diabetes, hypertension, and other types of co-morbidities). The overall response rate from all participating facilities is about 31%, with the range of response rates among participating facilities from 70% down to <1%.

One TEP member asked for clarification on whether they are being asked to consider whether or not to include the ICH CAHPS in the star ratings, or only where to report ICH CAHPS. The TEP co-chair answered that it is both. The favorable support to include it in the star ratings (based on the preliminary voting survey) was quite high, but the outstanding question is about how to incorporate the measure. As discussed, one of the main issues is that the measure is different from the other (clinical) measures in the star ratings, and the measure has a fairly low percentage of facilities that are eligible to receive a score. As a result, the focus of the discussion has been on how to report the measure on DFC, and whether to use top-box scores, linearized scores, or another scoring method. The TEP member responded that they were in favor of separate star ratings for patient experience because of the importance of that information to patients.

It was noted by another TEP member that ICH CAHPS does not assess patient experience in the pediatric dialysis population. They believed that is an important outstanding question and wanted more information on the long-term plan for pediatric dialysis units with respect to patient experience in the star ratings.

Dr. Andress (CMS) responded that they have worked to incorporate pediatric patients in the specification of measures to the extent possible. CMS has also been asked to look at options for stratified reporting of measure results (as a result of the Assistant Secretary for Planning and Evaluation (ASPE) report and the Impact Act). For example, CMS could potentially look at pediatrics as a sub-category for that kind of reporting. Julia Zucco (CMS) also noted that there are preliminary plans to explore additional populations that are not currently surveyed with ICH CAHPS, such as home dialysis.

One TEP member asked RTI to elaborate on what languages the survey is offered in, and whether a diverse patient mix is reflected in the responses.

Judy Lynch (RTI) clarified that the survey is offered in Spanish, Samoan, and different Chinese language versions. The surveys in Chinese and Samoan are only conducted via mail, while for Spanish and English they are conducted either by mail or by telephone. The response rate is lower for non-English speaking populations, except for Spanish. RTI often asked ICH CAHPS facilities to let CMS know if they are other languages that the survey should be translated into, but they have not received any additional requests at this time.

A TEP member asked to hear from patients on the TEP regarding survey burden. One TEP member stated that their personal experience is that they have never experienced survey burnout, and that they actually look forward to the ICH CAHPS survey because that is the only time a patient perspective will count. The survey is the only way this TEP member knows that their voice is heard.

Another TEP member noted that there were times when they were more than willing to complete the survey, but sometimes they were not at all interested simply because of mood swing reactions to dialysis. Sometimes patients are just not physically able to take the survey, but they also support offering the survey more often.

A TEP member expressed their concern with patients filling out the surveys multiple times and the effect on the accuracy of the measure. For the future, they would like further investigation into other ways to obtain results with just one response per patient for the ICH CAHPS measures.

The TEP co-chair closed the discussion by noting that when patients are filling out a survey, they should ask “to what end?” The point is that surveys are not always about self-interest; response rates may improve if it is made clear that those responses can help shape the experience of the next person that will sit in the dialysis chair today and two years from now.

Another TEP member asked CMS to consider whether response rates would improve if surveys were administered in the facility by facility staff, instead of by a third party vendor over the phone or by mail.

4.3c Standardized Readmission Ratio (SRR) Discussion

UM-KECC summarized the main outstanding issues with the SRR measure based on the pre-TEP teleconferences. The main concerns expressed by TEP members were facility attribution for readmissions, and whether readmissions were within or outside the facility’s control.

The TEP co-chair asked if there is a reasonable expectation that facilities can take action to manage readmissions, and relatedly, they asked the group to consider whether this measure is important enough to patients as to warrant including it in the star ratings.

One TEP member responded the SRR is related to the quality of life, so to them it is definitely important. Another TEP member noted that social workers are placed in facilities to examine the psychosocial dimensions of dialysis, and to determine how those factors impact someone's ability to follow through and manage their condition and being on dialysis. They felt that although a facility may not have control over particular aspects of patient treatment, it is expected that a facility works to help coordinate care. Another TEP member emphasized that while some readmissions may be outside of a facility's control, all providers (dialysis facilities, hospitals) must work collaboratively to address the issue of care coordination to prevent readmissions. Another TEP member suggested asking patients, perhaps through CAHPS, how well they are educated by their facility. Patient education is in the facility's control and could have an effect on readmissions.

The TEP co-chair asked CMS to clarify whether the adjustments in the measure are intended to control for things that could be outside the facility's control and thus impact the rate of readmissions across the facilities. Dr. Andress (CMS) confirmed that the TEP co-chair's statement is correct, and that the SRR has a complex risk adjustment strategy to attempt to control for factors outside of the facilities' control. CMS acknowledged that facilities do not have total control over readmissions but for any measure that assesses coordination of care it is implicit that multiple providers would work to coordinate care. CMS' strategy for care coordination measures is to try to adjust for factors that are clearly outside of the facility's control; for the SRR, this means excluding planned readmissions, and risk adjusting for a set of patient comorbidities (in the prior year) and high risk conditions, along with adjustment for other patient characteristics. Dr. Andress also suggested that the question be framed as, "Is it in the power of facilities to reduce the risk for readmission that patients face?"

A TEP member asked CMS to confirm whether SRR adjusts for geography, noting that it is widely understood that hospital admissions differ across the country. Dr. Andress (CMS) explained that SRR does not adjust for geography for various reasons. One reason is the concern that if there is a difference in practices that can affect the quality of care, CMS still wants that to be reflected in the assessment of care. One TEP member noted that local geographic adjustors may make more sense, considering the purpose of the star ratings, which is for patients to make a choice in their local community.

Another TEP member noted that the risk adjustment strategy in the SRR is not adequate to assess pediatric patients. They strongly recommend that pediatrics and adults be considered separately.

Dr. Andress (CMS) acknowledged the concern, and explained that they have met with the American Society of Pediatric Nephrologists (ASPN) to discuss the current approach to inclusion of pediatric patients in quality measurement. CMS also invited the TEP member to provide CMS with suggestions and modifications to the SRR that would make it more appropriate for measuring pediatric patients.

At the end of the discussion, TEP members were asked to vote on the SIR and SRR and whether these should be added to the star rating. They were also asked to vote for a scoring option for the ICH CAHPS.

4.4 Recommendations on Method for Inclusion of Current and Future Patient-Reported Outcomes in the Star Ratings Discussion

4.4a Opening Discussion

The TEP co-chair first provided a follow-up to the morning discussion of the NHSN SIR measure. The co-chair stated that the CDC sent an additional graph via email after the NHSN discussion, which provided further information that showed the actual number of BSIs (without standardized adjustments). It was displayed to TEP members for reference (see Appendix N).

Dr. Moon (CDC) explained that the graph was provided in response to TEP questions regarding the large number of facilities reporting zero BSIs. It was intended to inform the TEP members of the distribution of BSIs across facilities. Dr. Moon stated that the CDC does not have information on whether the data from facilities reporting zero BSIs is accurate or if they have underreported. Dr. Patel (CDC) stated that the graph illustrates that the number of facilities with zero BSIs may not be outliers. Dr. Patel stated that it appears the confusion results from the SIR model calculations and that a zero numerator will always result in an SIR of zero. The TEP co-chair thanked Dr. Patel and asked if there were any remaining questions regarding the NHSN SIR measure (there were not). The TEP co-chair stated that the graph would also be emailed to the TEP members after the in-person meeting.

The TEP co-chair asked Dr. Dahlerus (UM-KECC) to report the results of the second round of voting that occurred at the end of the morning session. TEP members were asked to vote on whether to add the NHSN SIR and the SRR to the star ratings, and scoring options for the ICH CAHPS measure (whether to use top-box or to investigate alternative methods). Thirteen TEP members were present for the vote and 13 members voted.

Table 2: In-Person Measure Voting Results

Measure	% Strongly Agree/Agree	% Neither Agree nor Disagree	%Strongly Disagree/Disagree
Standardized Readmission Ratio (SRR)	54%	0%	46%
The National Healthcare Safety Network Bloodstream Infection measure (NHSN SIR)	38%	15%	46%

Table 3: In-Person ICH CAHPS Voting Results

Measure	Percent of TEP Members in Favor of Investigating Alternative Scoring Methods	Percent of TEP Members in Favor of Using the Top-Box Results
In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)	85%	15%

Dr. Dahlerus stated that UM-KECC will meet with CMS after the in-person TEP to discuss next steps.

On the topic of the ICH CAHPS scoring, Dr. Dahlerus stated that 85% (11 TEP members) voted in favor of investigating alternative scoring methods, while 15% (2 TEP members) voted in favor of using the top-box results.

The TEP co-chair then moved to the next agenda topic, which was a discussion of how to incorporate the ICH CAHPS into the star rating. The TEP was presented with three options.

- Option 1: Report a single overall star rating (that combines the patient experience of care [ICH CAHPS] and clinical measures).
- Option 2: Report two separate star ratings (one for clinical measures and one for the patient experience of care measure [ICH CAHPS]).
- Option 3: Report separate clinical and patient experience of care [ICH CAHPS] star ratings; and report one combined overall star rating (i.e., Option 1 + Option 2).

The TEP co-chair stated that based on the TEP discussion so far, most TEP members were in support of reporting a separate patient experience of care star rating. The TEP co-chair further explained that many facilities are missing ICH CAHPS survey information, therefore, if the ICH CAHPS measure is incorporated into an overall star rating, the amount of missingness (i.e., missing data) poses a methodological problem for the star rating calculation.

The TEP co-chair asked the ICH CAHPS representatives how many facilities would not be able to receive an ICH CAHPS score. Julia Zucco, PhD (CMS/ICH CAHPS group) stated that for public reporting, there is a requirement for at least 30 completed surveys (for a facility) across the twice annual reporting period in order for the ICH CAHPS results to be reported and displayed on DFC. Approximately 40% to 45% of facilities currently have data reported on DFC. The CMS/ICH CAHPS representative noted that fewer than 40% of facilities would be eligible for a star rating using the current cutoff of 30 completed surveys.

One of the TEP co-chairs stated that the low percentage of facilities with ICH CAHPS data on DFC would imply that a separate star rating would be necessary in order to avoid methodological problems when calculating an overall rating. The TEP co-chair asked if the response rates are increasing or if they are relatively stable. The CMS/ICH CAHPS representative stated they have observed that response rates are decreasing over time across all the CAHPS surveys (i.e., Hospital CAHPS; other CAHPS survey settings). The CMS/ICH CAHPS group is investigating ways to address the decreasing response rates. They stated CMS considered increasing the reporting period to be over two years instead of one, however the disadvantage would be that facilities would not have a chance to act on survey results in a timely manner.

The TEP co-chair opened the discussion of the three reporting options to the workgroup. One TEP member expressed support for keeping the ICH CAHPS measure as a separate star rating from the star rating based on the clinical measures. The TEP member stated they wanted to hear more patient and statistician input on this topic, particularly regarding the value of the ICH CAHPS measures if 50% or 60% of facilities do not have a star rating. The TEP member questioned if patients would understand why a facility would not have ICH CAHPS results and wondered if that could lead to a negative perception of those facilities.

The TEP co-chair stated that the concern about possible negative perception of facilities without enough data to report the measure may be a presentation issue. The TEP co-chair asked for patient input as to whether this measure would be useful to patients.

One TEP member wondered how many patients are represented by the current results displayed on DFC. Another TEP member noted that the average number of completed surveys is 30, which may not be representative of a large facility that has 200 or more patients.

As an additional point of this discussion, the TEP co-chair asked if the ICH CAHPS group has considered also using the percentage of completed surveys instead (or minimum number of completed surveys) as additional criteria to determine what results are reported on DFC. For example, 30 completed surveys in a facility of several hundred patients would not necessarily be representative. The CMS/ICH CAHPS representative responded their group could investigate these criteria.

In responding to the issue of response rates, RTI/ICH CAHPS stated they are collecting a patient census right now, and that the overall facility-level response rate is still around 30%, and there is also variation in response rates by facility. They explained that many facilities are small facilities so if the required number of completed surveys is raised, then the results may be more accurate, but then fewer facilities (i.e., smaller ones) will be eligible to have their data publically reported. On the other hand, if the required minimum number of completed surveys is lowered, then more facilities will be eligible to have their data publically reported, but the measure results may be less accurate.

The TEP co-chair stated that a 30% response rate could be sufficient as long as it was representative of the patient population. Extending the comment, the co-chair noted that RTI, during their presentation earlier in the day, offered their interpretation that strong biases were not observed in the ICH CAHPS responses, which is encouraging as use of these data are being considered in the star ratings.

One TEP member emphasized that the low response rate suggests that patients do not think that what they say as patients makes a difference. The TEP member highlighted important ICH CAHPS questions, such as the question about whether patients were listened to or treated with respect. The TEP member strongly encouraged the incorporation of the patient voice, even if the method is imperfect. The TEP member stated that if patients see the patient voice is incorporated in public reporting of survey results, the survey response rate will increase. The TEP member also noted that the low response rate suggests that there may be barriers to completing the survey. Another TEP member also felt the TEP should not focus as much on the methodology for calculating scores, but to ensure that the patient voice is included (i.e., that the ICH CAHPS data are collected and reported).

One TEP member stated that they are supportive of trying to incorporate the smaller facilities in the ICH CAHPS. The TEP member stated that the required number of surveys could be lowered if there was a minimum response rate (percent completed) instead, which could mean that more patient voices are incorporated from the smaller facilities.

The TEP co-chair highlighted the original intent of the star ratings, which is for patients and consumers to have the ability to make decisions based on differences in performance between facilities, and reinforced the importance of hearing from patients, and noted that the patient voice is fundamental to star ratings (and public reporting). The TEP co-chair stated that from a policy perspective, a patient-reported metric must be included in star ratings. The TEP co-chair stated that the role of statisticians and methodologists is to figure out how to implement that measure into the ratings. The TEP co-chair explained that from industry's perspective, it is very important how a patient-reported metric is explained and displayed.

The other TEP co-chair stated that they were very supportive of a patient-reported metric in the star ratings, and they also wanted to see higher response rates and an increase in the number of facilities

with results displayed on the DFC website. The TEP co-chair also provided an example on small facility size: if 16 out of 20 patients in a facility complete the survey, then that would be representative for that facility, even if it is below the minimum number of required completed surveys. The TEP co-chair stated that combining the ICH CAHPS with the clinical measures will be difficult currently with all the missing ICH CAHPS data, therefore option 2 (report two separate star ratings) seems to make the most sense. The TEP co-chair further stated there should also be an emphasis on getting a higher response rate and to include as many facilities as possible.

Elena Balovlenkov (CMS) stated that CMS recognized the need to increase use of the DFC website by patients and consumers. For example, CMS has been working with the patient groups to determine if the current set of measures is meaningful to patients. CMS also explained they are partnering with patient advocacy groups such as the American Association of Kidney Patients (AAKP), the National Kidney Foundation (NKF), and Dialysis Patient Citizens (DPC) in order to investigate ways to better educate patients on the importance of the ICH CAHPS and the other quality measures reported. CMS is investigating the possibility of hosting public calls to discuss why the measures are important and to answer questions from the kidney community. CMS emphasized the importance of patients and facilities having conversations and being partners in care.

One TEP member asked if it would be beneficial to shorten the ICH CAHPS questionnaire. The TEP member stated that the survey has (about) 60 questions and it may be possible to shorten the survey, re-focusing the questionnaire to the questions that are most relevant to the patients, and so that patients understand what specific questions are covered in the ICH CAHPS questionnaire. The TEP member offered the question “Does the patient feel safe at their dialysis unit?” as an example.

In response to the discussion about the importance of the patient voice being reflected in quality measures, CMS stated they are holding an upcoming TEP on Patient Reported Outcomes (PRO) to look at other topics to consider for expanded reporting on patient-reported outcomes. The PRO TEP will cover topics such as quality of life, patient recovery time after dialysis, and investigate other patient-reported outcome instruments that have been developed. CMS stated that the TEP has the potential to propose a number of additional patient-reported outcome measures. CMS also noted that there are additional considerations such as survey and reporting burden (for patients and providers, respectively) and cost. CMS stated that right now the ICH CAHPS is the only PRO measure that is currently available for public reporting on DFC; and that the ICH CAHPS team is attentive to the TEP member concerns raised about the measure. CMS also acknowledged that the ICH CAHPS measure does not capture all dimensions of patient-reported outcomes.

Another TEP member stated that based on the discussion, Option 3 (reporting one overall star rating combining clinical measures and patient experience of care; one clinical quality measure star rating; one patient experience of care star rating) is not viable, and stated that Option 2 (separate clinical quality measure star rating and patient experience of care star rating) makes the most sense. The TEP member stated it would be beneficial to increase the number of facilities that have completed enough surveys to be reported on DFC. The TEP member stated that ICH CAHPS is a good tool that provides the patient voice, but it can be improved.

One TEP member asked a clarifying question about whether in Option 2 the ICH CAHPS would be labeled as the “overall patient experience of care” rating. The TEP member asked if there were alternatives to the label “overall” considering that many facilities will not have data for this score. For example, they referenced the 2015 Star Rating TEP discussion, where the patient/consumer workgroup expressed

strong interest in patients having a drop-down menu on DFC for the star ratings, which would allow patients to choose individual measures that were important to them to create their own custom star rating for a facility.

The TEP co-chair clarified that the ICH CAHPS individual measures will be combined to create an "overall" patient experience of care star rating. They explained that for the ICH CAHPS, facilities would either have all of the individual ICH CAHPS measures or have none of ICH CAHPS measures, in which case they would not have an overall patient experience of care star rating. The TEP co-chair further explained that even inclusion of half of the facilities would be a large enough group to establish cutoffs for the patient experience of care star rating.

Dr. Andress (CMS) stated that CMS has investigated looking at a drop-down box menu for patients to customize their own star rating, but it does not appear to be feasible in the near future. CMS further stated that they have conducted focus groups where patients have provided recommendations, but there are some limits on different approaches for how the data can be displayed.

One TEP member stated there are some very important patient questions in the ICH CAHPS survey, but the overall ICH CAHPS score may not provide patients with the specific information that patients are looking for (since only three composite and three global measures scores are currently reported on DFC).

The TEP co-chair offered a last chance for final questions on the topic and closed the discussion. The TEP co-chair opened the discussion on the re-baselining topic.

4.5 Re-Baselining Discussion

4.5a Opening Discussion

The TEP co-chair stated that in response to the TEP recommendations provided during the previous DFC Star Ratings TEP in 2015, the star rating methodology was updated. It established a baseline year, which was used to create score thresholds for the star rating categories. Star ratings were then assigned to facilities based on their scores in the reporting year compared to the score cutoffs established in the baseline year. The TEP co-chair stated that the result was that patients can see how the ratings change over time, and track how their facilities have improved or not improved compared to a common benchmark. The TEP co-chair proceeded to state that, however, when new measures are added or current measures are updated, there will be inherent changes to the star rating system. The TEP co-chair posed the question for discussion: if such changes are made, do the baseline thresholds have to be recreated, and if so, how should this be done?

The TEP co-chair stated that in the star rating methodology, correlated measures are identified and used to create domain scores; these domain scores are averaged to produce the overall final score and star ratings. If the measures used in the calculations are changed, the domains are changed. This essentially changes the variables that are used to create the star ratings (set in the baseline year). The TEP co-chair provided the example that if a hypothetical rating was based on height scores, but then the system changed and ratings were calculated based on weight scores, the cutoff values for height to evaluate weight would not apply since it is now a different measure. The TEP co-chair stated that if new measures are added to the star ratings (versus minor update to an existing measure), or if an existing measure is updated in a way that impacts the scoring, then there is a change to the whole variable set used to establish the baseline year cutoffs. As a result, the star rating cutoffs must be adjusted so they apply to

the new set of measure scores. The TEP co-chair suggested that one proposal is to establish the new baseline for using the distribution of star ratings: 10% - 1 Star, 20% - 2 Stars, 40% - 3 Stars, 20% - 4 Stars, 10% - 5 Stars (used for the current baseline year cutoffs, and denoted as 10-20-40-20-10). The TEP co-chair noted that this may not be the most favorable option because facilities may see a downward shift from their current star rating; this option also removes historical trending information about a facility's progress over time. The TEP co-chair concluded that it is open for the TEP's discussion as to how the star rating cutoffs are chosen, and that the cutoffs for a new measure set could be chosen to reflect any distribution of star (ratings), without being limited to the distribution of 10-20-40-20-10. The TEP co-chair posed two points of discussion for the TEP: (1) Are there instances in which there is a need to re-baseline? and (2) What is the best approach to re-baselining to reflect what patients want to know about (e.g., longitudinal, and current information)?

One TEP member stated that, based on the figure provided by UM-KECC (Slide #50) showing the distribution of the star ratings in the past three years, there has been improvement. The TEP member requested confirmation that changing the factors involved in calculating the Star Rating would mean that one can no longer compare between the years and that the star rating system would need to be adjusted, e.g., by setting a new baseline.

One of the TEP co-chairs confirmed this, stating that if different variables are used for the evaluation of facilities, then yes, the resulting ratings cannot be exactly comparable and resetting the baseline is needed. The TEP co-chair further clarified that cut-points determine the proportion of facilities in each of the star rating categories, but deciding where these thresholds are set is a separate issue from deciding whether the thresholds need to be reset, e.g., due to adding new measures.

To further address the reasons for re-baselining, the TEP co-chair clarified that a change in the measure set necessitates some form of re-baselining, due to not having complete comparability of measures between consecutive years. If a concern is frequent re-baselining (hence loss of historical comparability over time) an alternative is to change the measure set infrequently. They stated these two alternatives present a trade-off: adding new measures adds more information (about quality), but at a cost of loss in comparability between reporting years.

One TEP member asked for clarification on how this discussion differs from the TEP discussion for the original methodology (i.e., relative rating methodology implemented in 2015).

The TEP co-chair responded that under the original methodology every year the thresholds would be re-set to the current year's data to achieve a 10-20-40-20-10 bell-shaped distribution of facilities in the star ratings. As this approach does not provide year-to-year continuity of star rating shifts (a facility with constant absolute performance can receive a lower rating in subsequent years if average performance across all facilities improves over time) the 2015 TEP recommended developing a system in which facility improvement could be tracked over time. The TEP chair continued by saying that if the measure set remains unchanged, the cutoffs could remain unchanged as well. The TEP co-chair, however, noted that continuous improvement relative to a baseline year standard could result in most facilities eventually receiving a 5-star rating, which means that facility performance can no longer be distinguished. The TEP co-chair stated that this issue is separate from the current issue of re-baselining when adding new or updated measures, and in particular when adding patient-centered measures to the star rating system. The TEP co-chair posed the question of how to proceed with re-baselining given the trade-offs described earlier.

One TEP member stated that this is an issue of accuracy versus responsiveness to patient needs.

One of the TEP co-chairs agreed, stating the goal is that new information added to the star rating system is valuable, and that patients will have to know that new measures means more valuable information, but that direct comparability with previous years' star ratings will not be possible. The TEP co-chair asked for TEP input on the trade-offs: the importance of continually updating (and more re-baselining) to achieve the best possible set of measures versus having continuity in the star ratings over time (and staying with the same set of measures). One TEP member explained that after re-baselining, one could still track changes in the original measures that remained in the star rating system. This could be done by looking at the national average scores, which provide information on the longitudinal trend for these measures. The TEP co-chair made the caveat that this additional information would increase the complexity in reporting information to patients.

One TEP member asked if, from the perspective of the TEP co-chairs, adding a measure to the star rating system is the same as updating a current measure. The TEP member stated that these should be viewed as potentially separate pathways for re-baselining.

One of the TEP co-chairs responded that this would depend on the magnitude of the update. A minor update to a current measure would not affect the score for a facility, and thus would not warrant re-baselining. The TEP co-chair noted that, however, the need to update a measure is predicated on the assumption that the update provides improvement or added information to the existing measure. The TEP co-chair then asked UM-KECC if the currently proposed updated candidate measures set are substantial enough to warrant re-baselining. UM-KECC responded that this was correct.

One TEP member stated that the take-away of this part of the discussion on re-baselining is that adding new measures should not be done frequently, and further, that re-baselining every year will not likely align with the rate of production of new scientific evidence supporting changes to the clinical measures. They agreed that if new measures are added, re-baselining should be done by using the current star ratings distribution to maintain the continuity of the rating. The TEP member asked if there is an optimal number of metrics for this system, and proposed a "zero-sum" system, such that if a new measure is added, then the measure that has least impact or the measure that is the most topped out should be removed in order to retain a fixed number of metrics.

CMS responded that the best example of this is the proposed replacement of the current vascular access measures with the updated measures (two measures added and two removed). CMS further stated that it is difficult to argue for a method that assumes interchangeability between measures. For example, if there is a need to add an infection measure, this should not require removal of a dialysis adequacy measure. Dr. Andress stated that any choice about the balance of measures should be taken in the context of the specific change to the measure set, and that there is no ideal number of measures. CMS further stated that the main question to be considered is how effectively the measure set captures the quality and complexity of care patients receive. Therefore it is appropriate to consider if or when to remove measures, not with respect to what the current measure set is, but with respect to how facilities are performing on that individual measure. If there is still a considerable performance gap among facilities, then the measure should remain in the star rating system, and that the inclusion of a separate measure does not automatically require removing an existing measure.

One of the TEP co-chairs asked for additional comments pertaining to the discussion. There were no additional comments and the TEP co-chair moved to continue on to the next topic.

4.5b Hypercalcemia: Case Study on Re-Baselining

Dr. Dahlerus (UM-KECC) began the discussion by giving a high-level description of the hypercalcemia measure (the percentage of adult hemodialysis and peritoneal dialysis patients whose average 3-month calcium is greater than a value of 10.2; see slide 46 in appendix P). Dr. Dahlerus stated that the measure was updated (as part of the annual NQF update schedule) in late 2016 and accepted by the National Quality Forum (NQF) in early 2017; it was regarded as a non-substantive update by NQF. The update made was to include patients with missing calcium values in the numerator.

Dr. Dahlerus explained that since the update was not completed in time for CMS to announce it on the October 2016 CMS National Provider Call (NPC), where CMS announced the other updated and new candidate measures proposed for the DFC October 2018 Star Ratings release, this topic was brought to the TEP for input. This has implications with respect to the timing of adding the measure and the re-baselining timeline. Dr. Dahlerus stated that there are two options for which CMS would like the TEP's feedback: (1) include the updated hypercalcemia measure in the next update of the DFC Star Ratings, which would occur in October 2018, or (2) do not include the updated measure in the 2018 update to the DFC Star Ratings, and instead include it in the 2019 update to the DFC Star Ratings, to allow CMS to announce the measure in the October 2017 NPC. The drawback to option (1) is that the updated measure was not announced in the 2016 national provider call, and it is standard procedure for CMS to announce proposed changes during this call; the drawback to option (2) is that the star ratings would need to be re-baselined two years in a row (2018 and 2019). Given the options and their noted drawbacks, CMS needs input from the TEP on whether to add hypercalcemia to the 2018 release or to wait and add it in 2019 (and re-baseline both in 2018 and 2019).

One of the TEP co-chairs opened the hypercalcemia issue for TEP discussion.

One TEP member suggested that the new hypercalcemia measure be included as quickly as possible. With respect to public awareness, an announcement could be made (by CMS) explaining their reasons for including the updated hypercalcemia measure in the 2018 star rating.

One of the TEP co-chairs asked for clarification as to whether CMS would be under any legal or regulatory penalty for not announcing the update in the October 2016 national provider call.

Dr. Andress (CMS) stated that CMS is not under any legal requirement to provide notification of changes to the Dialysis Facility Compare site (including to the star ratings). CMS noted that this standard procedure is fairly new, and that this is an example of where the process would be violated. Dr. Andress asked for feedback on what the potential response to either option would be, stating that CMS intends to be transparent on this point, and this is an example of where the implementation timelines for DFC and the star ratings do not line up with the measure development/maintenance and NQF timelines. Dr. Andress emphasized that it was important for the TEP to be aware of this issue, and have the opportunity to suggest the appropriate pathway for moving forward with the hypercalcemia measure.

One of the TEP co-chairs wondered whether a third option would be possible: all of the currently proposed updated (and new) measures are held until 2019 to align with the timeline for announcing the updated hypercalcemia measure.

In response, CMS stated the proposed updates are improvements on the existing measures, thus there would need to be a good reason for not moving forward with them in 2018. Aligning with the timeline for the updated hypercalcemia measure is not justification enough.

One of the TEP co-chairs stated that the preference for infrequent re-baselining would warrant potentially holding a new (and updated) measure for a certain fixed number of years or until a critical

mass of measures is reached at which point the set would be proposed for adding to the star ratings. They also noted, however, the drawback of holding the measures that have improvements compared to the current measure set in the star ratings.

One TEP member stated that the second option is preferred (for when to add the hypercalcemia measure); the third option is favorable but does not seem realistic. The TEP member agreed that the CMS timeline (for measure announcement on the NPC) is a good process, and that this process should be maintained as it was developed to better help both CMS and dialysis providers prepare for future changes to the star ratings. The TEP member concluded that it is best to wait on adding the updated hypercalcemia measure to the star ratings (in 2019) in the interest of maintaining the current CMS process for announcing changes.

Another TEP member stated that they felt since the hypercalcemia measure has been updated it seems to imply that the previous version of the measure is no longer valid, more so due to the fact that the update was NQF endorsed, and NQF is seen as the gold standard for evaluating measures. They also felt the hypercalcemia measure is important to patients and thus a high priority (i.e., to have the updated version added). The new hypercalcemia measure should be included in the 2018 star ratings to prevent the need to re-baseline two years in a row. They felt that would be more understandable to patients. The TEP member also suggested CMS could make an announcement regarding the measure update in the interest of transparency.

Another TEP member agreed that the (updated) measure is important as reflected by having gone through the NQF process.

One of the TEP co-chairs stated the reason CMS was seeking input on whether to add the measure now versus later is in the interest of cohesive policy planning.

With respect to timing of adding measures and re-baselining, another TEP member stated the current consensus is to avoid re-baselining year after year.

One of the TEP co-chairs responded by explaining that this was the opinion expressed at the last TEP. Patient groups advocated for consistency and comparability (over time) in contrast to re-baselining the star ratings every year, which showed where, each year, a facility is ranked relative to every other facility. The TEP co-chair further stated that the current star rating method using a baseline year was intended to make it easier to track the progress of a facility over time. The TEP co-chair concluded that this is a valid point for not reverting to a bell curve when re-baselining, but noted that is also a slightly different question from whether to re-set the baseline cutoffs when there are new measure sets while maintaining the same star ratings distribution.

One TEP member asked if there would be a discussion of what the re-baselining methodology (determination of cutoffs) would be, given that re-baselining appears to be inevitable.

The TEP co-chair responded the TEP is being asked for input on how re-baselining should be done as a general policy, and that the general idea of how to re-baseline is in purview of this discussion.

Dr. Messana (UM-KECC) provided an illustrative example for TEP discussion: if a re-baselining strategy is recommended such that the distribution of star rating categories is set to reflect the prior year's distribution, where in that year 50% of facilities received either 4 or 5 stars, how is that continuity (in star rating placement) maintained? Dr. Messana noted that this addresses the strongly stated notion that patients want to view facility improvement over time, however the appropriateness of this in terms of active reporting depends on whether or not there is evidence that, for the new measures, dialysis facilities will perform the same (in the new measures) as they did with the current measures.

The TEP co-chair agreed, stating that if there is an update of an existing measure, facilities most likely will not have gotten worse, and should not be rescored (therefore re-baselining would not involve a new distribution for determining cutoffs for star ratings). The TEP co-chair stated that, on other hand, if there is a new measure on which performance is poor, then assigning 50% of facilities four or five stars may not be appropriate, and keeping the same star distribution may not reflect true facility performance.

The TEP co-chair further stated that there are two well-defined options thus far: (1) re-baselining by distribution (10-20-40-20-10), or (2) re-baselining to the current star ratings distribution. The TEP co-chair noted that if it is believed that for a new measure, facilities are performing 20% worse than on a current measure, the choice between options (1) and (2) is difficult. It depends on the nature and extent of the updates to the star rating measure set. The TEP co-chair concluded that there is no clear suggestion for an intermediate approach (between options 1 and 2), but whatever approach is used should be as interpretable as possible for the patients.

One TEP member stated that decisions should be made according to the importance of the measure being added or updated, thus determining how often the baseline should be set. Another TEP member agreed with the statement that simplicity is better for patients.

One of the TEP co-chairs asked if there were any other outstanding comments on this issue.

One TEP member asked how much change is expected to come from the updated hypercalcemia measure (i.e., the change to include missing values in the numerator which count against the facility) as they believed performance for that measure is already high.

Dr. Messana (UM-KECC) responded that hypercalcemia measure values are obtained from CROWNWeb. The number of missing values for calcium may be as high as 8% or 9% in any month. Dr. Messana stated that inclusion of the missing values changes the measure performance distribution. It was also noted that the measure is calculated as a three-month rolling average and the percent missing (all three months) is much less than the percent missing for an individual patient month.

One TEP member asked for clarification as to how data for a patient with an individual missing value is included in the measure numerator.

Dr. Dahlerus (UM-KECC) responded there are two ways in which a patient is included in the numerator: (1) if the patient has an average calcium value greater than 10.2, calculated as a rolling average across three months, or (2) if the patient is missing data in all three months.

The TEP member noted that the facility is therefore penalized for having patients with calcium levels that are too high or for not reporting measure values for three consecutive months. UM-KECC confirmed that the TEP member's understanding of the measure update was correct.

One of the TEP co-chairs asked for a point of further clarification, e.g., a patient with reported calcium values that are too high for two months, and missing on the third month. Dr. Messana responded that the rolling average is calculated using what data are available in the three months. If there is only one month of available data, that value is used; for two and three months of available data, the average is used. Thus, a facility is penalized if the average of all of the available months is too high or if no data are reported for all three consecutive months.

One TEP member asked if there are differences in the hypercalcemia measure results between smaller and larger facilities, or between home and in-center modalities.

Dr. Messana stated that they have not examined the data at these subpopulation levels.

The TEP member stated that there is no evidence of differences in the hypercalcemia measure based on differing facility size or modality as shown in her own data.

One of the TEP co-chairs asked if the TEP had any outstanding questions about the hypercalcemia measure.

One TEP member stated that after having listened to the TEP discussion, the implications for CMS not sticking to its original process (and timeline for announcing measures) are not as concerning as re-baselining two years in a row. The TEP member also suggested that re-baselining be done with respect to using the cutoffs from the current distribution, and not re-setting to a bell curve.

One of the TEP co-chairs asked if there were any final comments.

4.5c Clustering of Star Ratings Distribution and Loss of Discrimination

The TEP co-chair moved to discuss one additional re-baselining topic. The TEP co-chair stated another circumstance that may require a need for re-baselining is if facilities keep improving over time on all of the measures, and eventually most or all end up in the 5-star category. To further the example, they noted there has been an upward shift in the distribution of star ratings in recent years into the 4- and 5-star categories, reflecting overall facility improvement on the measures.

The TEP co-chair also explained that there is still substantial variation across facilities in terms of performance, however if the trend continues, eventually the vast majority of facilities will fall at the high end of the distribution. The TEP co-chair described two ways this could happen: (1) a ceiling on the measures is reached as some individual measure becomes topped out, in which all of the facilities can do no better on these measures and the measures are no longer distinguishing facilities, or (2) the scores are all shifted upward, but the gaps or variations between the facilities remain. That is, even though all of the facilities are in the 5-star category, there is still substantial variation among them. The TEP co-chair stated that, in the second instance, individual measures are still distinguishing facility performance, but the star ratings distribution cut-points are not, and, therefore, do not provide meaningful information to distinguish facility performance.

The TEP co-chair further stated, even though the system would be giving patients information that facilities have passed the original standards as set in the baseline year, the patients will lose the ability to distinguish among facilities when all facilities are receiving the same star rating.

The TEP co-chair then posed the question of whether the star ratings distribution should be re-baselined, so that the variation in the scores can inform patients about differences between very high performing facilities and high (or average) performing facilities. The TEP co-chair further stated that, in the first instance, when measures in the star ratings become topped-out, the measure(s) is no longer providing useful information to distinguish facility performance, or cannot be improved upon further. This would suggest reset by removal of the measure, rather than a reset of the whole star ratings distribution. The TEP co-chair concluded that both instances have consequences for patient interpretation, in that all or most facilities with a very high star rating would still be shifted (to different categories in the distribution). The TEP co-chair asked for input from the TEP.

One TEP member asked for clarification on the suggestion that if a topped out measure is not providing useful information, then the recommendation should be to remove the topped out measure.

One of the TEP co-chairs responded this is the easier of the two options, however there is a question of at what point can it be determined that a measure is topped out, (i.e., how little variation is needed to

consider a measure topped out). The TEP co-chair noted that this is both a technical question and a policy decision on when to remove topped out measures.

The TEP member asked if there is a formal definition of "topped out" which CMS uses.

Dr. Andress (CMS) stated that CMS has a common definition for "topped out," which it uses across all programs. This has been described, for example, in prior years of the QIP Rule. Dr. Andress noted that he can provide the specific criteria to the TEP, but that the definition is based on (1) how much interquartile variation is present in a given measure and (2) the coefficient of variation for that measure (the ratio of the standard deviation of the measure to its mean).

The TEP member asked if there are different methods used for determining if a measure on DFC is topped out than for other CMS programs.

Dr. Andress responded this issue has not come up yet, but that for individual measures, the same criteria would be applied. Dr. Andress further stated that there is more flexibility in what measures are included (or not) in the DFC Star Ratings as there are no statutory mandates as to what measures have to be in the program. This is in contrast to programs, such as the QIP, that have requirements for which measures must be included.

One of the TEP co-chairs responded that based on this the question is what to do when most measures are in the 5-star category but the individual measures are not topped out and there is still substantial variation in either the individual measures or in the overall scores among all the facilities. The TEP co-chair asked if the star ratings are still providing meaningful information if virtually every facility is at the top of the distribution or rated as 4 or 5 stars. If not, it is necessary to reset (re-baseline) so that patients can make distinctions among those facilities, which will aid patients when choosing facilities or requesting facilities to make further improvement. All facilities may be doing well with respect to the baseline, but some are doing better than others and this may be important information for patients. The TEP co-chair then posed a question of the TEP, asking how important is it that the star ratings reflect that there is still variation among the facilities, even if they are all in the same top 1- or 2-star rating categories?

One TEP member responded that it is a good sign if facilities are doing much better, but there should be information available to show differences among facilities. The TEP member recommended there would be a need to re-baseline (reset cutoffs) so differences in facilities are easier to understand.

Another TEP member agreed, stating that re-baselining is necessary in this circumstance.

One of the TEP co-chairs asked if there were additional comments.

One TEP member asked, given the current trend, when would all facilities be expected to have a 5-star rating?

The TEP co-chair stated that no projections have been made at this time.

Another TEP member noted that in the current distribution of star ratings 41% of facilities are in the 3-star category.

The TEP co-chair noted that the percentage of facilities at 1 and 2 stars is still low, and while there is the extreme case that all facilities will reach 5 stars, the more likely scenario is that there will no longer be facilities in the 1- and 2-star categories. They asked at what point then the TEP would suggest that the star ratings are no longer distinguishing quality between facilities.

One TEP member responded that it may be a long time until all facilities reach a 5-star rating.

The TEP co-chair agreed. They further asked for input on when to set a new threshold, noting that re-baselining (the reset) implies the new thresholds/cutoffs imply the system is encouraging facilities to aim for a higher performance bar on the measures.

One TEP member noted that the number of 1-star facilities has gone down by 3% each year, and the number of 2-star facilities has gone down even more (Slide #50 in Appendix P). The TEP member stated that reducing the number of 1- and 2-star facilities may be an alternative goal before re-baselining, and it will occur faster than the rate at which all facilities become 5-star facilities.

Another TEP member asked for the patient perspective on this discussion. The TEP member stated that 1- and 2-star facilities are viewed as below average or undesirable, and that a facility would want to be at least 3 stars or above. The TEP member agreed that no longer having facilities with 1- and 2-star ratings is key to showing this continuous progress.

One of the TEP co-chairs agreed, stating that it is better to frame this discussion in terms of the 1- and 2-star (low) end of the distribution, rather than the 4- and 5-star (high) end. The TEP co-chair asked UM-KECC and CMS for clarification as to whether or not 1- and 2-star facilities are viewed as 'bad.' The TEP co-chair noted a point of clarification: by the definition of the 2014 baseline year, 10% of facilities were baselined as 1-star and 20% were baselined as 2 stars. The TEP co-chair asked what the descriptive labels would be in the DFC language for a 1- or 2-star facility.

Dr. Andress (CMS) stated that the descriptions of the star rating levels are: (1 star) much below average, (2 stars) below average, (3 stars) average, (4 stars) above average, and (5 stars) much above average. Dr. Andress further stated that these descriptions were developed in the previous star rating system that used a relative system for measuring performance in the same year (i.e., all facilities rated relative to each other in the same year; there was no baseline standard).

Dr. Andress further noted that it is unclear for measures (such as the mortality rate) what is deemed 'okay' or 'not okay' performance in terms of setting an absolute standard. The implication (and perception) of the 1- or 2-star rating is going to be negative regardless of what description is used. Dr. Andress further stated that having 1- and 2-star facilities is inevitable after a re-baselining, unless there is a move to a system which does not have 1 and 2 stars, but this works against encouraging continuous improvement.

Dr. Andress posed the example of readmission rates, in that the national average readmission rate is 33%. Dr. Andress noted that if this average is reduced to 25% then this is marked improvement among the population of dialysis patients, but this is still not as low as the 15% national average for the rest of the population.

Dr. Andress concluded that re-baselining should occur in order to keep driving improvement, and at the same time, give better information to dialysis patients when they are getting information on DFC. Dr. Andress stated that there is a need (for patients) to be able to distinguish between higher and lower performing facilities. The focus should be on how to reach out to consumers, dialysis patients, and dialysis facilities, to ensure they have the relevant information they need.

One of the TEP co-chairs stated that the current descriptions reflect that a facility is above or below the 2014 average, and that re-baselining is necessary to differentiate between facilities.

One TEP member stated that the industry will improve in delivery of care, and that there should be an opportunity for clearer distinctions in quality among dialysis providers as information technology is

implemented, and coordination of care becomes easier and better. The TEP member stated that there will be push-back from the industry due to increased cost, but that this is understandable. The TEP member further stated that as information increases, there will be greater ability to make better and more focused decisions, and in a 3 to 5 year span, information will have a large impact, such that rapid improvements in performance (due to different factors) will be drivers of that change.

As a separate point, the TEP member noted that patients do not respond to surveys at a higher rate as part of the clinical culture, and expressed the notion that patients lose hope. The TEP member clarified that is not the facility's fault, but rather the nature of the disease. The TEP member concluded that as the industry changes, CKD patients representing 1% of the patient population for Medicare and 6%-7% of the cost should change as well.

Another TEP member stated that there are already effective tools that drive facility performance. The TEP member further stated that star ratings stand out in that they are for the patient's use. The TEP member concluded that all facility ratings shifting toward 5 stars is not a bad thing, and that high performing facilities should be receiving a high rating.

One of the TEP co-chairs stated that in fall 2014, a number of organizations had a meeting with CMS, and in this meeting an example was brought up comparing star ratings to Yelp or Kayak, which are consumer-driven sites for rating services. The TEP co-chair reinforced that how star ratings are communicated to the community, and the importance of how to approach patients and engage them in understanding the quality of their facility and care, is an important point of discussion.

Continuing the discussion about shifts in star ratings, one TEP member asked UM-KECC if there is a specific domain (or measure), for example, standardized measures that caused the shift in the distribution of facilities' star ratings.

Dr. Li (UM-KECC) stated that the standardized measures do not cause much shift in the star ratings distribution, but other measures might have been stronger forces to cause the shift. Dr. Li stated that the distributional shift is being driven by several factors, not by a single measure.

One of the TEP co-chairs asked if, since the standardized measures are being recalculated annually, they are re-baselining themselves.

Dr. Li clarified that when the standardized measures are calculated in the star ratings, they are multiplied by an adjustment factor, such that they contribute somewhat to the distributional shift (taking into account performance compared to the baseline year), but they are not contributing much to the overall star ratings shift.

Dr. Messana added that the adjustment factor is to take into account the baseline year average so that if there is a shift or improvement, the distribution is shifted by the difference in the mean of the current year and the baseline year.

One TEP member asked if the shift in the distribution of the star ratings is due to reporting measures.

Dr. Messana clarified that there are no reporting measures used in the calculation of the star ratings, and the other measures referenced besides the standardized measures are two vascular access type measures, hypercalcemia, and the overall Kt/V measure.

As a point for discussion, one TEP member asked if rating systems such as Yelp or Amazon rebased every three years, would it cause the consumer to lose credibility in the rating system. The TEP member further asked if the baseline is changing on a regular basis, does the rating become uninterpretable. The

TEP member stated that this is the issue with re-baselining to a bell curve shaped distribution (versus using the same distribution).

One of the TEP co-chairs stated that one of the differences, with a rating system such as Yelp, is that the standards (e.g., of restaurants) is going up, such that what it takes for a person to give 5 stars goes up every year. In this manner, the TEP co-chair stated that consumers still give favorite restaurants a 5-star, and the next a 4-star, therefore there is a gradual resetting of what it takes to earn a 5-star, but it is more relative than the methodology for the DFC Star Ratings.

The TEP member clarified that the science that re-baselining is necessary is well understood, but the question remains, from a consumer standpoint, how is re-baselining understood.

The TEP co-chair stated that the issue is that consumers view the DFC Star Ratings similarly to Yelp or other rating systems. The TEP co-chair asked how to get that point across, so that the consumer understands how to interpret the star rating system.

One TEP member agreed, stating a star rating system inherently implies 5 stars is good, and 1 is not.

Another TEP member asked for further clarification as to which domains are causing the shift in the star ratings distribution.

Dr. Messana stated that the slides in the appendix of the meeting presentation provide the three-year distribution shifts for each of the individual measures, and that this information was provided to the TEP for reference (see Appendix P beginning on slide #130).

One of the TEP co-chairs stated that these slides provide a means of visualizing the individual measure distribution shift over three years.

The TEP co-chair moved to conclude the baseline discussion.

Dr. Dahlerus (UM-KECC) stated that we still need to determine if there was a consensus on whether to add hypercalcemia to the star ratings in 2018 or 2019. If the TEP prefers, a vote will be taken.

One of the TEP co-chairs moved to take a vote.

The TEP moved to break and vote on hypercalcemia.

4.6 Summary of Recommendations and Discussion of Next Steps Discussion

4.6a Discussion

In follow-up to the end of the last discussion on hypercalcemia, UM-KECC provided the results from the vote regarding inclusion of the updated hypercalcemia measure. The TEP was asked to vote for either Option 1 (to include the updated hypercalcemia measure in the next update of the DFC Star Ratings) or Option 2 (do not include the updated measure in the 2018 update to the DFC Star Ratings, and instead include it in the 2019 update to the DFC Star Ratings, to allow CMS to announce the measure in the October 2017 NPC). Option 1 received 12 votes (92%), and Option 2 received 1 vote (8%). Thirteen TEP members voted in total.

The TEP co-chairs stated they would provide a summary of the afternoon discussion; a summary of the morning discussions was provided earlier in the meeting before the panel adjourned for lunch.

Scoring ICH CAHPS for Star Rating

The first topic discussed was whether there should be a separate star rating for ICH CAHPS and the clinical measures, or whether all measures should be part of a combined overall rating. The TEP co-chair stated there was consensus for creating separate ICH CAHPS and clinical measure star ratings. This is in part because many facilities will not be able to have an ICH CAHPS rating calculated (due to not enough surveys), which would likely mean they would not be able to have an overall star rating calculated. Additionally, having a separate measure providing the patient experience of care perspective was felt to be important. Therefore, the ICH CAHPS star rating and the clinical star rating should be calculated and reported separately using comparable methodologies. The TEP co-chair also stated there was a strong TEP preference for adding patient-reported outcome measures to the star ratings as soon as possible; and that the TEP's preference was to investigate using the linear mean scores for calculating the ICH CAHPS scores for the star rating versus using what is reported now, which are top-box results. Follow-up discussion for this topic of using linear means may be necessary once the data are available (timeline TBD).

Re-baselining (new and updated measures)

The TEP co-chair stated that it will be necessary to re-baseline when measures are added or removed from the star ratings, or if a measure has a significant update. Re-baselining too frequently would not be desirable, because it would interfere with the ability to observe changes over time, which is something patients and providers feel is important. The TEP co-chair stated that there is value in measuring facility improvement over time. The TEP co-chair stated that the TEP discussion indicated a preference for re-baselining infrequently, and doing so when new or updated measures are added. There were also some TEP concerns reverting to the original 10-20-40-20-10 distribution for re-baselining (where 10% of facilities would receive 1 star, 20% would receive 2 stars, 40% would receive 3 stars, 20% would receive 4 stars, and 10% would receive 5 stars). Going back to this distribution would possibly result in many facilities receiving a lower star rating than the prior reporting year, once the re-baselining distribution is reset. The TEP recommended considering re-baselining to a distribution that reflects the current distribution of the star ratings.

The TEP co-chair summarized the issue of whether to re-baseline when a majority of facilities cluster near the top of the distribution and are receiving 4 or 5 stars. The TEP co-chair stated that there was the least amount of TEP consensus on this issue of star rating shift (facilities clustering in the 4- and 5-star rating categories). Some TEP members stated that they were comfortable with a majority of facilities receiving 4 or 5 stars, while other TEP members stated it was important to re-baseline to be able to distinguish the differences between facilities, something that is difficult to do when facilities are clustered in one or two categories.

The TEP co-chair stated there was a clearer consensus that if most or all facilities were scoring very high on one specific measure, then that measure could be considered for removal, rather than re-baselining the whole star ratings distribution. They noted however, that there will be some measures on which most facilities score high but these measures should remain in the star ratings because they are critical to patients or may be necessary for maintaining a standard of care. The TEP co-chair invited others to add or clarify the summary that was provided.

The other TEP co-chair agreed that the summary accurately reflected the TEP discussion. The TEP co-chairs invited TEP members to add any other discussion points that were not covered during the summary. No TEP members provided any additional comments or additions to the summary.

UM-KECC stated that they will listen to the recording and produce a draft summary report to accurately record what the TEP discussed. The draft summary report will be sent to the TEP to review. UM-KECC strongly encouraged the TEP members to review the draft summary report and ensure that the report reflects the TEP opinions and discussion.

4.7 TEP Recommendations

4.7a TEP Recommendations

The TEP recommendations, derived from the teleconference and in-person meeting discussion, are summarized below:

1. Recommendations for updated and new candidate measures (final voting results):
    a. Greater than 60% of TEP members supported replacing the current versions of the vascular access (fistula, catheter), Standardized Mortality Ratio (SMR), Standardized Hospitalization Ratio (SHR), Standardized Transfusion Ratio (STrR) quality measures with the recently updated versions: Standardized Fistula Rate, Long-term Catheter Rate, SMR, SHR, STrR.
    b. Greater than 60% of TEP members supported replacement of the current hypercalcemia measure with the updated measure (vote held at the in-person meeting).
    c. 54% of TEP members supported the addition of Standardized Readmission Ratio (SRR) to the star ratings; 46% did not support adding SRR to the star ratings.
    d. No TEP members voted against adding the new Pediatric PD Kt/V measure to the star ratings. The Pediatric PD Kt/V measure would be included in the combined Kt/V measure (currently in the star ratings).
    e. The TEP did not achieve majority consensus to support addition of the National Healthcare Safety Network Bloodstream Infection (NHSN SIR) measure to the star rating.

2. The TEP recommended adding the ICH CAHPS to the star ratings. They also recommended:
    a. Calculating and reporting ICH CAHPS as a separate star rating from the overall star rating for clinical measures.
    b. Investigate alternative scoring (linear mean of ICH CAHPS measures) for the ICH CAHPS star rating.

3. The TEP provided potential next steps on re-baselining (see "Re-baselining (new and updated measures)" under section 4.6 Summary of Recommendations and Discussion of Next Steps).

Outstanding Issues:

Recommendations needed on the following:

1. If and when to re-baseline when there is star rating shift such that a majority of facilities receive 4 or 5 stars;
2. How often to re-baseline
3. When to re-set the baseline distribution

The issues listed above were discussed on the Star Rating Post-TEP Teleconference Call (see Appendix F for the Star Rating Post-TEP Teleconference Call Minutes).

4.8 Public Comment Period

A 15-minute public comment period was held at the conclusion of the In-Person TEP Meeting on February 21, 2017. Three public comments were received. The first public comment was provided by Jackson Williams, JD, of Dialysis Patient Citizens (DPC).

> "Hi, good afternoon everyone. I'm Jackson Williams with Dialysis Patient Citizens. I just wanted to-I'm not going to discuss the matters that were already voted on- I just wanted to call the panel's attention to a few items that weren't brought up today. On the issue of 1- and 2-star ratings for truly poor performers versus those who have been given descriptions on Dialysis Facility Compare-, I just wanted to call your attention to the consumer testing reports from NORC that were commissioned back in 2014, and they were apparently not circulated to the group, but they are available from Dr. Messana. They do indicate that when consumers, average people, were shown the star ratings of 1- and 2-star facilities, their belief or their impression was that they were in fact poor performers that they should avoid. So this is not just an impression among us, but it is confirmed by actual focus groups. Second, I do think that it's unfortunate that this group was not presented with options for re-baselining, meaning various alternatives that could be used. I think when CCSQ briefs the administrator on their recommendations they usually give the administrator the courtesy of alternative options even if one of them is recommended, and I do think it would be appropriate to present alternative options to this group. And third, there is one other projected opportunity for re-baselining, it's in the future that was just sort of briefly alluded to, and that is that Congress mandated a report on socio-economic or socio-demographic status adjustments to quality programs. The report came out in December during the Christmas break; it didn't get a lot of attention but the experts from the Assistant Secretary of Planning and Evaluation who were commissioned to write this report, did find that in fact that star ratings and pay-for-performance programs including the ESRD QIP are biased systematically by various socio-demographic factors. And the recommendation of the report is that these programs be redesigned so that they are measuring true quality and not disadvantages of the population being served. So what is up right now, is in the Cures Act mandate to redo the hospital readmissions penalty to place hospitals into peer groups so that they are competing against hospitals that have a similar population for instance in terms of the percentage of dual-eligibles that they are serving. I do think there will be a necessity to look at Dialysis Facility Compare Star Ratings and the QIP to hopefully create a peer grouping system where facilities are in a real competition with other facilities serving similar populations so that facilities serving skiers and snowboarders in Colorado or hedge fund managers in Connecticut are not pitted against those that are serving poorer people in Appalachia or the deep south. I don't know when it is feasible to re-baseline the star ratings for socio-demographic status, but I do hope that it would happen at the same time as the other re-baselining, so that there are not multiple re-baselinings, which I understand to be the sentiment of this group. Thank you very much."

Dr. Messana (UM-KECC asked a clarifying question): "Mr. Williams, one quick question, if any TEP members request the 2014 report- do you want all three files that you sent me in that email or just the two, I think there are two documents?"

Mr. Williams: "Yes, there are just two reports."

Dr. Messana: "Okay thank you, just for clarification."

Note: Mr. Williams sent the reports to Dr. Messana via email during the lunch break of the in-person TEP meeting.

The second public comment was provided by Kathy Lester, JD, on behalf of Kidney Care Partners (KCP).

> "Thank you. First, I want to thank CMS, UM-KECC, and all of you TEP members for allowing us to be part of this process. My name is Kathy Lester and I'm here on behalf of Kidney Care Partners. That is an organization that represents more than 35 dialysis related organizations including patient organizations, physicians, nurses, technicians, manufacturers, and then all types, sizes, and geographic locations of dialysis facilities across the country. We feel a very special obligation when it comes to quality, especially in terms of having accurate and timely data, because we were the group, KCP was the group that asked the Congress to put into place the first value-based purchasing program. Also, our members have been reporting measures for 20 years prior to that program going into place, and have used those measures to share with patients and share that data to be transparent in the quality of care provided. So we do take this very seriously and are very grateful to CMS for allowing us to participate in this program. The first principal we judge anything by is - are the measures accurate and valid and reliable? In other words, does it provide patients with accurate information on which they can and should make the decisions that are affecting their lives? If a measure does not meet these standards, it should not be added. It is as simple as that. So we are very pleased about the discussion today around NHSN Bloodstream Infection measure and the Standard Readmissions Ratio measure, which we do not believe provide that accurate data yet. We have consistently recommended and very willing to work with CMS and the TEP and others to get those measures so they can be used, because we do believe those are measures that are important topic areas. But we urge CMS at this point to keep those measures out of the 5-star program until they can be fixed, so they are representing accurate and reliable information. We also share the concerns that many TEP members raised about the decrease in response to the ICH CAHPS. And as you may have seen in the letter we sent in on Friday, we believe this is due to patient fatigue. There are ways to administer this survey less than twice a year and also with fewer questions. And that has been validated when the instrument was been created. So we again would welcome the opportunity to work with everyone to improve the administration of the ICH CAHPS so that it can be used in a meaningful way and that information shared.
>
> As we said in the letter, we support including the ICH CAHPS survey on Dialysis Facility Compare so long as it is not included in those 5-star ratings. I think members of the TEP articulated our reasons as to why that would be very well. Finally, in terms of the rebasing, I very much appreciate the point that Dr. Sugar made that nothing says we have to re-baseline to 10-20-40-20-10. We'd like to see what options could be established and we agree that you should not be

> re-baselining every single year. There needs to be some stability to the system over time. So we recommend that you take the comments that we have made over the last several years about structuring the baseline year, not on that forced distribution curve, but rather using thresholds and cut points that are determined in another way and happy to share those detailed comments with all of you again going forward. So again, thank you for the opportunity to comment. Appreciate the healthy dialogue today, and looking forward to see what comes out as the next steps of this process."

Dr. Li (UM-KECC) closed the session by thanking the TEP for their time and their input. Dr. Li stated that UM-KECC has documented the TEP input and will circulate the summary report for TEP members to review. Dr. Li stated that the team may schedule follow-up calls to address any remaining concerns. Dr. Li's vote of thanks for the TEP co-chairs' leading the discussion was passed by acclamation.

The TEP co-chairs thanked the TEP members for their participation and stated that they look forward to the continued work.

The CMS representatives, Dr. Andress and Elena Balovlenkov, thanked the TEP for their participation, emphasized that they will continually improving the DFC site that is devoted to patients, encouraged patients to ask questions of CMS and of their facilities.

The third public comment was provided by Barry Smith, MD, PhD of the Rogosin Institute.

> "This is Barry Smith from the Rogosin Institute. A very simple comment, just to be supportive of the effort and willing to be helpful in any way that we can, you know, from our own experience."

5. Post-TEP Teleconference Calls

After the in-person TEP meeting, a Post-TEP Teleconference Call was scheduled for March 22, 2017 as a follow-up to the re-baselining issues discussed during the DFC Star Ratings TEP In-Person Meeting (see Appendix F for the Star Rating Post-TEP Teleconference Call Minutes).

6. Appendices

List of Appendices

Appendix A: Technical Expert Panel Charter
Appendix B: Technical Expert Panel Composition Form
Appendix C: Star Rating Pre-TEP Teleconference Call #1 Minutes
Appendix D: Star Rating Pre-TEP Teleconference Call #2 Minutes
Appendix E: Star Rating Pre-TEP Teleconference Call #3 Minutes
Appendix F: Star Rating Post-TEP Teleconference Call Minutes
Appendix G: TEP Preliminary Voting Form for Candidate Measures
Appendix H: TEP Preliminary Voting Results and TEP Comments for Candidate Measures
Appendix I: In-Person TEP Voting Form for SRR and SIR Measures, and ICH CAHPS Scoring Options
Appendix J: In-Person TEP Voting Results and Comments for SRR and SIR Measures, and ICH CAHPS Scoring Options
Appendix K: TEP Star Rating Re-baselining Proposal Voting Form
Appendix L: TEP Star Rating Re-baselining Proposal Voting Results and TEP Comments
Appendix M: CDC NHSN SIR Distribution Figure 1: Distribution of Denominator - Facilities with 12 Months Reporting and 0 BSI - 2015
Appendix N: NHSN SIR Distribution BSI 12 Month Reporter Figure 2
Appendix O: KCP public comment letter
Appendix P: TEP In-person meeting presentation slides (the teleconference slides are in the appendix of the in-person presentation)

Appendix A: Technical Expert Panel Charter

Project Title:

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

Dates: January – June 2017

Project Overview:

The Centers for Medicare & Medicaid Services (CMS) has contracted with The University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) to convene a Technical Expert Panel (TEP) to obtain recommendations on potential quality measures to include in the Dialysis Facility Compare Star Ratings. The contract name is the ESRD Quality Measure Development, Maintenance, and Support contract. The contract number is HHSM-500-2013-13017I.

The Medicare Dialysis Facility Compare (DFC) website displays overall star ratings of dialysis facilities. The star ratings provide an overall summary of the quality of care delivered to patients with ESRD. CMS developed the overall star rating to help health care consumers (including patients and caregivers) understand how CMS measures quality and to more easily identify differences in overall quality when selecting a dialysis facility. CMS is considering adding additional quality measures to this overall star rating. The public reporting of these ratings is part of CMS' broader initiative for all of the Medicare Compare sites to make quality information more accessible to patients, caregivers, and other key audiences, including providers and policymakers.

To develop a TEP that focuses on star ratings, we adopt CMS' structured and standardized approach. A key step is to ensure input from experts and the public. We are now seeking input from individuals with relevant experience, expertise, and a variety of perspectives to serve on this TEP tasked with review of the proposed candidate measures for inclusion in the Dialysis Facility Compare Star Ratings.

Historically, a TEP was convened in 2015 to review the original star rating methodology and presentation of the star ratings on the DFC website. The 2015 TEP provided several recommendations that were implemented in the updated star rating methodology that was publicly released in October 2016. In the new TEP that is to be convened in 2017, TEP members are requested to provide recommendations on candidate measures proposed for inclusion into the DFC Star Ratings, and how current and future potential patient reported outcome measures could be reported in the Star Ratings. The TEP will also provide recommendations on how to reset the baseline year thresholds for the Star Ratings when new measures are added or old measures are removed, or when the categorization of star ratings is no longer informative, e.g., when a very high percentage of facilities are clustered in certain star categories.

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

We anticipate that the meeting will take place over one day and will include additional pre- and follow-up teleconference calls. TEP members' attendance at all these meetings is required.

Project Objectives:

The University of Michigan Kidney Epidemiology and Cost Center (UM-KECC), through its contract with the Centers for Medicare and Medicaid Services (CMS), will convene a technical expert panel to evaluate and make recommendations regarding the addition of new measures proposed for the October 2018 release of the DFC Star Ratings.

TEP Objectives:

The TEP will be expected to:

1. Develop recommendations on the inclusion of candidate measures reported on the Dialysis Facility Compare (DFC) into the DFC Star Ratings. Recommendations should take into account whether the potential addition of new measures to the star rating provides a more well-rounded depiction of the quality of dialysis facilities and whether the information is something patients can understand and want to see reported on DFC.
2. Develop specific recommendations on the method for inclusion of current and future potential patient reported outcomes in the Star Ratings (e.g. separate from or combined with clinical outcome measure sets, and how current and future potential patient reported outcome measures could be reported in the Star Ratings).
3. Develop recommendations on how to reset the baseline year thresholds when measures are added/retired or when the Star Ratings categories no longer reflect informative differences among facilities.

Scope of Responsibilities:

The role of the TEP and each member is to advise UM-KECC regarding the star ratings system.

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

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

Duties and Role of TEP members: As defined by CMS in the Measure Management System Blueprint, TEPs are advisory to the measure contractor. In this advisory role, the primary duty of the TEP is to review the proposed star rating candidate measures and supporting materials, and provide recommendations to UM-KECC regarding the addition of candidate measures.

In January and February 2017, TEP members will be expected to attend pre-TEP conference calls as necessary; and attend one in-person meeting in February of 2017 (dates to be determined) in Baltimore, MD; and attend additional follow-up teleconference meeting and provide follow-up written feedback and comments as needed (via e-mail).

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

During the In-Person Meeting, the TEP will review the proposed candidate measures to be included in the star rating. The key deliverable of the TEP in-person meeting includes a summary report documenting the discussions, decisions and proposed recommendations that are made during the In-Person Meeting.

At the end of the in-person meeting the TEP chair(s) and TEP members will prepare and present a summary of proposed recommendations. Subsequent to the in-person meeting, the TEP chair(s) will have additional contact with UM-KECC moderators to work through further discussion of proposed recommendations. After the In-Person Meeting (approximately February –June 2017): TEP members will be asked to review and provide input on a summary report of the TEP meeting discussions, proposed recommendations, and other necessary documentation forms.

Guiding Principles:

Potential TEP members must be aware that:

- Participation on the Technical Expert Panel is voluntary.
- Input will be recorded in the teleconference meeting minutes and the TEP in-person meeting summary report.
- Proceedings of the in-person meeting will be summarized in a report that is disclosed to the general public.
- Patient TEP participants may elect to remain anonymous in all TEP proceedings. They should notify UM-KECC if they choose to have their names omitted from the TEP roster, in-person meetings, and all meeting minutes.
- If a TEP member chooses to disclose private and personal data as part of TEP in-person or teleconference discussions or as part of other input, that information and communications are not covered by patient-provider confidentiality.
- All questions about confidentiality and anonymity will be answered by the TEP organizers.
- All potential TEP members must disclose any current and past activities that may pose a potential conflict of interest for performing the tasks required of the TEP.
- All potential TEP members must commit to the expected time frame and participation outlined for the TEP.
- The TEP summary report will include the results of TEP votes taken for specific decisions and recommendations.
- Written opinions and rationales for votes from individual TEP members will be included, if requested by the TEP member.

Estimated Number and Frequency of Meetings:

- TEP members should expect to come together for one to two teleconference meetings prior to the in-person meeting held February 2017, in Baltimore, MD.
- The in-person meeting February 2017 (final dates to be determined).
- After the in-person meeting, TEP members should expect to attend additional teleconference meetings as needed.

Date Approved by TEP:

TBD

TEP Membership:

TBD

Appendix B: Technical Expert Panel Composition Form

TECHNICAL EXPERT PANEL COMPOSITION (MEMBERSHIP) LIST

Project Title:

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

Dates:

January – June 2017

Project Overview:

The Centers for Medicare & Medicaid Services (CMS) has contracted with The University of Michigan Kidney Epidemiology and Cost Center (UM-KECC). The contract name is ESRD Quality Measure Development, Maintenance, and Support contract. The contract number is HHSM-500-2013-13017I. As part of its measure development process, CMS asks measure developers to convene groups of stakeholders and experts who contribute direction and thoughtful input to the measure developer during measure development and maintenance. The following individuals were selected and have agreed to serve as the Technical Expert Panel for this project.

Name, Credentials, and Professional Role	Organizational Affiliation, City, State	Consumer/ Patient Perspective	Clinical Expertise	Performance Measurement	Coding and Informatics	Conflict of Interest Disclosure
Paul T. Conway, BA *TEP co-chair* *President* *Board Member* *Board Member*	 American Association of Kidney Patients (AAKP) Mid-Atlantic Renal Coalition (MARC) Polycystic Kidney Disease Foundation (PKDF) Falls Church, VA	X				None

Catherine A. Sugar, PhD, MS *TEP co-chair* *Director* *Professor*	 Semel Institute Statistics Core, University of California, Los Angeles Los Angeles, CA Departments of Biostatistics, Statistics & Psychiatry University of California, Los Angeles Los Angeles, CA			X		None
Lorien Dalrymple, MD, MPH, *Vice President of Epidemiology and Research* *Volunteer Clinical Faculty, Associate Professor*	Fresenius Medical Care North America (FMCNA) Dept. of Medicine, Division of Nephrology, University of California, Davis		X	X		Employed by Fresenius Medical Care NA and member of the KCQA Steering Committee
Pasquale Dangelantonio, *Network 4 SME; former Patient Representative*	ESRD Network 4	X				None
Amanda Grandinetti, MPH, *Senior Specialist, Performance Measures and Analysis* *Kidney Action Committee Member*	American Academy of Dermatology National Kidney Foundation	X		X		None
Mark Joseph, MD, *Medical Director* *Clinical Assistant Professor*	Dialysis and CRRT Phoenix Children's Hospital Phoenix, Arizona Department of Pediatrics, University of Arizona Tucson, Arizona		X	X		Member of the Horizon Pharmaceutical speaker's bureau for nephropathic cystinosis

Richard Knight, MBA, *Vice President/Chair of Public Policy* *Board Member*	American Association of Kidney Patients (AAKP), New Carrollton, MD Mid-Atlantic Renal Coalition (MARC)	X				None
Jewell Kyle, RN, BSN, CNN, *RN Charge Nurse/Staff Educator*	Chattanooga Kidney Centers, LLC		X			None
J. Richard Landis, PhD, MS, *Professor of Biostatistics* *Professor of Statistics*	University of Pennsylvania Perelman School of Medicine Philadelphia, PA University of Pennsylvania (Wharton School) Philadelphia, PA			X		None
Allen R. Nissenson, MD, FACP, *Chief Medical Officer* *Emeritus Professor of Medicine*	DaVita HealthCare Partners El Segundo, California David Geffen School of Medicine at University of California, Los Angeles Los Angeles, CA		X	X		Full time employee, DaVita, Inc.
Chris Sarfaty, MSW, LICSW, *Patient-Centered Collaborative Care Coach* *Licensed Independent Clinical Social Worker*	Coaching for Health Professionals Whately, MA Baystate Franklin Medical Center	X				None
Nicole Stankus, MD, MSc, *Medical Director* *Associate Professor of Medicine*	DaVita Stony Island Dialysis Center The University of Chicago Chicago, IL		X	X		Member of the DaVita Physician Council

Sumi Sun, MPH, *VP of Analytics and Quality Strategy*	Satellite Healthcare San Jose, CA			X	X	None
Frederic Talton, BA, *Board Member*	Dialysis Patient Citizens (DPC)	X				None
David M. White, *Board of Director Member* *Acting Chair* *Chair on the Patient Advisory Committee and Medical Review Board member*	American Association of Kidney Patients (AAKP) Kidney Health Initiative's Patent and Family Partnership Council (PFPC) Mid-Atlantic Renal Coalition (MARC) Hillcrest Heights, MD	X				None
Anonymous Patient (anonymity requested)	Dialysis Patient	X				None
Contractor Staff	University of Michigan Kidney Epidemiology and Cost Center (UM-					
Yi Li, PhD, *Professor of Biostatistics*						
Joseph Messana, MD, *Interim Director/ Nephrologist*						
Richard Hirth, PhD, *Professor of Health Management and Policy*						
Claudia Dahlerus, PhD, MA, *Principal Scientist*						
Ji Zhu, PhD, *Professor of Statistics*						

Stephen Salerno, *Graduate Student Research Assistant*						
Karen Wisniewski, MPH, *Lead Research Analyst*						
Natalie Scholz, MPH, *Lead Research Analyst*						
Cindy Liao, MS, MPH, *Research*						
Casey Parrotte, PMP, *Project Manager/ Research Analyst*						
Jordan Affholter, BA, *Research Analyst*						

*One TEP member was unable to attend the in-person meeting.

**Another TEP member was unable to continue participation.

Appendix C: Star Rating Pre-TEP Teleconference Call #1 Minutes

End-Stage Renal Disease (ESRD) Quality Measure Development, Maintenance, and Support Project
ESRD Dialysis Facility Compare (DFC) Star Ratings
Star Rating Pre-TEP Teleconference Call #1 Minutes
February 2, 2017 1:00pm – 2:30pm (EST)

TEP Members	UM-KECC	CMS
Paul Conway	Yi Li	Joel Andress
Catherine Sugar	Joseph Messana	Elena Balovlenkov
Lorien Dalrymple	Richard Hirth	Jesse Roach
Amanda Grandinetti	Claudia Dahlerus	
Richard Knight	Cindy Liao	**CMS/RTI**
Jewell Kyle	Stephen Salerno	Elizabeth Goldstein (CMS/CM)
Allen Nissenson	Jordan Affholter	Julia Zucco (CMS/CM)
Chris Sarfaty	Casey Parrotte	Judy Lynch (RTI)
Nicole Stankus	Karen Wisniewski	Amy Hendershott (RTI)
Sumi Sun	Tempie Shearon	Celia Eicheldinger (RTI)
David White	YiFan Wu	
	Amy Jiao	**CDC**
	Natalie Scholz	Shunte Moon
		Priti Patel
		Daniel Pollock

Introductions
Dr. Yi Li welcomed everyone to the Pre-TEP conference call, and thanked the TEP members for their time and for serving on the TEP. Dr. Li stated the call was open to the public, being recorded, and that the last five minutes were set aside for public comments.

Dr. Li introduced Paul Conway and Dr. Catherine Sugar, who were selected to serve as TEP co-chairs, and thanked them for their willingness to lead the discussion. Dr. Li explained that the TEP would take place over a one day in-person meeting on February 21, 2017, in Baltimore, MD. Three teleconference calls were set up to begin TEP discussion of the primary topics prior to the in-person meeting.

Dr. Li stated that the University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) is the Centers for Medicare & Medicaid Services (CMS) contractor tasked with developing End-Stage Renal Disease (ESRD) quality measures. Dr. Li introduced himself as the Principal Investigator on this project and a Professor of Biostatistics at the School of Public Health at the University of Michigan, and provided introductions to the rest of the UM-KECC team: Dr. Joseph Messana (Clinical Nephrologist, University of Michigan; Dr. Richard Hirth (Professor of Health Management and Policy at the University of Michigan); and Dr. Claudia Dahlerus (Principal Research Scientist at UM-KECC). Dr. Li, Dr. Messana, Dr. Hirth, and Dr. Dahlerus are the TEP facilitators for the Star Rating TEP. Jordan Affholter is the contact person for the TEP members.

Dr. Joel Andress of CMS was also in attendance. Dr. Andress is the CMS Contracting Officer Representative for the ESRD Quality Measure and Public Reporting contract. Dr. Andress thanked all of the TEP members for volunteering their time to this project.

Elena Balovlenkov of CMS is the Dialysis Facility Compare (DFC) lead for Public Reporting and has been working in ESRD for 38 years. She welcomed and thanked everyone for their time and commitment. She thanked the large number of patient experts for joining the TEP and the TEP call.

Dr. Jesse Roach was introduced as a nephrologist working for CMS with Dr. Joel Andress and Elena Balovlenkov.

RTI (Research Triangle Institute), CDC (Centers for Disease Control and Prevention), and other CMS representatives were also present for the call.

Each TEP member gave a brief introduction.

Paul Conway: Paul Conway is the President of AAKP (American Association of Kidney Patients). He served as a co-chair of the 2015 Star Rating TEP. In addition, he has experience in patient advocacy positions and has personal experience with ESRD.

Dr. Catherine Sugar: Dr. Catherine Sugar is a Professor (Statistics, Biostatistics, and Psychiatry) at UCLA (University of Southern California, Los Angeles). She served as a co-chair of the 2015 Star Rating TEP.

Dr. Lorien Dalrymple: Dr. Lorien Dalrymple is a Nephrologist and Epidemiologist and is the Vice President of Epidemiology and Research for Fresenius Medical Care North America (FMCNA).

Amanda Grandinetti: Amanda Grandinetti is a Kidney Action Committee Member from the National Kidney Foundation (NKF). She is also a Senior Measure Developer at the American Academy of Dermatology. She has personal experience with ESRD.

Richard Knight: Richard Knight is the Vice President of the American Association of Kidney Patients (AAKP). He is also the Chair of Public Policy for the AAKP. He has personal experience with ESRD.

Jewell Kyle: Jewell Kyle is the Charge Nurse and Staff Educator for Chattanooga Kidney Centers, LLC.

Dr. Allen Nissenson: Dr. Allen Nissenson is a clinical Nephrologist and the Chief Medical Officer of DaVita Kidney Care. He is also a Professor Emeritus of the David Geffen School of Medicine at UCLA (University of California, Los Angeles). He also served on the 2015 Star Rating TEP.

Chris Sarfaty: Chris Sarfaty is a Patient-Centered Collaborative Care Coach with a Master's in Social Work. She has worked in dialysis for 25 years and has familial experience with ESRD. She also served on the 2015 Star Rating TEP.

Dr. Nicole Stankus: Dr. Nicole Stankus is a clinical Nephrologist and Medical Director of a dialysis facility in Chicago, Illinois. She is also an Associate Professor of Nephrology at the University of Chicago. She also served on the 2015 Star Rating TEP.

Sumi Sun: Sumi Sun is an epidemiologist and the Vice President of Analytics and Quality Strategy at Satellite Healthcare, which is a non-profit dialysis provider focused in California. She also served on the 2015 Star Rating TEP.

David White: David has personal experience with ESRD. He is a member of the Board of Directors for the American Association of Kidney Patients (AAKP). He is the acting chair of the Kidney Health Initiative's Patient and Family Partnership Council (PFPC). He also served on the 2015 Star Rating TEP.

Dr. Li gave brief introductions for Pasquale Dangelantonio, Dr. Mark Joseph, Dr. J. Richard Landis, and Frederic Talton, who were not able to attend this call.

Pasquale Dangelantonio: Pasquale Dangelantonio is a Subject Matter Expert with ESRD Network 4. He has personal experience with ESRD.

Dr. J. Richard Landis: Dr. J. Richard (Dick) Landis is a Professor of Biostatistics and Statistics at the University of Pennsylvania Medical School. He also served on the 2015 Star Rating TEP.

Dr. Mark Joseph: Dr. Mark Joseph is a Pediatric Nephrologist and Medical Director for Phoenix Children's Hospital.

Frederic Talton: Frederic Talton is a Board Member of Dialysis Patient Citizens (DPC). He has personal experience with ESRD.

TEP Overview and Introduction to Methodology

Dr. Li gave a brief overview of the TEP Member Role and Responsibilities. Dr. Li encouraged TEP members to share their opinions and experience. Dr. Li brought attention to the TEP objectives (from the TEP charter) and the TEP agenda for this call. Dr. Li encouraged any TEP members that had questions regarding the charter or agenda to contact UM-KECC.

Dr. Li gave a brief overview of the Dialysis Facility Compare (DFC) Site Star Ratings Description. A star rating is used to show how well a dialysis center delivers care and to make data on the quality of patient care easier to understand and use. Each dialysis center receives a rating between 1 and 5 stars. Currently, patient survey results are not currently included in the star ratings.

Dr. Li gave a brief overview of how measures are scored in order to create a final score and overall star rating. He explained that the star rating measures are standardized in order to create measure scores so that all scores are on comparable scales. Factor analysis grouped the reporting measures into various domains. Measure scores within each domain are averaged to determine a domain score; domain scores are averaged to determine a final score. The distribution of final scores across all dialysis facilities is used to establish the cutoffs. The cutoffs are used to determine the star rating category.

Dr. Li gave a brief explanation of the purpose of setting a baseline year and how data are used to define star category cutoff values in the baseline year. He explained how facility performance in subsequent years is scored against cutoff values defined in the current baseline year. Establishing the "baseline year" allows reporting of changes in facility performance over time. This recommendation emerged out of the prior 2015 Star Rating TEP.

He further provided definitions of *baseline year* and *current year* that would be used in future discussion. Baseline year refers to the collection year of data which is analyzed to set scoring standards for the DFC Star Ratings. Current year refers to the collection year of data being analyzed to evaluate facilities for the DFC Star Ratings compared against the cutoffs set in the baseline year.

Dr. Li presented an example that shows how a facility is scored. Measures values are translated into measure scores which then get grouped into domain scores. The respective domain scores are averaged to create final scores. The final score cutoffs determine the facility star rating categories.

Dr. Li also gave an example of how using a baseline year to set cutoff values allows one to see upward shifts in star ratings over time when the overall performance improves. Dr. Li offered a chance for TEP questions before the next section before turning the presentation over to Dr. Claudia Dahlerus.

Proposed Updated Measures and Candidate Measures for Star Rating

Dr. Claudia Dahlerus (UM-KECC) briefly explained that the layout of this section of the presentation is divided into a section on the proposed updated measures and a section on new candidate measures. Dr. Dahlerus clarified that UM-KECC would answer technical questions about the measures but that the TEP co-chairs would lead the TEP discussion.

Dr. Dahlerus explained that after the second teleconference call, UM-KECC will distribute a survey to the TEP members asking them to rate the measures considered for inclusion in the star ratings and to provide a brief rationale for the ratings given. She explained the survey feedback will be used to identify areas of consensus and areas where there may be a need for more discussion.

The TEP was first presented with the list of the current measures included used in the DFC Star Ratings:

- Standardized Transfusion Ratio (STrR)*
- Standardized Mortality Ratio (SMR)*
- Standardized Hospitalization Ratio (SHR)*
- Percentage of adult hemodialysis (HD) patients who had enough wastes removed from their blood during dialysis
- Percentage of pediatric hemodialysis (HD) patients who had enough wastes removed from their blood during dialysis
- Percentage of adult peritoneal dialysis (PD) patients who had enough wastes removed from their body during dialysis
- Percentage of adult dialysis patients who had hypercalcemia
- Percentage of adult dialysis patients who received treatment through arteriovenous (AV) fistula*
- Percentage of adult patients who had a catheter left in vein longer than 90 days for their regular hemodialysis treatment*

*Measures with an asterisk (*) are measures that have been updated.

Dr. Dahlerus then provided an overview of updated quality measures proposed to replace the current measures used in the star ratings; she then moved to an overview of the new candidate measures. First, she gave a brief overview on the following proposed measures that were updated as part of the National Quality Forum (NQF) process:

- Vascular Access: Standardized Fistula Rate
- Vascular Access: Long-term Catheter Rate
- Standardized Mortality Ratio (SMR)
- Standardized Hospitalization Ratio (SHR)
- Standardized Transfusion Ratio (STrR)

Dr. Dahlerus briefly explained the technical updates on the measures (specific details on the updates can be found on the Star Rating TEP Teleconference slides). After presenting the proposed updated measures, Dr. Dahlerus turned the discussion over to the TEP co-chairs to lead discussion.

One of the TEP co-chairs noted that efforts by UM-KECC and CMS reflect their interest in hearing patient and consumer feedback. The co-chair explained that the TEP calls would focus on building TEP knowledge of methodology and the measures in order to facilitate TEP in-person discussion. The co-chair highlighted the importance of translating technical data into understandable results. The comments that are received by the TEP will be used to develop the DFC Star Ratings.

The TEP co-chair asked a clarifying question, asking UM-KECC to confirm whether the updated measures are already part of the star ratings, and if the updates were improvements to the measures. UM-KECC confirmed that the proposed updated measures reflect improvements to the measures, but the proposed updated measures have not been implemented in the star ratings.

After the discussion of measure updates, the TEP co-chairs opened up the discussion to the TEP members.

One TEP member asked UM-KECC technical questions about the standardized measures (including SMR, SHR, STrR, and Standardized Fistula Rate). The TEP member brought up a concern of reliability for the standardized measures (SMR, SHR, STrR, and SRR), specifically referencing testing results showing low reliability for small facilities. For example, for SMR, he indicated small facilities had a reliability of about 0.40, noting it is lower than a certain acceptable target of above 0.70. The TEP member invited other members and methodologists in particular to weigh in on this issue. The TEP member was in favor of the Vascular Access measure changes.

One TEP member asked about how fistulas are counted in the updated Standardized Fistula Rate (SFR) measure. UM-KECC clarified that SFR only counts patient-months where the fistula is being used to provide dialysis treatments, as reported by the facilities. Non-functional fistulas are not counted as fistulas for that measure.

Regarding the change to the updated SMR measure, which is limited to the Medicare population, the TEP co-chair asked about how to find out which facilities have a higher private insurance population. Dr. Dahlerus responded that UM-KECC could look into this issue. In response to this question about the inclusion of prevalent comorbidities in the SMR, Dr. Dahlerus also clarified that the adjustment for comorbidities are from the Medicare claims data. She also noted that UM-KECC does not have access to comorbidity data for privately insured patients as these are non-Medicare patients. The other TEP co-chair asked if facilities have different SMR depending on private insurance or Medicare. In response, Dr. Dahlerus stated that looking into this would take time. Dr. Dahlerus also clarified that the proposed updated measures have been reviewed and adjudicated through the National Quality Forum (NQF) process. NQF has vetted the measures for reliability and validity and recommended these measures for endorsement. Dr. Andress (CMS) stated that Dr. Jack Wheeler (UM-KECC) may have analyses that addressed the private insurance versus Medicare patient mortality issue as part of the work performed for updating the SMR with prevalent comorbidities. Dr. Andress asked UM-KECC to look into this further.

The TEP co-chair asked for clarification regarding the definition of a transfusion event. Dr. Dahlerus clarified that the updated transfusion definition now includes only procedure codes or value codes but no longer uses revenue-center dates on inpatient claims. It was noted this is a more restricted definition and reduces the number of observed transfusion events.

After the TEP Measure Update Discussion, Dr. Dahlerus (UM-KECC) presented the new candidate measures proposed for inclusion into the DFC Star Ratings:

1. Standardized Readmission Ratio (SRR)
2. Pediatric Peritoneal Dialysis Adequacy: Achievement of Target Kt/V (Pediatric PD Kt/V)
3. In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)
4. National Healthcare Safety Network (NHSN) Bloodstream Infection

Only the SRR and Pediatric PD Kt/V measures were presented and discussed on this call. The ICH CAHPS (developed by RTI and CMS) and the NHSN Bloodstream Infection measure (developed by the Centers for Disease Control, [CDC]) discussion will occur on the second pre-TEP teleconference call on February 13, 2017. CDC, RTI, and CMS will be available to answer technical questions regarding their respective measures on the second pre-TEP teleconference call.

Dr. Dahlerus provided a brief overview of the SRR and the Pediatric PD Kt/V measures (the specific details on the updates can be found on the Star Rating TEP Teleconference slides). After the overview of the candidate measures, the TEP co-chairs led a discussion of these measures.

The TEP co-chair asked about the measures being calculated as ratios but reported as rates. UM-KECC explained that CMS made a policy decision to report the ratios as rates since rates are simpler to interpret by consumers. To obtain the rate, each ratio is multiplied by the national average for that measure.

One TEP member asked if the SRR on DFC aligns with the SRR measure used in QIP. UM-KECC clarified that the SRR measure is harmonized between the DFC and QIP programs. The TEP co-chair stated how measure alignment between the two programs was also a discussion topic during the 2015 DFC Star Ratings TEP. In response, Dr. Joel Andress (CMS) clarified that DFC and QIP are two separate CMS programs. The QIP program implements a payment penalty for facilities based on quality measure performance. A number of the DFC measures are also in the QIP program, but the two programs are separate. Dr. Andress explained that based on stakeholder feedback, effort has been made to harmonize measures between the two programs. Dr. Andress explained that a Measures Manual has been created to show how the measures are similar or different between the two programs.

One of the TEP co-chairs asked a question about the SHR and SRR. The co-chair asked how the SHR and SRR measures work together. Dr. Dahlerus (UM-KECC) clarified that the measures are complementary. The SHR focus is on how facilities perform on overall patient care at the facility to assess how often patients are being admitted to a hospital. The SRR looks only at the facility's patients who have been discharged and whether they are readmitted within 4-30 days. The focus of the SRR is on how facilities are managing those patients post-discharge, with the goal of encouraging better care coordination after discharge.

The TEP co-chair asked if a readmission from 4-30 days after discharge would be counted in both the SHR and SRR. Dr. Dahlerus clarified that the readmission would be counted in both the SHR and SRR.

One TEP member stated that the SRR measures overall readmissions instead of a cause-specific readmission. The TEP member stated that a cause-specific readmission measure would be preferred.

Another TEP member stated concerns about the SRR, including that other factors influence readmissions.

One TEP member asked a clarifying question regarding new dialysis patients that were recently hospitalized. They wanted to know if a patient just started at the dialysis facility whether they are counted in the SHR measure for that facility. The TEP member also asked if a new patient was readmitted within the 30-day period, would they would also be counted in the SRR as well? Dr. Dahlerus stated UM-KECC would send a follow-up to that question before the next TEP call.

One TEP member stated that the SRR may be viewed as a coordination of care measure.

One of the TEP co-chairs stated that it would be helpful to know what adjustments are included in the SRR. Dr. Dahlerus clarified that the specific adjustments for the standardized measures were in the Measure Information Forms provided in the background materials folder that was sent to the TEP. Dr. Dahlerus provided a basic overview of the risk adjustment factors for SRR, referencing the Measure Information Form.

Another TEP member stated that some admissions and readmissions are preventable and some may not be preventable. The TEP member felt that there are some items facilities do not have control over.

Dr. Joseph Messana (UM-KECC) thanked the TEP member for the comment. Dr. Messana provided a clarifying statement regarding the focus of the TEP. Dr. Messana stated that UM-KECC is asking for TEP discussion about the measures as specified. Dr. Messana stated these measures have been vetted by the NQF measure evaluation process. Dr. Messana clarified that this TEP is not charged with measure development, but is charged with the question of whether the measures as specified are appropriate for use in the DFC Star Ratings.

Public Comments
There were no public comments received during this TEP call.

Closing Remarks
Dr. Joel Andress (CMS) thanked the TEP members for attending the teleconference.

Elena Balovlenkov (CMS) thanked the TEP members for attending the teleconference, and stated the public reporting contractors (NORC and Ketchum) were on the call listening to the TEP discussion.

Jordan Affholter (UM-KECC) reminded the group of the upcoming TEP calls on February 13, 2017 from 3:30-5:00pm (EST) and on February 14, 2017 from 11:00am-12:30pm (EST).

Appended Answers

During the teleconference call, one TEP member asked if a new dialysis patient (that was not previously cared for in a facility) was recently hospitalized, and then goes to a facility to start dialysis at a facility, is the hospitalization attributed to this facility? Additionally, if the new patient was readmitted within the 30-day period would they be counted in the SRR as well?

UM-KECC provided the response below after the TEP call.

Response: For SHR, if the hospitalization happened before the patient was dialyzing at this facility, then this hospitalization would not be attributed to this dialysis facility. A patient needs to be dialyzed in the facility for 60 days before a hospitalization gets counted against the facility.

See Measure Information Forms for description of how patients are assigned to a facility.

For SRR, if the index discharge happened after the patient started dialyzing at this facility, and the patient got readmitted to the hospital within 4-30 days after the index discharge, then the readmission counts against the facility.

See Measure Information Forms for a list of exclusion criteria for index discharges.

Appendix D: Star Rating Pre-TEP Teleconference Call #2 Minutes

End-Stage Renal Disease (ESRD) Quality Measure Development, Maintenance, and Support Project
ESRD Dialysis Facility Compare (DFC) Star Ratings
Star Rating Pre-TEP Teleconference Call #2 Minutes
February 13, 2017 3:30pm – 5:00pm (EST)

TEP Members	UM-KECC	CMS
Paul Conway	Yi Li	Joel Andress
Catherine Sugar	Joseph Messana	Elena Balovlenkov
Lorien Dalrymple	Claudia Dahlerus	Jesse Roach
Amanda Grandinetti	Cindy Liao	Celeste Bostic
Mark Joseph	Stephen Salerno	**CMS/RTI**
Richard Knight	Jordan Affholter	Debra Dean-Whittaker (CMS/CM)
Jewell Kyle	Casey Parrotte	Julia Zucco (CMS/CM)
J. Richard Landis	Ji Zhu	Judy Lynch (RTI)
Chris Sarfaty	Natalie Scholz	Amy Hendershott (RTI)
Nicole Stankus	Jennifer Sardone	Celia Eicheldinger (RTI)
Sumi Sun		**CDC**
David White		Shunte Moon
		Priti Patel
		Daniel Pollock
		Christi Lines

Introductions

Jordan Affholter (UM-KECC) welcomed everyone to the second Pre-TEP conference call, and thanked them for joining the call. He reminded the group that the call was open to the public, and was being recorded.

Dr. Yi Li (UM-KECC) welcomed everyone to the second Star Rating TEP conference call. Dr. Li thanked everyone for attending the previous call. Dr. Li administered an ordered roll call to take TEP member, CMS, CMS/RTI, and CDC attendance.

CMS's public reporting contractors NORC and Ketchum were also listening in on the call.

Dr. Li identified the two main agenda items for the call. The first item was to review the In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) and National Healthcare Safety Network (NHSN) Bloodstream Infection Standardized Infection Ratio (SIR) measure. The second item was how to include potential Patient Experience of Care measures in the star ratings.

TEP Presentation and Discussion
Dr. Dahlerus (UM-KECC) provided a brief summary of the ICH CAHPS measures (developed by RTI and CMS) and the NHSN Bloodstream Infection measure (developed by the Centers for Disease Control, [CDC]). Dr. Dahlerus explained that RTI/CMS and CDC were on the call to address any technical questions regarding their measures.

Dr. Dahlerus stated that after the call, TEP members will be sent a voting form asking TEP members to rate all the candidate measures presented on the first and second call. TEP members will be asked to provide feedback for their rating and to submit their survey by close of business on Wednesday February 15, 2017. Dr. Dahlerus stated that the feedback will help determine areas of consensus or areas that need more discussion at the in-person meeting.

Dr. Dahlerus provided a brief overview of the ICH CAHPS measures. ICH CAHPS has three composite measures and three global measures. The three composite measures are (1) kidney doctors' communication and caring, (2) quality of dialysis center staff care and operations, and (3) providing information. The three ICH CAHPS global (individual) ratings are (1) rating of your kidney doctors, (2) rating of the dialysis center staff, and (3) rating of the dialysis center. Top-box results only are currently reported on the public Dialysis Facility Compare (DFC) website for the composite and global ratings. The ICH CAHPS survey results are risk adjusted.

Dr. Dahlerus provided a brief overview of the NHSN Standardized Infection Ratio (SIR) measure. The SIR of bloodstream infections is calculated among patients receiving hemodialysis at outpatient hemodialysis centers, and shows the number of observed infections to the number of predicted infections at the facility. Dr. Dahlerus briefly described the SIR measure distribution among dialysis facilities. After Dr. Dahlerus provided a brief summary of the two measures, TEP members were given time for questions and discussion.

One of the TEP co-chairs asked several technical questions about the composite scores from the ICH CAHPS measures. The ICH CAHPS representative stated that multiple questions compromise the composite scores. The ICH CAHPS representative stated the raw scores are patient-mix adjusted. The ICH CAHPS representative stated that each of the composites have five or more questions. "Yes" and "No" answers are averaged to get the final score for the composite measures.

The other TEP co-chair stated that if a patient answers the questions on the survey, patient input is counted, and therefore these scores are patient driven. The ICH CAHPS representative verified that this statement was correct.

The TEP co-chair stated that each individual question contributes to the respective composite scores.

One TEP member asked about the composite scores. They asked that if a unit had 50% "always" responses and 50% "never" responses, and a second unit had 50% "always" and 50% "usually," whether both centers would be rated the same because they both had 50% of "always" responses, and that is only what is reported.

The ICH CAHPS representative clarified that for the top-box results, those two facilities would show the same top-box results. However, the other boxes (middle and bottom-box) reflect the full results for the composites.

One of the TEP co-chairs asked which boxes are reported on DFC. The co-chair asked if the top, middle, and bottom boxes are reported on DFC. The ICH CAHPS representative stated that there are three levels for the ICH CAHPS measures on DFC. UM-KECC next stated they would follow-up to clarify exactly what information is currently publicly reported on DFC.

One TEP member asked for clarification regarding the ICH CAHPS boxes. The ICH CAHPS representative stated that "Always" is reflected in the top-box, "Usually" is reflected in the middle box, and

"Sometimes" and "Never" are reflected in the bottom box. The ICH CAHPS representative referred to the ICH CAHPs website that explains the measure methodology in greater detail.

The TEP co-chair asked for TEP input on the factors captured on the ICH CAHPS and what was important to patients, and if those factors were reflective of the quality of the dialysis facility.

One TEP member asked for clarification on the NHSN Standardized Infection Ratio (SIR) distribution (on the slide), which shows a large number of facilities with a value of zero for SIR. The CDC stated that the graph with a large number of facilities with an SIR of zero includes a number of facilities that did not report all months, but also includes facilities that reported no bloodstream infections for 2015.

One of the TEP co-chairs stated that the large number of facilities where SIR is zero does not fit the trend of the whole distribution and asked for some clarification. CDC stated that there are some facilities that would have no bloodstream infections in a year. CDC representatives stated they would examine the results for those facilities.

One TEP member asked if the CDC could comment on the underreporting of bloodstream infections. CDC stated that they know underreporting happens, but to what extent underreporting occurs is unknown. CDC clarified that the studies referenced are CDC studies, and gave an example of one of the studies that measured the data in one dialysis organization's electronic medical record. CDC stated that the study demonstrated that the electronic medical record does not cover all the bloodstream infections. A large percentage of bloodstream infections would be missed if electronic medical records were the only source of identification. The second study looked at a specific subset of data for a small number of facilities, which might be taken out of context. CDC stated that the recent validation studies are small, and there are no nationally representative data to show the levels of underreporting across facilities.

One of the TEP co-chairs asked, how does the patient population benefit from the NHSN Bloodstream Infection measure? CDC responded that it is important for patients to know about the frequency of bloodstream infections at facilities, because they affect mortality and hospitalization. CDC responded that bloodstream infections are also an indication of patient quality of life. CDC stated that some of bloodstream infections can be prevented. Facilities that have adopted safety practices observe fewer bloodstream infections.

One TEP member asked whether we were moving on or continuing the discussion about items in the patient survey (ICH CAHPS) that are important to patients. One of the TEP co-chairs stated that this item will be addressed in the next section. The TEP co-chairs finished this TEP discussion and asked Dr. Li to present the second agenda item.

Dr. Li presented the three options for how to include patient-reported outcomes into the DFC Star Ratings. Option 1 involves having one overall star rating that combines patient experience of care and clinical quality measures in one rating. Option 2 involves two separate star ratings, one for patient experience of care rating and one for clinical quality measures. Option 3 would include an overall star rating, an overall patient experience of care rating, and a clinical quality measure rating.

Dr. Li explained that UM-KECC uses a statistical approach called factor analysis in order to generate groupings of related individual measures. The domains are used to compute the final rating for each facility.

Dr. Li summarized the results which show a strong correlation among the six ICH CAHPS measures, and a lack of correlation between the ICH CAHPS and clinical quality measures. He indicated these results may suggest creation of 1-2 ICH CAHPS domains separate from the clinical measure domains. For the clinical measure factor analysis, three clinical domains were identified.

Dr. Li completed his presentation and asked the TEP co-chairs to lead the discussion about the three options for incorporating the patient-reported outcomes in the star ratings.

One of the TEP members asked how measures reflecting other aspects important to patients would be presented in the star ratings. In particular, he wanted to know about how a dialysis facility deals with itching and how that would be reflected in the grouping (i.e., which factor would represent that issue). The TEP member stressed the importance of individual patient goals.

Dr. Li stated that the DFC Star Ratings were based on the measures reported on the DFC website. He explained that factor analysis groups similar measures to prevent one measure from dominating the whole star rating.

The TEP member stated that some centers are more aware of issues like itching than others. The TEP member stated that there are some things that patients may not be aware of so it may be hard for them to articulate what aspects of quality and care are important to them.

Dr. Messana stated many of those patient perceptions of care are not yet available to be included in the DFC Star Ratings. UM-KECC will be convening a Patient Reported Outcomes TEP in a few months to develop recommendations about patient-reported outcome measures. Dr. Messana stated that the current data may not be able to fully address those concerns about patients' perception of quality.

One of the TEP co-chairs stated that the ICH CAHPS measures may be reporting different information than the clinical measures from a statistical point of view. The TEP co-chair asked whether patients would prefer a separate rating for patient experience of care measures and for clinical measures.

One TEP member recommended separate DFC Star Ratings for the patient experience of care measures and the clinical measures. The TEP member stated a preference for the second option (separate star ratings), but might even prefer a fourth option: the overall star rating (combining the clinical star rating and the patient experience star rating) is reported, and then the patient experience of care star rating is reported separately. In other words, no separate clinical measure star rating. The TEP member stated that the fewer star ratings the better.

One of the TEP co-chairs asked the TEP member to confirm whether it was important to include a star rating dedicated to patient input/patient experience. The TEP member confirmed that it was.

Dr. Messana stated there might be more options than the three presented so far. The three options were chosen to help start the discussion, but more options could be considered. Dr. Messana stated that both options 2 and 3 have a separate patient experience of care star rating.

One TEP member stated that more weight should be put on the patient experience of care star rating. One of the TEP co-chairs asked the panel if the measures going into the patient star rating are more important than the clinical measures. One TEP member responded that the second option would allow patients to indicate what is more important to them, whether that is the patient experience measure or a clinical measure. The TEP member stated if the different types of ratings are combined, then the TEP has to decide whether the clinical or patient measures carry more weight.

One TEP member stated there is importance in having a balance between the clinical measures and the patient experience/patient-reported outcomes. Patients may want clinical and patient-reported information, therefore they would want to have a separate patient star rating.

Another TEP member stated that keeping the clinical and patient star ratings separate would be the simplest solution. If the TEP uses the option with three star ratings including the overall star rating, there will always be a question of weighting, or specifically how to weight each in the overall measure. Having the two separate star ratings would be the most valuable.

One of the TEP co-chairs agreed that it would be ideal to have two separate categories as they are measuring different things.

There was also further brief discussion about the new candidate measures. One TEP member asked what percent of dialysis facilities are unable to get the minimum number of surveys for the ICH CAHPs measures. The TEP member asked, how is the statistical precision of the score impacted, and, the ability to distinguish differences that are meaningful to patients in the ICH CAHPs composite scores when there are less than 200 returned surveys from a facility over the reporting period?

The ICH CAHPS representative reported that approximately 50% of the dialysis facilities do not have ICH CAHPS measures publicly reported on DFC. The ICH CAHPS representative stated that some facilities may not be required to administer the survey. The ICH CAHPS representative clarified that if facilities do not have 30 completed surveys, the rule is that the publicly reported data is suppressed.

One of the TEP chairs stated that a high percentage of missing ICH CAHPS results would be problematic for combining clinical and patient experience care domains, as this would impact scoring. Therefore, with missing data, it might make sense to keep the two categories separate.

As part of the wrap-up, Dr. Messana stated that the discussions so far are preliminary. The discussions along with the survey will give UM-KECC a sense of how much time to budget for the items at the in-person meeting. For additional discussion about patient-reported outcomes, there will be more time for discussion at the in-person meeting.

Public Comments
There were no public comments received during this TEP call.

Closing Remarks
Dr. Li (UM-KECC) thanked the TEP members for attending the teleconference and reminded the group of the upcoming TEP call on February 14, 2017 from 11:00am-12:30pm (EST).

Addendum

Dr. Nissenson was unable to attend the call and provided these comments:

1. Very concerned by NHSN. Issue here is one of accurately capturing BSIs. The best way to get a low BSI rate is to have a low blood culture rate. Since NHSN includes blood cultures drawn by hospitals in the first 24 hours after admission this can be problematic for dialysis facilities since hospitals are slow to share data. Unless there is some way of determining that all facilities have equal access to these, and

report them, there is not a level playing field. Once again the well performing, diligent facilities will be penalized.

2. Several issues with ICH-CAHPS:

a. Every 6 month sampling is a burden for patients and because of lag in availability of results does not allow for a "corrective" response by the facility

b. There is a real question if this survey is a true representative of PROs- that is, do we really think that what is important to an individual patient is reflected in an aggregated facility score?

c. The survey contains a domain about satisfaction about physicians- this would be a physician metric, not one for which the facility should be held accountable

d. Finally, to the extent that PROs are desirable in this area, there is currently so much work being done by KCQA, NQF, ESCOs (the latter looking at KDQOL)- there is an opportunity for working in silos, duplicating effort and even working at cross-purposes. Need to clearly articulate if CMS believes that using ICH-CAHPS represents PROs in 5-Star- I would be opposed to this.

3. Regarding the options for reporting if this moves forward- option 2 seems the clearest

Appendix E: Star Rating Pre-TEP Teleconference Call #3 Minutes

End-Stage Renal Disease (ESRD) Quality Measure Development, Maintenance, and Support Project
ESRD Dialysis Facility Compare (DFC) Star Ratings
Star Rating Pre-TEP Teleconference Call #3 Minutes
February 14, 2017 11:00am – 12:30pm (EST)

TEP Members	UM-KECC	CMS
Paul Conway	Yi Li	Joel Andress
Catherine Sugar	Joseph Messana	Elena Balovlenkov
Lorien Dalrymple	Claudia Dahlerus	Jesse Roach
Mark Joseph	Cindy Liao	Delia Houseal
Richard Knight	Stephen Salerno	
J. Richard Landis	Jordan Affholter	**CMS/RTI**
Allen Nissenson	Casey Parrotte	Elizabeth Goldstein (CMS/CM)
Nicole Stankus	Karen Wisniewski	Julia Zucco (CMS/CM)
David White	Jennifer Sardone	Judy Lynch (RTI)
		Amy Hendershott (RTI)
		Celia Eicheldinger (RTI)
		Scott Scheffler (RTI)
		CDC
		Shunte Moon
		Christi Lines

Introductions
Jordan Affholter (UM-KECC) welcomed everyone to the third Pre-TEP conference call, and thanked them for joining the call.

Dr. Yi Li (UM-KECC) welcomed everyone to the Star Rating TEP conference call. Dr. Li thanked everyone for attending the previous call. Dr. Li reminded the group that the call was open to the public, and being recorded. Dr. Li administered an ordered roll call to take TEP member, CMS, CMS/RTI, and CDC attendance.

CMS's public reporting contractors NORC and Ketchum were also listening in on the call.

Dr. Li stated that the main agenda item would be re-baselining. Dr. Li explained that the baseline year was established due to the recommendations from the previous Star Rating TEP in 2015.

TEP Presentation and Discussion
Dr. Li explained that scoring with a baseline year allows for establishing the final score cutoffs for DFC Star Ratings that facilities can aim to achieve (in future reporting years). Dr. Li stated that the same performance in different years would result in the same star rating based on the fixed thresholds set in the baseline year. Dr. Li stated that improved performance over time means that facilities would not move down in their star rating.

Dr. Li explained that substantive changes to the measures used in DFC Star Ratings have consequences for interpretation. Dr. Li explained that if a measure is added to the star rating which was not previously

available, that does not allow for continuing to use the current baseline year. Additionally, baseline year cutoffs may not reflect quality of care measured with a revised measure set.

Dr. Li stated that by using cutoffs for scoring in a baseline year, improvements in individual measure values over time will always improve a facility's final score. One consequence is that over time it may be difficult to distinguish very good facilities from other facilities. Dr. Li presented a DFC Star Ratings distribution table to illustrate the point of shifting over time, which shows more facilities are receiving 4 and 5 stars over time.

Dr. Li stated that because of changes in measures or shifts over time, there are decisions that need to be made about re-baselining. He identified three potential triggers for re-baselining such as when measures are added or removed, when measures are updated, or when 50% of more of facilities receive either a 4 or 5 star rating. Dr. Li turned over the discussion to the TEP co-chairs.

One of the TEP co-chairs asked a clarifying question for UM-KECC: how often do we expect star rating measures to be added, removed, or updated? The co-chair further explained that if changes to star ratings are planned every year, then having a baseline is not as helpful because the cutoffs are always being reset. The co-chair asked for further clarification: is the discussion about whether a baseline is necessary or about how to re-baseline?

Dr. Andress (CMS) explained that the baseline year was addressed at the last TEP (in 2015). Dr. Andress explained CMS is not trying to determine whether a baseline is appropriate, but how to re-baseline. Dr. Andress stated that adding measures has implications for the baseline year. Dr. Andress explained they are seeking TEP input for when re-baselining is appropriate and how it should be implemented.

One of the TEP co-chairs asked, how often does CMS expect updates or additions to the Star Rating measures?

Dr. Andress (CMS) explained that measure updates and maintenance is an ongoing process. Dr. Andress explained that there are a lot of considerations. Dr. Andress explained this is one of the things on which CMS is seeking input, that is, when to add measures. Dr. Andress explained that measures considered for implementation will be announced ahead of time to give time for the public to comment, and for implementing the new changes.

One of the TEP co-chairs stated that it would be expected that if the measures are changed, it might change the factor analysis, which would also change the domains (groupings), so re-baselining might be necessary. The TEP chair asked UM-KECC how much the factors (for groupings) change when adding new measures or updated measures. Dr. Li explained that the groupings might change with the addition of new measures or removal of current measures, as the underlying correlation structure will likely change when the new measures are added or removed.

The TEP co-chair asked if a practical example could be provided for adding a new measure and asked how re-baselining would work.

The other TEP co-chair explained that the factor analysis would be re-run with the new measures. The domains would be identified and star ratings would be created based on those new domains. It would not be applicable to use the original baseline cutoffs because the scores would no longer be on the same scale, therefore, it would be necessary to re-baseline in that situation, based on the new set of measures. In other words, the same procedure would be used, but the methodology would be based on

an updated set of domain scores. If these measures are updated, it is possible that domains and cutoffs may still be applicable depending on the scope of the measure change.

Dr. Li confirmed that this was an accurate explanation.

The TEP co-chair went on to state that it is important that star ratings are understandable to patients. The star rating system should be able to show improvement over time, where improvements are being aimed for in quality care. The TEP co-chair stated that when there are clusters around 4 and 5 stars, then the next question is what is the next level of excellence. The TEP co-chair asked what would need to be done to make sure facilities are always aiming to improve. The TEP co-chair stated that they were thankful that CMS's public reporting contractor was on the call, because it will be important to communicate how the star ratings are displayed and explained once changes are made.

The other TEP co-chair stated if there is a lot of bunching at the top, then it would be necessary to either subdivide ratings into different categories or do a reset (of the baseline) to ensure differences can be shown in quality of care. The TEP co-chair stated that the point of the reset to a more normal distribution may be a problem for facilities and patients, in terms of explaining the change in ratings. For example, they further explained that if a reset happens, many facilities will go down in the star rating. The co-chair asked, from the patient perspective, if it would be helpful to have a regular reset of the baseline or have a reset when too many facilities have high star ratings.

One TEP member responded, saying that what is important to patients is that they are receiving good care.

Another TEP member commented on the considerations for baselining. The TEP member stated that showing the new and previous thresholds may be helpful.

The TEP co-chair stated that what may be meaningful is to also show scores at the individual measure level.

The TEP member stated that the factor score would not be intuitive.

The TEP co-chair stated that showing the actual individual measure values would be intuitive. It would be possible to show that the measure scores/values are increasing over time. However they emphasized that when adding new measures to the star rating, the star ratings would have to be reset to some extent, therefore, showing individual measure scores would still provide useful information.

One TEP member stated that patients need to understand that the numbers (Star Ratings and measures scores) mean something relative to the patients' own experiences. The TEP member stated concerns that the TEP was getting too involved in the process instead of focusing on building a system for patients where they can get the best care.

Another TEP member asked what level of care is good enough. The TEP co-chair asked, is there a cutoff for the measures where a certain score means excellence? The TEP co-chair asked UM-KECC or CMS to respond to the comment.

Dr. Andress (CMS) responded that in some cases a standard of excellence may not be clear. Dr. Andress provided an example and asked how much mortality is acceptable. Dr. Andress further explained that a standard of excellence is not available for certain measures. Dr. Andress provided an example for the SRR measure. Dr. Andress stated that one out of three dialysis patients discharged from the hospital are

readmitted within 30 days. Dr. Andress stated that some facilities have much less than one-third of patients who are readmitted, while other facilities report that much greater than one-third of patients are readmitted. Dr. Andress explained that it is possible to improve on current outcomes, but it may not be possible to get a natural clinical cutoff for the measures at this time.

Dr. Messana (UM-KECC) made a final explanatory point about re-baselining. He explained that re-baselining has a few considerations. One consideration is how to add new measures. Another consideration is when new measures are added, how to re-baseline or how do patients interpret the re-baselining. Dr. Messana stated that patients on the TEP have provided feedback that the current star rating is not ideal, and more patient-centric measures are needed. Dr. Messana pointed out that the downside of adding new measures is that it would affect the baseline calculation. Dr. Messana explained that in a scenario where all facilities score 5 stars, and all facilities are doing their very best, re-baselining would not be helpful. Dr. Messana then provided a second scenario where all facilities score 5 stars, but while some facilities are doing very well, others are not doing well. Dr. Messana explained that in the second scenario, failure to re-baseline would potentially give false information to patients. Dr. Messana asked the TEP if that is an important issue for patients.

One of the TEP members responded that it is important to patients. The TEP member stated that there are situations where the numbers have context. The TEP member stated that is why it is important for patient-centric measures to be included, so that patients understand the ratings include their perspective. The TEP member acknowledged that this is a difficult process to consider, and clarified that the goal of the TEP is to find ways to make the star ratings better. The TEP member stated that until the TEP is looking at more than physical aspects, the (ICH CAHPS) survey may not capture all of the patient experience. The TEP member highlighted issues important to patients such as whether facilities are addressing other patient concerns such as depression, anxiety, deterioration of cognitive skills, itching, and quality of life related to sex. The TEP member stated that there are tools to capture these items. The TEP member stated that as the TEP continues to develop the DFC Star Ratings, it is important to capture more of the items mentioned previously if they are going to be used in the star ratings.

The TEP co-chair stated that in terms of patient interest, it is not a matter of taking the measures and making them patient-driven at the cost of reporting clinical measures. Patients want to know that a facility can make an impact on clinical outcomes in a positive way. The patient-driven measures are relevant to patient experience. The TEP co-chair stated that patients need to be able to look at measures and understand the context for ease of use. As new measures are added, it will be important to explain how the process (of rating facilities) works so that is it is relevant to patients and to facilities for continuing improvement. The co-chair stated that, if 5-star facilities adopt certain practices, then patients might want to ask their facilities if they would adopt similar practices. The TEP co-chair stated the balance between qualitative and quantitative (measures) is important.

The TEP co-chair explained that the three teleconference calls have been used to establish a base understanding of the issues before the in-person meeting discussion. The TEP co-chair asked if there were other remaining questions from the previous call.

Dr. Li confirmed that the topics were covered well and agreed that the remainder of the call should be open for any TEP member questions.

One TEP member asked what year would be used for re-baselining if re-baselining was implemented.

Dr. Messana (UM-KECC) stated that it will be important to have data for all the measures for re-baselining. Dr. Messana stated that the measures would need to fit the CMS public reporting

requirements for potential inclusion in the star rating. Dr. Messana offered that one option would be for the baseline year to be the first year that a measure is publically reported. Dr. Messana asked the TEP member if they had any recommendations regarding the year to use for re-baselining.

The TEP member asked the methodology experts whether there is a standard for re-baselining. The TEP member acknowledged that having enough time to collect the baseline data would be necessary. The TEP member also asked if that topic is open to debate.

Dr. Messana reinforced that if TEP members have recommendations about how to best re-baseline they should provide them. The discussion should not be constrained.

The TEP co-chair stated that the most important thing from an analytical point of view is to have a measure that is validated, tested, and being reported. The co-chair explained in terms of a re-baselining standard, there is nothing about the star rating scoring methodology that requires a certain amount of time. The TEP co-chair explained that the scoring procedure would stay the same. The TEP co-chair stated that each measure to be used when re-baselining is a stable measure that is being reported. One of the TEP members thanked the co-chair for the explanation.

One TEP member asked a question of the patient TEP members on the call. If the re-baselining happens, then a large percentage of facilities will go down in stars. The TEP member asked the patient TEP members if that would have any effect on patients as consumers if they are already satisfied with their care.

One TEP member responded that they would discuss the star rating change with their facility administrator to understand the reason for the change.

Another TEP member responded that they would have analyzed their facility's star rating and review why it decreased. If there were issues in the center, then the TEP member stated they would ask about the changes. The TEP member stated that in that case it would be okay for the star rating to go down, as long as it was outside of facility control and that as a patient they were getting appropriate care.

One of the TEP co-chairs stated that the star ratings are for patients, and it is important for patients to be engaged. The TEP co-chair stated that patients want to be engaged, but they want information that is relevant and timely.

Elena Balovlenkov (CMS) stated that CMS has heard from patients that patients are interested in patient-reported outcomes. CMS is looking for ways to incorporate patient-reported outcomes, including convening a future TEP on patient-reported outcomes. CMS has heard that there is an expectation that a CMS-certified facility would provide good care. Part of CMS' responsibility is to educate patients and help define what good care is.

One TEP member asked a clarifying question about pediatric dialysis facilities. The TEP member stated that pediatric facilities have trouble meeting the minimum number of patient requirements for many of the measures. The TEP member asked how the pediatric units are scored in the DFC Star Ratings.

Dr. Messana (UM-KECC) answered that many pediatric units do not get scored in the DFC Star Ratings because of the small number of patients. In addition, some measures exclude pediatric patients. As an example, Dr. Messana stated that pediatric patients are not administered the ICH CAHPS survey at this time.

The TEP member also asked whether pediatric units that have a star rating are being compared against adult facilities or other pediatric facilities. The TEP member stated if the TEP is considering adding the Pediatric PD Kt/V measure, they had questions about how it would be used because it appears to be incomplete.

One TEP member thanked Dr. Andress for previously highlighting the variance in the SRR measure. The TEP member asked whether a patient could view that SRR metric on DFC and how they can view the measure.

Dr. Andress (CMS) instructed the TEP member on how to view the SRR measure on the DFC site.

One TEP member asked when would be the appropriate time at the in-person meeting to discuss the issue regarding the NHSN Bloodstream Infection data validity.

One of the TEP co-chairs stated they anticipate that will be covered on the in-person agenda. The TEP member highly recommended that the NHSN Bloodstream Infection validity issue be included on the agenda.

Dr. Li clarified that UM-KECC is still finalizing the agenda.

Jordan Affholter (UM-KECC) asked TEP members to complete the preliminary voting form by close of business on Wednesday February 15, 2017.

Public Comments
There were no public comments received during this TEP call.

Closing Remarks
Casey Parrotte (UM-KECC) stated that every TEP member should have received travel information for the in-person meeting. Casey Parrotte asked that if TEP members have not received that information to contact UM-KECC.

Dr. Li (UM-KECC) thanked the TEP members for attending the teleconference. Dr. Li stated that UM-KECC will summarize the meeting minutes and send them to the TEP. Dr. Li asked for TEP members to complete the voting form. Dr. Li stated that he looks forward to speaking with the TEP at the in-person meeting on February 21, 2017.

Appendix F: Star Rating Post-TEP Teleconference Call Minutes

End-Stage Renal Disease (ESRD) Quality Measure Development, Maintenance, and Support Project
ESRD Dialysis Facility Compare (DFC) Star Ratings
Star Ratings Post-TEP Teleconference Call #1 Minutes
March 22, 2017 1:00pm – 2:00pm (EST)

TEP Members	UM-KECC	CMS
Paul Conway	Yi Li	Joel Andress
Catherine Sugar	Joseph Messana	Elena Balovlenkov
Lorien Dalrymple	Claudia Dahlerus	Jesse Roach
Richard Knight	Rich Hirth	
J. Richard Landis	Ji Zhu	
Allen Nissenson	Cindy Liao	
Chris Sarfaty	Jordan Affholter	
Nicole Stankus	Casey Parrotte	
Sumi Sun	Jennifer Sardone	
David White		

Introductions
Jordan Affholter (UM-KECC) welcomed everyone to the Star Ratings Post-TEP conference call, and reminded the group that the call was open to the public and being recorded.

Dr. Yi Li (UM-KECC) welcomed everyone to the Star Ratings TEP conference call. Dr. Li stated that the TEP co-chairs would lead the discussion and that there will be a public comment period at the end of the call. Dr. Li administered an ordered roll call for TEP member and CMS attendance.

CMS's public reporting contractors NORC and Ketchum were also listening in on the call.

Dr. Li handed over the leadership of remainder of the call discussion to the TEP co-chairs, Catherine Sugar and Paul Conway, to lead the presentation and discussion.

TEP Presentation and Discussion
The TEP co-chair explained that there was no clear consensus on several re-baselining issues raised at the Star Rating TEP in-person meeting, though the TEP did recognize that re-baselining is necessary when updating measures and when new measures are added. The TEP co-chair stated that there were still questions about how often re-setting the star ratings distribution should occur and what cut-offs would be established for re-setting the star ratings. The TEP co-chair explained that there is a trade-off between maintaining continuity for measuring and tracking improvement over time versus the necessity of potentially re-setting the distribution when the star rating distribution becomes compressed. The TEP co-chair stated that facilities receiving higher star ratings over time is a positive outcome, reflecting improvements in absolute performance relative to the baseline year. However, if the star ratings become compressed or topped out, then it may be necessary to reset the star ratings in order to establish a new standard to meaningfully differentiate between facilities and encourage continuing improvement. The TEP co-chair asked the question of how these considerations can be balanced.

The TEP co-chair stated that in the in-person meeting there was discussion of re-setting to the 10-20-40-20-10 distribution (10% - 1 Star, 20% - 2 Stars, 40% - 3 Stars, 20% - 4 Stars, 10% - 5 Stars). During that

meeting, some TEP members expressed concern that constantly re-setting would not allow improvements to be measured. Additionally, this may be confusing for consumers and raise concerns about the quality of their facility, especially if their facility received a lower star rating after re-setting. During the in-person meeting, there was discussion that when updating or adding measures, the TEP generally preferred using the cut-offs from the current star ratings distribution instead of re-setting to the 10-20-40-20-10 distribution. When there are too many 4-star and 5-star facilities or few 1-star and 2-star facilities (termed compression of distribution), the issue is whether re-setting of the star ratings is necessary to encourage continued improvement.

The other TEP co-chair emphasized the importance of maintaining consistency of star rating information for patients so that they can use this to make decisions about their care, and explained that proper messaging for patients and stakeholders will be necessary . Proper messaging will prevent confusion and the loss of patient engagement in the star ratings. The TEP co-chair also stated that the star ratings should have some flexibility so that improvement can be made, while properly reflecting a facility's quality.

The TEP co-chair explained it is important to have a planned schedule and set of guidelines for re-setting the star ratings in order to ensure patients and facilities properly understand the changes. The star ratings re-setting should happen in a predictable, clear, and understandable way so that stakeholders can plan for the reset.

One TEP member asked CMS if they have programs other than DFC/DFC Star Ratings to measure (and report) continued facility improvement over time that is available to patients and broader renal community. Joel Andress, PhD (CMS) stated that, to his knowledge, there is no other mechanisms dedicated to showing facilities' longitudinal improvement. Dr. Andress further explained that the ESRD Quality Incentive Program (QIP) does show improvement, but only improvement from one year to the next year. Additionally facilities must post their ESRD QIP certificates in the facility. The certificate reports the overall score, but Dr. Andress stated that to his understanding the certificate does not show information by each measure.

Another TEP member asked the TEP co-chair if there are other methodological approaches that could be considered for re-setting the star ratings distribution, referencing options from the 2015 DFC Star Rating TEP. For example, at the 2015 TEP there was a presentation that displayed different options for re-setting the star ratings distribution.

The TEP member additionally asked for more clarity on the "Re-baselining is necessary when updating measures" bullet-point (Star Rating Post TEP Presentation slide 7), wanting to know if there were specific criteria to assess when to re-baseline when a measure has been updated. The TEP co-chair replied that minor updates to measures would not drastically change the scores (and domains from factor analysis) used to calculate the star rating, and thus re-baselining would not be necessary and using the previous distribution would be applicable. However, substantial measure updates would make the measure a new "variable" (i.e., new metric), and can no longer be treated as the same variable (metric).

The TEP co-chair further explained the distinction between re-baselining and re-setting. Re-baselining refers to re-scoring of measures when establishing a new baseline year, while re-setting refers to determining new cut-points for the entire star ratings distribution. The TEP co-chair explained that re-baselining is necessary whenever a substantive change is a made to a measure. With respect to re-setting, the TEP co-chair stated that there is flexibility in how to (re)set the star ratings distribution. The TEP co-chair explained that a compromise position would be to continue to show improvement

overtime while also increasing the standards. Therefore when measures are updated or added, measure scores are re-baselined, but the previous star ratings distribution (same cut-offs) can be kept in order to prevent a large shift (up or down) in the star ratings (no re-setting of distribution). However, if the star ratings distribution becomes compressed, then the potential need for re-setting of the star ratings could be reviewed on a fixed schedule of medium to long intervals. This compromised proposal (described on Star Rating Post TEP Presentation slide 8) is the starting point and scope of discussion for the call.

The other TEP co-chair verified the intent of the call. The TEP co-chairs explained that they are presenting a proposal to develop TEP consensus on a system that could work.

One TEP member stated that there appear to be two choices for this call. The TEP member asked for the definition of what would qualify as medium or long term intervals for evaluating if the star ratings distribution becomes compressed.

The TEP co-chair stated that definitions could be created for determining at what point the star ratings distribution has become too compressed to distinguish facilities (i.e., in terms of facilities' variation in their performance). The TEP co-chair asked what time interval would be ideal and reasonable in evaluating the star ratings. For example, both patient groups and providers will need time to prepare communications in order to explain any potential changes to patients and other consumers about changes to the star rating due to re-setting (i.e., develop messaging). The TEP co-chair stated that the time interval can be longer if the scheduled evaluation reveals that the star ratings are not compressed. The TEP co-chair further clarified that re-setting is to be evaluated only at these set intervals.

The TEP co-chair asked for TEP member input on the presented re-baselining (compromise) proposal (described above).

One TEP member stated that they agreed with the proposal.

Another TEP member stated they were in favor of the proposal as long as it was transparent to the public.

The TEP co-chair responded that the goal of the proposal is to be transparent and have a regular pattern.

Another TEP member agreed with the proposal and stated that they were in favor of the emphasis on the positive messaging about achievement of higher quality. The messaging helps the community understand any downward shift in stars is due to the quality standards being raised, because facility care has improved.

The TEP co-chair clarified that the re-setting of the star ratings distribution would happen because the field has achieved a new level of quality improvement. It sends the message that it is time to raise the bar. Therefore the planned re-setting is different from an unplanned re-set due to, for example, individual measure changes.

One TEP member stated that they are in favor of maintaining the star ratings distribution when measures are added or updated in the star rating system; and they would favor evaluating re-setting the star ratings distribution at some interval, but that would also allow the star ratings to continue without being reset if compression is not observed. The TEP member asked if there were other issues for the TEP to provide feedback during the call. The TEP co-chair confirmed that those were the two main issues, noting they also wanted TEP input on the related question of (1) how often would be ideal for evaluating

re-setting the star ratings distribution and (2) what star ratings distribution should be used when re-setting.

There was further discussion about the minimum time interval for evaluating a potential re-set. The TEP co-chair also noted that if the star ratings were not compressed (i.e., not topped-out) that would then allow for a longer period of time to elapse without re-setting the distribution. In terms of planning, the TEP co-chair asked how much time CMS, UM-KECC, and facilities, and patient groups would need to prepare for a reset. The TEP co-chair again emphasized the importance of messaging so that everyone is aware of the reset.

The other TEP co-chair stated that it would likely take 9 to 12 months (with some flexibility) to inform the community of the star rating reset, in order to make sure that message is very clear across multiple communication platforms with stakeholders. The TEP co-chair further stated that it is key to have a general expectation of the timeline, while allowing for some flexibility.

One TEP member stated that aside from whether re-baselining happens in a given year, it would be helpful to have information communicated about where improvement was and was not observed (e.g., on specific measures). This information would help organizations prepare for a future re-baselining. This would require having a standard approach for measuring improvement that is accepted by the community.

The TEP co-chair agreed this is a good point and such an approach can help determine if the star ratings are approaching a reset.

With respect to messaging to the community, Elena Balovlenkov, RN (CMS) stated that CMS hosts national calls about the DFC website twice a year. The focus of the calls is to announce changes and improvements to the DFC website. She stated that additional calls could be scheduled as well that could communicate changes (e.g., re-baselining or re-setting).

The TEP co-chair then asked if the 9 to 12 month time period for re-setting the star ratings was feasible. Dr. Andress (CMS) stated that CMS would need to look at their operational timeline as there are several timeline considerations such as messaging, analyses, and making a final decision. Because of this, Dr. Andress explained that CMS is not able to provide a decision about the length of time needed during this call. He also emphasized that providers and patients would be asked for input on the timeline for communicating a star rating reset to the public.

The TEP co-chair clarified there are multiple timelines (both sequential and concurrent): The first timeline ensures the methodology for re-setting the distribution is fully tested; after that there is a timeline for industry to work through their processes, as well as a timeline for patient groups and stakeholder organizations to understand and communicate information on a re-setting of the star ratings; this helps ensure time is taken for appropriately communicating the information to all stakeholders.

Dr. Andress further stated that CMS will need to discuss the operational timeline with UM-KECC before being able to develop a finalized timeline that includes considerations for the implementation of a re-setting of the star ratings (e.g., messaging to the community, analyses, making a final decision, final implementation).

The TEP co-chair clarified that they did not expect to obtain the final timeline on this call. For example, as part of the evaluation period, UM-KECC would need to evaluate data at specific time intervals to

determine if the star ratings are becoming compressed and if so when it would reach that point. The TEP co-chair stated that it will be important for the community to have an estimation of the time it takes from determining a re-set is necessary to rolling out and implementing a re-setting of the star ratings distribution. The TEP co-chair stated that if it takes a long period of time to implement the distribution reset, then it may be important to consider a different starting time point, because star rating compression may continue while the process for implementing the re-setting is completed.

One TEP member asked Dr. Andress (CMS) to clarify that 1) they would not re-set the star ratings distribution more often than once every certain number of years and 2) that when they evaluate at the fixed time period that CMS would not automatically reset the star ratings distribution. The TEP member acknowledged compression would be a trigger for re-setting the star ratings and asked how compression would be defined, for example if there is a set amount of star rating shift that must occur in order for the star ratings to be reset.

In response to these questions, the TEP co-chair reiterated the difference between the terms re-baselining and re-setting. Re-baselining refers to re-scoring of measures when there are measure additions or updates; re-setting refers to changing the entire star ratings distribution (establishing new cut-offs). The TEP co-chair stated that the TEP appears to be in favor of re-baselining when measure change or are added. The current question is around re-setting the star ratings distribution (i.e., the evaluation of re-setting at fixed time intervals). The TEP member thanked the TEP co-chair for clarifying.

Dr. Andress (CMS) stated that CMS does not have a specific definition of star rating compression at this point but that different criteria could be investigated (using the example of the "topped out" definition used for the QIP). Dr. Andress also noted at the in-person meeting at the Star Rating TEP, there was discussion on whether compression could be defined as the number of 4 or 5 star facilities or the disappearance of 1 and 2 star facilities (all facilities in the 3, 4, or 5-star categories only). Dr. Andress stated that CMS would be open to input from the TEP on the topic of compression and that analyses will be needed to evaluate different methods for determining compression. For example, Dr. Andress stated that compression may be determined if the star rating system is no longer providing information on improvement and if it is no longer distinguishing facilities for patients. Compression is expected to reflect a change in level of quality of care and therefore the star ratings need to be reset because the community has achieved the current level of care quality, and in response CMS wants to continue pushing the standard further to achieve even higher quality care.

The TEP co-chair explained that there are various considerations surrounding the definition of compression such as statistical issues, operational distinctions, and patient interpretation. The TEP co-chair stated that it was not feasible to arrive at consensus of a definition of compression during this TEP call. The TEP co-chair asked for additional comments on this and other topics discussed during the call.

One TEP member asked how often other programs reset their star ratings, such as Hospital Compare. The TEP member also asked if the re-baselining proposal has been tested.

The TEP co-chair stated that the first question is most appropriate for CMS and UM-KECC. In response to the latter question, the TEP co-chair stated that the underlying scoring system of the star ratings has already been tested (i.e., during development and implementation of the current system). The TEP co-chair clarified that re-setting is about re-setting the cut-points which results in a particular distribution, and is not as much a methodological change.

The TEP member asked how adding new or updated measures impacts the star ratings distribution itself.

The TEP co-chair responded that an example was provided for when new measures are added (presented on slides during the second teleconference call). For example, when measures are added, the measure domains from factor analysis may change with a different set of measures. The TEP co-chair stated that the star ratings distribution would likely remain stable unless facilities had particularly poor or strong performance in the new measure(s).

In response to the question how often other CMS programs re-baseline/re-set the star ratings, Dr. Andress (CMS) stated that the original plan for Nursing Home Compare was to conduct a re-baselining/re-setting of the star ratings at set time intervals but that CMS did not reset the star ratings distribution for several years. It was the eventual compression of the Nursing Home Compare Star Ratings that triggered the reset of the distribution. The program experienced a rapid upward shift in the star ratings distribution and there was concern this was not related to quality improvement but related to measure changes and data submission. Nursing Home Compare is now tracking the star rating in order to assess how to re-set the star ratings distribution in the future, but Dr. Andress stated he is not aware if they have a regular schedule for re-setting the star ratings.

Jordan Affholter (UM-KECC) reminded the TEP members that the Star Ratings TEP Summary Report was sent to the TEP for review, noting UM-KECC carefully drafted the report based on the Star Ratings TEP In-person meeting discussion. Jordan Affholter stated that UM-KECC encourages the TEP members to review the TEP summary report and ensure that it reflects an accurate representation of the discussion. TEP members were asked to provide feedback on the TEP Summary Report by close of business on Wednesday April 5, 2017.

The conference call reached the time for public comments.

Public Comments

One public comment was provided by Kathy Lester, JD, on behalf of Kidney Care Partners (KCP).

> "Hi, this is Kathy Lester. I'm with Kidney Care Partners. First, I want to thank you for the open dialogue, the chance to listen, and to provide comments. I think as you know- KCP has been part of this process from day one, so we're very pleased to be able to continue to participate and you know excited to see that it appears there is a movement away from returning to the bell curve, the 10-20-40-20-10 system. And assume that when you are talking now about the term re-setting, that you are not returning to that forced distribution either. So would encourage you to you know stay far away from that. Also, we wanted to let you know we are working with technical experts to discern how to understand the impact of these models and are testing scientific reliability and validity of different ways to do this, because we agree that it's really important to know that reliability and validity of what's going on if you are going to make sure that consumers can understand the changes that are going to take place. We also agree that before anything is decided on we really have to understand the definitions of compression, and what we're talking about in terms of frequency. So please there will be other opportunities to comment and provide information on and that will also be part of our analysis of the proposal. You know bottom line for us is that we want to get this process right. We have been strong supporters of quality whether it's in the QIP or trying to get the five star program right. But it's important that these technical questions get answered. So again, we appreciate the opportunity for the comment. We're happy to share our ongoing work in hopes that the TEP and CMS and UM-KECC will consider these comments and ensure that the program is the best that they can be. So thank you."

Closing Remarks

Dr. Li concluded the call by thanking the TEP members for the discussion. Dr. Li asked the TEP members to let UM-KECC know if they had any further questions or comments, and encouraged TEP members to review the Star Ratings TEP Summary Report.

Appendix G: TEP Preliminary Voting Form for Candidate Measures

ESRD DFC Star Rating TEP: Preliminary Voting Form for New Candidate and Updated Measures for DFC Star Rating Release in October 2018

Name _____Date_____

New Measures: Please rate each new candidate measure for addition to the ESRD DFC Star Rating System. Select only one rating for each measure using the five-point scale. Please use the free-text box to provide reasons for your rating.

1. Patient experience of care measure, In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH-CAHPS) measure should be added to the DFC Star Ratings.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments:*

2. The National Healthcare Safety Network (NHSN) Bloodstream Infection measure should be added to the DFC Star Ratings.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments:*

3. The Standardized Readmission Ratio (SRR) measure should be added to the DFC Star Ratings. This measure reports how often the facility's patients are readmitted to the hospital within 4-30 days.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments:*

4. The Pediatric Peritoneal Dialysis Adequacy: Achievement of Target Kt/V (Pediatric PD Kt/V) measure should be added to the DFC Star Ratings. This measure reports how many of the facility's pediatric PD patients achieve dialysis adequacy.

1	2	3	4	5

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments:*

Updated Measures: Please rate each updated measure to replace the current version of that measure currently used in the ESRD DFC Star Rating System. Select only one rating for each measure using the five-point scale. Please use the free-text box to provide reasons for your rating.

5. The updated Vascular Access measure, Standardized Fistula Rate, should replace the current Vascular Access Fistula measure used in the DFC Star Ratings. This measure reports how many patients have a functioning fistula for their hemodialysis treatment, while taking into account clinical and patient risk factors.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments:*

6. The updated Vascular Access measure, Long-term Catheter Rate, should replace the current Catheter measure used in the DFC Star Ratings. This measure reports how many patients at the facility have a catheter in place for 90 days or more.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments:*

7. The updated Standardized Mortality Ratio (SMR) measure should replace the current SMR measure used in the DFC Star Ratings.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments*

8. The updated Standardized Hospitalization Ratio (SHR) measure should replace the current SHR used in the DFC Star Ratings.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments*

9. The updated Standardized Transfusion Ratio (STrR) measure should replace the current STrR used in the DFC Star Ratings.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

-*Comments*

Appendix H: TEP Preliminary Voting Results and TEP Comments for Candidate Measures

Table H1: Preliminary Measure Voting Summary Results (%)

Measure	% Strongly Agree/Agree	% Neither Agree nor Disagree	%Strongly Disagree/Disagree
Standardized Fistula Rate	92%	8%	0%
Long-term Catheter Rate	92%	8%	0%
Standardized Mortality Ratio (SMR)	84%	8%	8%
Standardized Hospitalization Ratio (SHR)	92%	0%	8%
Standardized Transfusion Ratio (STrR)	75%	8%	17%
Pediatric Peritoneal Dialysis Adequacy: Achievement of Target Kt/V (Pediatric PD Kt/V)	58%	42%	0%
The National Healthcare Safety Network Bloodstream Infection measure (NHSN SIR)	50%	8%	42%
In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)	70%	15%	15%
Standardized Readmission Ratio (SRR)	50%	17%	33%

Table H2: Preliminary Measure Voting Summary Results by Rating Category (N)

Measure	# Strongly Agree Votes	# Agree Votes	# Neither Agree nor Disagree Votes	# Disagree Votes	# Strongly Disagree Votes
Standardized Fistula Rate	8	3	1	0	0
Long-term Catheter Rate	8	3	1	0	0
Standardized Mortality Ratio (SMR)	3	7	1	1	0
Standardized Hospitalization Ratio (SHR)	3	8	0	1	0
Standardized Transfusion Ratio (STrR)	5	4	1	2	0
Pediatric Peritoneal Dialysis Adequacy: Achievement of Target Kt/V (Pediatric PD Kt/V)	4	3	5	0	0
The National Healthcare Safety Network Bloodstream Infection measure (NHSN SIR)	6	0	1	3	2
In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)	4	5	2	0	2
Standardized Readmission Ratio (SRR)	4	2	2	3	1

Preliminary Candidate Measure Voting Results: TEP Comments *This vote took place from February 13-16, 2017 (before the in-person TEP meeting).*

The following comments were provided by TEP members about each candidate measure on the preliminary voting form. The comments have been de-identified to maintain TEP member anonymity.

Table H3: In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) Measure: TEP Member Comments

Patient satisfaction w/care is key factor
I believe it would add value to the star ratings to have these initial measures of patient experience included. However, as others have expressed on the call, I think additional consideration should be made how these are presented to the public (both patients & staff). Currently, the website only shows top box.
I feel it is critical to not combine the ICH-CAHPS into an overall star rating score. My concern regarding adding this involves the patients' survey fatigue from ongoing twice yearly surveys. CMS should re-evaluate this requirement and consider alternative approach or approaches to gathering this valuable information. Perhaps once annual testing and 90 day surveys for new starts would be more beneficial to having a more robust survey completion. Additionally, Pediatric units are not adequately assessed by ICH CAHPS. There are different surveys, such as a Press Gainey or similar survey which may provide more benefit for a Pediatric Unit comparison and rating.
The addition of this measure will serve to better integrate – from a consumer standpoint – one of several measures that seek to capture patient experience. Although different from the Star ratings in terms of intent – it would take a familiar tool to patients and embed it into a comparative model that patients are being encouraged to utilize (5 Star)
[sic] 30 patients maybe a good measure for smaller facilities, but I feel that a percent participation should be the criterion. I would ask our statistical colleagues to opine what rate of participation in surveys would make the data reliable. We also know that patients who do complete the survey may be very different from the ones who do not. What could we do to minimize this bias? Questionnaire also does not include questions about APNs and PAs who provide significant part of care at many dialysis units. Patients in fact may be receiving information about their treatments and getting care from these practitioners, but they may not necessarily associate them with physician, who in fact may be rounding at the facility separately and maybe also less often than physician extenders.
The ICH-CAHPS survey, while inadequate, is the most patient oriented information available that reflects how patients perceive their quality of care, facilities, staff, care by nephrologists and other factors affecting patients.
I feel PROs are critically important. However, I have concerns about the large percent of facilities that are ineligible or do not complete at least 30 ICH CAHPS surveys, the unknown statistical precision when < 200 surveys are completed per 12 months, and I would like more information as to whether the distribution of performance will allow for meaningful differences to be detected within a 5 Star paradigm. In addition, there continues to be

concern about the survey burden of ICH CAHPS. If incorporated into 5 Star Ratings, I would favor separate reporting from clinical measures given the large number of clinics without a sufficient number of surveys, the correlation findings, and TEP teleconference discussion.
I do not think the ICH-CAHPS captures all the important Patient Experience information; I look forward to the third TEP where this is looked at specifically, and closely. And, I look forward to reading the white paper you have produced on patient experience of care measures.
I feel that this particular rating will be used by the patient/consumer more so that some of the other questions. Patients/consumers rely on what others have to say about a facility/product.
I am still interested in hearing further discussion of the measure and in particular the extent to which it gets at the concepts of most concern to patients but it seems very clear there is strong sentiment for including a patient perception component to the ratings and this is the best candidate currently available for that purpose.
This is clearly an important data set, but very complex since it includes evaluation not only of the dialysis facility but also the physicians at the facility. Making this data public would be reasonable as something to include in DFC but because of the complex shared accountability I would be opposed to using this in 5-Star. Finally, the fact that this is done every 6 months and results are not available in real time, the fact that the % of patients who complete the form varies by facility and is a significant minority of all patients, along with the focus on physicians as well as the facility, makes this a measure set for which action to improve scores is difficult and validity is not established.
This measure should be reported out to consumers separately.

Table H4: National Healthcare Safety Network (NHSN) Bloodstream Infection Measure: TEP Member Comments

Critical safety signal
While I agree with the importance of measuring infections, I would like to better understand the distribution/validity of this measure.
Although I understand the CDC claims that the underreporting of infections was isolated to a single sample that was not validated when re-checked, I continue to have significant concerns about the data. It is my opinion that the distribution charts of NHSN Dialysis Center infection rate data demonstrate the underreporting of infection with an unrealistic majority of units with a 0 (zero) rate and an almost separate distribution curve in the same graph. Until we can be certain as to the accuracy of data, I cannot support this measure. There are only a few co- morbidity measures that apply to Pediatric Dialysis patients, so comparison for Pediatric patients to adult rates is not equitable. There are non-measurable risks (e.g. pulling at lines) that can effect pediatric patients that cannot be equitably measured comparing to an adult population.
The issue of bloodstream infections is critical – and the work across Federal agencies to educate and protect patients should be cross-wired into the 5 Star measure as a means of reinforcing this issue - and the connection to health outcomes – among both patients and providers.

I was very interested in the part of discussion why there was such a dichotomous distribution between facilities in the reporting of this measure. BSI reporting is very dependent on the data that dialysis facilities obtain from the outside institutions. The accuracy of such reporting likely would explain why data distribution we see is so dichotomous. As far as the measure is concerned, counting a second positive blood culture as a new episode more than 21 days after the first may not be always accurate. Often dialysis facilities provide long-term antibiotic therapy for chronic infections which lasts longer than 21 days, and repeat blood cultures are done 7 days after this therapy is complete to make sure that infection has indeed been cleared. Which current methodology if the culture would show that infection with the same organism is still present it would be counted as a new infection— which it is not, rather it is just a failure of antibiotic therapy. CDC assured that previously quoted CDC data showing inaccuracy of such measure may indeed not be a reliable data, and that BSI capture perhaps is better than shown in those studies. However, I would suggest running a pilot program to ascertain accuracy of the data provided by the facilities before this measure would be added to the Star Ratings.
Critical measure that is of great importance in making a decision as to which facility is best suited for a patient.
Further discussion and consideration of underreporting and validity would be helpful. I am concerned by the large number of zeros shown during the TEP pre-conference calls.
Important measure; not sure how to rate its importance in relation to other measures, without hearing other TEP and TEP team opinions. I know it was discussed some on the call, but I don't feel it gave me as strong a grasp as I would like to have to be able to state an opinion, so I abstain from casting a vote.
Infection control should be a major topic for rating. The patient/consumer can compare facilities so that a more informed decision can be made.
This seems like a reasonable clinical measure to me but given the mixture of opinions I have heard expressed I think it merits further discussion. An issue which wasn't discussed much but about which I would be interested in hearing more is whether it is seen as being at an equal level of performance and an equal degree of specificity as the other clinical measures, several of which seem broader in [scope] to me. I worry a little methodologically about mixing in a lot of detailed measures with a few general measures as even with the factor analysis approach the specialty measures can end up getting a lot of weight if they split off into their own domains. I am not enough of a clinical expert to judge where this measure fits on that list.
This measure has been shown to rely on data for which validation studies have shown there are huge gaps. CMS itself has acknowledged this, and small studies have shown this as well. The response of CDC that they don't believe CMS or the studies and "trust us this is valid" should not be accepted. Because of the lack of validity using this measure is more likely to mislead patients than inform them about facility quality.
No

Table H5: Standardized Readmission Ratio (SRR) Measure: TEP Member Comments

Strong motivator for provider to weigh clinical care issues

I have concerns about the denominator for the ratio calculation. It still appears that there may be influence on the ratio which has a component of random chance. Until there is more firm data, I cannot support this measure. If this were reported as a rate, then I would be more likely to agree with its addition. Still if pediatrics is included, it needs to encompass co-morbidities that are applicable to those children. Given the underenrollment into Medicare, this population will be undersampled.
This measure will be – in my opinion and if properly contexed through the DFC website – of great importance not just to patients but also and especially among care-givers. Caregivers are intensely interested in the practical issues of insurance, medical specialist coordination – and the elephant in the room is often this "how far away is the next crisis or next incidence?" This tool – perhaps not perfect – give you a clear sense of what can be expected.
I am concerned that this measure may not be reflecting the care provided by the dialysis facilities and may in fact not be something that facilities would be able to impact. These are the readmission diagnoses/conditions that Medicare is targeting: 30days readmission for AMI, stroke, COPD exacerbation, pneumonia and after CABG or total hip replacement. I just do not see how dialysis facility could help minimize those readmissions. Only one diagnosis, readmission for heart failure, is likely to be impacted by improved communications between Hospital and dialysis center. I am not sure how the hospital-wide unexpected readmission rate would be applied to the dialysis facility. Would inner city facilities located close to safety net hospitals (which many times have worse outcomes as patients use them instead of regular outpatient care for chronic conditions) be inadvertently impacted? More discussion should follow. Proposed measure of combining SRR and SHR has drawbacks. Facility which sends patients to the hospital ER for non-functioning catheter may not incur any hospitalization penalty, since the patient will likely be "observation" status (if admission is shorter than 2 midnights). This is not a good practice, but it would not be captured by either SRR or SHR measures. Perhaps a measure like ER visits would be something to consider. Incorrect evaluation of patient at the dialysis center can trigger ambulance transport for issues that can be managed either at the facility, or as outpatient, without calling an ambulance. That is not only costly, but also negatively impacts the patient.
Important measure because it reflects a standard of care a patient receives in the hospital and after the hospital visit. The primary factor is that the focus is on the care delivered to a patient from the team/health care delivery system. Instead of one entity (hospital) having a dispute with the dialysis facility - the patient's welfare is the focus.
The SRR may be influenced by a number of care processes, healthcare settings, geography, and provider practices that are not a reflection of the dialysis facility care. If the intent is to communicate information about care that is more clearly attributable to the facility then I would favor other measures. In addition, SHR is already in the 5 Star ratings (which includes hospitalizations that are readmissions).
Important measure; not sure how to rate its importance in relation to other measures, without hearing other TEP and TEP team opinions. I know it was discussed some on the call, but I don't feel it gave me as strong a grasp as I would like to have to be able to state an opinion, so I abstain from casting a vote.
DRG's and insurance companies often dictate how long a patient's hospital stay will be paid for. This can lead to readmission to the hospital because the patient was not ready to be discharged. Also, from a clinical standpoint, hospital communication with the dialysis facility

is much to be desired. I know that I have to either call many times for discharge information or fax and re-fax the hospital for the needed information.
Same basic comments as to the blood infection measure though this measure seems broader in scope.
I am aware that the purpose of this TEP is not to relitigate the validity of measures. However, there are considerable concerns about this measure that have been repeatedly raised and have not been addressed including the poor reliability for small facilities. If the task is to judge the metric as it exists, it is not sufficiently reliable to be included in 5-Star.
No

Table H6: Pediatric Peritoneal Dialysis Adequacy: Achievement of Target Kt/V (Pediatric PD Kt/V) Measure: TEP Member Comments

Although it is not clear at this time how relevant this measure will actually be (it is unlikely many Pediatric units will meet criteria to receive a DFC star rating), I feel that it is a tangible measure of a Pediatric Dialysis Units quality of care and perhaps another step toward rating Pediatric dialysis units.
I can go either way on this one – given the overall number of facilities, patient populations and the mixed population pool of many of these facilities – it may make sense to have these strata of facility depicted separately – thus the measure would not have to be included in this round.
None
We must begin to capture this information on the population under 18 years old.
I do favor quality measure reporting for pediatric measures, but in a way that allows one to discern the performance on pediatric measures. I favor further discussion as it is unclear to me how many facilities would have a pediatric PD measure and I think the Star Ratings would still primarily reflect care of adults. It would be helpful to walk through implementation and inference as a group.
Important measure; not sure how to rate its importance in relation to other measures, without hearing other TEP and TEP team opinions. I know it was discussed some on the call, but I don't feel it gave me as strong a grasp as I would like to have to be able to state an opinion, so I abstain from casting a vote.
I do not deal with pediatric patients so I do not have any input on this question.
This sounds like an important measure. My question, which I didn't get a chance to come back to on the calls, is whether enough of the facilities have a sufficient pediatric population that we will be able to get stable ratings. If this will end up being missing for a lot of facilities I would consider that problematic but otherwise would favor including it. During the calls it was clear that there are facilities that are much more pediatric centered and that they on average are less likely to have complete data for some of the current metrics. Has there been any consideration of separating those facilities off and rating them on more pediatric-centered metrics? I am not sure that there are enough of them for this to be worthwhile or that even separated in that way they would be large enough to get good ratings data--this just occurred to me based on the recent discussions.
No

Table H7: Standardized Fistula Rate Measure: TEP Member Comments

I generally agree with the intention for more robust capture of adjusters, however, I am concerned about the tradeoff – impact to centers that have fewer Medicare patients. (Has this analysis been presented?)
I appreciate how this is a true rate based measure (taking into account some comorbidities and risk factors) that allows unit to unit comparison.
Favor the update
other severe comorbidities, such as advanced lung disease or advanced heart failure, advanced dementia may also need to be added to limited life expectance exceptions
An improvement in reporting on vascular accesses. I believe that this measure will allow patients and caregivers to better understand how much emphasis [is] placed on maintaining functioning fistulas. Catheters and grafts are not the best option for patients and will cause increased hospital visits. Facilities are motivated to encourage AV fistulas which make the dialysis process much safer for the patient.
I think it is important to incorporate updated measures when the measure specifications have been refined and re- specified after initial use and implementation and this seems like the expected course of healthcare quality measure development.
Important measure; not sure how to rate its importance in relation to other measures, without hearing other TEP and TEP team opinions. I know it was discussed some on the call, but I don't feel it gave me as strong a grasp as I would like to have to be able to state an opinion, so I abstain from casting a vote.
The Fistula First initiative was a great concept, but not everyone can have a well-developed fistula. Unfortunately, I have taken care of many patients that have lost digits and hands related to fistulas that have impeded circulation to the distal upper extremity. This particular question is of great importance to me.
The updates to all of the measures sound technically strong to me so I would favor including them. My only concerns lie around the implications for re-baselining. It will not make sense to "re-baseline" every year or it will have no meaning, so I think there will need to be a general rule about only making updates every so often, either of new measures or updates to old measures. It is possible (and this is something I wouldn't mind discussing further with the UM-KECC team) that a minor measure update would in fact not affect the factor analysis groupings of the measures and would have sufficiently little impact (after rescaling) on the actual standardized scores that existing cutoffs could be retained. If that were true it would be very helpful.
There are some improvements in definitions that would strengthen this metric, but I would still support the current version if changes are not possible.
No

Table H8: Long-term Catheter Rate Measure: TEP Member Comments

Key safety signal
I generally agree with the intention for more robust capture of adjusters, however, I am concerned about the tradeoff – impact to centers that have fewer Medicare patients. (Has this analysis been presented?)

I appreciate how this is a true rate based measure (taking into account some comorbidities and risk factors) that allows unit to unit comparison
Favor the update
other severe comorbidities, such as advanced lung disease or advanced heart failure, advanced dementia may also need to be added to limited life expectance exceptions
An improvement in reporting on vascular accesses. I believe that this measure will allow patients and caregivers to better understand how much emphasis [is] placed on maintaining functioning fistulas. Catheters and grafts are not the best option for patients and will cause increased hospital visits. Facilities are motivated to encourage AV fistulas which make the dialysis process much safer for the patient.
I think it is important to incorporate updated measures when the measure specifications have been refined and re- specified after initial use and implementation and this seems like the expected course of healthcare quality measure development.
Important measure; not sure how to rate its importance in relation to other measures, without hearing other TEP and TEP team opinions. I know it was discussed some on the call, but I don't feel it gave me as strong a grasp as I would like to have to be able to state an opinion, so I abstain from casting a vote.
As I said, not everyone can have a functioning fistula. Some patients have no choice but to have a long-term catheter.
The updates to all of the measures sound technically strong to me so I would favor including them. My only concerns lie around the implications for re-baselining. It will not make sense to "re-baseline" every year or it will have no meaning, so I think there will need to be a general rule about only making updates every so often, either of new measures or updates to old measures. It is possible (and this is something I wouldn't mind discussing further with the UM-KECC team) that a minor measure update would in fact not affect the factor analysis groupings of the measures and would have sufficiently little impact (after rescaling) on the actual standardized scores that existing cutoffs could be retained. If that were true it would be very helpful.
No

Table H9: Standardized Mortality Ratio (SMR) Measure: TEP Member Comments

I generally agree with the intention for more robust capture of adjusters, however, I am concerned about the tradeoff – impact to centers that have fewer Medicare patients. (Has this analysis been presented?)
I feel similarly to this as I do to the SRR and the SHR. I have concerns about the denominator for the ratio calculation. It still appears that there may be influence on the ratio which has a component of random chance. Until there is more firm data, I cannot support this measure. If this were reported as a rate, then I would be more likely to agree with its addition.
Favor the update
Although I generally favor using updated measures, I would like to further discuss SMR IUR for small facilities and implications when revised SMR used for 5 Star ratings – it would also be helpful to further discuss the impact on discerning 1 through 5 Stars. Do we have any information on how the updated measure as compared to prior measure influences classification of facilities in the 5 Star Ratings?

Important measure; not sure how to rate its importance in relation to other measures, without hearing other TEP and TEP team opinions. I know it was discussed some on the call, but I don't feel it gave me as strong a grasp as I would like to have to be able to state an opinion, so I abstain from casting a vote.
The updates to all of the measures sound technically strong to me so I would favor including them. My only concerns lie around the implications for re-baselining. It will not make sense to "re-baseline" every year or it will have no meaning, so I think there will need to be a general rule about only making updates every so often, either of new measures or updates to old measures. It is possible (and this is something I wouldn't mind discussing further with the UM-KECC team) that a minor measure update would in fact not affect the factor analysis groupings of the measures and would have sufficiently little impact (after rescaling) on the actual standardized scores that existing cutoffs could be retained. If that were true it would be very helpful.
The use of claims data as well as 2728 data to establish incident and prevalent co-morbidities is a step forward. I do have concern, however about the lack of reliability for smaller facilities which does not allow meaningful conclusions about mortality in these facilities. In fact only 55-70% of a facilities score is due to differences in performance for small and medium size facilities.
No

Table H10: Standardized Hospitalization Ratio (SHR) Measure: TEP Member Comments

I generally agree with the intention for more robust capture of adjusters, however, I am concerned about the tradeoff – impact to centers that have fewer Medicare patients. (Has this analysis been presented?)
I feel similarly to this as I do to the SRR and the [SMR]. I have concerns about the denominator for the ratio calculation. It still appears that there may be influence on the ratio which has a component of random chance. Until there is more firm data, I cannot support this measure. If this were reported as a rate, then I would be more likely to agree with its addition.
Favor the update
Unless there are specific reasons an updated measure should not be used in 5 Star ratings – methodologic implications, etc., then I generally favor the incorporation of updated measures. Similar to above, do we have any information on how the updated measure as compared to prior measure influences classification of facilities in the 5 Star Ratings?
Important measure; not sure how to rate its importance in relation to other measures, without hearing other TEP and TEP team opinions. I know it was discussed some on the call, but I don't feel it gave me as strong a grasp as I would like to have to be able to state an opinion, so I abstain from casting a vote.
The updates to all of the measures sound technically strong to me so I would favor including them. My only concerns lie around the implications for re-baselining. It will not make sense to "re-baseline" every year or it will have no meaning, so I think there will need to be a general rule about only making updates every so often, either of new measures or updates to old measures. It is possible (and this is something I wouldn't mind discussing further with the UM-KECC team) that a minor measure update would in fact not affect the factor analysis

groupings of the measures and would have sufficiently little impact (after rescaling) on the actual standardized scores that existing cutoffs could be retained. If that were true it would be very helpful.
The use of claims data as well as 2728 data to establish incident and prevalent co-morbidities is a step forward. I do have concern, however about the lack of reliability for smaller facilities which does not allow meaningful conclusions about mortality in these facilities. In fact only 57% of a medium sized facility's score and only 46% of a small facility's score relates to actual performance.
No

Table H11: Standardized Transfusion Ratio (STrR) Measure: TEP Member Comments

Difficult for facilities to influence or know when patients receive a transfusion.
If this were reported as a rate, then I would be more likely to agree with its addition. I have similar issues since it is reported as a ratio. Furthermore, most outpatient units do not give transfusions, so tracking exactly who gets a transfusion may be problematic since the majority are handled inpatient (unless there is a guaranteed way to capture all of those).
Favor the update – for context for site users – may want to explain whether or not this measure can be interpreted as related in any way to health outcomes and viability for future [eligibility] for transplant.
I note this is not a question of whether STrR should be in the 5 Star ratings but whether the updated specifications should be used. The revised measure attempts to address concerns about original specifications and from my perspective reflects efforts to continuously evaluate and re-specify quality measures when needed.
Important measure; not sure how to rate its importance in relation to other measures, without hearing other TEP and TEP team opinions. I know it was discussed some on the call, but I don't feel it gave me as strong a grasp as I would like to have to be able to state an opinion, so I abstain from casting a vote.
I am seeing a slight increase in transfusions. One reason is that there are more patients on dialysis with cancer. If they are taking cancer treatments whether it be chemotherapy, radiation, or both, they will eventually need a transfusion. This is related to the cancer itself as well as the chemotherapy and radiation.
The updates to all of the measures sound technically strong to me so I would favor including them. My only concerns lie around the implications for re-baselining. It will not make sense to "re-baseline" every year or it will have no meaning, so I think there will need to be a general rule about only making updates every so often, either of new measures or updates to old measures. It is possible (and this is something I wouldn't mind discussing further with the UM-KECC team) that a minor measure update would in fact not affect the factor analysis groupings of the measures and would have sufficiently little impact (after rescaling) on the actual standardized scores that existing cutoffs could be retained. If that were true it would be very helpful.
No

Appendix I: In-Person TEP Voting Form for SRR and SIR Measures, and ICH CAHPS Scoring Options

ESRD DFC Star Rating TEP: Voting Form for New Candidate Measures for DFC Star Rating Release in October 2018 (SIR, SRR, ICH CAHPS)

Name _____Date_____

Please rate each measure below (SIR, SRR, ICH CAHPS). Select only one rating/response for each measure.

1. The National Healthcare Safety Network (NHSN) Bloodstream Infection measure should be added to the DFC Star Ratings.

5	4	3	2	1
Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree

2. The Standardized Readmission Ratio (SRR) measure should be added to the DFC Star Ratings. This measure reports how often the facility's patients are readmitted to the hospital within 4-30 days.

5	4	3	2	1
Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree

3. Select ONLY one of the options below for scoring the patient experience of care measure, In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS), for the DFC Star Rating:

_____ Investigate alternative scoring (e.g., using a continuous score with all ICH CAHPS data)

_____ Use top-box results only, as publicly reported on Dialysis Facility Compare

Appendix J: TEP In-Person Voting Results and TEP Comments: SRR and SIR Measures, ICH CAHPS Scoring Options, and Hypercalcemia Measure Update Options

Thirteen TEP members were present for the vote and 13 members voted.

Table J1: In-Person Measure Voting Summary Results (%)

Measure	% Strongly Agree/Agree	% Neither Agree nor Disagree	%Strongly Disagree/Disagree
Standardized Readmission Ratio (SRR)	54%	0%	46%
The National Healthcare Safety Network Bloodstream Infection measure (NHSN SIR)	38%	15%	46%

Table J2: In-Person Measure Voting Summary Results by Rating Category (N)

Measure	# Strongly Agree Votes	# Agree Votes	# Neither Agree nor Disagree Votes	# Disagree Votes	# Strongly Disagree Votes
Standardized Readmission Ratio (SRR)	4	3	0	4	2
The National Healthcare Safety Network Bloodstream Infection measure (NHSN SIR)	5	0	2	2	4

Table J3: In-Person ICH CAHPS Voting Summary Results (% and N)

Measure	Percent of TEP Members in Favor of Investigating Alternative Scoring Methods	Percent of TEP Members in Favor of Using the Top-Box Results
In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)	85% (11 Votes)	15% (2 Votes)

TEP In-Person Voting Results: TEP Comments

This vote occurred at the in-person TEP meeting on February 21, 2017.

The following comments were provided by TEP members about the SRR and SIR Measures, and ICH CAHPS Scoring Options on the in-person voting form. The comments have been de-identified to maintain TEP member anonymity.

Table J4: National Healthcare Safety Network (NHSN) Bloodstream Infection Measure: TEP Member Comments

I would really like to see this included based on importance but the data feel incomplete. On balance, I would say yes, include it, but it's really a close call.
Concerned about validity and differential underreporting
A necessary measure that reflects patient safety based on following best practice protocols
Measure needs better validation. Perhaps claims data would help or diagnostic codes for blood cultures to be then ascertained by facilities
Agree that there has been directional improvement in reporting/surveillance over time, but don't feel like it has met the bar for inclusion in the star ratings. Continued visibility and understanding/examination is warranted.

Table J5: Standardized Readmission Ratio (SRR) Measure: TEP Member Comments

Similar concerns as previously outlined (concerned about validity and differential underreporting)
There are many factors related readmissions. Dialysis facilities are behind on evidence based practice. Ex: Crit-line recall, use of mannitol over albumin to move fluid across, and staff not properly educated
Validity/reliability not adequate for smaller facilities
In order to improve QOL readmissions should be minimized so that all preventable/unnecessary ER visits should be counted. There are workflow changes in facilities that could significantly impact this

Table J6: Options for Scoring the Patient Experience of Care Measure, In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS), for the DFC Star Rating: TEP Member Comments

There is also the question of whether to keep it separate from clinical measures or combine into an overall score. I vote for keeping separate.
I would discourage inclusion of ICH CAHPS data into 5 star as it is currently administered.
Would omit the two domains (composite + global) related to kidney doctors.
Top box is not as effective + shows the only the cream that has risen to the top.
Keep patient experience scores and clinical measures separate please.
If ICH CAHPS is included in 5 star ratings, I think it should be reported separately from clinical measures for the following reasons: 1) large percent of clinics do not have greater than or equal to 30, 2) low response rate, 3) lack of correlation with clinical measures, 4) resampling same patients.
(In reference to option 1) Will they [top-box scores] still be included on facility compare?
(For option 1) I would like to see an overall score with access (via a link) to the more detailed data. (For option 2) Will top-box results preclude access to more detailed information?

Fundamentally, the 5 star must move asap practically within process to include patient experience
Patient Experience inclusion is vital, no matter how imperfect it might be at the moment. (On option 1)-Yes, but not at the loss of including what we have now as soon as possible. In other words, I don't want the investigating alternate scoring to mean that no patient experience score is added now. If to add something now means use top box score fine, but only while investigation of alternative scoring goes on.

In-Person TEP Voting for the Hypercalcemia Measure Update

TEP members were asked to vote on one of two options for including the updated Hypercalcemia measure in the star rating. This vote occurred at the in-person TEP meeting on February 21, 2017.

- Option 1: include the updated hypercalcemia measure in the next update of the DFC Star Ratings, which would occur in October 2018, or
- Option 2: do not include the updated measure in the 2018 update to the DFC Star Ratings, and instead include it in the 2019 update to the DFC Star Ratings, to allow CMS to announce the measure in the October 2017 NPC.

Table J7: In-Person Hypercalcemia Voting Summary Results (% and N)

Measure	Percent of TEP Members in Favor of Option 1	Percent of TEP Members in Favor of Option 2
Hypercalcemia	92% (12 Votes)	8% (1 Vote)

Appendix K: TEP Star Rating Re-baselining Proposal Voting Form

Name _____Date_____

Re-baselining Proposal: Using the 5-point scale below, please rate the ***re-baselining proposal*** discussed on the March 22, 2017, TEP teleconference for implementation in the ESRD DFC Star Rating System. The proposal description you are voting on is based on the TEP discussion corresponding with teleconference slide 8.

- *Use the prior year DFC Star Rating distribution to set cut-offs to maintain continuity when updating or adding measures.*
- *Evaluate the entire DFC Star Rating distribution for a potential re-set of the distribution at predictable time intervals.*

Note:

Re-baselining and re-setting are different.

- ***Re-baselining** refers to re-scoring of measures when establishing a new baseline year;*
- ***Re-setting** refers to determining new cut-points for the entire star rating distribution.*

*** The specific time interval for scheduled re-setting of baseline and whether or not to re-baseline is a CMS policy decision*

Select only one rating. Please use the "comments" area to provide reasons for your rating.

1. The re-baselining proposal as discussed during the TEP should be implemented in the DFC Star Rating System.

5	4	3	2	1
Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree

-*Comments:*

Appendix L: TEP Star Rating Re-baselining Proposal Voting Results and TEP Comments

TEP members were asked to vote on the Re-baselining Proposal that was presented during the Post TEP Teleconference Call on March 22, 2017. 12 TEP members voted on the Re-baselining Proposal.

Table L1: Re-baselining Proposal Voting Summary Results (%)

	% Strongly Agree/Agree	% Neither Agree nor Disagree	%Strongly Disagree/Disagree
Re-baselining Proposal	67%	8%	25%

Table L2: Re-baselining Proposal Voting Summary Results by Rating Category (N)

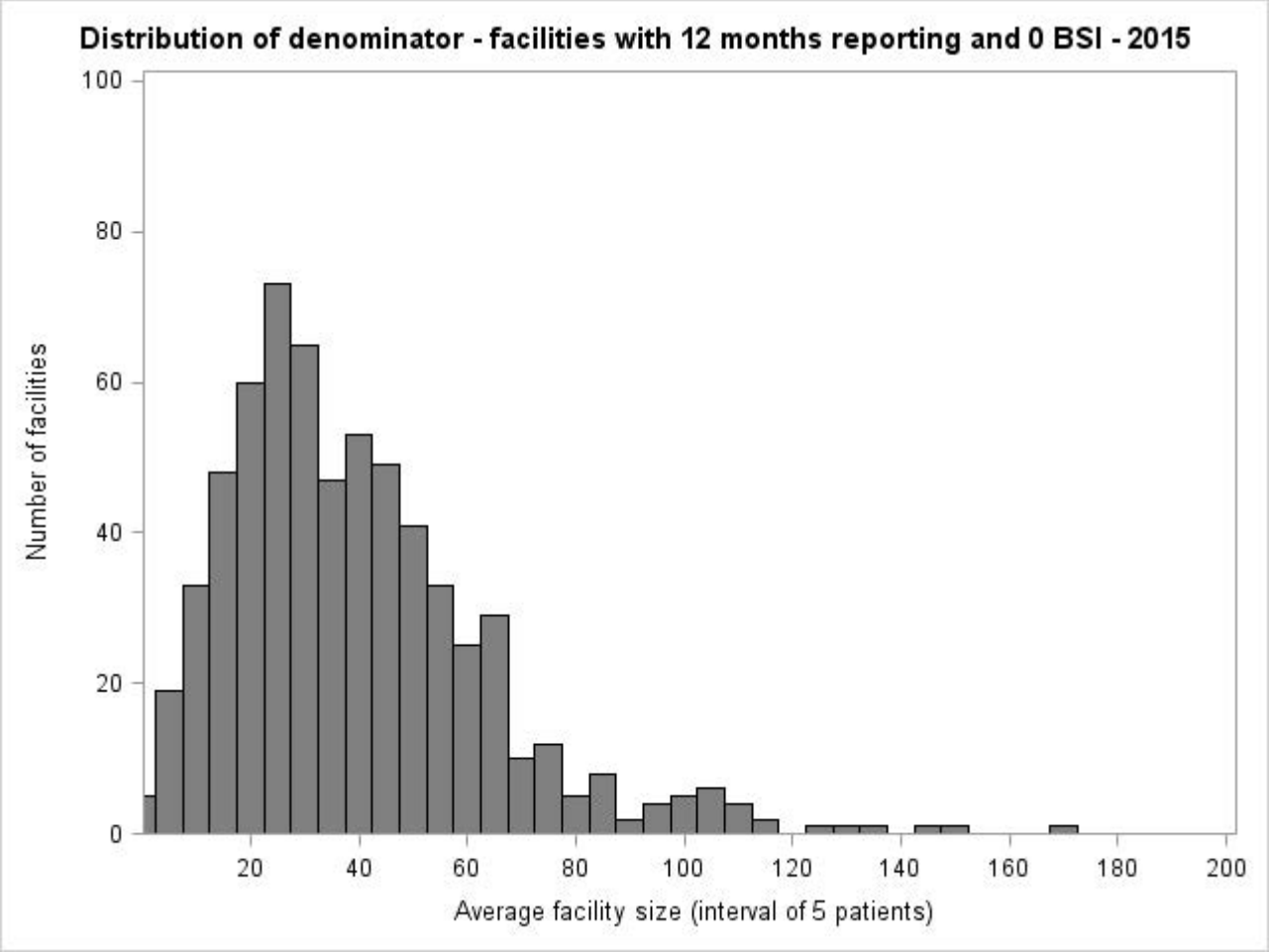
Response	# of Votes
Strongly Agree	4
Agree	4
Neither Agree Nor Disagree	1
Disagree	1
Strongly Disagree	2

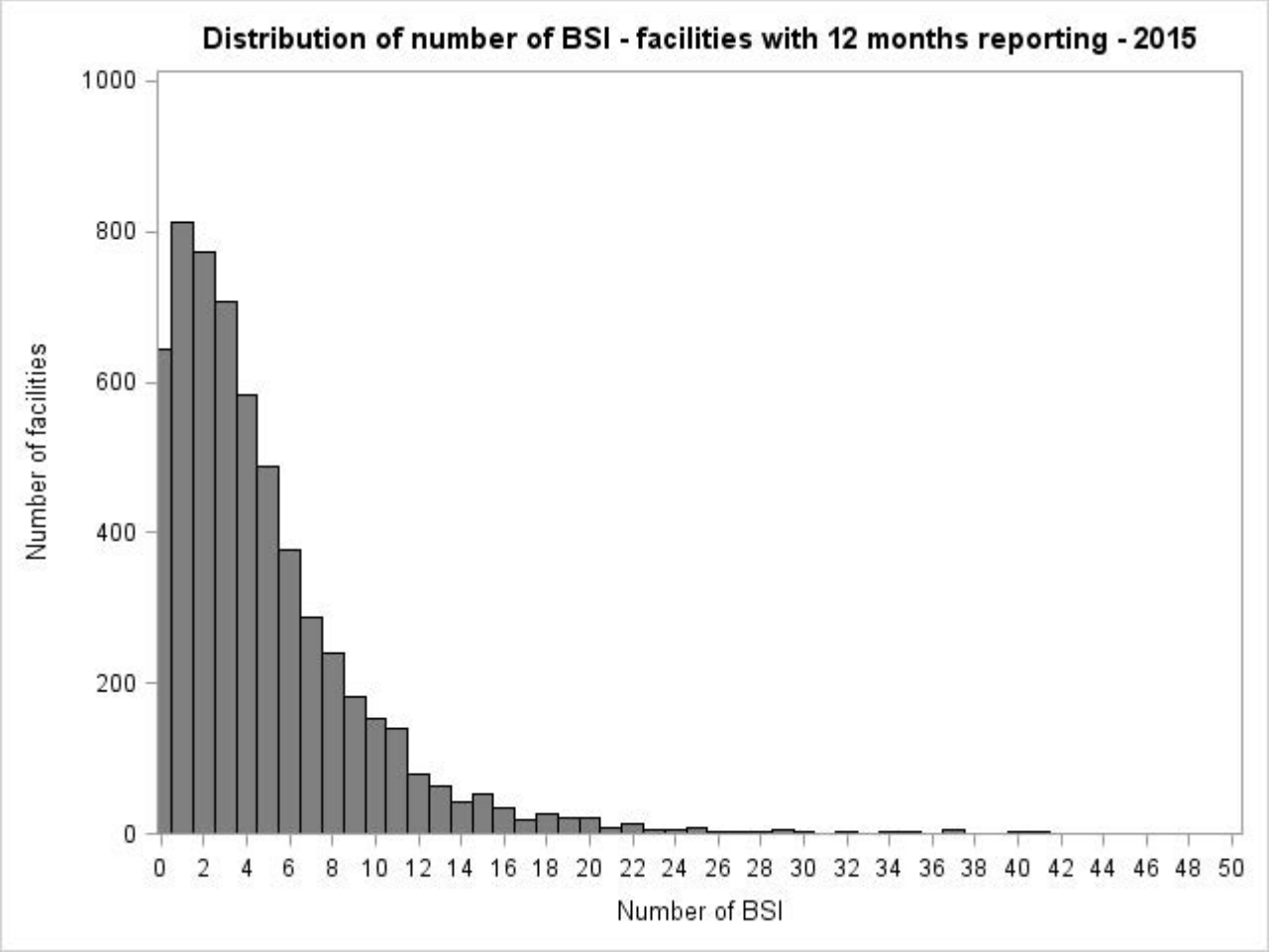
Table L3: Re-baselining Proposal Voting: TEP Member Comments

10 out of 12 TEP members provided comments on the Re-baselining Proposal Voting Form. The comments have been de-identified and are listed below.

I agree w/ this plan, especially as it promotes continuity w/ recent data and incorporation of any new measures.
I am not in favor of any methodology that creates arbitrary cut off points putting facilities into established buckets of any configuration. I would instead support a methodology that is consistent for when new measures are added or old measures are removed. I would prefer a fixed threshold methodology meaning something that tells me the actual performance of the facilities, not necessarily the marginal differences for facilities.
Agree with the proposal to favor continuity of the ratings – by using the prior year's distribution. Also am in favor of having predictable AND known minimum intervals where data are evaluated for possible reset. As a provider, frequent changes & resets make education of staff/physicians/public challenging as we leverage the public rating systems to promote behavior & system changes.
Additional specifications about the timing of re-baseline should be provided, such as % of facilities within 4/5 range reaches ..x, and/or facilities in the ½ range reaches ..y% would necessitate re-baselining

I favor using the prior year distribution to set cut-offs when adding or updating measures. In terms of the re-setting, I think it is reasonable to outline a predictable interval that reflects a minimum number of years before re-setting can occur, but specifies that re-setting is not required if there is no evidence of compression. I think there should be substantive discussion about what is an appropriate distribution when re-setting occurs.
Major caveat to this vote- would not agree if the re-setting vote would include re-setting using forced bell curve. In addition, any re-setting would not automatically be done "…at predictable time intervals" but a minimum frequency of this would be set and then the need would be evaluated as that time approached and the criteria for the decision would have been agreed upon
Good process and this approach makes sense.
My opinion is that re-baselining to reset (some of) the measures to a strict statistical distribution between 1 and 5 stars is counterintuitive to the many common rating systems consumers use regularly. So I am in favor of leaving the distributions as is and allowing "compression of stars" as facilities improved.
I strongly agree with the proposal as discussed on the March 22 TEP teleconference and as captured in writing on this ballot. I believe that the proposal allows multiple stakeholder equities to be acknowledged - from providers to patients. Once again – the original objective of the Dialysis Facility Compare Five Star Rating was to allow consumers to make well-informed choices about their care – and I believe that this method will move us closer to that objective.
Efforts to maintain "continuity" and transparency are a very good idea. My comments below, pertain to the question above, but in a more general way. They speak to why I selected Agree instead of Strongly Agree. I include them in case they might be helpful to I liked [Redacted] comment on the call that referred to an "annual message of here is improvement" and "a standard way of how to present that on DFC". I'm not entirely sure of what [they] meant, but it made me think of an annual summary message on DFC (and in other ways or places also) that reflects on improvement over the last year. After [Redacted] said that, [Redacted] commented "could show a running tracker in between baseline resets" (my quotes might not be exact). Again, I'm not entirely sure of what [they] meant, but it made me think of a summary way to show improvement. Both of their comments make me think that they might be ideas that might help with the "messaging" idea of keeping the point of view "positive" for the times things are readjusted and everyone's stars slip down. [Redacted] had asked "is there another mechanism to show continued progress to patients other than 5 stars?" and [Redacted] asked "are there really no other choices? So, just choose between these two?" Again, I don't know entirely what they meant by their questions….but they caught my ear because, I wonder, how could progress be shown so that it doesn't ever look like a center is slipping on its improvement if it is not slipping. I know this is the big thing the TEP 1 and TEP 2 struggled with, and a tremendous amount of good and qualified hard thought has gone into it. I have to trust that the experts know best. Having said that, I will say that no matter how much positive messaging you put into it, no matter how much preparation you put into preparing folks for that readjustment….it will be very, very hard, (I think, however I could be wrong) for dialysis centers and staff to not see it as negative. A lot will depend on how you lay it out visually…..


Distribution of denominator - facilities with 12 months reporting and 0 BSI - 2015
Number of facilities
0
20
40
60
80
100
20
40
60
80
100
120
140
160
180
200
Average facility size (interval of 5 patients)


Distribution of number of BSI - facilities with 12 months reporting - 2015
Number of facilities
1000
800
600
400
200
0
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50
Number of BSI

KIDNEY CARE PARTNERS

February 17, 2017

Joseph Messana, M.D.
Interim Director
KECC
University of Michigan
1415 Washington Heights
Suite 3645 SPHI
Ann Arbor, MI 48109-2029

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

Dear Dr. Messana and Dr. Goodrich,

In light of the End Stage Renal Disease (ESRD) Five Star Technical Expert Panel's (TEP) discussions, I am writing on behalf of Kidney Care Partners (KCP) and its members to reiterate for the TEP the comments that KCP provided to you on November 14, 2016. These comments were in response to the changes presented on October 5, 2016 for ESRD Five Star Rating Program (ESRD Five Star) methodology and for new measures for Star Ratings of dialysis facilities. KCP also commented at that time on the baseline methodology used in ESRD Five Star. We appreciate the Agency's efforts to work with the kidney care community to revise the Star Rating methodology and the opportunity to comment on measures under consideration for Dialysis Facility Compare (DFC)/ESRD Five Star Rating Program. 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.

This letter provides an overview of our comments on both the modified measures and the new measures the TEP is considering for ESRD Five Star. It also expresses our deep concerns about the proposed approach to re-baselining the ESRD Five Star program that has been discussed during the recent TEP teleconferences.

I. KCP supports the process for considering candidate measures and facility review and provides comments on the candidate measures.

KCP supports the goal of DFC/ESRD Five Star to provide information about the performance of dialysis facilities to empower patients, family members/caregivers, and consumers. As we have described in the past, the

cornerstone of achieving this goal is to ensure meaningful community input and transparency.

A. Measures that Matter

As CMS continues to consider modifying the measures in DFC/ESRD Five Star, we ask that the Agency work with KCP to make sure that these programs focus on valid and reliable measures that will have the greatest impact on improving patient outcomes. As the attached chart (see Appendix B) demonstrates, we are concerned that some of the measures being considered are not reliable and that the NHSN Bloodstream Infection measure is not valid, missing 60-80 percent of the events as noted by CMS. We are disappointed that during the recent calls, it was suggested that TEP members not take into account that the underlying specifications for a measure may not be valid or reliable and that the measures should be taken "as is." This is an unacceptable position, given that validity and reliability are needed to ensure that the measures included in this public evaluation program provide accurate data to patients. The federal government should not support any public reporting system that promotes measures that have not been shown to be accurate for all facilities and that, therefore, could mislead patients and consumers.

The number of measures should also be limited to prevent the dilution of their impact on the overall star rating. We echo MedPAC's concerns and its recommendation that "[t]he set of measures should be small to minimize the administrative burden on providers and CMS."[1] We ask that CMS work closely with KCP and others in the kidney care community to create a parsimonious set of measures that will further the Triple Aim, rather than compromise it.

B. KCP supports adding the Pediatric Peritoneal Dialysis Adequacy, Standardized Fistula, and Long-Term Catheter Measures to ESRD Five Star.

KCP supports adding the pediatric peritoneal dialysis and new fistula and long-term catheter measures to ESRD Five Star. We have separately recommended additional refinements to the fistula measure and look forward to working with CMS to improve that measure's risk adjustment in the future.

Specifically, KCP believes the specifications should be clarified as to whether facilities would receive credit for patients using an AVF as the sole means of access, but who also have in place a catheter that is no longer being used. The measure definition of autogenous AVF "as the sole means of vascular access" is imprecise as to whether facilities would receive credit for patients using an AVF as the sole means of access, but who also have in place a catheter that is no longer being used. In previous letters we have described how patients with catheters remain at risk for

[1]MedPAC, *Report to the Congress*, "Chapter 3: Measuring Quality of Care in Medicare" 41 (June 2014).

infection and other adverse sequellae, so credit should not be given when a catheter is present, even if an AVF is being used. A numerator that specifies the patient must be on maintenance hemodialysis "using an AVF with two needles and without a dialysis catheter present" would remove ambiguity. In contrast, removal of an AV graft is complex and not without risk of complications, so KCP believes credit should be received for a patient who is using an AVF as the sole means of access, but who also may have a non-functioning AV graft present.

We remain supportive of the removal of the 90-day ESRD requirement from the denominator statement. Additionally, we commend the developer for adding an exclusion for patients with limited life expectancy and for now unambiguously identifying the four subcategories, both approaches that KCP had recommended.

While we appreciate that the developer has removed the covariate alcohol dependence from the model's risk variables, we continue to believe two additional vasculature risk variables would strengthen the model: A history of multiple prior accesses and the presence of a cardiac device. The validity testing yielded an overall c-statistic of 0.71, which raises concerns that the model will not adequately discriminate performance—particularly that smaller units might look worse than their actual performance really is. A minimum c-statistic of 0.8 is a more appropriate indicator of the model's goodness of fit and validity to represent meaningful differences among facilities and encourage continuous improvement of the model.

C. KCP supports including the ICH CAHPS measure on a separate DFC webpage and keeping it out of the star ratings, but also encourages CMS to address the burden of implementing the ICH CAHPS measure on patients.

KCP has consistently supported using the ICH CAHPS measure as a reporting measure for the Quality Incentive Program (QIP); we similarly support providing the results of the ICH CAHPS survey on DFC, so long as it is not included in the ESRD Five Star overall rating and so long as CMS includes the response rate (*i.e.,* how many patients were eligible to respond and how many actually responded) with the results. The response rate is critically important to allow patients and caregivers to understand and interpret the information they are seeing. Before it can be included in the ratings, the burden it places on patients needs to be resolved to ensure that the majority of patients are able and willing to complete the survey tool. Therefore, we ask that CMS work with KCP and the kidney care community.

In previous letters, we have suggested that CMS decrease the burden on patients and facilities of the twice-yearly administration. The American Institutes for Research/RAND *et al.* have described in detail the difficulties in translating the results from ICH CAHPS into interventions resulting in meaningful improvement

when administered more frequently than once a year.[2] We continue to believe that reducing the frequency and eliminating Network duplication in administration will decrease the burden on patients, increase their participation and survey completeness rates, decrease costs, and increase facilities' capacity to respond to survey results. Given our previous recommendations, we would like to better understand why CMS considers administering the survey once each year inadequate so that we can work to find a viable solution.

In previous letters, we have raised concerns about patients being unable to finish the complete survey because of its length and recommended that CMS divide the survey into the three sections that were already independently validated. If there is a reason why this suggestion is not workable, we would like to better understand the concern and work with CMS to find another alternative that promotes the completion of the survey by patients.

There is clearly survey fatigue with regard to ICH CAHPS among patients. One of our members, the American Society of Nephrology (ASN), has observed that the response rate has fallen significantly since the inclusion of ICH CAHPS in the ESRD Quality Incentive Program (QIP) and even more so since CMS added the requirement to administer it twice a year, with a national response rate of only 33% in 2015 (https://ichcahps.org/ICHCAHPS_2015_NatlStateAvgs.pdf) and a response rate of only approximately 30% on the mail-only administration of ICH CAHPS in the Mode Experiment used to generate adjustment coefficients for survey responses (https://ichcahps.org/Portals/0/PublicReporting/ICHCAHPS_coefficients_for_2016_Surveys.docx). Everyone agrees that we need better response rates as low response rates equate to bias and, if the same patients are among the few responding to both administrations of the ICH CAHPS, that bias is enhanced, further threatening validity. To achieve that goal, we encourage CMS to work with the community to strike the right balance in terms of the burden on patients and the length and frequency of administration of the survey.

We also recommend that CMS ensure the accuracy of the administration of the survey. First, it is critically important to have a mechanism, which does not appear to exist currently, for facilities to ensure that patients' contact information is as accurate and up-to-date as possible. Because response rates necessarily depend on accurate contact information, we recommend inclusion of an opportunity for facilities to ensure that the primary survey and/or any follow-up is delivered to the most current contact (phone or mail) given the penalty that applies for non-responsiveness.

---

[2] American Institutes for Research, RAND, Harvard Medical School, Westat, Network 15. Using the CAHPS® In-center Hemodialysis Survey to Improve Quality: Lessons Learned from a Demonstration Project. Rockville, MD: Agency for Healthcare Research and Quality (Dec. 2006).

As we have noted previously, we also suggest that the Agency update the survey to include home dialysis patients as well.

KCP agrees that it is important to provide information about patient experience. While ICH CAHPS may not be perfect, it is a reasonable starting point from which progress can be made to address and resolve the concerns KCP and most significantly patients have raised with the burden of completing the survey and the accuracy in its administration. We are sincere in our request to work with CMS to resolve these problems in the near term.

D. KCP continues to have significant concerns about the reliability of the Standardized Mortality Ratio (SMR), Standardized Hospitalization Ratio (SHR), Standardized Transfusion Ratio (STrR), and Standardized Readmission Ratio (SRR).

KCP applauds CMS for moving away from ratios and transition to rates. We were also pleased to see prevalent co-morbidities incorporated into the SMR and SHR measures as well.

Despite these positive steps, KCP remains concerned about the reliability of these measures. It is simply not clear what value these measures provide patients when a clear majority of measure's reliability score is due to random chance. For example, CMS's testing data indicates 60-70 percent of a small facility's score is due to chance. Similarly for the SHR, 43 percent of a medium-sized facility's score is due to noise and not a signal of quality while 54 percent is due to noise for small facilities. Similar poor reliability exists for the 4-year SMR, where 55-70 percent of a facility's score is due to differences in performance for small- and medium-sized facilities. Rather than providing the accurate information patients, family members/caregivers, and consumers need to make decisions, these measures present random data that can be misleading and confuse patient decision-making. We recommend that CMS describe how it will address these short-comings before adding these measures to the ESRD Five Star ratings.

Additionally, concerns about several of the technical details of the SMR, SHR, STrR, and SRR unfortunately remain unresolved. We have conveyed those concerns separately and have included them in the appendix to this letter.

E. Because the data show that the NHSN Blood Stream Infection Measure is not valid, KCP cannot support including it in DFC/ESRD Five Star.

Finally, as we have communicated in our most recent letter, KCP recognizes the vital importance of reducing infections and strongly supports efforts to do so.

However, we cannot support use of the NHSN BSI Measure for inclusion on DFC and in ESRD Five Star because the Centers for Disease Control and Prevention's (CDC) research and CMS's data have demonstrated that the measure is not valid. For example, CMS has stated that its review of data reported for the PY 2015 NHSN Dialysis Event Reporting Measure and results from the PY 2014 NHSN data validation feasibility study suggest that as many as 60-80 percent of dialysis events are under-reported.[3] Simply put, this high under-reporting rate demonstrates the measure is not valid. A lack of validity means that we cannot be certain that the measure results in accurate findings. Reporting inaccurate findings on Dialysis Facility Compare and including it in the Five Star ratings misleads patients who are trying to use measures to make informed decisions about their care.

II. KCP commends CMS for modifying the methodology to move away from the forced distribution and seeks clarity regarding re-baselining to ensure that the problems of the past do not recur.

KCP continues to believe that ESRD Five Star ratings should align as closely as possible with actual facility performance. Therefore, we were pleased when CMS announced the revised ESRD Five Star methodology that moves to a z-score model to score most of the individual measures, as KCP had suggested previously and that patient organizations strongly support.

Additionally, we are pleased that CMS has changed the methodology to use fixed performance benchmarks for the Star Rating cut points. This will allow facilities to demonstrate performance changes over time and eventually would allow the distribution of Star Ratings to shift based on overall improvement trends. Both of these results are aligned to the overall program goals on conveying accurate information to consumers.

We remain concerned, however, that the improvements in the methodology could be undermined by using the 10-20-40-20-10 distribution when the stars are "re-baselined." We strongly urge CMS not to use this distribution and instead continue to rely upon the fixed performance benchmarks to address the concerns expressed by the kidney care community and, in particular, the patient organization participants in the previous ESRD Star Rating TEP.

Discussions during the recent TEP calls seem to indicate that CMS may re-baseline the Five Star program whenever new measures are added or other measures are modified. Given that CMS is adding or modifying measures annually, this means practically that the improvement methodology the previous TEP strongly urged CMS to adopt will rarely be applied.

---

[3]ESRD QIP Proposed Rule Display Copy 90.

KCP has raised concerns previously about too frequently re-baselining the ESRD Five Star program for two reasons. First, the base year would still be determined on a forced bell curve, which KCP and others in the kidney care community, including the patient members of the TEP that considered the ESRD Five Star methodology, strongly believe is not an appropriate way to determine the star ratings. Second, re-baselining too frequently will make it impossible to compare dialysis facilities year-over-year.

During the recent TEP calls, the CMS representative indicated to TEP members that re-baselining is necessary, especially when new measures are added. If CMS plans on adding or updating measures annually, then the first two possible triggers discussed during the TEP teleconference (when measures are added or modified) should be rejected. Otherwise, the system maintains the forced distribution, which the previous TEP sought to eliminate. Under this approach, patients will rarely, if ever, be able to see actual performance data to make informed choices. The TEP also discussed re-baselining when 50 percent or more of facilities receive either a 4 or 5 star rating. If this trigger were adopted, re-baselining would occur less frequently, but could be more disruptive to consumers. For example, if the distribution of stars were 5-15-20-25-35, when re-baselined, some facilities would go from 5 stars to 3 stars just due to the re-basing, rather than because of any change in their performance.

The only reason re-baselining is necessary is because of the underlying forced distribution for establishing the baseline year star ratings. Therefore, KCP continues to recommend that CMS use a methodology in the base year that awards stars on actual performance and not a forced distribution curve.

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

Appendix A: Technical Comments on Selected Candidate Measures

STANDARDIZED MORTALITY RATIO (NQF #0369)

KCP recommends working with the kidney care community to address concerns about the current Standardized Mortality Ratio (SMR) measure.
KCP believes mortality is an important outcome to measure, but has on several occasions expressed concern about the current SMR. We appreciate the CMS's recognition in 2013 that it needed to "properly take into account the effect that comorbidities have on hospitalization and mortality rates in the ESRD population,"[4] as well as its movement away from exclusively relying on the 2728 data. However, we remain concerned about the testing data, which indicate significant reliability issues with the SMR for small- and medium-sized facilities—even with the 4-year measure. Empirical testing has demonstrated that for the 4-year SMR, on average, less than 60% of a facility's score is attributable to between-facility differences; testing results specifically for small- and medium-sized facilities indicate very poor reliability, with IURs of 0.30 and 0.45, respectively. Given the poor reliability testing results, KCP believes the specifications must explicitly require a minimum sample as identified through the developer's empirical testing.

Additionally, we note the SMR specifications indicate the measures can be expressed as a rate, but is calculated as a ratio. KCP continues to support the use of rate measures because they allow patients and facilities to see year-over-year differences between normalized rates (deaths per 100 patient years) for mortality and hospitalization. We believe comprehension, transparency, and utility to all stakeholders is superior with a scientifically valid *rate* methodology.

STANDARDIZED HOSPITALIZATION RATIO (NQF #1463)

KCP would like to support the Standardized Hospitalization Ratio (SHR), but cannot until its reliability has been demonstrated.

KCP concurs that hospitalization is an important quality domain, and we appreciate and approve that the SHR now accounts for prevalent comorbidities. We would like to support a hospitalization measure, but continue to be concerned about the significant reliability issues for the 1-year SHR for small facilities and do not support incorporation of the SHR until its reliability at the proposed facility size is demonstrated.

Specifically, for facilities with <=50 patients, more than half (54%) of a facility's score is due to random noise; even for medium facilities, 43% of a facility's score

[4] "End-Stage Renal Disease Prospective Payment System, Quality Incentive Program, and Durable Medical Equipment, Prosthetics, Orthotics, and Supplies; Proposed Rule" 78 *Fed. Reg.* 40836, 40861 (July 8, 2013).

attributable to random noise and is not a signal of quality. Given the poor reliability testing results, KCP also did not support CMS's proposal to include it in the Quality Incentive Program (QIP) for Payment Year 2020.

Additionally, we are concerned that only facilities with <5 patient-years at risk during the performance period are not eligible for the measure. As we have noted elsewhere, KCP believes the standardized ratio measures should be harmonized—currently the SHR uses a <5 patient-years at risk threshold, but the standardized mortality ratio and standardized transfusion ratio use <10 patient-years at risk.

Finally, the SHR specifications indicate the measures can be expressed as a rate, but is calculated as a ratio. KCP continues to support the use of rate measures because they allow patients and facilities to see year-over-year differences between normalized rates (hospitalizations per 100 patient years) for mortality and hospitalization. We believe comprehension, transparency, and utility to all stakeholders is superior with a scientifically valid *rate* methodology.

STANDARDIZED TRANSFUSION RATIO (NQF #2979)

KCP continues to have significant concerns about the reliability of the Standardized Transfusion Ratio (STrR) measure.

KCP again expresses our concern about the reliability of the STrR for small facilities. Specifically, testing yielded IURs of 0.30-0.41 for small facilities for each of 2011, 2012, 2013, and 2014, indicating approximately 60-70% of a small facility's score is due to random noise. KCP believes the specifications must specifically require a minimum sample as identified through the developer's empirical testing. Additionally, we again note that physicians independently (or following hospital protocols) make decisions about whether or not to transfuse a specific patient; the measure does not adjust for such hospital- and physician-related transfusion practices.

Finally, while KCP is pleased that CMS has decided to evaluate the impact of the STrR on access to care through the SRR/Standardized Transfusion Ratio Impact Study, we again question the appropriateness of using the measure until the results of the study are known remains. If CMS is unclear about whether these measures will have a positive or negative impact on dialysis patients and the care they receive, the Agency should not use these measures until it has such clarity. We again also recommend evaluating the effectiveness of the STrR in measuring the actual care provided in dialysis facilities.

STANDARDIZED READMISSION RATIO (NQF #2496)

KCP continues to have significant concerns about the reliability of the inclusion of the Standardized Readmission Ratio (SRR) measure.

KCP again expresses our concern about the reliability of the SRR. CMS presented reliability data to NQF for which even for large facilities with >121 patients, the IUR was only 0.61. Additionally, for SRR implementation in the QIP, CMS proposes an adjuster of 11-41 index discharges, but offers no rationale for this value. This lack of transparency undermines our ability to assess the proposed use of the measures. KCP believes that the values are too low, and will result in random volatility that the Small Facility Adjuster, as proposed, cannot fully offset.

Finally, while KCP is pleased that CMS has decided to evaluate the impact of the SRR on access to care through the SRR/Standardized Transfusion Ratio Impact Study, we again question the appropriateness of using the measure until the results of the study are known remains. If CMS is unclear about whether these measures will have a positive or negative impact on dialysis patients and the care they receive, the Agency should not use these measures until it has such clarity. We again also recommend evaluating the effectiveness of the SRR in measuring the actual care provided in dialysis facilities.

Appendix B: RELIABILITY STATISTICS FOR STrR, SMR, SHR

(Empirically derived, CMS testing information to NQF)

- An Inter-Unit Reliability (IUR) statistic of >=0.70 is generally considered "acceptable" in the statistical literature; 0.60-0.69 is "questionable", 0.50-0.59 is "poor", and <=0.49 is "unacceptable".[5] NQF also generally uses 0.7 as the acceptable threshold. At 0.7, 70% of the variation among measured entities is attributable to quality signal and 30% to random noise.

Standardized Transfusion Ratio

	2011		2012		2013		2014	
Size	***IUR***	***N***	***IUR***	***N***	***IUR***	***N***	***IUR***	***N***
All	0.64	5,142	0.66	5,319	0.65	5442	0.60	5651
<=46 pts	0.41	1,714	0.41	1,828	0.39	1,840	0.30	1,934
47-78 pts	0.55	1,699	0.56	1,753	0.55	1,823	0.50	1,941
>=79 pts	0.78	1,729	0.79	1,738	0.79	1,779	0.78	1,776

Standardized Mortality Ratio (4-year; NQF pushed back on the 1-year and refused to recommend until CMS committed to 4-year).

Size	***IUR (2010-2013)***	***N***
All	0.59	5,935
<=135 pts	0.30	1,242
136-305 pts	0.45	2,320
>=306 pts	0.73	2,373

Standardized Hospitalization Ratio

	2010		2011		2012		2013	
Size	***IUR***	***N***	***IUR***	***N***	***IUR***	***N***	***IUR***	***N***
All	0.72	5,407	0.71	5,583	0.70	5,709	0.70	5,864
<=50 pts	0.54	1,864	0.51	1,921	0.48	1,977	0.46	2,028
51-87 pts	0.65	1,702	0.63	1,785	0.58	1,825	0.57	1,930
>=88 pts	0.81	1,841	0.81	1,877	0.81	1,907	0.82	1,906

[5]George, D and Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference.* 11.0 update (4th ed.). Boston: Allyn & Bacon; Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London: Routledge, page 13; DeVellis, RF. (2012). *Scale development: Theory and applications*. Los Angeles: Sage. pp. 109–110. Adams, JL. (2009). The reliability of provider profiling. RAND Health.

CMS Dialysis Facility Compare (DFC)
Star Rating TEP
In-person Meeting
February 21, 2017

SCHOOL OF PUBLIC HEALTH
KECC
UNIVERSITY OF MICHIGAN

University of Michigan Kidney Epidemiology and Cost Center DFC Star Ratings TEP Team

Yi Li, PhD: Principal Investigator

Joseph Messana, MD

Claudia Dahlerus, PhD, MA

Richard Hirth, PhD

Ji Zhu, PhD

Casey Parrotte, PMP

Karen Wisniewski, MPH

Natalie Scholz, MPH

Cindy Liao, MS, MPH

Stephen Salerno, BS, Graduate Student Research Assistant

Contact Person:
Jordan Affholter, BA
affjorda@med.umich.edu
(734)764-2189

SCHOOL OF PUBLIC HEALTH
KECC
UNIVERSITY OF MICHIGAN

CMS Representatives

Centers for Medicare & Medicaid Services, Division of Quality Measurement

- Joel Andress, PhD, ESRD Measures Development Lead
- Elena K. Balovlenkov, MS, RN, CHN Technical Lead, Dialysis Facility Compare
- Jesse Roach, MD, Nephrologist, Clinical Subject Matter Expert
- Golden Davis, Communications Liaison

CDC, ICH-CAHPS and CMS Contractor Representatives

- Centers for Disease Control & Prevention (CDC)
- CMS and RTI for In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)
- NORC
- Ketchum

In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) Representatives

- Elizabeth Goldstein, PhD, CMS, ICH CAHPS Division Director
- Julia Zucco, PhD, CMS, ICH CAHPS Co-Project Officer
- Debra Dean-Whittaker, PhD, CMS, ICH CAHPS Co-Project Officer
- Judith Lynch, RTI, ICH CAHPS Project Director
- Celia Eicheldinger, RTI, Public Reporting Task Lead

Centers for Disease Control & Prevention (CDC) Representatives

- Daniel Pollock, MD, Surveillance Branch Chief
- Priti Patel, MD, MPH, Medical Officer
- Duc Nguyen, MD, Medical Epidemiologist
- Shunte Moon, PhD, MPH, Dialysis Epidemiologist

TEP Member Introductions and Conflict of Interest Disclosures

TEP Members

- **Paul Conway**, BA, TEP co-chair, President, American Association of Kidney Patients (AAKP)
  - Conflict of Interest Disclosure: None
- **Catherine Sugar**, PhD, MS, TEP co-chair, Professor, Departments of Biostatistics, Statistics & Psychiatry University of California, Los Angeles
  - Conflict of Interest Disclosure: None
- **Lorien Dalrymple**, MD, MPH, Vice President of Epidemiology and Research, Fresenius Medical Care North America (FMCNA)
  - Conflict of Interest Disclosure: Employed by Fresenius Medical Care NA and member of the KCQA Steering Committee
- **Amanda Grandinetti**, MPH, Senior Specialist, Performance Measures and Analysis, American Academy of Dermatology, Kidney Action Committee Member, National Kidney Foundation
  - Conflict of Interest Disclosure: None

TEP Members

- **Mark Joseph**, MD, Medical Director, Dialysis and CRRT, Phoenix Children's Hospital
  - Conflict of Interest Disclosure: Member of the Horizon Pharmaceutical speaker's bureau for nephropathic cystinosis
- **Richard Knight**, MBA, Vice President/Chair of Public Policy, American Association of Kidney Patients (AAKP)
  - Conflict of Interest Disclosure: None
- **Jewell Kyle**, RN, BSN, CNN, RN Charge Nurse/Staff Educator, Chattanooga Kidney Centers, LLC
  - Conflict of Interest Disclosure: None
- **J. Richard (Dick) Landis**, PhD, MS, Professor of Biostatistics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA
  - Conflict of Interest Disclosure: None

TEP Members

- **Allen Nissenson**, MD, FACP, Chief Medical Officer, DaVita Kidney Care
  - Conflict of Interest Disclosure: Full time employee, DaVita, Inc.
- **Chris Sarfaty**, MSW, LICSW, Patient-Centered Collaborative Care Coach, Coaching for Health Professionals
  - Conflict of Interest Disclosure: None
- **Nicole Stankus**, MD, MSc, Medical Director, DaVita Stony Island Dialysis Center
  - Conflict of Interest Disclosure: Member of the DaVita Physician Council
- **Sumi Sun**, MPH, VP of Analytics and Quality Strategy, Satellite Healthcare
  - Conflict of Interest Disclosure: None

TEP Members

- **Frederic Talton**, BA, Board Member, Dialysis Patient Citizens (DPC)
  - Conflict of Interest Disclosure: None
- **David White**, Board of Director Member, American Association of Kidney Patients (AAKP)
  - Conflict of Interest Disclosure: None

TEP Member Role and Responsibilities

- TEP members will assist UM-KECC in developing recommendations to CMS about measure addition, approach to inclusion of patient reported outcomes, and how to re-baseline the star ratings

- Share your expert opinions and experience

- Allow for and encourage the sharing of opinions of all TEP members to be heard

TEP objectives

- Develop recommendations on:
  - Inclusion of candidate measures reported on Dialysis Facility Compare (DFC) into the DFC Star Ratings, taking into account the extent to which the new measures would provide a more well-rounded depiction of the quality of dialysis facilities, and whether the information is understandable and important to patients
  - Methods for inclusion and reporting of current and future patient reported outcomes in the Star Ratings (e.g. separate from or combined with clinical outcome measures)
  - Resetting the baseline year thresholds when measures are added/retired or when the Star Rating categories no longer reflect meaningful differences among facilities

TEP Agenda

9:00 – 10:30 Voting Results and New Candidate Measure Discussion

10:30 – 10:45 BREAK

10:45 – 11:30 Continue Candidate Measure Discussion

11:30 – 12:00 Finalize Candidate Measure Recommendations

12:00 – 1:00 LUNCH

TEP Agenda Part 2

1:00 – 1:30 Recommendations on Method for Inclusion of Current and Future Patient Reported Outcomes in the Star Ratings

1:30 – 3:00 Re-baselining

3:00 – 3:15 BREAK

3:15 – 3:45 Wrap-up: Summary of Recommendations and Discussion of Next Steps

3:45 – 4:00 Public Comment Period

Preliminary Measure Voting Results

Measure Voting Results for Standardized Fistula Rate

Strongly Agree/Agree	92%
Neutral	8%
Strongly Disagree/Disagree	0%

Measure Voting Results for Long-term Catheter Rate

Strongly Agree/Agree	92%
Neutral	8%
Strongly Disagree/Disagree	0%

Measure Voting Results for Standardized Mortality Ratio (SMR)

Strongly Agree/Agree	84%
Neutral	8%
Strongly Disagree/Disagree	8%

Measure Voting Results for Standardized Hospitalization Ratio (SHR)

Strongly Agree/Agree	92%
Neutral	0%
Strongly Disagree/Disagree	8%

Measure Voting Results for Standardized Transfusion Ratio (STrR) measure

Strongly Agree/Agree	75%
Neutral	8%
Strongly Disagree/Disagree	17%

Measure Voting Results for the Pediatric Peritoneal Dialysis Adequacy: Achievement of Target Kt/V (Pediatric PD Kt/V) measure

Strongly Agree/Agree	58%
Neutral	42%
Strongly Disagree/Disagree	0%

Measure Voting Results for The National Healthcare Safety Network Bloodstream Infection measure (NHSN SIR)

Strongly Agree/Agree	50%
Neutral	8%
Strongly Disagree/Disagree	42%

Measure Voting Results for the In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS) measure

Strongly Agree/Agree	69%
Neutral	15%
Strongly Disagree/Disagree	15%

Measure Voting Results for The Standardized Readmission Ratio (SRR)

Strongly Agree/Agree	50%
Neutral	17%
Strongly Disagree/Disagree	33%

Discussion

1. National Healthcare Safety Network Bloodstream Infection (NHSN SIR)
2. Standardized Readmission Ratio (SRR)
3. In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

Discussion: NHSN SIR (Bloodstream Infection Standardized Infection Ratio)

- Main concern: Under-reporting and subsequent validity of the measure
  - High number of zero infections shown in the distribution.

Star Ratings of CAHPS Survey Results

The In-center Hemodialysis CAHPS Survey

- The ICH CAHPS Survey is conducted on a semi-annual basis with samples of hemodialysis patients
- CMS began reporting ICH CAHPS Survey results on Dialysis Facility Compare on www.medicare.gov in October 2016
- "Top-box" (the most positive) scores are currently reported on Dialysis Facility Compare

Publicly Reported ICH CAHPS Measures (cont'd)

Three composite and three individual measures are reported on the Dialysis Facility Compare

Composite Measures

- Kidney doctors' communication and caring, (6 survey items)
- Dialysis center staff, care and operations, (17 survey items)
- Providing information to patients, (9 survey items)

- Three global ratings (individual survey items)
  - Rating of kidney doctors
  - Rating of dialysis center staff
  - Rating of dialysis center

General Information about Star Ratings on CAHPS Surveys

- Star Ratings summarize performance using symbols (stars) to help consumers quickly and easily understand quality of care information
- CMS is currently using star ratings on other CAHPS Surveys, including the Hospital CAHPS and the Home Health CAHPS Surveys, and on the CMS Part C and Part D Star Ratings Program

General Information about Star Ratings on CAHPS Surveys (cont'd)

- Star ratings can be presented in different ways.
  - On Hospital CAHPS, a single summary star rating is calculated and presented
  - Home Health CAHPS calculates and present star ratings separately for clinical and CAHPS measures
- Programs also vary on how star ratings are calculated.
  - Methods used to calculate clinical measures might differ from those used to calculate star ratings based on survey data
- Methods used to calculate star ratings for CAHPS Surveys are similar
  - On most CAHPS Surveys, a star rating is calculated for each composite measure and for each individual (global) rating survey item
- In general, CAHPS star ratings are based on linearized scores that encompass all of the response options
- CAHPS star ratings use statistical clustering models to create five clusters (stars)

Home Health Compare – Star Ratings Display

General Information about Star Ratings on CAHPS Surveys (cont'd)

- CMS has not yet decided how star ratings will be calculated for the ICH CAHPS Survey
  - Methods used might be similar to those used on other CAHPS Surveys
- CMS will decide which in-center dialysis centers will be eligible for reporting star ratings on Dialysis Facility Compare
  - On other CAHPS Surveys, the health care provider must have a minimum number of completed surveys to receive star ratings for survey data
  - Also, the number of completed surveys required for star ratings might be different from the number required for public reporting

Creating Star Ratings on Another CAHPS Survey

This is an example of how linearized scores are calculated on another CAHPS Survey.

- Individual survey responses are converted into linear scores on a 0-100 point scale
- Composite scores are based on the mean of the linearized responses to the questions that are included in each composite measure

Creating Star Ratings on Another CAHPS Survey (cont'd)

- After linearized scores are created, they are adjusted for mode effects (if any) and patient mix
- Patient mix= level playing field among providers by adjusting for patient characteristics that affect response tendencies

Patient Mix Adjusters

- Patient adjustment factors vary on CAHPS Surveys, but they typically include age, gender, self-reported overall health status, education, and selected diseases and conditions
- Most of the adjustment factors come from the patient survey, but some are supplied by the health care provider

Creating Star Ratings on Another CAHPS Survey

Converting Linearized Scores to Star Ratings

- A statistical clustering technique is applied to the adjusted facility-level scores
- Clustering identifies groups so that
  - differences **between** groups are **maximized** and
  - differences **within** groups are **minimized**.
- A 1, 2, 3, 4, or 5 star is assigned to each CAHPS measure based on cluster assignments
  - There are no predetermined quotas on the number of health care providers that would be included in any star category

Star Ratings Cut Points on Another CAHPS Survey

- The cut points (boundaries) for star assignments are derived from the range of individual measure Star Ratings in each cluster.
- In each public reporting period, the cut points are reestimated and made available to health care providers in a Preview Report.

Discussion: ICH CAHPS

- Main concern: Low response rate and number of facilities not eligible to receive ICH CAHPS score due to not enough completed surveys

Discussion: SRR

- Main concern: Facility attribution
  - Readmissions are outside the facility's control

Finalize Recommendations for NHSN SIR, SRR, and ICH CAHPS

Lunch

12:00-1:00

Recommendations on Method for Inclusion of Current and Future Patient Reported Outcomes in the Star Ratings

Patient Experience of Care measures and Star Rating

Three Options:

1. One overall star rating combining Patient Experience of Care and Clinical Quality Measures
2. Two separate star ratings
    a. Overall Patient Experience of Care Rating
    b. Clinical Quality Measures
3. Three star ratings
    a. Overall Star Rating
    b. Overall Patient Experience of Care Rating
    c. Clinical Quality Measures

Re-baselining

- Necessity of adding new measures, updating existing measures, and removing older measures

Hypercalcemia: Input on Re-baselining

- This measure shows the percentage of adult hemodialysis and peritoneal dialysis patients treated at each center whose average (3 month) calcium was greater than 10.2 mg/dL
- Updated to include patients with missing calcium values in the numerator
- Lower values are better

Hypercalcemia: Input on Re-baselining

- The Hypercalcemia Measure was updated in late 2016, and updates accepted by NQF in early 2017
  - Non-substantive update
- The non-substantive update to the measure was not announced during the National Provider Call (NPC) for Public Comments in October 2016
  - Implications on when to implement in Star Rating and Star Rating re-baselining

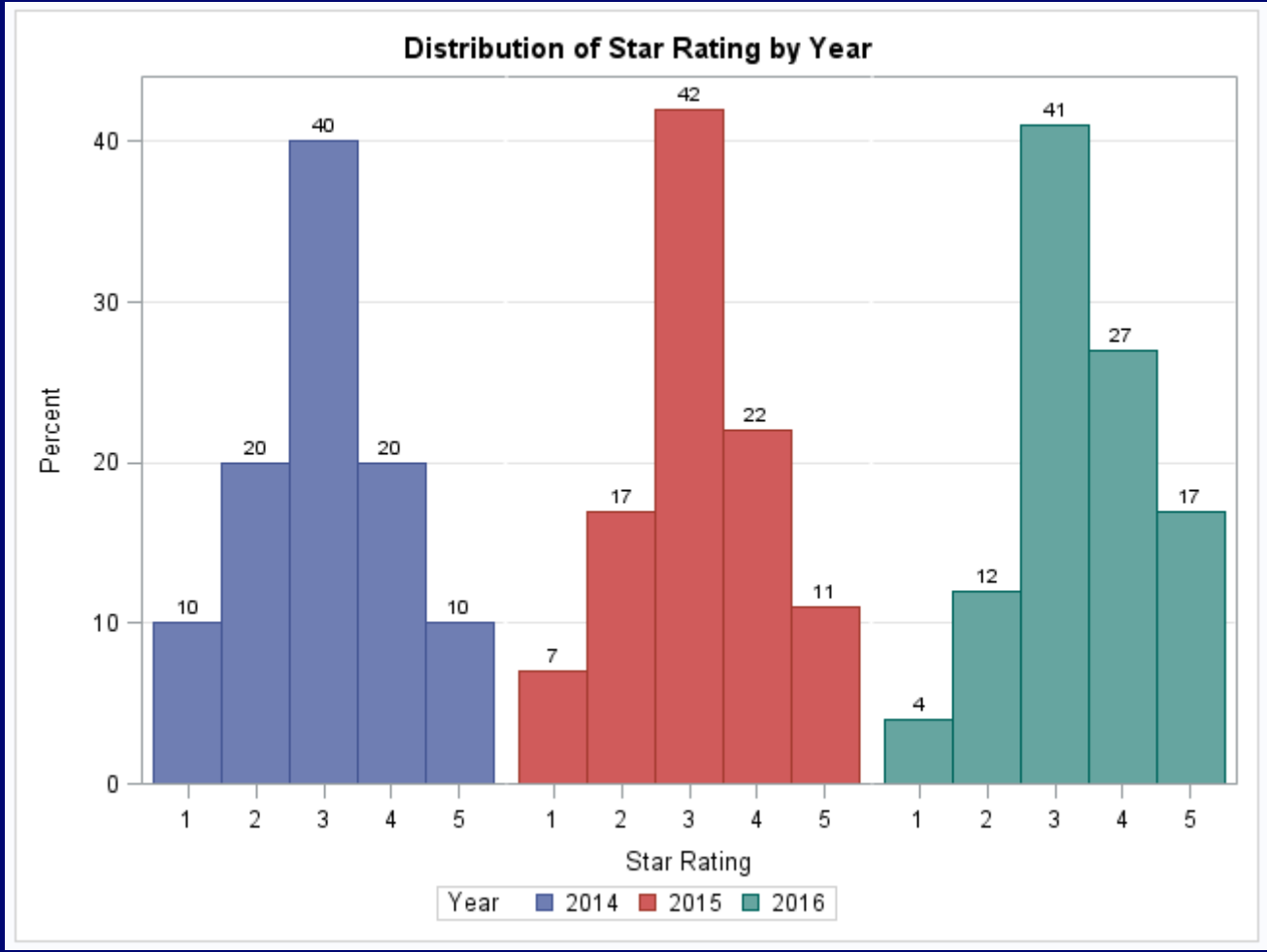
Discussion: Hypercalcemia and Input on Re-baselining

- Option 1
  - Include the updated version of Hypercalcemia in the next update of the Star Rating which will be implemented in October 2018
    - Issue: Update was not announced during NPC call
- Option 2
  - Do not include the updated version of Hypercalcemia in the 2018 update of the Star Rating, include the revised hypercalcemia in October 2019
    - Issue: Need to re-baseline star rating two years in a row (2018 and 2019)

Re-baselining

- Loss of ability to distinguish facilities' performance

Star Rating Shift



Wrap-up: Summary of Recommendations and Discussion of Next Steps

Public Comment Period

3:45pm – 4:00pm (EST)

Appendix

Overview of the DFC Star Rating Methodology

Dialysis Facility Compare Site Star Rating Description

- Intended to show how well a dialysis center delivers care and to make data on the quality of patient care easier to understand and use.
- Each dialysis center receives a rating between one and five stars.
- Patient survey results are not currently included in the star rating.

Rating Methodology Overview

- Highly related measures are grouped into domains using a technique called factor analysis so that no one measure overly influences the final score.
- Measure values are standardized to create measure scores so that they are on comparable scales
- Measure scores within each of the 3 resulting domains are averaged to determine a domain score.
- Domain scores are averaged to determine a final score.
- The final scores for all U.S. dialysis facilities are used to determine the cut points for different star categories.

Baseline Year

Purpose of setting up baseline year

- Baseline year data define star category cutoffs.
- Facility performance in subsequent years is scored against cutoffs defined in this baseline year.
- Definition of the "baseline year" allows reporting of changes in facility performance over time, which was requested by the prior TEP.

Baseline year vs current year

Baseline Year

- The collection year of data analyzed to set scoring standards for the DFC Star Rating

Current Year

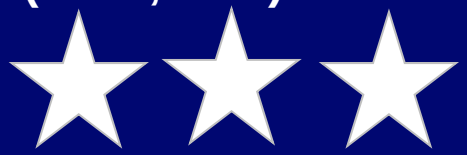
- The collection year of data being analyzed to evaluate facilities for the DFC Star Rating compared against scoring standards set in the baseline year

Scoring a current year using baseline year score cutoffs

Measure	SMR*	SHR*	STrR*	Fistula	Catheter	Kt/V	Hypercal
Values:	1.04	0.81	0.72	54.5%	11.8%	95.7%	1.0%
Measure Scores*:	-0.10	0.54	0.58	-0.92	-0.30	1.04	0.52
Domain Scores*: (-2.58, 2.58)			0.34		-0.61		0.78

Final Score*: (-2.58, 2.58) 0.17

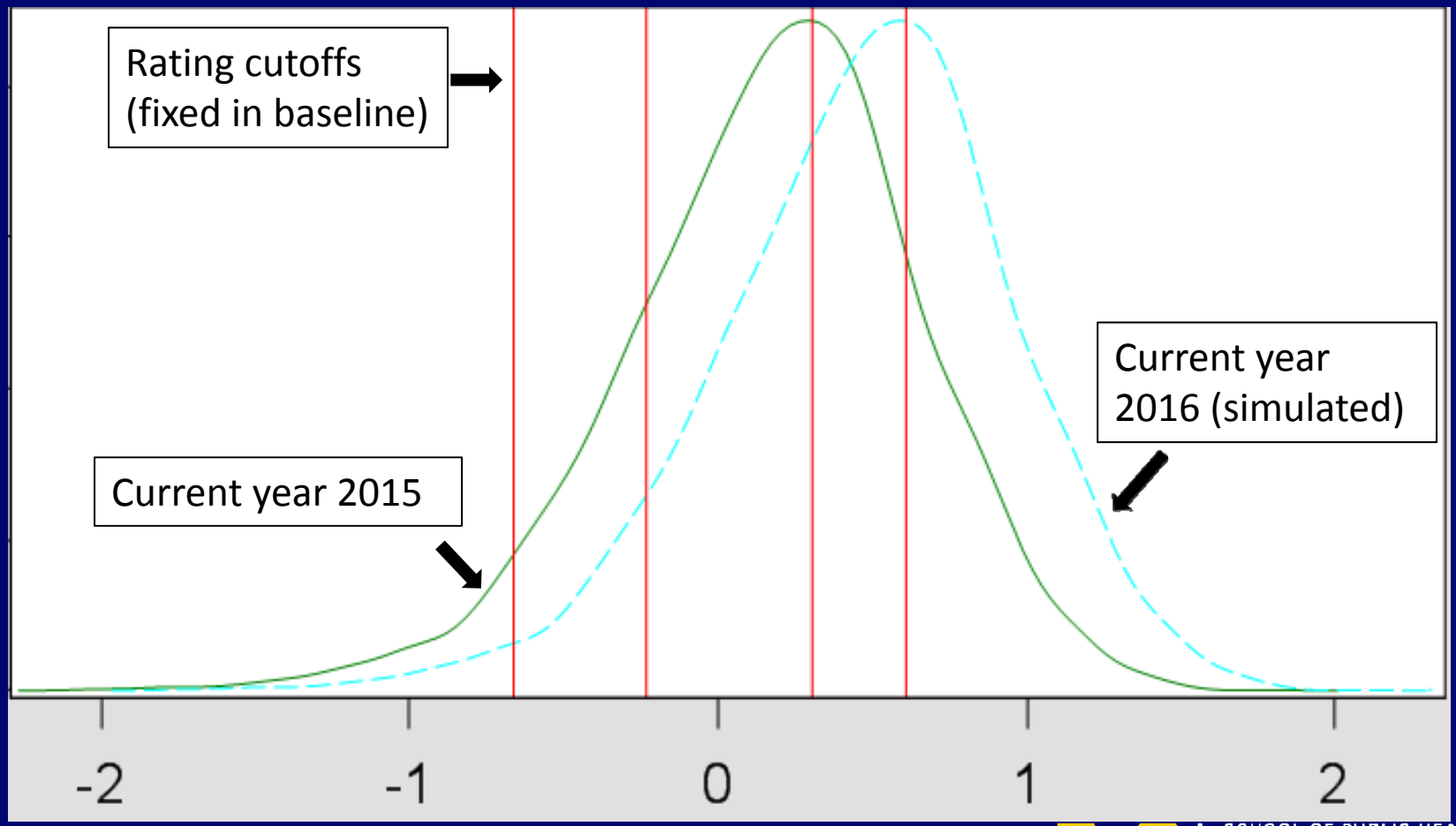
Star Rating*



***Values after application of adjustment factor**

SCHOOL OF PUBLIC HEALTH
KECC
UNIVERSITY OF MICHIGAN

Why was the baseline year implemented?



Quality Measures (QM) Used in Calculation of Star Ratings

- Standardized Transfusion Ratio (STrR)*
- Standardized Mortality Ratio (SMR)*
- Standardized Hospitalization Ratio (SHR)*
- Percentage of adult hemodialysis (HD) patients who had enough wastes removed from their blood during dialysis
- Percentage of pediatric hemodialysis (HD) patients who had enough wastes removed from their blood during dialysis
- Percentage of adult peritoneal dialysis (PD) patients who had enough wastes removed from their body during dialysis
- Percentage of adult dialysis patients who had hypercalcemia
- Percentage of adult dialysis patients who received treatment through arteriovenous (AV) fistula*
- Percentage of adult patients who had a catheter left in vein longer than 90 days for their regular hemodialysis treatment*

Notes:

The 3 dialysis adequacy quality measures were combined into a single weighted average of the 3 individual measures. This was needed because facilities with a very small number of PD patients or pediatric patients would not have enough data to calculate the individual measure score.

Adequacy as measured by Kt/V is reported on Dialysis Facility Compare separately for three groups of patients (children on HD, adults on HD, adults on PD)

*Updated Measure

DFC October 2018 Release
Star Ratings: Measure Updates

- Vascular Access: Standardized Fistula Rate
- Vascular Access: Long-term Catheter Rate
- Standardized Mortality Ratio (SMR)
- Standardized Hospitalization Ratio (SHR)
- Standardized Transfusion Ratio (STrR)

Measure Update: Vascular Access (Standardized Fistula Rate)

- This measure shows the adjusted rate of adult hemodialysis patient-months using an autogenous arteriovenous fistula (AVF) as the sole means of vascular access.
- Updates include calculation as a rate and adjustments for patient risk factors where fistula placement may be either more difficult or not appropriate
- Higher values are better.

Measure Update: Vascular Access (Long-term Catheter Rate)

- Adult patients who had a catheter (tube) left in a vein for 90 days or longer, for their regular hemodialysis treatments.
- Updates include exclusions for patients for whom other vascular access types may be either more difficult or not appropriate
- Lower values are better.

Measure Update: Standardized Mortality Ratio (SMR)

- The rate of patient deaths shows whether patients who were being treated regularly at a certain dialysis center lived longer than expected, don't live as long as expected, or lived as long as expected, compared to similar patients treated at other facilities. This measure is calculated as a ratio but is expressed as a rate on DFC.
- Updates include changes to incident comorbidity adjustments, additional adjustments for 210 prevalent comorbidities, and limiting the measure population to Medicare patients
- Lower values are better.

Measure Update: Standardized Hospitalization Ratio (SHR)

- The rate of hospitalizations shows whether patients who were being treated regularly at a certain dialysis center were admitted to the hospital more often, less often, or about the same, compared to similar patients treated at other centers. This measure is calculated as a ratio but is expressed as a rate on DFC.
- Updates include changes to incident comorbidity adjustments and additional adjustments for 210 prevalent comorbidities
- Lower values are better.

Measure Update: Standardized Transfusion Ratio (STrR)

- The rate of transfusions shows whether patients who were treated regularly at a dialysis center were transfused more often, less often, or about as often, compared to patients at other facilities. This measure is calculated as a ratio but is expressed as a rate on DFC.
- Updates include revision to the definition for a transfusion event
- Lower values are better.

Measure Update Discussion

DFC October 2018 Release
Star Ratings Candidate Measures

1. Standardized Readmission Ratio (SRR)
2. Pediatric Peritoneal Dialysis Adequacy: Achievement of Target Kt/V
3. In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH-CAHPS)
4. National Healthcare Safety Network (NHSN) Bloodstream Infection

Standardized Readmission Ratio: SRR

- The rate of (hospital) readmission shows whether patients who were being treated regularly at a certain dialysis center were readmitted more often, less often, or about the same, compared to similar patients treated at other dialysis centers. This measure is calculated as a ratio but is expressed as a rate on DFC.
- Lower numbers are better.

Pediatric PD Kt/V

- This shows the percentage of children (under 18 years of age) getting regular peritoneal dialysis treatments at a certain center whose average Kt/V was 1.8 or higher.
- Higher percentages are better.

New Measure Discussion

Wrap Up Discussion

- Wrap Up Discussion
- Any additional materials needed from UM-KECC or other measure developers?

Star Rating TEP
Teleconference Call #2

Teleconference #2 Agenda

- **Topic #1**: Review of Candidate Measures (continued from teleconference #1)
- **Topic #2**: How should Patient Experience with Care measure (ICH CAHPS) be Reported in the Star Ratings (Separate from or combined with Clinical Quality Measures)

In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

- The ICH CAHPS Survey:
  - Asks dialysis patients about their experiences while getting hemodialysis in their dialysis facility (center).
  - Focuses on patients who have gotten hemodialysis in their current facility for 3 months or longer.
  - Focuses on topics that are important to patients, such as how well their kidney doctors and dialysis center staff communicate with them.
  - Is conducted twice each year in spring and fall.
- An ICH CAHPS facility must have a minimum of 30 completed surveys over two survey periods for results to be reported on the DFC.

In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

- Three composite measures and three global ratings are reported on DFC.
- Composite measures
  - Kidney doctors communication and caring
  - Quality of dialysis center staff care and operations
  - Providing information

- Individual survey questions included in each composite measure are available on the ICH CAHPS website at https://ichcahps.org/Portals/0/ICH_Composites_English.pdf

In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

- The three ICH CAHPS global (individual) ratings are:
  - Rating of your kidney doctors
  - Rating of the dialysis center staff
  - Rating of the dialysis center

In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

- Top-box results are reported on DFC for the composites and global ratings
  - Global ratings: Top-box shows the percentage of patients who gave the most positive response: the percentage of patients who gave a rating of 9 or 10
  - Composite measures: Top-box shows the percentage of patients who gave the most positive response: the percentage of patients who responded “Always” or “Yes”

In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

- ICH CAHPS Survey results are statistically adjusted for type of survey administration (paper; telephone) and patient-mix at the facility.
- Patient-mix variables used for adjustment include patient demographics, patient clinical characteristics, and health status.

Bloodstream Infections (BSIs) in Hemodialysis Patients

- In 2014, there were approximately 29,500 BSIs in hemodialysis patients reported to NHSN
  - Three-quarters of these infections are considered access-related
- Bloodstream infections often lead to hospitalizations, can result in severe complications, and are a precursor to the most life-threatening infectious disease syndrome, namely sepsis.
- CDC and others have demonstrated preventability of these infections through improved infection prevention practices
- Several national quality improvement and prevention initiatives are focused on this outcome (e.g., CDC Making Dialysis Safer for Patients Coalition)

Bloodstream Infections (BSI) Standardized Infection Ratio (SIR)

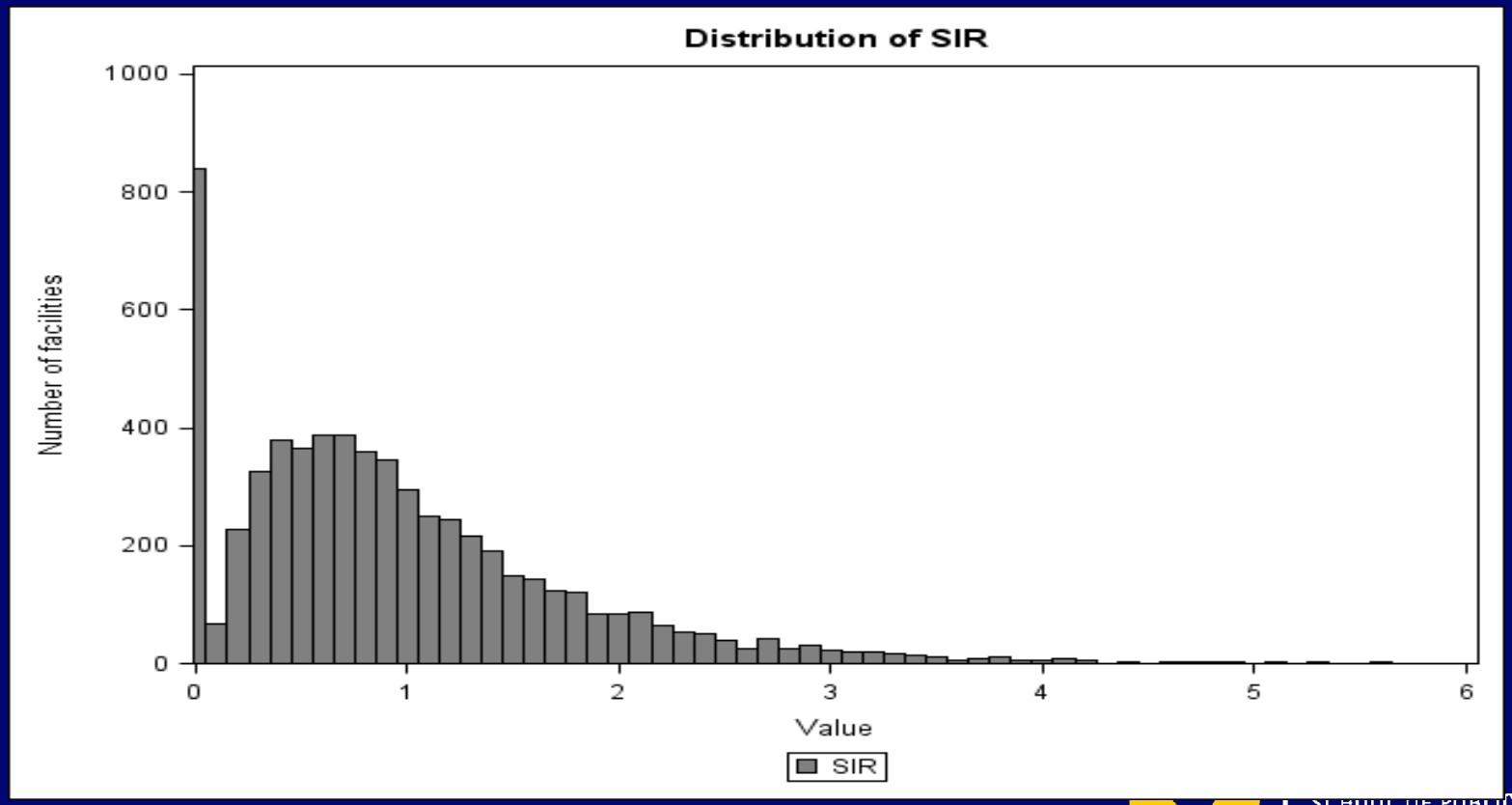
- The NQF-endorsed blood stream infection measure provides data for analysis and action that can drive improvements in infection prevention and patient safety.

National Healthcare Safety Network (NHSN) Bloodstream Infection Standardized Infection Ratio (SIR)

- The SIR of bloodstream infections is calculated among patients receiving hemodialysis at outpatient hemodialysis centers.
- The measure compares the observed number of bloodstream infections at a facility to the number of infections predicted for that facility based on the national aggregate data and vascular access types of patients in the facility.
- Lower numbers are better.

NHSN BSI SIR Distribution, 2015

- 6218 dialysis facilities reported data to NHSN in 2015



Candidate Measure Discussion

Topic #2

How to include potential Patient Experience of Care measures in the Star Ratings?

Results of Preliminary Analyses

Patient Experience of Care measures and Star Rating

Three Options:

1. One overall star rating combining Patient Experience of Care and Clinical Quality Measures
2. Two separate star ratings
   a. Overall Patient Experience of Care Rating
   b. Clinical Quality Measures
3. Three star ratings
   a. Overall Star Rating
   b. Overall Patient Experience of Care Rating
   c. Clinical Quality Measures

Correlations Between ICH CAHPS and Clinical Measures

	Nephrologist Communication/Caring	Staff/Facility Quality	Providing Information	Nephrologist Rating	Staff Rating	Facility Rating
Nephrologist Communication/Caring	1.00	0.48 (<.0001)	0.38 (<.0001)	**0.82 (<.0001)**	0.42 (<.0001)	0.43 (<.0001)
Staff/Facility Quality	0.48 (<.0001)	1.00	**0.54 (<.0001)**	0.43 (<.0001)	**0.84 (<.0001)**	**0.82 (<.0001)**
Providing Information	0.34 (<.0001)	**0.54 (<.0001)**	1.00	0.34 (<.0001)	**0.54 (<.0001)**	**0.52 (<.0001)**
Nephrologist Rating	**0.82 (<.0001)**	0.43 (<.0001)	0.34 (<.0001)	1.00	0.48 (<.0001)	0.48 (<.0001)
Staff Rating	0.42 (<.0001)	**0.84 (<.0001)**	**0.54 (<.0001)**	0.48 (<.0001)	1.00	0.89 (<.0001)
Facility Rating	0.43 (<.0001)	**0.82 (<.0001)**	**0.52 (<.0001)**	0.48 (<.0001)	0.89 (<.0001)	1.00
Standardized Transfusion Ratio	0.05 (0.0088)	0.05 (0.0083)	0.05 (0.0041)	0.02 (0.2401)	0.06 (0.0002)	0.07 (<.0001)
Standardized Hospitalization Ratio	0.04 (0.0164)	0.05 (0.0056)	0.09 (<.0001)	0.04 (0.0304)	0.06 (0.0006)	0.08 (<.0001)
Standardized Mortality Ratio	0.08 (<.0001)	0.10 (<.0001)	0.04 (0.0099)	0.07 (<.0001)	0.11 (<.0001)	0.11 (<.0001)
NHSN Bloodstream Infection	0.02 (0.3075)	-0.02 (0.1874)	-0.05 (0.0054)	-0.02 (0.2496)	-0.04 (0.0328)	-0.03 (0.0690)
Standardized Readmission Ratio	0.05 (0.0016)	0.07 (<.0001)	0.10 (<.0001)	0.08 (<.0001)	0.08 (<.0001)	0.10 (<.0001)
Fistula	0.04 (0.0099)	0.14 (<.0001)	0.10 (<.0001)	0.06 (0.0002)	0.15 (<.0001)	0.14 (<.0001)
Catheter	0.03 (0.0574)	0.00 (0.9955)	0.05 (0.0031)	0.02 (0.2460)	0.01 (0.5732)	0.02 (0.2081)
Hypercalcemia	0.01 (0.4681)	0.08 (<.0001)	0.10 (<.0001)	0.02 (0.1646)	0.09 (<.0001)	0.08 (<.0001)
Total Kt/V	0.08 (<.0001)	0.19 (<.0001)	0.11 (<.0001)	0.07 (<.0001)	0.19 (<.0001)	0.20 (<.0001)

Quality Measure Data from DFC Release: October 2016

Factor Analysis

- Factor analysis groups quality measures into domains based on relatedness
- Domains are statistically unrelated to each other
- Domains will be used to compute the final rating for each facility

Summary of Results

- Full factor analysis on the current measures, clinical candidate measures, and ICH CAHPS measures revealed strong correlation presenting two domains within the ICH CAHPS measures
- Subsequent separate factor analyses were performed on (1) the ICH CAHPS measures and (2) the clinical candidate measures + current measures

Summary of Results

- Results show strong correlation within the six ICH CAHPS measures, and a lack of correlation between the ICH CAHPS and clinical measures
- Factor analysis would support creation of 1-2 ICH CAHPS domains separate from the clinical measure domains
  - One Domain: All (6) ICH CAHPS Measures
  - Two Domains: (2) Nephrologist-Specific, (4) Others

Factor Analysis Results: Clinical Measures and ICH-CAHPS

The top three domains as determined by factor analysis:

- Domain 1:
  - Staff/Facility Quality
  - Providing Information
  - Staff Rating
  - Facility Rating
- Domain 2:
  - Nephrologist Comm./Caring
  - Nephrologist Rating
- Domain 3:
  - Standardized Transfusion Ratio
  - Standardized Hospitalization Ratio
  - Standardized Mortality Ratio
  - Standardized Readmission Ratio

Factor Analysis Results: ICH CAHPS Measures

The top two domains as determined by factor analysis:

- Domain 1:
  - Staff/Facility Quality
  - Providing Information
  - Staff Rating
  - Facility Rating
- Domain 2:
  - Nephrologist Communication/Caring
  - Nephrologist Rating

Factor Analysis Results: Clinical Measures

The top three domains as determined by factor analysis:

- Domain 1:
  - Standardized Transfusion Ratio
  - Standardized Hospitalization Ratio
  - Standardized Mortality Ratio
  - Standardized Readmission Ratio
- Domain 2:
  - Fistula
  - Catheter
- Domain 3:
  - NHSN Bloodstream Infection
  - Hypercalcemia
  - Total Kt/V

How to include potential Patient Experience of Care measures in the Star Ratings?

Option 1: One Overall Star Rating

- One overall Star Rating that combines Patient Experience of Care measure with Clinical Quality Measures

Option 2: Two Star Ratings

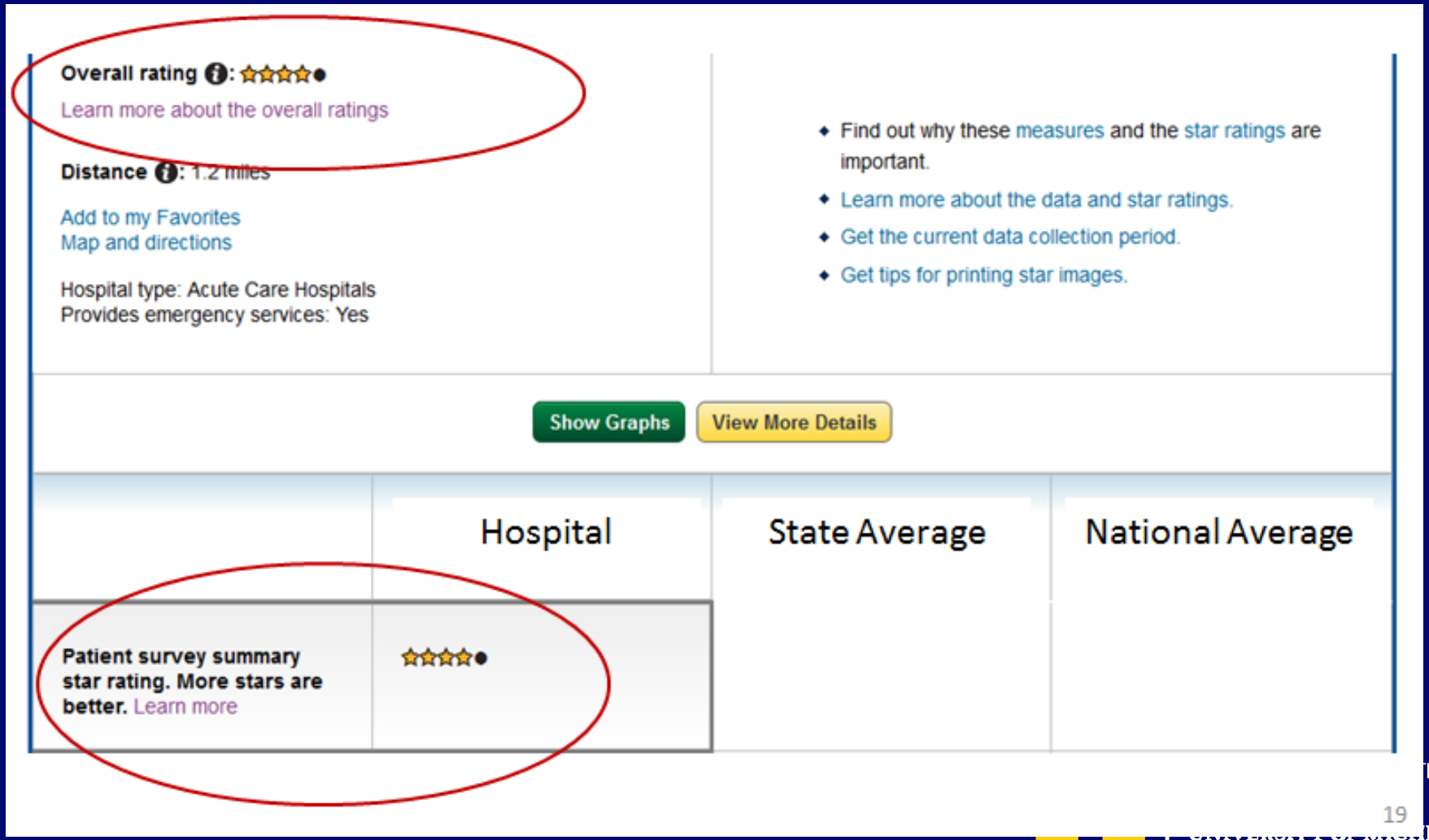
- Separate Star Ratings for ICH-CAHPS and the Clinical Quality Measures

Option 3: Three Star Ratings

- One overall Star Rating that combines Patient Experience of Care measure with Clinical Quality Measures
- Separate reporting of Star Ratings for ICH-CAHPS and the Clinical Quality Measures

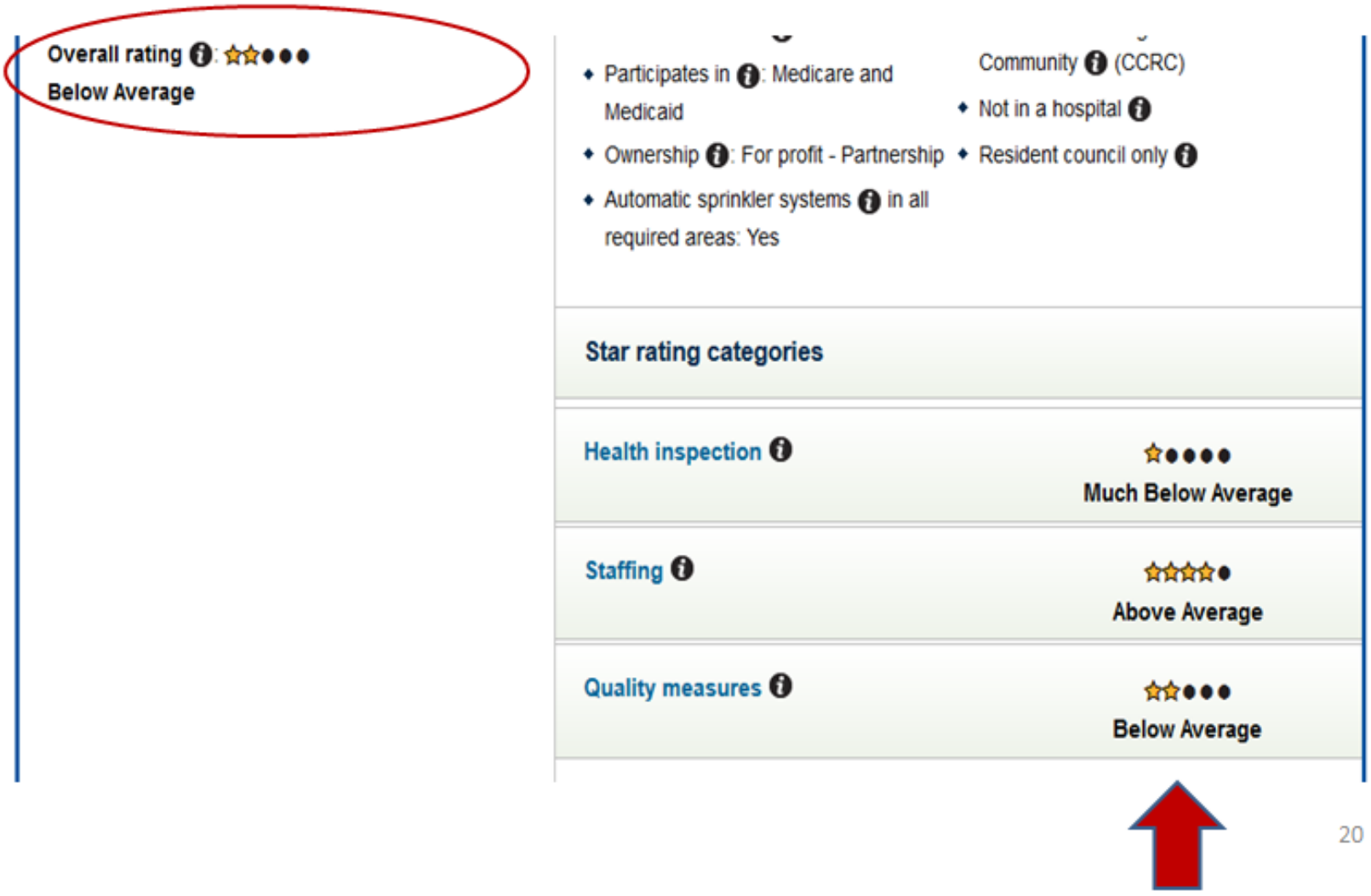
Hospital Compare

- There is one overall star rating and one star rating dedicated to patient survey summary



Nursing Home Compare

- There is an overall star rating and separate star ratings for three other categories of interest



Discussion: How to include potential Patient Experience of Care measures in the Star Ratings?

Wrap Up Discussion

- Wrap Up Discussion
- Identify issues and questions to be discussed at the in-person meeting or next TEP call
- Any additional materials needed from UM-KECC or other measure developers?

Star Rating TEP Teleconference Call #3

Teleconference #3 Agenda

- Brief review of baselining
- Re-baselining
- Discussion and wrap-up

Background: Baseline Year Implemented to Measure Performance of Facilities Over Time

Scoring with a baseline year allows:

- Establishment of final score cutoffs for star ratings that facilities can aim to achieve
- Same performance in different years to result in the same star rating
- Improved performance over time meaning facilities will not move down in their star rating
- **Summary**: Facilities that maintain same overall performance in reporting years will receive same star rating as in the baseline year

Changes in Measures in the Star Rating

- Substantive changes to the measures used in star ratings have consequences for interpretation
- Why?
  - Adding a quality measure that was not previously available does not allow recalculation of the baseline year
  - New measures may change grouping of measures into new domains
  - Baseline year cut-offs may not reflect quality of care measured with a revised measure set

Rating Shift Discussion

- By using cut-offs for scoring in a baseline year, improvements in individual measure values over time will always improve a facility's final score
  - This makes it harder to discriminate very good performing facilities from others
  - For example, if most facilities are clustered in the highest star rating categories, it is difficult for consumers to understand potential differences in quality among 4 or 5 star facilities.
- Must ask what quality information we want the Star Rating to convey to consumers

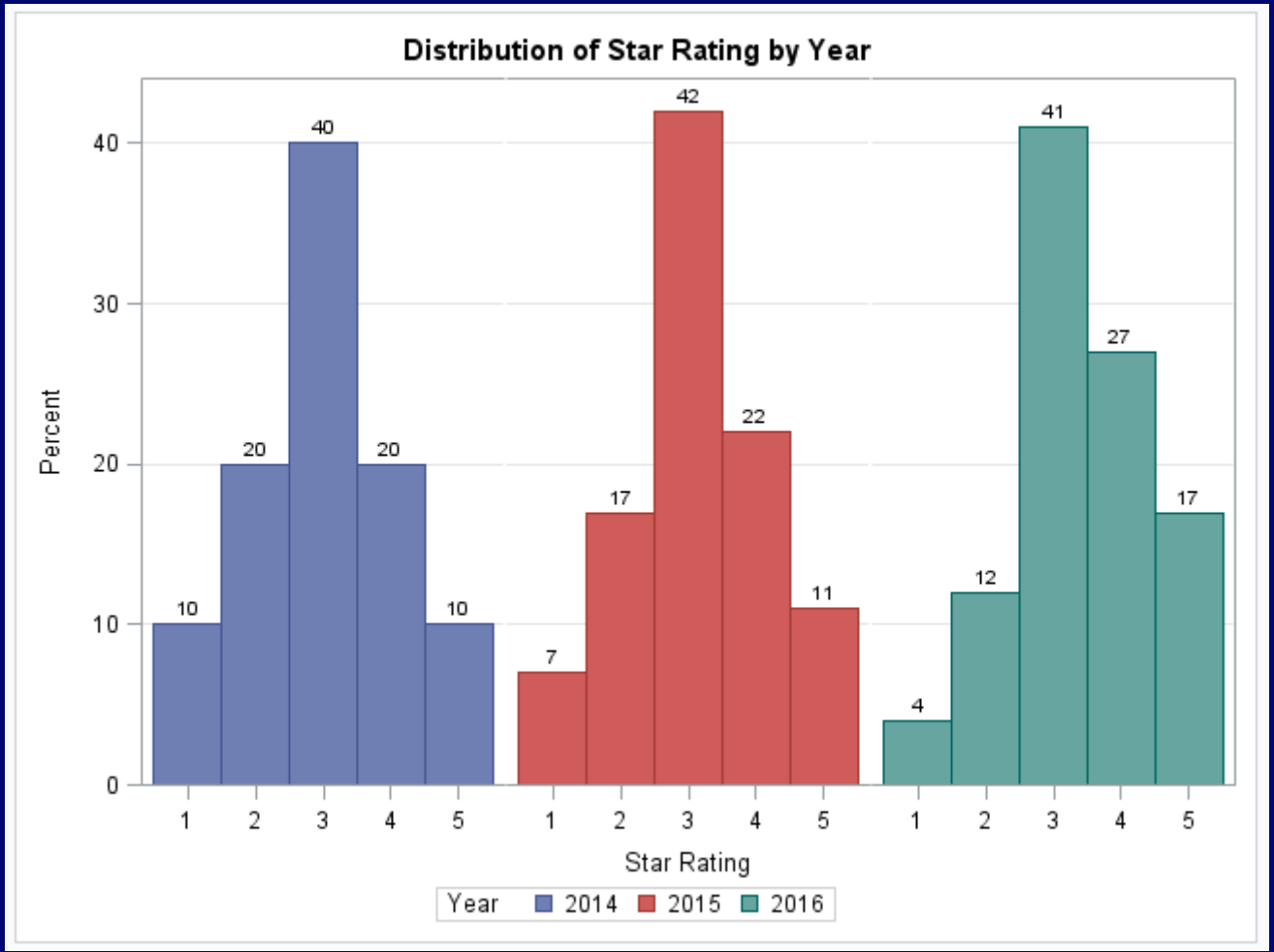
Star Rating Distribution by Year

Reporting Year	Star Rating					
	1 Star	2 Star	3 Star	4 Star	5 Star	Total
2014	**568 (10%)**	1138 (20%)	2274 (40%)	**1138 (20%)**	**568 (10%)**	5686 (100%)
2015	**419 (7%)**	1011 (17%)	2442 (42%)	**1312 (22%)**	**672 (12%)**	5856 (100%)
2016	**228 (4%)**	724 (12%)	2470 (41%)	**1624 (27%)**	**1015 (17%)**	6061 (100%)

Frequency Missing 1570

Table Cells Report: Number of Facilities (%)

Star Rating Shift



Discussion: When to establish a new baseline year?

- Potential Triggers for Re-baselining
  - When new measures are added or removed?
  - When measures are updated?
  - When 50% or more of facilities receive either a 4 or 5 star rating?

**Re-baselining means that final scores may correspond with different star rating categories compared to the previous baseline.*

Wrap-Up Discussion

- Identify issues and questions to be discussed at the in-person meeting
- Any additional materials needed from UM-KECC or other measure developers?
- Reminder to fill out preliminary voting form by end of day Wednesday February 15, 2017

Dialysis Facility Compare Website

www.medicare.gov/dialysisfacilitycompare/

Star Rating on DFC

Dialysis facility profile

Back to Results

General information | Survey of patients' experiences | Quality of patient care

UNIV OF MI DIALYSIS CLINICS - LIVONIA

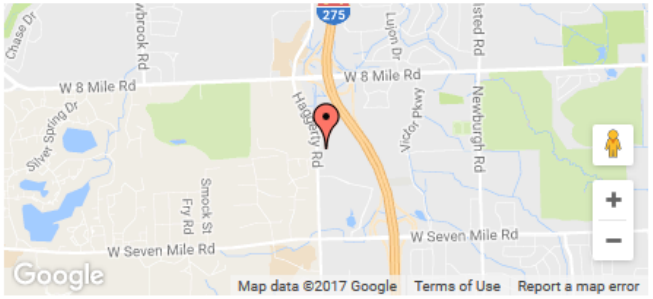
19900 HAGGERTY RD, SUITE 106
LIVONIA, MI 48152
(734) 432-7870

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

Learn more about the quality of patient care star rating

Distance: 7.5 miles

Add to my Favorites
Map and Directions



Dialysis center information

Learn why these characteristics and services are important.

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

Factor Analysis Results: Clinical Measures and ICH CAHPS

Rotated Factor Pattern										
	Factor1		Factor2		Factor3		Factor4		Factor5	
Nephrologist Comm./Caring	33	*	82	*	2		2		3	
Staff/Facility Quality	87	*	21		2		5		4	
Providing Information	56	*	20		11		6		-8	
Nephrologist Rating	36	*	81	*	3		2		0	
Staff Rating	92	*	18		3		9		3	
Facility Rating	89	*	19		6		8		6	
Standardized Transfusion Ratio	1		1		40	*	22		24	
Standardized Hospitalization Ratio	2		0		67	*	14		17	
Standardized Mortality Ratio	7		3		25	*	18		32	*
NHSN Bloodstream Infection	-4		-1		5		-7		25	*
Standardized Readmission Ratio	6		2		56	*	3		1	
Fistula	11		2		6		53	*	1	
Catheter	-3		3		12		55	*	-6	
Hypercalcemia	8		-3		7		12		-9	
Total Kt/V	20		0		22		-2		29	*

Printed values are multiplied by 100 and rounded to the nearest integer. Values greater than 0.25 are flagged by an '*'.

Quality Measure Data from DFC Release: October 2016

Factor Analysis Results: ICH CAHPS Measures

Rotated Factor Pattern					
	Factor1		Factor2		Factor3	
Nephrologist Comm./Caring	27	*	84	*	5	
Staff/Facility Quality	84	*	29	*	9	
Providing Information	56	*	25	*	9	
Nephrologist Rating	29	*	84	*	-4	
Staff Rating	90	*	25	*	-6	
Facility Rating	88	*	27	*	-9	

Printed values are multiplied by 100 and rounded to the nearest integer. Values greater than 0.25 are flagged by an '*'.

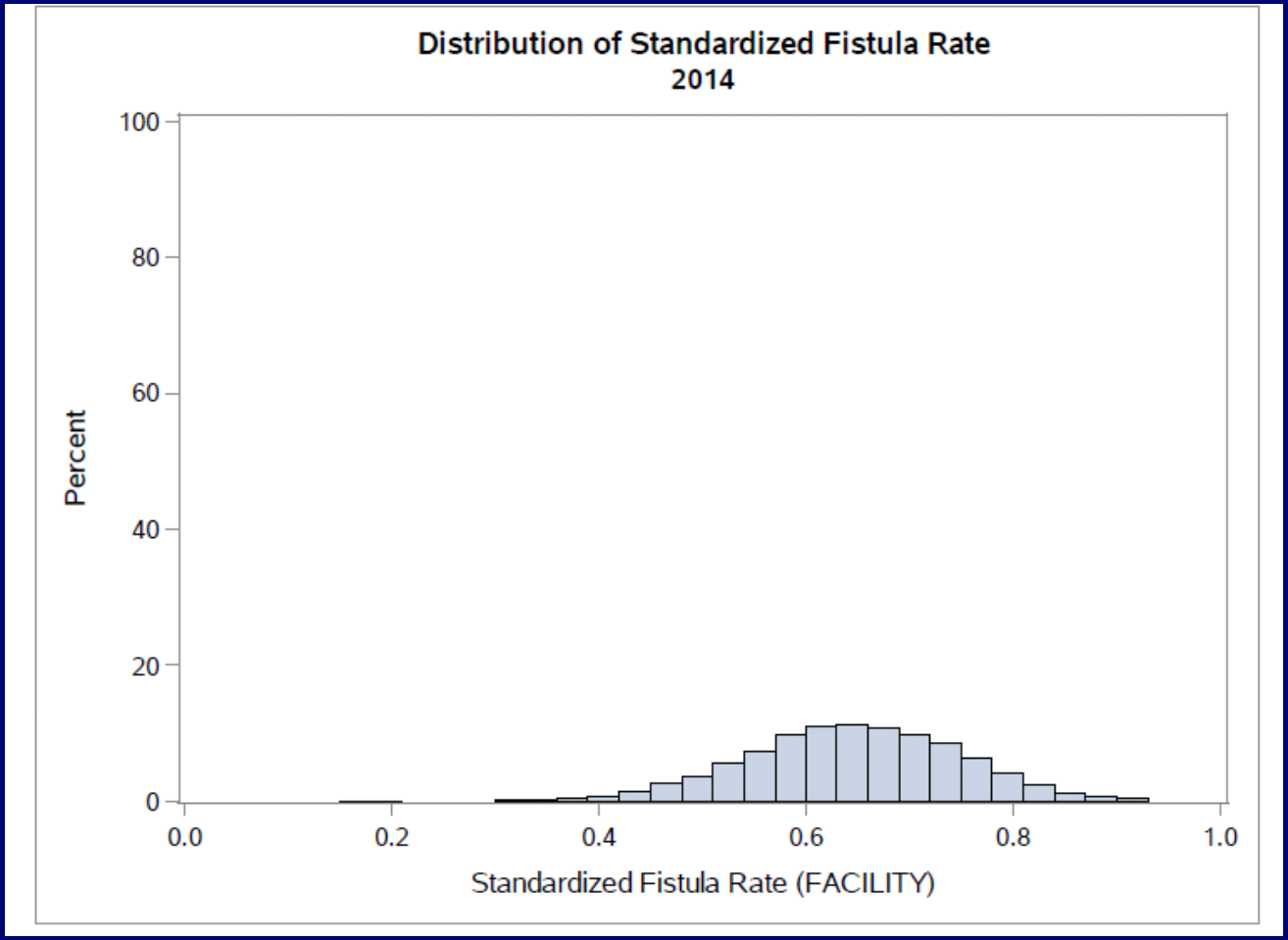
Factor Analysis Results: Clinical Measures

Rotated Factor Pattern								
	Factor1		Factor2		Factor3		Factor4	
Standardized Transfusion Ratio	43	*	18		24		6	
Standardized Hospitalization Ratio	67	*	16		19		10	
Standardized Mortality Ratio	26	*	13		23		20	
NHSN Bloodstream Infection	5		-5		21		-4	
Standardized Readmission Ratio	55	*	4		3		5	
Fistula	9		52	*	-1		11	
Catheter	11		54	*	-4		11	
Hypercalcemia	3		11		-4		18	
Total Kt/V	22		2		21		19	

Printed values are multiplied by 100 and rounded to the nearest integer. Values greater than 0.25 are flagged by an '*'.

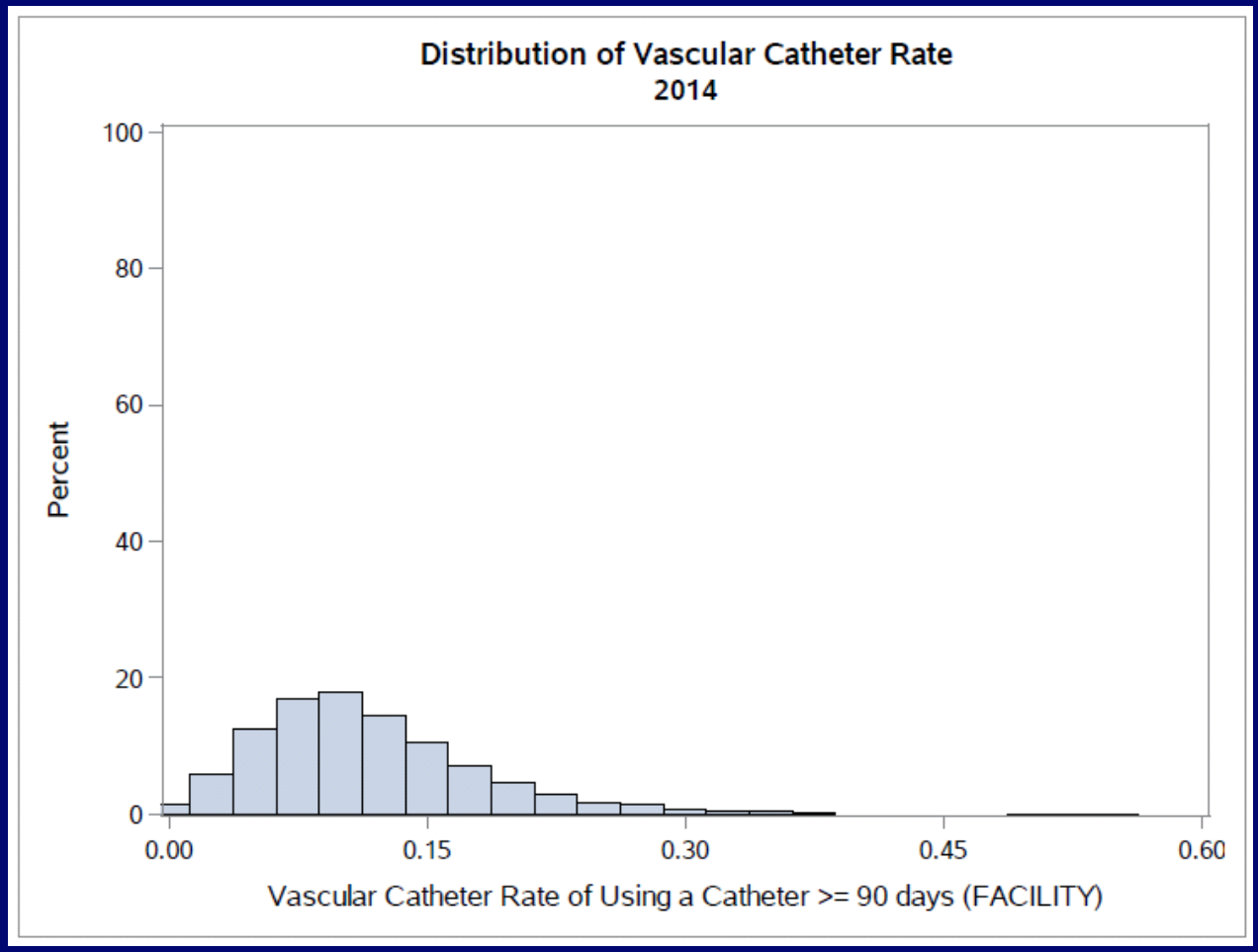
Quality Measure Data from DFC Release: October 2016

Distribution: Standardized Fistula Rate*



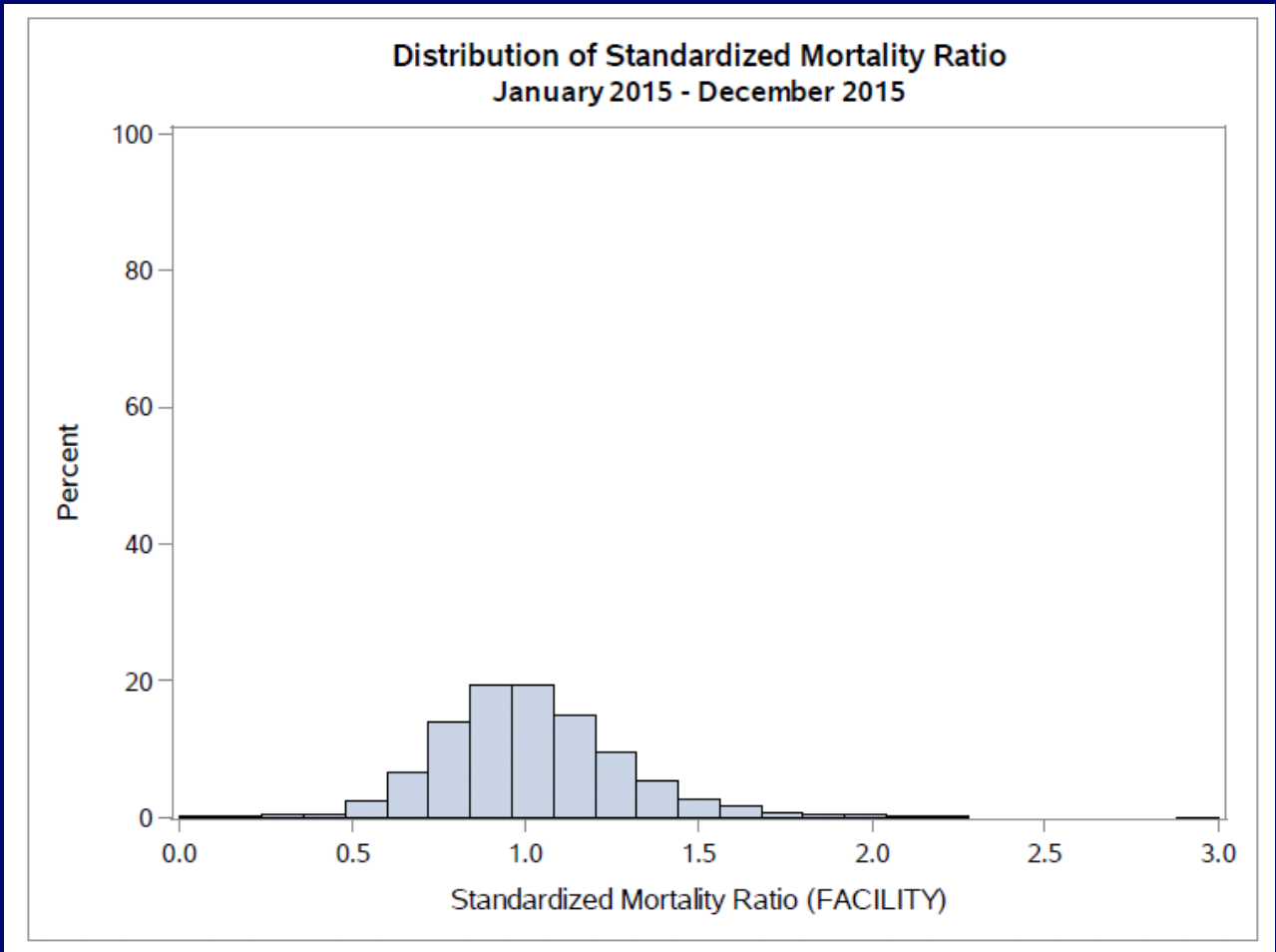
*Excludes facilities with < 11 eligible patients

Distribution: Long-term Catheter Rate*



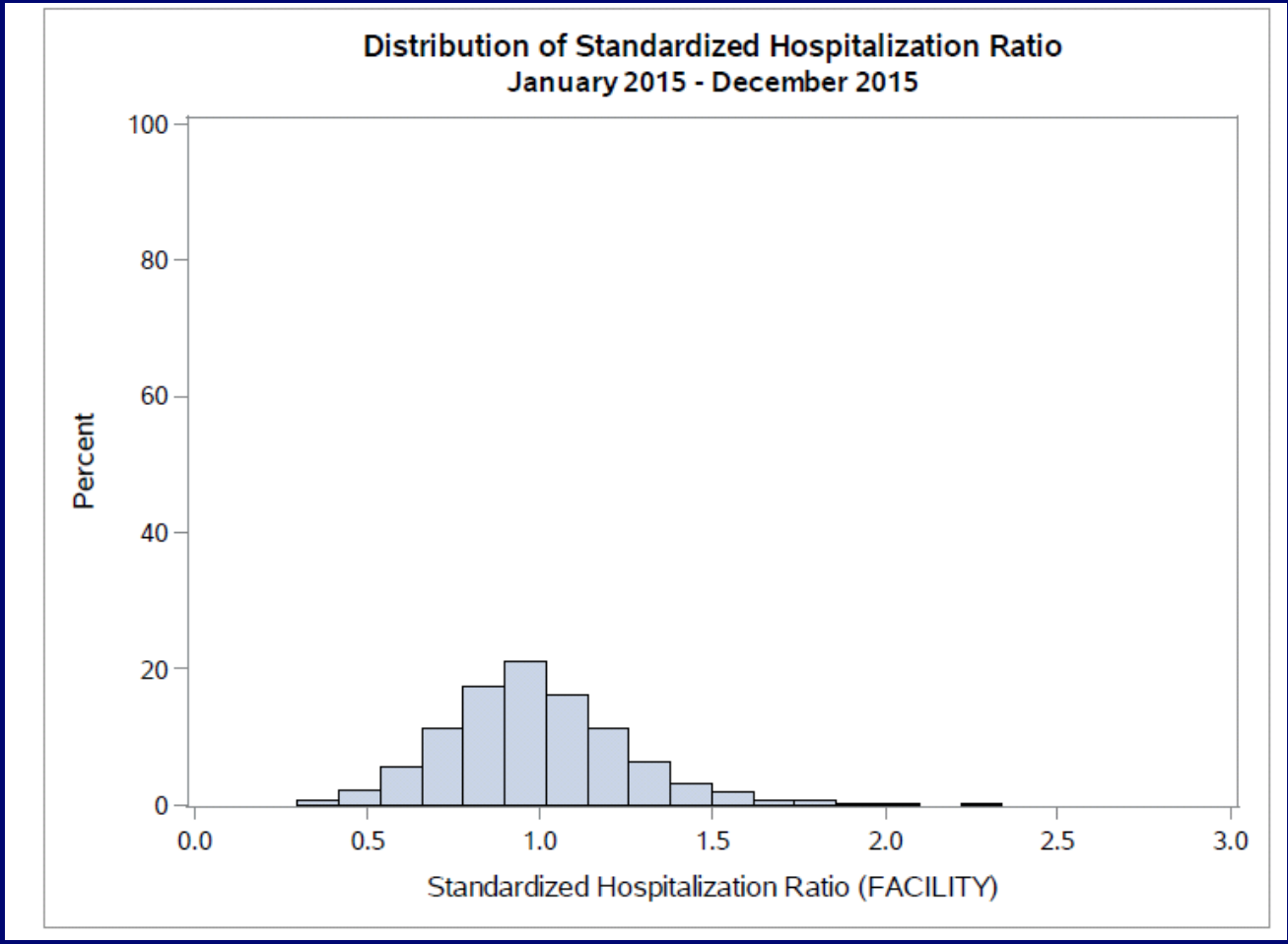
*Excludes facilities with < 11 eligible patients

Distribution: Standardized Mortality Ratio (SMR)*



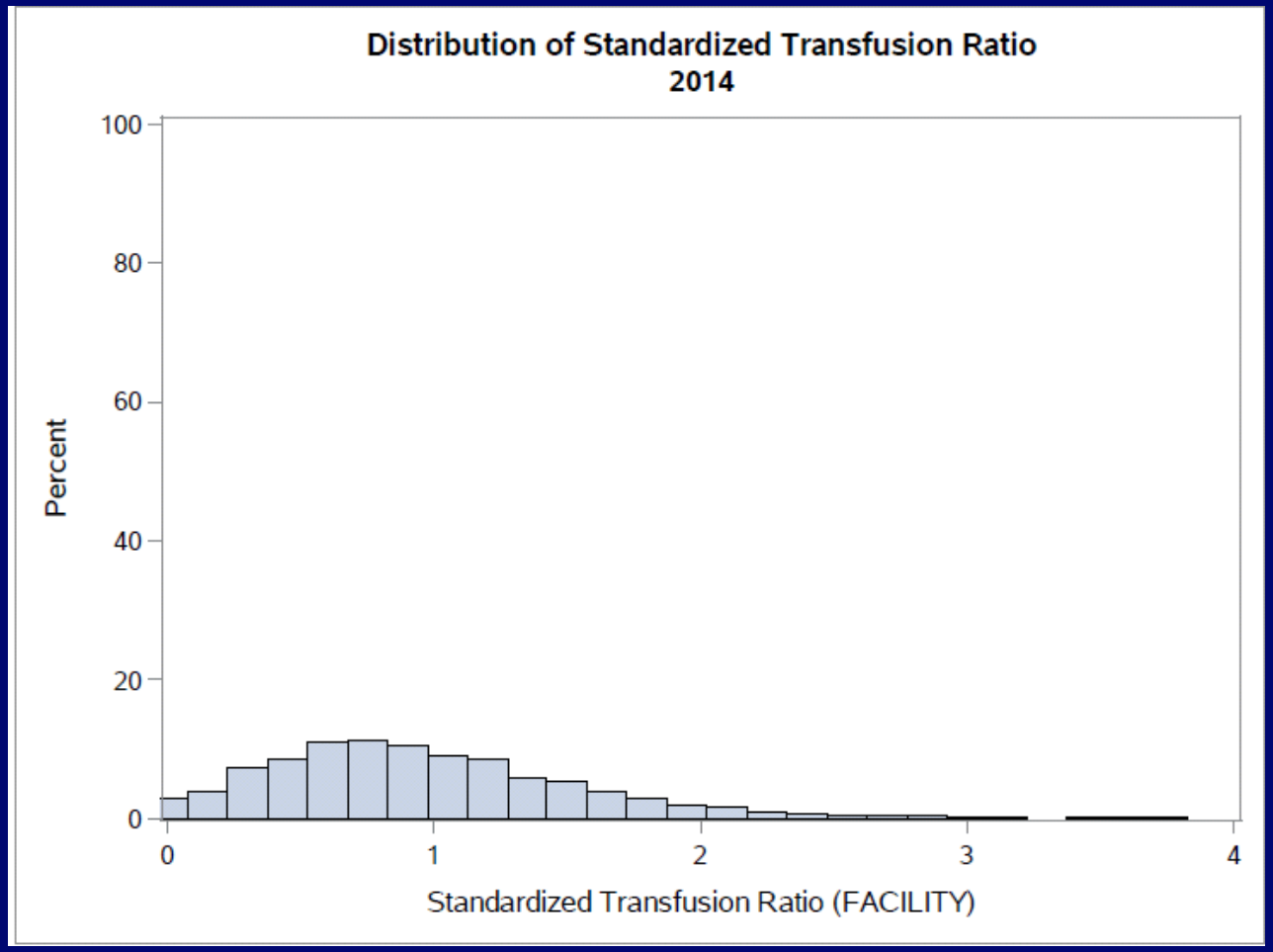
*Excludes facilities with < 3 expected deaths

Distribution: Standardized Hospitalization Ratio (SHR)*



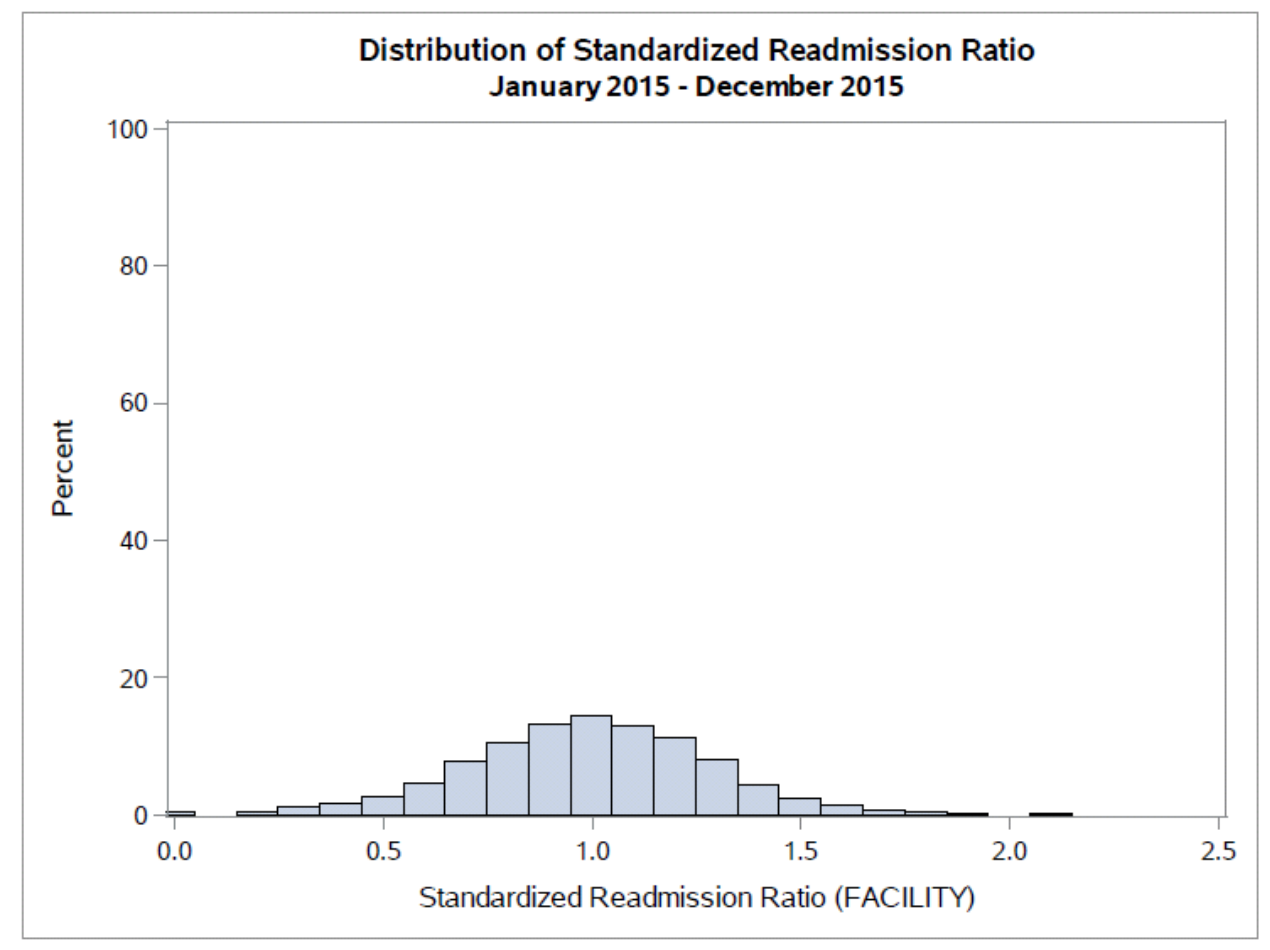
*Excludes facilities with < 5 patient-years at risk

Distribution: Standardized Transfusion Ratio (STrR)*



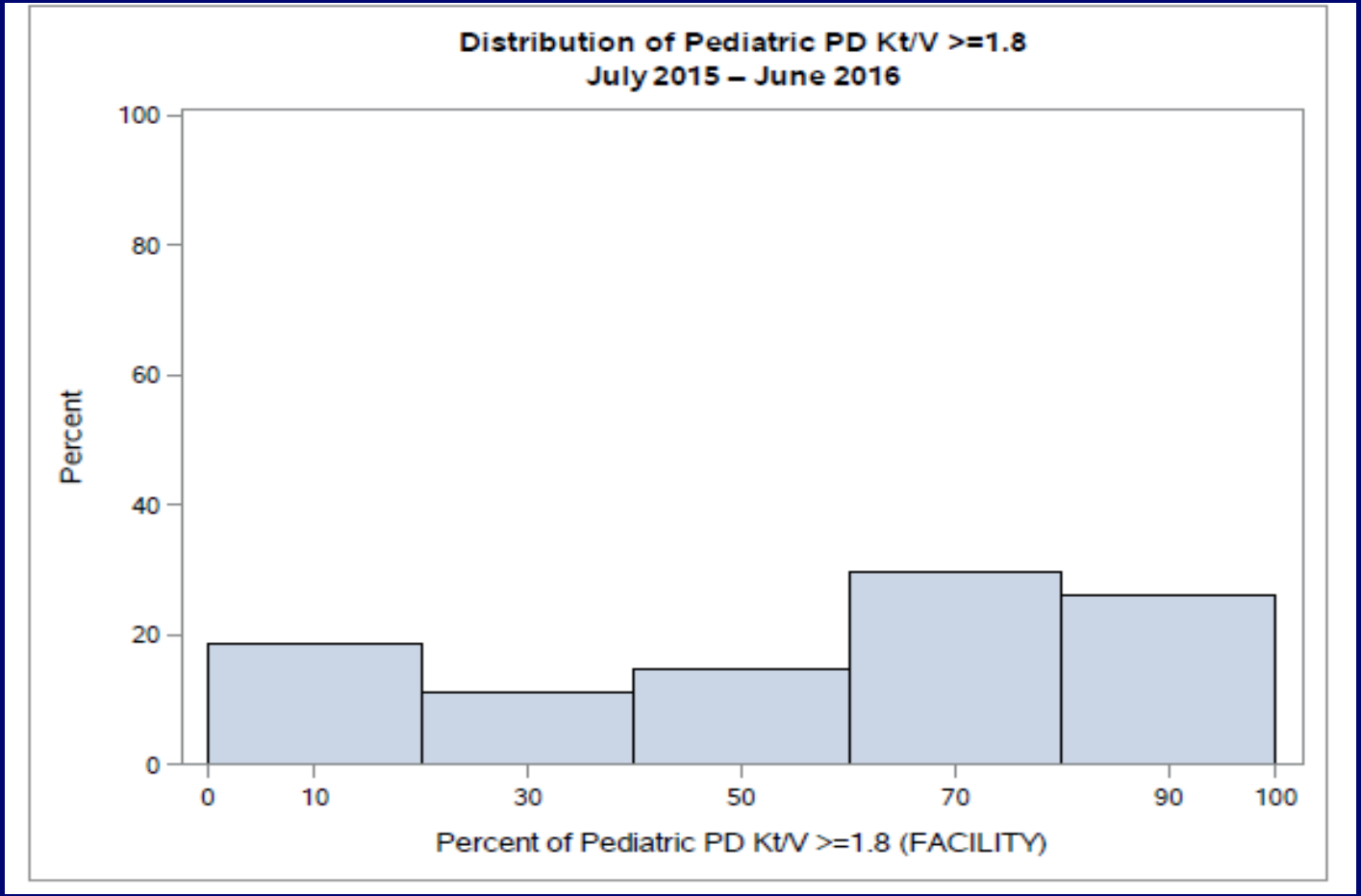
*Excludes facilities with < 10 patient years at risk
* Outlier facilities not shown (N=13)

Distribution: Standardized Readmission Ratio (SRR)*



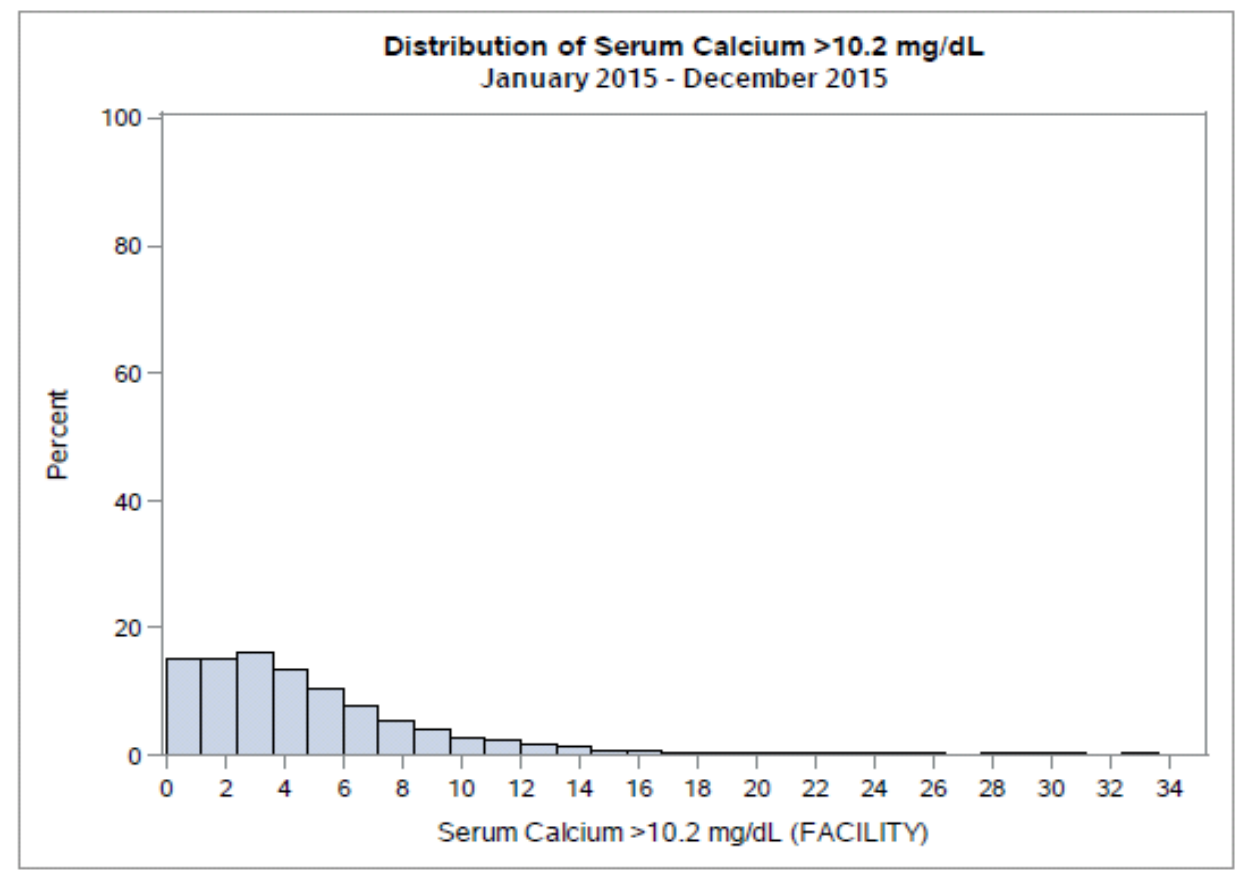
*Excludes facilities with <11 index discharges

Distribution: Pediatric PD Kt/V*



*Analysis restricted to facilities with at least 11 pediatric PD patients (n=25)

Distribution: Hypercalcemia*



*Excludes facilities with < 11 eligible patients

* Outlier facilities not shown (N=131)

In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

The 14 adjusters are:

- Mode of survey administration
- Overall health
- Overall mental health
- Difficulty dressing or bathing
- Age
- Sex
- Education
- Heart disease

In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

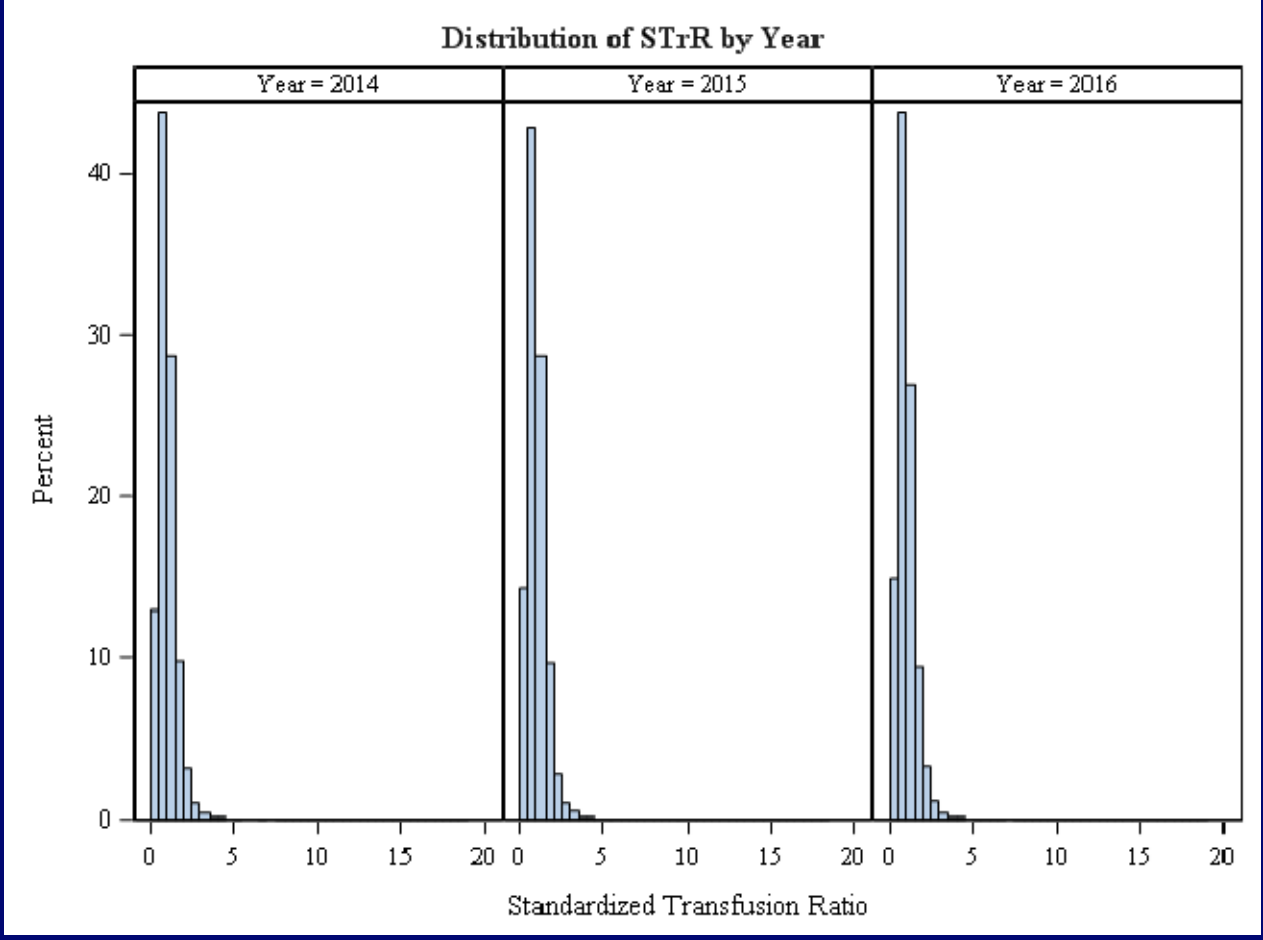
The 14 adjusters (cont'd):

- Difficulty concentrating, remembering, or making decisions
- Deaf or serious difficulty hearing
- Blind or serious difficulty seeing
- Speaking a language other than English at home
- Received help completing the survey
- Total number of years on dialysis

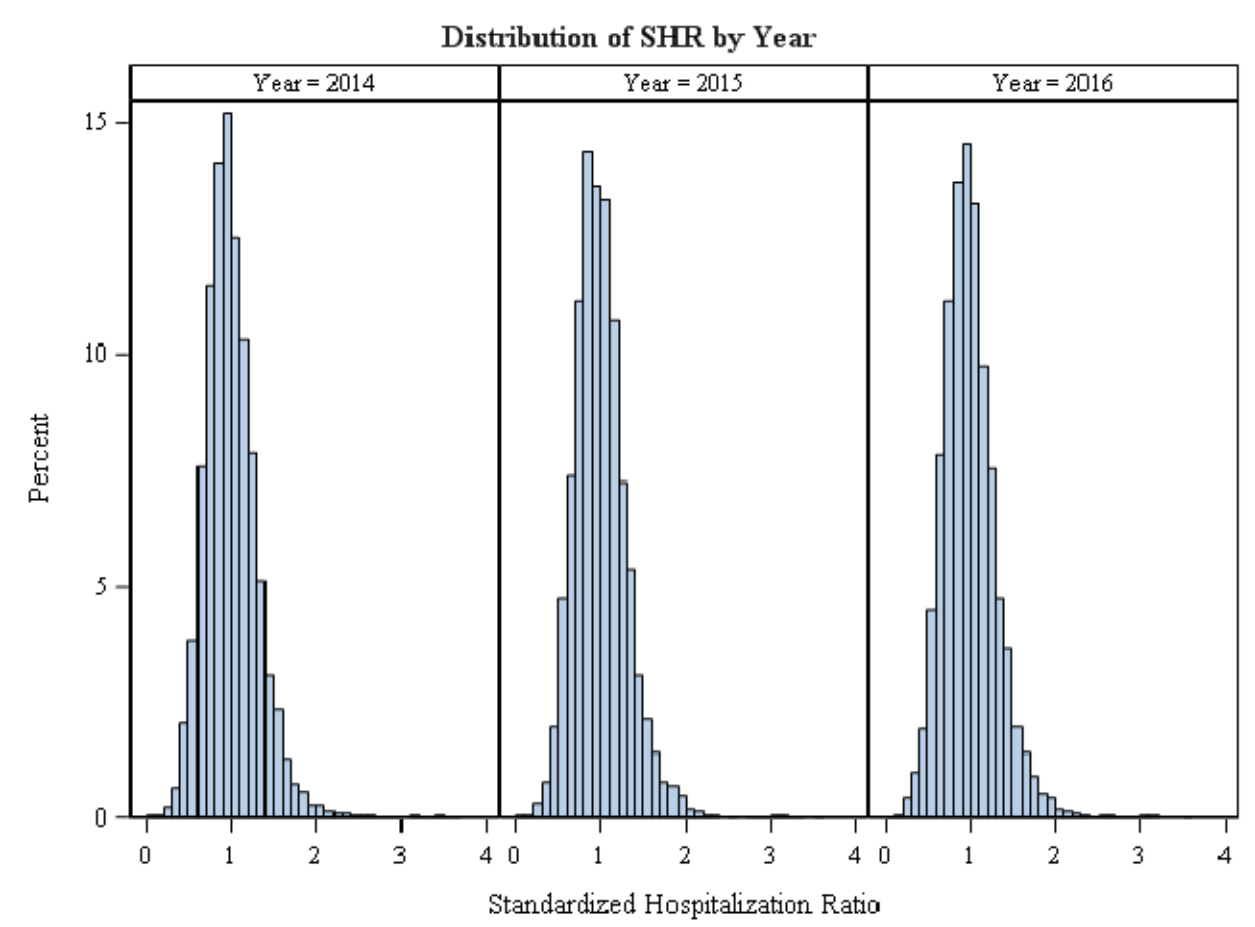
In-Center Hemodialysis Consumer Assessment of Healthcare Providers and Systems (ICH CAHPS)

ICH CAHPS Percentiles						
Facility Percentile	Nephrologists' Communication and Caring	Quality of Dialysis Center Care and Operations	Providing Information to Patients	Rating of the Nephrologist	Rating of the Dialysis Center Staff	Rating of the Dialysis Facility
	TOP-Box Score					
95th (near best)	80	74	87	80	81	84
90th	77	71	86	77	77	81
75th	72	66	82	70	71	74
50th	66	61	79	62	62	65
25th	60	56	75	54	54	56
10th	54	51	71	47	46	48
5th (near worst)	50	49	68	42	41	42
	BOTTOM-Box Score					
5th (near best)	7	8	13	4	3	2
10th	9	10	14	6	4	4
25th	12	13	18	10	8	6
50th	17	17	21	15	12	11
75th	22	21	25	20	18	17
90th	27	26	29	27	23	22
95th (near worst)	31	29	32	31	27	27

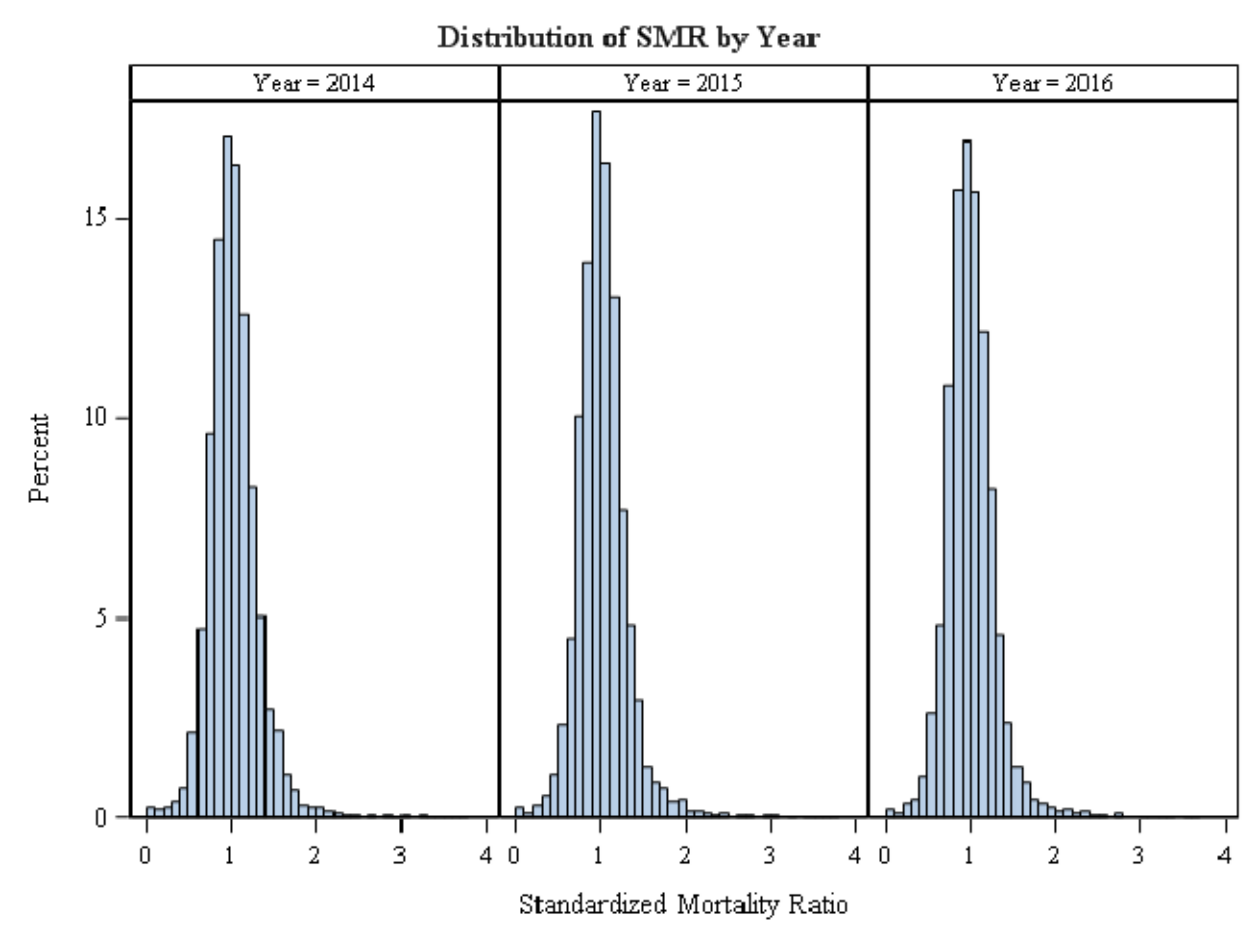
Three-Year Distribution Shift: Standardized Transfusion Ratio



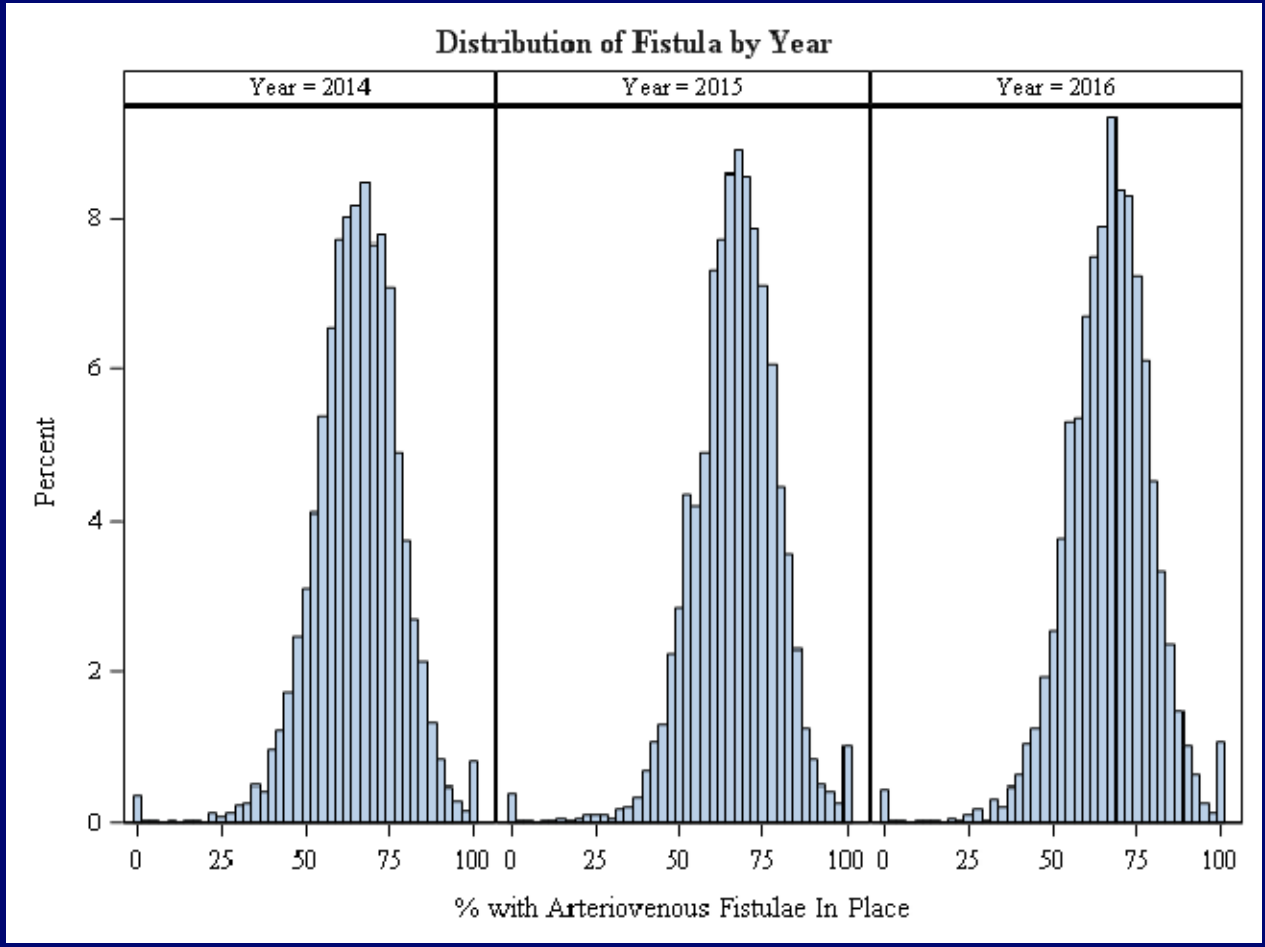
Three-Year Distribution Shift: Standardized Hospitalization Ratio



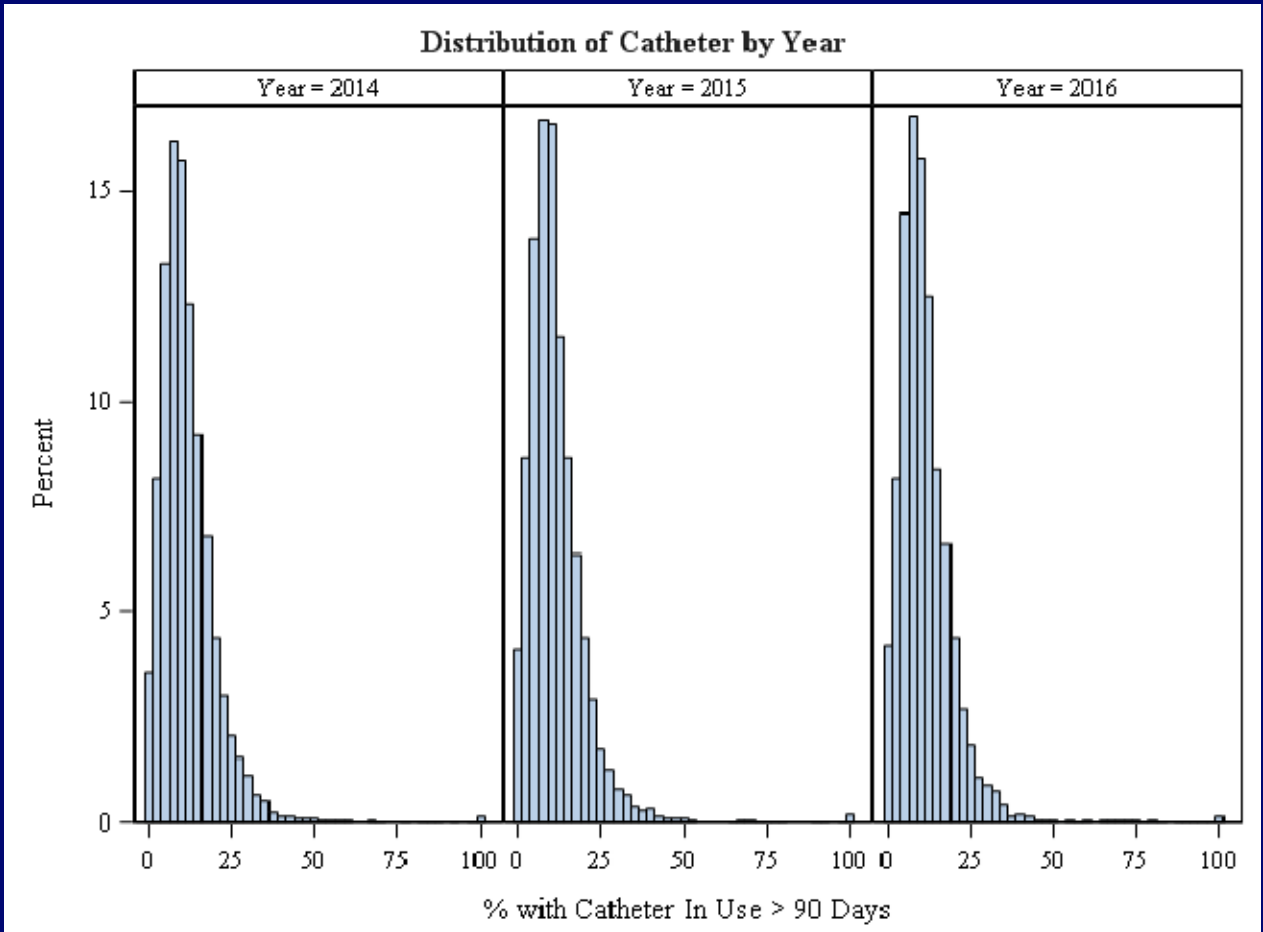
Three-year Distribution Shift: Standardized Mortality Ratio



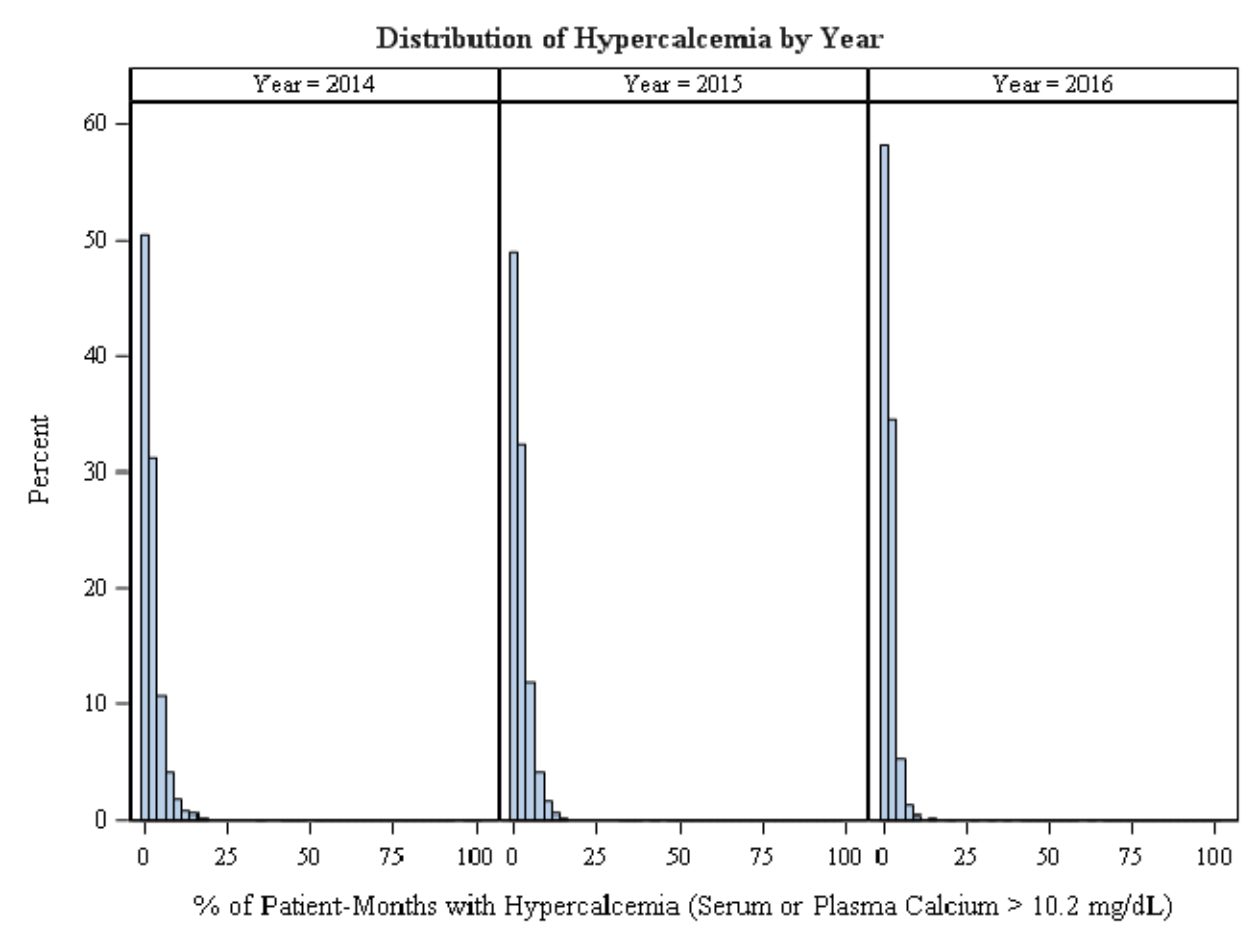
Three-Year Distribution Shift: Fistula



Three-Year Distribution Shift: Catheter



Three-Year Distribution Shift: Hypercalcemia



Three-Year Distribution Shift: Total Kt/V