



December 2, 2016

Measure Applications Partnership
c/o National Quality Forum
1030 15th Street, NW - Suite 800
Washington, DC 20005

Subject: Pre-Meeting Comment on Measures Under Consideration

Thank you for the opportunity to comment on the Measures Under Consideration (MUCs) prior to the Workgroup and Coordinating Committee meetings. Kidney Care Partners (KCP) is a coalition of members of the kidney care community that includes the full spectrum of stakeholders related to dialysis care – patient advocates, health care professionals, dialysis providers, researchers, and manufacturers and suppliers – organized to advance policies that improve the quality of care for individuals with chronic kidney disease and end stage renal disease (ESRD). We greatly appreciate the MAP undertaking this important work.

Three MUCs submitted to the MAP by the Centers for Medicare and Medicaid (CMS) (dated December 1, 2016) are proposed for use in the ESRD Quality Incentive Program (QIP), and consequently are of particular interest to KCP. While no corresponding NQF measure numbers were indicated in the MUC List, we note that the measures' specifications are identical to those recently reviewed in NQF's Renal 2016-2017 Project and thus make the presumption that the technical details of the associated risk models, where applicable, also are unchanged. In reviewing these measures, we offer the following comments.

- **MUC 16-305 – Standardized Transfusion Ratio for Dialysis Facilities (STrR; NQF 2979).** KCP opposes this measure. As for the NQF Renal 2016-2017 Project, KCP again expresses 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. CMS does not identify a minimum sample in the specifications, although the developer's empirical testing clearly demonstrates poor reliability in the facilities with small samples in which the testing was conducted. 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 the hospital- and physician-related transfusion practices that are out of dialysis facility control.
- **MUC 16-308 – Hemodialysis Vascular Access: Standardized Fistula Rate (NQF 2977).** KCP supports MUC16-308, but recommends the developer consider modifications to improve the measure prior to incorporation into the ESRD QIP portfolio of measures:
 - KCP believes the specifications are 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 graft or catheter that is no longer 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.

- KCP believes two additional vasculature risk variables would strengthen the model: a history of multiple prior accesses and the presence of a cardiac device.
- **MUC 16-309 – Hemodialysis Vascular Access: Long-Term Catheter Rate (NQF 2978).**
KCP supports MUC16-309.

In addition to the above, concerns about several of the technical details of the STrR and Standardized Fistula Rate measure remain unresolved. We include these in an attachment to this letter.

KCP again thanks you for the opportunity to comment on this important work. If you have any questions, please do not hesitate to contact Lisa McGonigal, MD, MPH (lmcgon@msn.com or 203.530.9624).

Sincerely,

A handwritten signature in cursive script that reads "Sara-Love Rawlings". The signature is written in black ink and is positioned below the word "Sincerely,".

Sara-Love Rawlings
Executive Director

ATTACHMENT 1: TECHNICAL COMMENTS ON MUC 16-305 AND MUC 16-308

MUC 16-305 – Standardized Transfusion Ratio for Dialysis Facilities (STrR; NQF 2979)

Again, while details of the measure's risk model and code sets do not appear to have been included with the MUCs List, we note that the measures' specifications are identical to those recently reviewed in NQF's Renal 2016-2017 Project and thus make the presumption that the technical details of the measure are also unchanged. We note during the prior (2015) NQF renal project, the Standing Committee reviewed the STrR (then as NQF 2699) and did not recommend it. As we discuss further in the section on Validity, we do not believe the new measure (MUC16-305/NQF 2979) sufficiently addresses our and the Committee's concerns about hospital- and physician-related factors. We comment on the specifications, reliability, validity (risk model), and harmonization issues.

- **SPECIFICATIONS.** In MUC13-305 (NQF 2979), CMS has revised the original (NQF 2699) measure specifications to more "conservatively" define transfusion events, such that all inpatient transfusion events must include, at a minimum, an appropriate ICD-9 Procedure Code or Value Code to be captured in the measure – inpatient transfusion events for claims that include only 038 or 039 revenue codes without an accompanying procedure or value code are not captured in the numerator. The specifications also specify a maximum of one event per day and that an event not be defined by the number of units of blood transfused.

KCP supports and appreciates the need to refine and tighten how transfusion events are counted and applauds CMS's intent in undertaking these revisions, but we do not believe the proposed solution is a valid representation of transfusion events.

Importantly, there is no existing coding requirement that procedure or value codes be used, which means valid transfusion claims that include only revenue codes will be missed. KCP believes the proposed specification changes result in a measure with significant threats to validity.

Current transfusion coding practices clearly vary by hospital,¹ and hospital coding practices are beyond dialysis facilities' sphere of control. For example, we are aware of hospitals that exclusively use revenue codes and do not use the procedure or value codes. In-patients at this type of hospital will appear to have no transfusion events assigned to the dialysis facility, whereas those at a hospital that uses the procedure and/or value codes will have recorded events. Simply put, facilities within given catchment areas will be differentially affected by hospital coding variations, which clearly impact measure scoring. We are particularly concerned that the revisions, if implemented, will result in increased variability in performance across dialysis facilities wholly due to external factors and not performance. Facilities will appear to have "poor" performance because of higher than expected numbers of transfusions – and will expend time and resources to improve – when in fact the score is merely a reflection of coding practices.

Again, KCP strongly supports the need to refine how transfusion events are defined, and we urge the MAP to recommend the developer continue considering alternative models to define transfusion events. Alternatively, the MAP could suggest that CMS consider revising hospital transfusion coding rules to require that the ICD-9/ICD-10 procedure and value codes necessary for the validity of the proposed methodology be

¹ Weinhandl ED, Gilbertson DT, Collins AJ. Dialysis facility-level transfusion rates can be unreliable due to variability in hospital-level billing patterns for blood. Chronic Disease Research Group poster, ASN. 2014.

universally included in claims.

Additionally, the testing documentation submitted for review in the NQF Renal 2016-2017 Project notes that facilities with 10 or fewer patients were excluded, but we note the specifications did and still do not state this. Again, KCP believes that a minimum size exclusion should be indicated and, as the developer's results document, and we discuss in the following section, reliability is poor even when the facility size is significantly greater than 10 patients.

The NQF submission also indicated the minimum data requirement for the STrR is 10 patient-years at risk, which differs from CMS's Standardized Hospital Ratio (SHR) measure, which uses 5 patient-years at risk. No justification or empirical analyses are offered to justify the selected threshold or the difference.

Finally, the STrR specifications submitted to NQF indicate the measure can be expressed as a rate, but is calculated as a ratio. KCP prefers normalized rates or year-over-year improvement in rates instead of a standardized ratio. We believe comprehension, transparency, and utility to all stakeholders is superior with a scientifically valid *rate* methodology.

- **RELIABILITY.** In addition to our concerns that the specifications pose a threat to the validity of the updated STrR, KCP also has concerns about the reliability testing for these revised specifications, as submitted to NQF.

KCP notes a reliability statistic of 0.70 is often considered as "good" reliability, though the characterization also depends on the analytic method. Reliability testing, overall, for the STrR yielded IURs of 0.60-0.66 across all facilities for each of 2011, 2012, 2013, and 2014. Such values indicate about 65% of the variation in a score can be attributed to between-facility differences (signal) and about 35% to within-facility differences (noise)—a moderate degree of reliability. However, when looking exclusively at small (defined as ≤ 46) and medium (47-78) facilities, the IURs are substantially lower. Specifically, the IURs ranged from 0.30-0.41 and 0.50-0.56 for small and medium facilities, respectively, over the same time period. KCP thus believes the specifications must specifically require a minimum sample as identified through the developer's empirical testing.

- **VALIDITY.** In addition to KCP's concerns about the specifications and the threat to validity of variable capture of transfusion events depending on hospital coding practices, KCP has several concerns about the covariates (or lack thereof) and risk model.

NQF did not endorse the STrR in 2015, in part because the Renal Standing Committee raised concerns that the measure did not adjust for hospital- and physician-related transfusion practices. Physicians independently, or following hospital protocols, make decisions about whether or not to transfuse a specific patient, so it is important to account for the variability these factors create. The revised measure does not incorporate these factors into the risk model, so KCP's concurrence with the Renal Standing Committee's original concern remains.

KCP notes that while CMS's Standardized Mortality Ratio (SMR) and SHR measures have been revised to incorporate prevalent co-morbidities into their risk models, the STrR has not been so revised; only incident co-morbidities, derived from the Medical Evidence Form (CMS 2728), are considered. This approach means the STrR risk model only reflects those conditions present upon when the patient initiates dialysis; failure to

appropriately account for prevalent co-morbidities is a threat to validity. In the harmonization section, we also note that CMS adjusts for 2728-derived co-morbidities for SHR and SMR differently than it does for the STTrR. Finally, as we have noted before, we continue to be concerned about the validity of the 2728 as a data source and urge that the Committee recommend that CMS assess this matter.

KCP notes that the validity testing for the measure, as submitted to NQF, indicated an overall c-statistic of 0.65. We are concerned the model will not adequately discriminate performance – particularly that smaller units might look worse than reality. We believe 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.

- **HARMONIZATION ISSUES.** The new SMR and SHR risk models adjust for each incident co-morbidity (from the 2728) separately, instead of using a “co-morbidity index.” The models also approach diabetes as a single co-morbidity rather than four separate indicators (currently on insulin, on oral medications, without medications, diabetic retinopathy). The STTrR has not been similarly revised. KCP urges the MAP to recommend that the developer harmonize the STTrR with these other measures so that each incident co-morbidity is examined separately (i.e., unbundled, as compared to the current measure) and diabetes is approached as a single co-morbidity (i.e., bundled, as compared to the current risk model).

The risk models for the groupings used for patient age and duration of ESRD differ among the SMR, SHR, and STTrR. For example, the age groups for the SMR is n=3, but for the SHR and STTrR the age groupings are the same, but n=6. Similarly, the number of groups for ESRD duration for the SMR (n=4) differs from that for the SHR (n=6). No justification or empirical analyses are offered to justify these differences.

There also are significant inconsistencies in how facility size is defined when assessing reliability for the SMR, SHR, and STTrR. Specifically, for the SMR, the definitions were <=45, 46-85, >=86 for the 1-year reliability analyses, but were <=135, 136-305, and >=306 for the 4-year analyses. For the SHR, <=50, 51-87, and >=88 were used. Finally, for STTrR reliability analyses, small, medium, and large facilities were defined as <=46, 47-78, and >=79, respectively. We understand reliability for a given measure depends on sample size, but find the varying demarcations analytically troubling. We posit a more appropriate analytic approach would be to analyze reliability using consistent “bins” of size (i.e., small, medium, and large are consistently defined) and identify the facility size at which reliability for that particular measure can be confidently inferred – and then reflect the minimum size in the actual specifications.

MUC16-308 – Hemodialysis Vascular Access: Standardized Fistula Rate (NQF 2977)

We offer the following technical comments on the specifications and validity of the Standardized Fistula Rate measure:

- **SPECIFICATIONS.** The language in the prior version of the measure (NQF #0257) that specifically defines an autogenous AVF as using two needles has been replaced in MUC16-308 (NQF 2977) with an autogenous AVF “as the sole means of vascular access.” KCP believes the specifications are 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 graft or catheter that is no longer being used. We note patients with catheters remain at risk for infection and other adverse sequelae, so credit should not be not 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.

- **VALIDITY.** KCP believes MUC16-308 (NQF 2977) improves on NQF 0257. While details of the risk model do not appear to have been included in the MUCs List, our review of the measure as submitted to NQF in its recent Renal 2016-2017 Project indicated that the developer accepted KCP’s recommendation in previous comments to remove the covariate alcohol dependence from the model variables; we commend the developer for this revision. 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.

However, KCP notes that the validity testing results submitted to NQF indicated an overall c-statistic of 0.71. We are concerned the model will not adequately discriminate performance – particularly that smaller units might look worse than reality. We believe 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.