NATIONAL QUALITY FORUM
Driving measurable health improvements together

Renal Fall 2022 Submissions Cycle: Pre-evaluation Comments

Contents

Funded by the Centers for Medicare & Medicaid Services under contract HHSM-500-2017-00060I Task Order HHSM-500-T0001

https://www.qualityforum.org
1099 14th Street NW, Suite 500 | Washington, DC 20005 | M 202.783.1300 F 202.783.3434

NQF #3719 Prevalent Standardized Waitlist Ratio (PSWR) (Centers for Medicare & Medicaid Services)

Pre-evaluation Standing Committee Comments

1a. Evidence

- The measure is listed as an outcome measure but only one element of the numerator is an outcome - LRD tx. Waitlisting is a process leading to tx. In order to be transplanted a patient must be on a waitlist for DD kidneys. If waitlisting requires optimizing patient health status for some patients it may reflect an intermediate outcome.
- I am not aware of any new investigations that would change the evidence for this study. From a patient's perspective, this is a sufficient study for providing transparent transplant waitlist information.
- Most of the evidence provided shows the benefits of kidney transplantation compared to ongoing dialysis and that there are a variety of factors (some related to the dialysis facility itself) that influence transplant rates. The data presented is all rather tangential as to whether the dialysis facility or the dialysis physician has any significant influence on the actual wait-listing of a patient or decision to move forward with a living transplant other than the need to be referred for transplant to start the process going forward, but this is such a very proximal step to a very distal outcome, and the factors with more immediate weight on successful listing or transplantation are not controlled by the dialysis facility or physician but seem more firmly rooted in the realm of the transplant facility and transplant physician.
- Low. Evidence does relate directly to the outcome of waitlisting. However - - I disagree strongly with the premise that all patients as defined in the denominator would be best treated with a kidney transplant. Patient choice is not considered. For many patients, a choice of dialysis is a better one than transplant. I therefore believe the evidence is applied to the wrong denominator of patients.
- New health outcome measure. Evidence applies directly.
- Measure well supported by evidence.
- Strong evidence noted.
- No.
- Ok.
- Evidence supports that waitlisting is important for transplant and transplant has better outcomes than dialysis.
- The evidence supports that there is at least one healthcare action associated with transplant waitlisting (submitted as an outcome measure).
- This is a new outcome measure at the clinician group/practice level that tracks the number of prevalent dialysis patients in a practitioner group who are under the age of 75 and were listed on the kidney or kidney-pancreas transplant waitlist or received a living donor transplant. For each practitioner group, the Prevalent Standardized Waitlist Ratio (PSWR) is calculated to compare the observed number of waitlist events in a practitioner group to its expected number of waitlist events. The developer provided evidence that the provider can impact if a patient is on the kidney-pancreas transplant list or receives a transplant by educating on transplantation, referring them to a transplant center, assisting with the transplant evaluation process, and optimizing patient's health and functional status.
- The proposed measure is classified as an OUTCOME MEASURE. I do not agree with the classification but if accepted as an outcome measure i think the evidence to support is

moderate.

- Yes, adequate evidence was provided to warrant this measure.
- Evidence is clear that education and engaged staff promote waitlisting.
- Passes on evidence. We can do better at educating qualified patient on transplantation. More education is needed from the dialysis facility team.
- Tangential - measures nephrologists' ability to get patient listed; nephrologist does not control most of the process.
- Measure seems well researched. It may be good to assess if the 75 years of age cutoff is still the right number or not in today's world.

1b. Gap in Care/Opportunity for Improvement and Disparities

- Yes. Gender, ethnicity, API v White. IQ range 63%
- Disparities were addressed through the risk assessment.
- A variability in care as it relates to this measure was provided as well as some data on disparities based on ethnicity and sex.
- There is clearly wide variability in wait listing. There is a performance gap - - if there is value in comparing wait listing with the current denominator. I do not believe the current denominator serves patient best care...so the measured performance gap is not relevant to my commitment to best patient care.
- Developer evaluated PSWR performance scores for all practitioner groups that had at least 11 pts and 2 expected events in the evaluation period between 1/1/2017 and 12/31/2019 resulting in the inclusion of 2,022 groups totaling 362,093 patients. Mean value of PSWR was 103 percent with an interquartile range of 63 percent. PSWR scores by race, ethnicity and sex for sample. Other categories 437 percent; Asian Pacific Islander 424 percent, White 121 percent. Non-Hispanics PSWR 107 percent; Hispanics 164 percent. Females PSWR 118 percent; males 98 % Opportunity for improvement is moderate.
- Large performance gap.
- There is room for improvement in this area.
- No.
- Ok.
- Data was provided to show that some groups have lower rates of waitlisting and there was also data demonstrating disparities, particularly based on race.
- Timeframe examined 2017-2019; Number of Patients: 362,093; Number of Groups: 2,022; Median 0.94 (25th percentile 0.67, 75th percentile 1.30); Performance differed by race, ethnicity, and gender. In examination of disparities, notable differences between some of the mean and median PSWRs.
- The developer did provide performance gap data for race, ethnicity and sex and does show disparities.
- Moderate.
- Information about subgroups included race, ethnicity and sex. The care gaps between white and other groups, non-Hispanics and Hispanics and males and females are wide enough to warrant this measure.
- A 103 percent main and 63 percent range in performance warrants consideration for a performance measure.
- Current data was provided with a large majority of practice groups meeting the "as expected" waitlist rates. Disparities were addressed which show some disparity in care.
- Modest variability; less than 10% of practices are outside the expected performance.
- Yes, there is a moderate variability gap.

2a. Reliability - Specifications

- All ok.
- There are no concerns about the implementation of this study.
- The measure is intended for both adults (< 75 years old) and children of any age. I assume adults at least 75 years old are excluded due to a higher proportion of those individuals not being transplant candidates. I would point out that in children it is also less likely for kidney transplant to take place if the child is < 10 kg in weight often due to technical surgical issues. It would seem just as reasonable to exclude that population as the older individuals.
- The denominator presumes that all patients < 75 years old, not in nursing homes, in hospice care, would benefit by wait listing. I disagree with that premise; thus the specifications have a fatal error.
- Measures specifications are clear and precise. No concerns. Reliability testing done at inter-unit reliability with bootstrap approach. The approach utilizes a resampling scheme to estimate the within dialysis provider group practice variation that cannot be directly estimated by ANOVA. IUR values of 0.56 calculated for the measure which indicates that 56 percent of the variation in the measure can be attributed to the between-dialysis practitioner group practice differences and 44 percent to the within dialysis practitioner group differences. Developer noted that this IUR implies moderate degree of reliability. Dialysis practitioner group practices with less than 11 eligible patients and two expected events were excluded from this calculation. Rating for reliability moderate.
- None.
- No concerns.
- No.
- Ok.
- No concerns about consistent implementation.
- Data elements are clearly defined.
- Data elements are clearly defined and the steps in the algorithm are clear.
- Moderate, IUR is 0.56 (moderate to low) and agree with comment by SMP that use of Pt - months in calculations means one patient can be counted 12 times for a given year for an entity and raises questions about the independence of the observations going to calculating the performance.
- Moderate reliability since inter-unit reliability was 56% and differences within the dialysis practice group was 44%.
- A 0.56 IUR equals moderate degree of reliability.
- High reliability due to EQRS data collection and CMS forms required for dialysis patients.
- Adequate.
- Moderate reliability.

2a. Reliability - Testing

- Moderate reliability.
- I have no questions about the reliability.
- IUR of 0.56 seems to be getting rather close to point (0.50) where the variation in measure outcome would evenly affected by variation between provider practice and within provider practice, rather than more heavily favoring between provider practice.
- Not as constructed - - - there is reliability of a fatally flawed metric.
- No.
- No.

- No concerns.
- No.
- Ok.
- No concerns about testing.
- IUR 0.56.
- No concerns.
- See above.
- No.
- No issues.
- No.
- No.
- No.

2b. Validity - Testing

- Mortality relationship was weak. It is not surprising that tx rates correlate with waitlisting. No analysis demonstrates that it is the activity of the practitioner v something else (tx center criteria, role of facility) since all 3 are linked to the patient.
- There are no concerns with validity.
- Validity of higher rates of transplantation and lower mortality clearly demonstrated -- how that directly relates to the measure in question may not be as transparent.
- Validity is low - - - the denominator does not recognize the value of patient choice, patient autonomy, and clinical judgment. Transplantation is not the best therapy for all in this denominator - - it is for many, but not all.
- No.
- No.
- No concerns with results from my perspective.
- No.
- Ok.
- I think children < 2 should also be excluded as they may not be big enough to refer and be waitlisted for transplant.
- Examined association with mortality: Spearman correlation coefficient was -0.02 (p=0.264). Examined association with transplant rate: Spearman correlation coefficient was 0.41 (p<.001).
- I do not have any concerns.
- Low - the assignation of the measure is misdirected and therefore it is NOT an acceptable measure of the quality of care provided by a nephrologist, or by the nephrologist's group. Even with all the data manipulations suggested, the facts remain the same, the NEPHROLGISTS REFER INTERESTED patients for consideration by the transplant team and the TRANSPLANT TEAM decides who to place on the Transplant waitlist. The criteria used by the transplant team are not PUBLIC, or STANDARDIZED, but instead are based on the specific issues, goals and limits of that (a) particular transplant center. As such, A TRANSPLANT WAITLIST RATIO IS NOT A VALID measure of a nephrologist quality.
- No.
- The practice group does not control the actual waitlisting. Appreciate that they included the transplant center characteristics but have concern over lack of control for the patients listing.
- Transplant rates improving (lowering) mortality rates was no significant. Other than that I believe validity testing had moderate value.
- No.
- No.

2b2-2b3. Other Threats to Validity (Exclusions, Risk Adjustment)

- All ok.
- A sufficient risk assessment is presented.
- Risk adjustment in the model presented excludes patient sex, race, and ethnicity (despite evidence that they are significant factors in performance differences) because using these factors would reinforce disparities in current system. Yet the risk adjustment model still uses Area Deprivation Index and dual eligibility, and couldn't this adjust away social risk associated with these factors along the same lines as the factors not being considered? Shouldn't the adjustment be meant to get at medically legitimate reasons for forgoing waitlisting or transplant?
- Risk adjustment variables are incomplete - need to "adjust" for patient choice, or to change denominator definition.
- Exclusions included 1. age greater than or equal to 75 24 percent 2. nursing home patients 8.6 percent 3. hospice 0.7 percent 4. dementia 3.9 percent. Overall, measure scores were changed modestly by the exclusions. Though performance scores were modestly affected by exclusions fair degree of variation in the percentage of patients excluded across practitioner groups. Exclusions as important as they represent a group of patients highly unlikely to be suitable for transplant waitlisting.
- No concerns.
- Exclusions were tested for impact on measure outcome. They appear appropriate as they would not be suitable for transplant.
- No.
- Ok.
- I think the groups excluded are appropriate and as listed above, there needs to be an added exclusion of children < 2 years of age.
- Exclusions include age 75 and older, SNF, hospice in prior 365 days, and dementia in the prior year (26.7% of patients were excluded). Additional exclusions could be considered (e.g., diagnosis of cancer receiving chemotherapy/radiation). Model risk adjustment includes age; DE; ADI; prior waitlisting and transplant; transplant center characteristics (weighted waitlist mortality ratio, weighted transplant rate ratio); 2728 incident comorbidities, ambulatory status, use of tobacco or drugs; missing 2728 medical evidence form; and 64 prevalent comorbidities (IP and OP claims). C-statistic 0.72.
- Exclusions are appropriate. Risks adjustments are also appropriate. I do not have any concerns on the validity of the measure. No measure is perfect.
- Numerous data points and statistical adjustments are used to try to account for the myriad of reasons that patients are not listed for transplantation by transplant centers. However, the measure has NO LINE OF SIGHT into the factors ultimately used to determine suitability for listing by the centers. As such, the validity of these risk adjustments is unknown. 2. incident data from the 2728 form filed at the start of dialysis are being used to calculate prevalent risk without any updating. For example patient frailty as suggested by inability to ambulate or transfer may change significantly after the start of dialysis, but those changes are NOT factored into the risk model used. 3. The rate of Death on the transplant list has not been adequately tested to determine if it a valid indicator of the "ease of listing" at one center versus another center. 3. there is nothing in the measure that allows for patient choice or for the fact that some patients do not want to have a transplant, Patients may be referred for EVALUATION BY THE NEPHROLOGIST, and ultimately decide they do not want to undergo a transplant.
- Race, sex and ethnicity were not include in the final risk model. Developer indicated they could create or reinforce disparities. Is this a valid reason?

- Still have issues with lack of control for actual waitlisting. Not uncommon for the transplant center to inform the practitioner or facility of patient being removed from list.
- Exclusions appropriate. I agree with the rationale of the exclusions. Risk adjustments were appropriate and acceptable.
- There are additional exclusions that should be considered.
- Yes.

2b4-2b6. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)

- Developer assessed impact of missing data. Developer model categorized statistical differences as better or worse than expected.
- No.
- No concerns about current exclusion criteria/missing data and effects on performance score changing.
- With current denominator, meaningful differences are not relevant to best care.
- The developer hypothesized that for high PSWR performance transplant rates would be higher and for higher PSWR performance mortality would be lower.
- No concerns.
- No missing data noted, one set of specifications used.
- No.
- Ok.
- If enough missing data exists, could be a potential threat to validity.
- Minimal missing data (missing 2728 forms), missing data does not appear to constitute a threat to validity.
- I do not believe there are threats to validity. The developer did note that higher waitlist performance correlated with higher transplant rates, and the relationship with mortality was also as expected by the developer, though not statistically significant, with the numerically highest mortality in the lowest performance tertile on the PSWR measure. Missing data should not be a significant threat to validity.
- The assignation is not correct; thus, the measure fails both 2b4 and 2b5.
- Missing data stated as 0.51%. Missing data does not appear to be a threat to the validity of the measure.
- 98% of the 3 groups were better than expected or as expected. Need to discuss more. The Spearman correlation coefficient was not significant for mortality.
- Threats to validity - transplant rate did not significantly lower mortality rate. Meaningful differences: rated high for improving quality of life and improved health. Missing data: no
- It is not clear that the differences are meaningful or alterable.
- No.

3. Feasibility

- No concerns.
- There are no concerns with data usage.
- Data collection seems feasible.
- The measure would be feasible of the denominator were adjusted.
- All data elements are in defined field in electronic data. Developer did not identify any difficulties with data collection, data availability, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection or any other feasibility/implementation issues. No concerns on data collection. Feasibility rates high.
- No concerns.
- High feasibility. Data elements are collected regularly.

- No.
- Ok.
- No concern.
- Based on data sources, feasible to implement. Measure would be restricted to Medicare beneficiaries given the data sources.
- The data elements are routinely generated and available.
- Will be possible to collect these data from the dialysis facility. However the significance of the ability of the dialysis facility to impact the data or the ability to correctly attribute that impact or change to the dialysis facility is problematic.
- No concerns about collection of data.
- No issue with feasibility.
- High feasibility.
- Adequate.
- No.

4a. Use

- New measure.
- Others have had the opportunity to provide feedback on the measure's performance.
- N/A -- new measure.
- I have not seen measure performance or implementation data; nor feedback.
- New measure not publicly reported. Yes a credible plan for implementation provided. 1. those being measured have been given performance results or data as well assistance with interpreting the measure results an data. Those being measured and been given an opportunity to provide feedback on the measure performance or implementation and 3. the feedback has been considered when changes are incorporated into the measure.
- New measure.
- Not in use since it is a new measure. potential for public reporting or payment program is noted.
- No.
- Ok.
- No concerns.
- Measure is not in use.
- This is a new measure and is not currently in use. The developer stated the measure may be considered for use in a public reporting/payment program in the future.
- NQF endorsed measured are expected to be used in at least one accountability application.
- New measure so data not currently reported. There are two existing waitlist measure at the facility level.
- No issues with use.
- Waitlisted patient data is not publicly reported at this time. I agree with the TEP that waitlisting is an important to get into the transplant process. I agree that the Nephrologist and Advanced Providers are key to original inspiration/education on transplant but have little or no influence in the waitlisting process. I would make this a facility measure since the dialysis team has more consistent contact with patients.
- It is not reported publicly.
- Yes.

4a. Usability

- Insofar as tx centers make the determination, practitioners will require robust communication with tx center to understand where patient is on waitlisting journey.

- There are no unintended consequences related to this measure.
- Potential usability but this is predicated on accepting argument that dialysis providers have significant influence on successful waitlisting or transplant.
- Usability is a major concern - - the potential harm to patients who choose not to be transplanted has not been considered, or measured. If clinicians are pressured to request that patients go on a transplant list even if the patient clearly refuses this procedure, I can envision substantial harm to trust, to the integrity of the clinicians and to the autonomy of the patient.
- Results can be used toward achieving the goal of high quality, efficient health care for individuals. Not used in a public reporting program so improvement could not be evaluated. Developer states that progress toward achieving high quality efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations. Did not identify any unexpected findings and did not report any potential harms. Usability is rated as moderate.
- Potentially of great benefit.
- No harms noted.
- No.
- Ok.
- Unintended consequence would be patient may feel coerced to go through transplant evaluation if not interested, but transplant has better outcomes than dialysis.
- Potential future use in public reporting/payment program.
- Once implemented dialysis practitioner group practice performance on the measure can be evaluated to determine if the measure has supported and detected quality improvement in waitlisting rates among the target population. The developer did not identify any unexpected findings, or report any potential harms.
- The nephrologist does not control the decision to waitlist patients and therefore applying this to QIPs and other dialysis-linked quality measures is inappropriate.
- I believe the benefit would outweigh any potential unintended consequences.
- Great benefits of increasing the waitlist. Waitlisting the patient is the responsibility of the entire team.
- Transplantation improves quality of life for many CKD patients. No significant impact on mortality. No unintended consequences were found with high benefit to improving health.
- Not clear that it can improve healthcare.
- Difficulty accounting for patient choice.

5: Related and Competing Measures

- Yes. No. No.
- There are no additional steps needed for this measure.
- None apparent.
- A far better measure would be one excluding patients who choose not to consider transplant. Another measure better constructed would be the referral rate for transplant among consenting patients.
- Standardized first kidney transplant waitlist ratio for incident dialysis patients SWR; percentage of prevalent patients waitlisted. Developer stated that they are harmonized this measure with other related measures to the extent possible.
- One measure, harmonized.
- Related measures are noted by the developer. There seems to be a growing list.
- No.
- No certain.

- Don't think so.
- NQF #3695 PPPW is a related measure that has been endorsed; unendorsed measures that are related include SWR and PPPW.
- Other related/competing measures were identified. The developer stated that they harmonized this measure with other related measures to the extent possible.
- None.
- One NQF-endorsed measure and two non-NQF-endorsed measures were identified as related measures. NQF #3695 Percentage of Prevalent Patients Waitlisted (PPPW), Standardized First Kidney Transplant Waitlist Ratio for Incident Dialysis Patients (SWR), Percentage of Prevalent Patients Waitlisted. The developer states they harmonized this measure when possible.
- No issues.
- The percentage of prevalent patients waitlisted is a similar measure.
- No.
- No.

Pre-evaluation Public and Member Comments

Lisa McGonigal, Kidney Care Partners

Comment ID#: 8304
Council / Public: Public
Level of Support: Member Does not support

Comment 1 by: Lisa McGonigal, Kidney Care Partners; Submitted by Dr. Lisa McGonigal, MD, MPH, Kidney Care Partners

NQF 3719: Prevalent Standardized Waitlist Ratio (CMS) KCP recognizes the tremendous importance of improving transplantation rates for patients with ESRD, but does not support the attribution of successful or unsuccessful waitlisting to individual practitioners or group practices and thus cannot support this measure. KCP believes that while referral to a transplant center and initiation or even completion of the waitlist evaluation process might be appropriate measures for these levels of analysis that could be used in CMS's quality programs, the newly proposed clinician/group level Prevalent Standardized Waitlist Ratio (PSWR) measure is not. Waitlisting per se is a decision made by the transplant center and is beyond the locus of control of the providers targeted in this measure. In reviewing the details of the measure, we offer the following comments: Attribution. As above, we strongly object to attributing successful/unsuccessful placement on a transplant waitlist to individual clinicians or group practices and believe this is a fatal structural flaw with the measure. The transplant center decides whether a patient is placed on a waitlist, not the practitioner or group practice. KCP patient members who are transplant recipients have noted there are many obstacles and delays in the evaluation process with multiple parties that have nothing to do with the treating nephrologist or group. For instance, one patient noted their private pay insurance changed the locations where they could be evaluated for transplant eligibility on multiple occasions, repeatedly interrupting the process mid-stream. Penalizing a clinician/group practice each month through the PSWR for these or other delays is not only inappropriate; it is fundamentally misaligned with NQF's first "Attribution Model Guiding Principle" that measures' attribution models should fairly and accurately assign accountability. KCP emphasizes our commitment to improving transplantation access, but we believe other measures with an appropriate sphere of control should be pursued. For example, our sister organization, the Kidney Care Quality Alliance (KCQA), has developed a dialysis facility-level Transplant Access Measure Set that will be submitted to NQF for endorsement consideration later this year. The set pairs a referral rate metric with a measure assessing the waitlisting rate specifically among those patients who were referred by the facility within the preceding three years. Because the KCQA waitlisting measure denominator is limited to those patients specifically identified as appropriate transplant candidates and

deliberately referred by the dialysis facility within a defined time period, facilities have considerably more agency over the measure than less precise metrics like the PSWR; this construct will also provide a counterbalance to the referral measure, curbing the tendency to indiscriminately refer patients who are not appropriate transplant candidates, preventing unnecessary patient and transplant center burden. The same approach could be applied at the clinician/group level. Variation in Transplant Center Eligibility Criteria. We also note that criteria indicating a patient is "not eligible" for transplantation can differ by geographic location. For instance, one center might require evidence of an absence of chronic osteomyelitis, infection, heart failure, etc., while another may apply eligibility exclusions differently or have additional or different criteria. The degree to which these biological factors influence waitlist placement must be accounted for in any model for the measure to be a valid representation of waitlisting. Measure Reliability. Finally, the overall IUR of the PSWR is 0.56, interpreted as "questionable" reliability by statistical convention. Thus nearly half of the observed variation in the measure could be attributed to random noise rather than true performance differences between providers. Additionally, as reliability statistics were not stratified by facility size, we are unable to discern how widely reliability varies across the spectrum of practitioner and group practice sizes. As has been the case with other CMS standardized ratio measures, we are concerned that the reliability for small providers might be substantially lower than the overall IURs. To illustrate our point, CMS's Standardized Transfusion Ratio for Dialysis Facilities (STrR) measure (NQF 2979) was found to have an overall IUR of 0.60; however, the IUR for small facilities (defined by CMS as <=46 patients for the STrR) was only 0.3 ("poor" reliability). Without evidence to the contrary, KCP is concerned that PSWR reliability is similarly lower for small groups, effectively rendering the metric meaningless for use in performance measurement in this substantial subset of providers. Notably, the many such providers that treat small rural or low-income communities could be disproportionally impacted, resulting in the imposition of random and specious penalties on the most financially vulnerable clinician groups treating the most socially and medically disadvantaged patients. KCP believes it is incumbent on CMS to demonstrate reliability for all providers by stratifying data by practice size.

Lisa McGonigal, Kidney Care Partners

Comment ID#: 8305
Council / Public: Public
Level of Support: Member Does not support

Comment 2 by: Lisa McGonigal, Kidney Care Partners; Submitted by Dr. Lisa McGonigal, MD, MPH, Kidney Care Partners

NQF 3719: Prevalent Standardized Waitlist Ratio (CMS) KCP recognizes the tremendous importance of improving transplantation rates for patients with ESRD, but does not support the attribution of successful or unsuccessful waitlisting to individual practitioners or group practices and thus cannot support this measure. KCP believes that while referral to a transplant center and initiation or even completion of the waitlist evaluation process might be appropriate measures for these levels of analysis that could be used in CMS's quality programs, the newly proposed clinician/group level Prevalent Standardized Waitlist Ratio (PSWR) measure is not. Waitlisting per se is a decision made by the transplant center and is beyond the locus of control of the providers targeted in this measure. In reviewing the details of the measure, we offer the following comments: Attribution. As above, we strongly object to attributing successful/unsuccessful placement on a transplant waitlist to individual clinicians or group practices and believe this is a fatal structural flaw with the measure. The transplant center decides whether a patient is placed on a waitlist, not the practitioner or group practice. KCP patient members who are transplant recipients have noted there are many obstacles and delays in the evaluation process with multiple parties that have nothing to do with the treating nephrologist or group. For instance, one patient noted their private pay insurance changed the locations where they could be evaluated for transplant eligibility on multiple occasions, repeatedly interrupting the process mid-stream. Penalizing a clinician/group practice each month through the PSWR for these or other delays is

not only inappropriate; it is fundamentally misaligned with NQF's first "Attribution Model Guiding Principle" that measures' attribution models should fairly and accurately assign accountability. KCP emphasizes our commitment to improving transplantation access, but we believe other measures with an appropriate sphere of control should be pursued. For example, our sister organization, the Kidney Care Quality Alliance (KCQA), has developed a dialysis facility-level Transplant Access Measure Set that will be submitted to NQF for endorsement consideration later this year. The set pairs a referral rate metric with a measure assessing the waitlisting rate specifically among those patients who were referred by the facility within the preceding three years. Because the KCQA waitlisting measure denominator is limited to those patients specifically identified as appropriate transplant candidates and deliberately referred by the dialysis facility within a defined time period, facilities have considerably more agency over the measure than less precise metrics like the PSWR; this construct will also provide a counterbalance to the referral measure, curbing the tendency to indiscriminately refer patients who are not appropriate transplant candidates, preventing unnecessary patient and transplant center burden. The same approach could be applied at the clinician/group level. ▯ Variation in Transplant Center Eligibility Criteria. We also note that criteria indicating a patient is "not eligible" for transplantation can differ by geographic location. For instance, one center might require evidence of an absence of chronic osteomyelitis, infection, heart failure, etc., while another may apply eligibility exclusions differently or have additional or different criteria. The degree to which these biological factors influence waitlist placement must be accounted for in any model for the measure to be a valid representation of waitlisting. ▯ Measure Reliability. Finally, the overall IUR of the PSWR is 0.56, interpreted as "questionable" reliability by statistical convention. Thus nearly half of the observed variation in the measure could be attributed to random noise rather than true performance differences between providers. Additionally, as reliability statistics were not stratified by facility size, we are unable to discern how widely reliability varies across the spectrum of practitioner and group practice sizes. As has been the case with other CMS standardized ratio measures, we are concerned that the reliability for small providers might be substantially lower than the overall IURs. To illustrate our point, CMS's Standardized Transfusion Ratio for Dialysis Facilities (STrR) measure (NQF 2979) was found to have an overall IUR of 0.60; however, the IUR for small facilities (defined by CMS as <=46 patients for the STrR) was only 0.3 ("poor" reliability). Without evidence to the contrary, KCP is concerned that PSWR reliability is similarly lower for small groups, effectively rendering the metric meaningless for use in performance measurement in this substantial subset of providers. Notably, the many such providers that treat small rural or low-income communities could be disproportionally impacted, resulting in the imposition of random and specious penalties on the most financially vulnerable clinician groups treating the most socially and medically disadvantaged patients. KCP believes it is incumbent on CMS to demonstrate reliability for all providers by stratifying data by practice size.

NQF #3722 Home Dialysis Rate (Kidney Care Quality Alliance

Pre-evaluation Standing Committee Comments

1a. Evidence

- Observational data from other countries suggest US uptake of home therapies could be greater. one study suggests patients felt they lacked info on home therapies.
- New health outcome measure. Evidence applies directly.
- Low. This measure at the level of the facility, measures the percent of dialyzed patients treated by home dialysis. The presumption is that home dialysis is the best choice for all dialyzed patients, except those excluded in the denominator statement. However, the denominator statement does not exclude patients who choose not to dialyze at home, those with home

environments not conducive to best care, and social determinants of health that have a heavy impact on outcomes.

- Evidence does not apply directly and is tangential. Desired outcome is to increased home dialysis unit; this measure incentivizes home dialysis use (which may or may not lead to improved clinical outcomes).
- Empiric evidence only.
- The ETC model identifies increased use of home dialysis as a major performance metric of that program (a goal which will presumably eventually extend to all facilities). Currently, no clinical guidelines, recommendations, or RCT exist to address home dialysis uptake. The developers present TEP comments, and observational studies, to support need for, and evidence for, the measure. Moderate evidence rating.
- Agree with Moderate Rating for Evidence. There is sufficient evidence noted by the developer as this is a major CMS objective. Numerous points were also made and supported by observational studies.
- Adequate.
- Historical evidence shows improved quality of life, lower mortality with home therapies. The workgroup unanimously approved this measure. Pass on evidence.
- This is a new process measure at the facility level that examines the percentage of patients dialyzing via a home dialysis modality among patients assigned to a given dialysis facility and/or Hospital Referral Region (HRR) within a given measurement year. A logic model was depicted that provided the relationship between the individual measure components, process interventions, and the desired health outcomes which include reduced cardiovascular risk, mortality, hospitalization, and cost in addition to improved quality of life. It was note that adoption of the home dialysis measure would incentivize the facility to identify appropriate home dialysis candidates and to implement process interventions such as effective modality education and appropriate patient preparation, training, and support to increase home dialysis uptake and retention among those candidates. Increasing home dialysis utilization is now a major CMS objective to reduce costs while preserving or enhancing the quality of ESRD care.
- No systematic reviews available or grading of evidence; there is empirical evidence presented. Home dialysis utilization is a CMS goal without any evidence data available. Without evidence based data and only empirical information the rating is Moderate. I would support this.
- There was no evidence provided to support the measure. The developers used observational studies and industry experts in lieu of evidence.
- Moderately well supported.

1b. Gap in Care/Opportunity for Improvement and Disparities

- Gap in adoption of home therapies demonstrated.
- Yes current performance data was provided. Developer calculated 2021 performance scores during measure testing at facility level and aggregate parent Dialysis Organization Hospital Referral Region unit and reports that testing encompassed 543,155 patients; 4,937,405 patient-months 5,792 facilities and 295 HRR with 26HRRs for each of the participating DO. Facility level found a mean performance of 14.49 % a standard deviation of 25 and a interquartile range of 19.84 %. HRR unit level found a mean performance of 16.58% a standard deviation of 6.57 % and an interquartile range of 6.9 Developer also states home dialysis utilization rates are lower in the US compared to other developed nations. Yes data on the measure by subgroups was provided. Found that Bl acks and Hispanics have a 3.4 fold and 1.3 fold greater risk of high risk kidney failure compared with non-Hispanic white individuals. Black and Hispanics are less likely to be treated with home dialysis. Moderate for opportunity for improvement.
- There is clearly wide variation in % of home dialysis. However, it is hard to draw meaning from these data without appropriate selection of denominator patients - - so while there is a

performance gap, I do not find it informative with the measure definitions.

- Yes there is a gap; there was also data on disparities and how home dialysis is less frequently used in certain patient populations (Blacks, Hispanics).
- Gap and disparities well recognized.
- Difficult to judge, while rates of home dialysis use are lower in the USA ,than in other countries there are several reasons for this, many of which are NOT linked to the quality of care provided within a dialysis facility. Developers did demonstrate moderate gap in home dialysis use among dual eligible patients, and among different racial /ethnic groups. Moderate gap.
- Moderate evidence provided there is significant gap identified. Disparities are also noted.
- Adequate.
- Other countries were found to have higher rates of home therapies. Disparity review finds the highest rate of home therapy is in the "other" race (non-white, non-black). Blacks and Hispanics are less likely to be treated with home therapy.
- The developer did provide performance gap data for race/ethnicity and does show disparities.
- Home dialysis utilization in the US is 15% which is lower than other countries. There seems to be a gap between whites and black patients. There is also a gap between dual-eligible patients than non-dual eligible patients. The gap is moderate.
- Performance gaps and disparities exist in support of this measure.
- A gap exists. Yes, that data was provided.

2a1. Reliability - Specifications

- None.
- Measure specifications are clear and concise. Reliability testing was conducted at the facility and HRR levels signal to noise analysis; the beta binomial model.
- Specifications of the denominator are insufficient to serve best patient care - - autonomy, respect for family structure and other social determinants of care.
- No concerns.
- None.
- Moderate to high.
- Supports High reliability for reporting data elements.
- Adequate.
- The data should be reliable if reported honestly by the facility. The SMP is satisfied with the reliability testing.
- Data elements are clearly defined.
- SMP passed the measure on Reliability.
- The mean reliability at the facility level was 0.999 which seems high considering we are evaluating a measure to increase home therapies.
- High reliability.

2a2. Reliability - Testing

- LDOs have an advantage in being able to credit facilities via HRR calculations.
- No.
- Not as constructed - - there is reliability of a very flawed measure. Thus, I do not find the reliability relevant to best patient care.
- No concerns.
- None.
- No.
- No concerns. High reliability.
- Adequate.

- No.
- It is unclear to me if there may be some problems with reliability-- I will defer to the rest of the group.
- The SMP looked at this and had concerns but passed it.
- Same as above.
- No.

2b. Validity

- No.
- No.
- Validity is low - - - the denominator does not recognize the value of patient choice, patient autonomy, and clinical judgment. Home dialysis is not the best therapy for all in this denominator - - it is for many, but not all.
- No concerns.
- No.
- As a PAIRED measure (with NQF 3725) measure appears to have reasonable validity score, as NON paired measure, the validity of this measure as a marker of facility quality is much lower.
- No concerns. The SMP provided Moderate preliminary rating.
- Adequate.
- No.
- I do not believe there are threats to validity.
- Face validity with 9 TEP members; 8 members passed the measure on face validity. They also agreed that this measure should be paired with 3722. Concerns were raised by the SMP around risk stratification.
- Question the testing with facilities with zero months.
- No.

2b2-2b3. Other Threats to Validity (Exclusions, Risk Adjustment)

- Effect of supply shortages not factored in. effect of staffing shortages not factored in. Housing insecurity availability of a partner not factored in.
- Developer stratified the measure by age, gender, race/ethnicity and dual eligible status. Explored markers of functional risk and clinical variable for stratification but were not included due to data availability. Stratified analyses at both the facility and HRR level demonstrate clear trend by age, differences by race and differences on home dialysis less among dual eligible than non-dual eligible. Yes an appropriate risk adjustment strategy included.
- Risk adjustment variables are incomplete - need to "adjust" for patient choice, or to change denominator definition.
- Question about how do you account for patients that tried home dialysis but failed for whatever reason. Should that be an exclusion?
- Will inclusion of race in risk adjustment reinforce disparities in access to home care?
- 2b2 exclusions clear. 2b3 minimal risk adjustment provided other than dual eligibility, an inherent risk of financially rewarding facilities for increased home dialysis use is that patients who are not appropriate candidates for home modalities may be overly aggressively encouraged to accept home dialysis therapies. Linking this measure to the retention measure (NQF3725) provides a "guard rail" against the over-aggressive use of home modalities in an "at -risk-of failure" patient. Thus, implementation of this a part of a linked - retention measure set can serve as a "surrogate" for a more complicated (and perhaps less valid) statistical approach to patient level risk-adjustment.
- No concern.

- Adequate.
- Exclusions are appropriate. No risk adjustment needed for this measure.
- Exclusions are appropriate. Risks adjustments are also appropriate.
- Exclusions seem reasonable; since no evidence presented could not measure against evidence. Differentiation by age, race and insurance coverage were identified.
- Measure results are stratified by 5 risk factor groups: age, gender, race, ethnicity, dual-eligibility. The description of how they were tested seemed weak.
- Yes, no large identified.

2b4-2b6. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)

- No concerns note that tables referring to deciles express data in percentiles instead.
- Missing data does constitute a threat to validity.
- With current denominator, meaningful differences are not relevant to best care.
- Don't think so.
- No concerns.
- See above.
- No concerns.
- Adequate.
- Over half of dialysis facilities do not offer home therapies. The SMP had concerns that the risk stratification approach used by the developers. Noting a lack of detail.
- I do have some concerns that the measure may not adequately distinguish.
- >50% of the facilities reported 0 patient months on home dialysis. Missing data for hospice, LTCF reported.
- Zero months vs non-zero months warrants more discussion.
- No.

3. Feasibility

- No concerns.
- All data elements are in defined field in electronic data. Developer did not identify any difficulties with data collection, data availability, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection or any other feasibility/implementation issues. No concerns on data collection. Feasibility rates high.
- The measure would be feasible of the denominator were adjusted.
- No concerns.
- No concerns.
- High.
- Moderate. The data elements are routinely collected by healthcare professionals.
- Adequate.
- The data elements are in the EHR. Data collection should be easily obtained without undue burden.
- The data elements are routinely available.
- Data is part of the CMS ESRD Quality Reporting System (EQRS) and so it is feasible.
- No issues with feasibility.
- No concerns. Moderately rated feasibility.

4a. Use

- New measure.
- Performance results are used at least one accountability application after initial endorsement within 3 years of initial endorsement and publicly reported 6 years after initial endorsement. Its

a new measure not publicly reported. Developer wants CMS to add measure to ETC model and to KCC models for implementation in 2024. Yes feedback has been considered when changes are incorporated into the measure.

- I have not seen the evidence that facilities told their proportion of patients at home are lower compared with other facilities, improves outcomes.
- No comments as this is a new measure.
- New measure- no feedback but no concerns.
- 41a not currently in use, planned use in an accountability program. Developer plans to submit to CMMI to add measure to ETC model and may submit to KCC model. May propose adoption by broader ESRD program. 4a2. performance scores were shared with pilot facilities for feedback and input. Pass.
- Not currently reported publicly.
- Adequate.
- The developer plan to engage CMS to this measure which should be implemented in 2024 and available by end of 2025. The developer is also submitting the measure to MUC for adoption into the ESRD Network programs.
- The measure has been tested in in 2022 and participating facilities did not report issues with data availability.
- New measure. Developed has plans to discuss collect with CMMI and potential KCC Models in 2024. Measure was tested using 2021 data with 2 large dialysis organizations. Pass on accountability and transparency.
- No issue with use.
- No identified concerns.

4a. Usability

- No allowance made for impact of patient refusal on results. Patients may be pressured.
- Results can be used toward achieving the goal of high quality, efficient health care for individuals. Not used in a public reporting program so improvement could not be evaluated. Developer states that progress toward achieving high quality efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations. Did not identify any unexpected findings and did not report any potential harms. Usability is rated as moderate.
- Usability is a major concern - - the potential harm to patients who choose not to receive dialysis at home has not been considered, or measured. If clinicians are pressured to urge patients to go on home treatment, even when they explicitly choose not to, patient trust may be violated, clinician judgment ignored, and substantial harm could result.
- Unintended consequence is that patients who do not want to pursue home modality may be pressured by dialysis facility to switch; patients may not have housing/resources to be able to do home modality.
- Beneficial, especially if publicly reported.
- Would be acceptable as quality measure if done as a paired measure.
- Would encourage accountability to conduct dialysis at home. Which would be a shift in our current healthcare system, the infrastructure to provide a higher home dialysis rate would need to keep up with the intended trend.
- None.
- Will improve patient education on modality. Will inspire clinic to have a home therapy training program or collaborate with a HT program. The benefits outweigh the possible unintended consequences.
- The measure has not been implemented so improvement cannot be evaluated.

- Not part of a public reporting program. No potential harm reported.
- No issue.
- None were identified.

5: Related and Competing Measures

- ETC has a metric.
- None.
- A far better measure would be one excluding patients who choose not to consider home dialysis.
- None.
- No competing measures.
- None.
- None. Intended to be implemented with a Retention measure.
- None.
- None.
- None.
- None.
- None.
- No.

Pre-evaluation Public and Member Comments

Lisa McGonigal, Kidney Care Partners

Comment ID#: 8306
Council / Public: Public
Level of Support: Member Supports

Comment 1 by: Lisa McGonigal, Kidney Care Partners; Submitted by Dr. Lisa McGonigal, MD, MPH, Kidney Care Partners

Home Dialysis Rate (NQF 3722, KCQA) Home Dialysis Retention (NQF 3722, KCQA) KCP supports endorsement of KCQA's Home Dialysis Measure Set. As noted by the developer, home dialysis modalities favorably impact both clinical and patient-reported outcomes and enhance patient autonomy and quality of life—but remain underutilized in the US. Accordingly, increasing home dialysis is a major objective of the new ESRD Treatment Choices (ETC) Payment Model, and CMS has identified home dialysis utilization as one of the performance metrics to be used within the program. Of note, the ETC model provides significant financial incentives—and penalties—to improve home dialysis utilization. In the absence of appropriate safeguards and a sufficiently robust infrastructure to support the anticipated rapid increase in home modalities use, there is concern among our patient and advocate stakeholders that the current unilateral focus on home growth will certainly lead to increased technique failure rates, may subject many patients to a treatment modality for which they have not received adequate education or training, and may even inadvertently infringe on patient choice. As such, we and our constituent members appreciate that KCQA's "Home Dialysis Measure Set" was developed and designed to promote steady, deliberate performance improvement over time by addressing both sides of the home dialysis utilization equation—uptake and retention. The Home Dialysis Retention Measure (NQF 3722) provides a counterbalance to the Home Dialysis Rate Measure (NQF 3725), minimizing the potential adverse consequences of unchecked home dialysis growth. As noted by the Developer, the retention measure will also allow providers to more readily assess the success of their efforts to create a sustainable home program through appropriate patient education, preparation, and support, and to apply targeted quality improvement interventions when and where needed. Finally, both measures have been demonstrated as statistically reliable, valid, and feasible during measure testing and have

successfully met the Scientific Methods Panel's rigorous assessment criteria. For the above reasons, KCP supports the KCQA Home Dialysis Measure set and urges the Renal Standing Committee to recommend endorsement of both measures.

NQF #3725 Home Dialysis Retention (Kidney Care Quality Alliance)

Pre-evaluation Standing Committee Comments

1a. Evidence

- This is a balancing measure. Data on usual dropout rates presented.
- Moderate evidence based on observational data and expert opinion. The developer noted that there are no relevant clinical practice guidelines, United States Preventive Services Task Force (USPSTF) recommendations, systematic reviews, or formal randomized controlled studies addressing home dialysis modalities uptake or retention.
- Moderate. In this facility-level outcome measure, evidence relates directly to home dialysis retention.
- Evidence does not apply directly and is tangential. Desired outcome is to increased home dialysis unit; this measure incentivizes home dialysis use (which may or may not lead to improved clinical outcomes).
- No evidence submitted just observational studies. Measure impacts both clinical and patient reported outcomes. The basic premise of the measure is set to incentivize prescription of and preparation for home modalities for all clinically appropriate patients in accordance with patient preference. The logic model illustrates the relationship between the individual measure components process interventions and the desired health outcomes which includes lowering patient mortality hospitalization and cardiovascular risk improving patients quality of life reducing cost of care.
- Based on empirical evidence. Advisory panel only included one patient.
- Agree with Moderate Rating for Evidence. There is sufficient evidence noted by the developer as this is a major CMS objective. Numerous points were also made and supported by observational studies.
- Adequate.
- This is a new measure that provides for retention of home hemodialysis. There are no relevant clinical practice guidelines. The developer does provide observational study evidence. As home dialysis is now a metric that dialysis facilities and organizations will be measured on – this measure seems appropriate.
- There was no empirical evidence but the data suggest we can improve outcomes if we educated patients better, provided enough qualified training nurses and did better at supporting patients for improve longevity. Passes without empirical evidence.
- No systematic reviews available or grading of evidence; there is empirical evidence presented. Observational studies and TEP is cited. Without evidence based data and only empirical information the rating is Moderate. I would support this.
- No evidence was submitted.

1b. Gap in Care/Opportunity for Improvement and Disparities

- Age and other race, Hispanic relate to better rates.
- Moderate.
- There is a demonstrated performance gap.

- Agree with moderate gap.
- Developer calculated 2021 performance scores during measure testing. Testing encompassed 30,549 new home dialysis patients regardless of patient age, vintage or payer. only 2812 of the 5781 facilities across the two participating DO had new home dialysis patient to contribute to the denominator and were included in the analysis, developer found a mean performance 74.7 percent a std deviation of 30.5 percent and an interquartile range of 31.2 percent. Noted variations in performance across demographic groups, with trends being identified by age, race, ethnicity and insurance status. Regarding age, patients under 18 were achieving 90 or more days of home dialysis more consistently than older age groups. Developer reports that concerning race, there was a higher performance among "other" races 81.8 percent than in black 78.9 percent or white 77.3 patients. Hispanics 82.8 per formed more than 7 percent higher than non-Hispanics 75.2 percent. Dual eligibility patients 79.1 performing slightly better than non-dual eligible 76.5 percent.
- There appears to be a gap though limited data makes disparities hard to discern.
- There is a gap and room for much improvement in this area. Moderate rating for Gap.
- Adequate.
- The developer did provide performance gap data for race/ethnicity and does show disparities.
- Performance data was provided from two large dialysis organizations. The data showed disparities. The Scientific Methods Panel passed on reliability. I pass this measure on reliability.
- >30,00 patients were in the test group. Younger patients (<18) vs older patients, other races vs black or white, and dual eligible patients performed better. There is a gap that rates moderate.
- Performance gaps and disparities exist to warrant the need for the proposed measure.

2a1. Reliability - Specifications

- No concerns.
- Moderate to high.
- Specifications are appropriate, logic algorithm appropriate.
- No concerns.
- Measure was previously submitted to the SMP under NQF 3697 as a clinical intermediate outcome measure, it has been resubmitted as a clinical intermediate outcome measure under NQF 3725 reliability testing was conducted at the facility level using signal to noise analysis; the beta binominal model. HHR level aggregation is not necessary for this measure because it only includes incident patients and does not need to account for facilities that do not offer home dialysis. No concerns. Reliability rated as moderate.
- Agree that two year construct more reliable given small numbers.
- Data Definitions are supplied and clear. No concerns.
- Adequate.
- Data elements are defined - however reliability may be overestimated. In addition 2 year data would be better.
- The question about "late in the year home therapy starts" is a good question. A two year rolling average is acceptable. The measure should be able to be consistently measures.
- SMP passed measure on Reliability and Validity. This measure was previously submitted as 3697. Change made after receiving feedback to 3697 lead to only facility level of analysis in this measure. The specifications are clear.
- Committee needs to discuss the one year vs two year rolling requirement.

2a2. Reliability - Testing

- No.
- No.

- No concerns.
- No concerns.
- No.
- Small numbers in one year construct, especially in smaller units.
- I see there is concern with this being a calendar year measure and ability to capture 90 days if in home dialysis is started late in year. Is there consideration for a look back and look forward period to be included in the numerator statement?
- Adequate.
- Reliability may be overestimated.
- I do not.
- Is doubling to 2021 data to get two year data OK to use to test reliability?
- Same as above.

2b. Validity - Testing

- No.
- Measures that encourage an Increase in the rate of home dialysis uptake as an indicator of dialysis facility quality, carry the risk of an unintended consequence of encouraging patients to Initiate home dialysis even though they may be ill suited for this therapy due to their own life choices or because of underlying comorbidities. Linking this retention measure to gather with an NQF3722 home dialysis rate will help to provide a guardrail against these types of unintended consequences.
- No concerns.
- No concerns.
- No.
- No.
- No concerns on validity. Agree with moderate rating.
- None.
- No concerns.
- 8 of 9 Scientific panel members agree that the measure had passed validity testing. I agree.
- No concerns with face validity results.
- No issue.

2b2-2b3. Other Threats to Validity (Exclusions, Risk Adjustment)

- No concerns.
- Adjustments proposed are appropriate. However, agree with TEP regarding concerns raised with late year starts (October and beyond) and 90 day time of measure as planned. This should be discussed in more detail.
- Social risk factors potentially have a major effect on the outcome of retention.
- Wonder if patients who surgically do PD due to abdominal issue should be considered an exclusion? What about if in temporary foster care/group home?
- Exclusions include 1. pt months with hospice 0 percent; patient months in nursing home or ltcf 1 percent; patient discharge secondary to transplant .4 percent; death 1.8 percent, discontinuation of dialysis .5 percent; recovery of renal function .1 percent. mean facility level performance before exclusions was 72.4 percent and with them applied, 74.8 percent. these exclusions were needed to avoid creating disincentive for home dialysis trials by penalizing providers for unanticipated events beyond their control that prevent a patient from achieving the 90 day numerator criterion. Risk stratified as age, gender, race/ethnicity and dual eligible status.

- I am concerned that adjustment which includes race may reduce utility of measure by reinforcing existing disparities in access to home care.
- No concern.
- Adequate.
- Exclusions and risk adjustment appears to be appropriate.
- Exclusions are appropriate for the measure. No risk adjustment for this measure and risk assessment showed clear trends.
- Ok.
- No issues.

2b4-2b6. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)

- No concerns.
- As above.
- No significant treats to validity.
- Don't think so.
- Developer notes that while they believe their observed percentage of patient months excluded secondary to hospice enrollment is not accurate they believe those patients are captured in other exclusions. Developers believe their observed percentage of patient-months excluded due to nursing/ltcf residence is an underestimate. However if they were to use the highest exclusion rate reported, there is only a difference of 0.4 percent in the overall facility level score. When patient months were excluded from denominator due to missing values in the stratification variables i.e. age, sex, race, ethnicity and dual eligibility status the mean facility level was 74.7 percent before exclusions and 74.8 % excluding missing values. Validity rated moderate.
- No.
- No concerns.
- No.
- Exclusions and missing data was appropriate.
- To make it work they had to use more than one year of data. There was an insignificant amount of missing data.
- Missing months were described. When the months were excluded the mean facility-level performance was essentially unchanged.
- No issue.

3. Feasibility

- No concerns.
- High.
- Moderate - - data should be easily available.
- No concerns.
- Data elements are generated or collected by and used by healthcare personnel during the provision of care. No concern about data collection since the data is already collected by facilities and submitted to CMS. Feasibility rated as moderate.
- None.
- Data is captured in EQRS. No concerns.
- None.
- Data elements are defined and accessible.
- The data is easily retrievable from EHR if data are entered in a timely manner. No concerns about data collection.
- Data is intendent to be used by CMS ESRD Quality Reporting System (EQRS) and is already collected by the facility, so it is feasible.

- No issue.

4a. Use

- No concerns.
- New intermediate outcome measure. Expert panel reviewed. Data were shared and feed back was obtained from pilot facilities.
- Use - - unclear if feedback on variability in retention will have an effect on the outcome.
- No concerns.
- New measure and is therefore not currently in use. Developer wants CMS to add measure to the ETC model and possibly to Kidney Care Choices for implementation in 2024 results then would become available to the public by end of 2025. Developer also plan to submit this measure to the Measures Under Consideration list for adoption into ESRD program. 3 criteria demonstrate feedback 1. those measured have been given performance results or data as well assistance in interpreting the results 2. those being measured and other users are given opportunity for feedback on performance/implementation 3. feedback has been considered when changes are incorporated into measure.
- New measure- no feedback.
- Not currently reported publicly.
- None.
- This is a new measure and therefore there is no feedback as of yet.
- The large and medium sized dialysis facilities already track this data for the quality improvement programs.
- New measure. Developer has plans to discuss collect with CMMI and potential KCC Models in 2024. Measure was tested using 2021 data with 2 large dialysis organizations. Pass on accountability and transparency.
- No issue.

4a. Usability

- No concerns.
- Ultimately will be applied as a performance measure and developer plans to propose that it be applied to all ESRD facilities as well as to those in the ETC model now. For full effectiveness this measure should be paired with NQF# 3722.
- Usability - - Potential for this metric to assist facilities in examining their procedures to increase retention is present - - need data to confirm.
- No concerns.
- Measure has not been implemented in a public reporting program, so improvement cannot be evaluated. However developer states the adoption of the measure in the ETC Model would increase the utilization of home dialysis among individuals or populations. Benefits in facilitating progress toward achieving high quality, efficient healthcare for individuals or populations outweigh evidence on unintended negatives consequences to individuals or populations. Did not report any unexpected findings as not implemented yet. No reported potential harm.
- Some concern about unintended consequences, perhaps skewing access to home care.
- Important to use this in a Measure Set which includes the Home Dialysis Start.
- None.
- The adoption of the measure would increase utilization of home dialysis. There are no reported any unexpected findings or potential harm noted.
- Facilities with low retention rates should perform a root cause analysis to determine how to improve their training/retention programs (or how they choose patients). There should be no unintended consequences to this measure.

- Not part of a public reporting program. No potential harm reported.
- No issue.

5: Related and Competing Measures

- ETC metric is not looking at retention.
- No competing measures, NQF 3722 is related measure.
- None.
- None.
- Noted one non-NQF endorsed measure and -NQF stated they harmonized this measure with other non NQF endorsed measures to the extent possible.
- No.
- A Home Dialysis Rate measure was noted. But again this measure is to be used as a Measure Set with the Home Dialysis Start measure.
- None.
- There is one non NQF measure that is related to this measure and has been harmonized to it.
- None.
- Home Dialysis Rate is related to this measure. The developer harmonized the measure with other non-NQF endorsed measures when possible.
- None.

Pre-evaluation Public and Member Comments

Lisa McGonigal, Kidney Care Partners

Comment ID#: 8307
Council / Public: Public
Level of Support: Member Supports

Comment 1 by: Lisa McGonigal, Kidney Care Partners; Submitted by Dr. Lisa McGonigal, MD, MPH, Kidney Care Partners

Home Dialysis Rate (NQF 3722, KCQA) Home Dialysis Retention (NQF 3722, KCQA) KCP supports endorsement of KCQA's Home Dialysis Measure Set. As noted by the developer, home dialysis modalities favorably impact both clinical and patient-reported outcomes and enhance patient autonomy and quality of life—but remain underutilized in the US. Accordingly, increasing home dialysis is a major objective of the new ESRD Treatment Choices (ETC) Payment Model, and CMS has identified home dialysis utilization as one of the performance metrics to be used within the program. Of note, the ETC model provides significant financial incentives—and penalties—to improve home dialysis utilization. In the absence of appropriate safeguards and a sufficiently robust infrastructure to support the anticipated rapid increase in home modalities use, there is concern among our patient and advocate stakeholders that the current unilateral focus on home growth will certainly lead to increased technique failure rates, may subject many patients to a treatment modality for which they have not received adequate education or training, and may even inadvertently infringe on patient choice. As such, we and our constituent members appreciate that KCQA's "Home Dialysis Measure Set" was developed and designed to promote steady, deliberate performance improvement over time by addressing both sides of the home dialysis utilization equation—uptake and retention. The Home Dialysis Retention Measure (NQF 3722) provides a counterbalance to the Home Dialysis Rate Measure (NQF 3725), minimizing the potential adverse consequences of unchecked home dialysis growth. As noted by the Developer, the retention measure will also allow providers to more readily assess the success of their efforts to create a sustainable home program through appropriate patient education, preparation, and support, and to apply targeted quality improvement interventions when and where needed. Finally, both measures have been demonstrated as statistically reliable, valid, and feasible during measure testing and have

successfully met the Scientific Methods Panel's rigorous assessment criteria. For the above reasons, KCP supports the KCQA Home Dialysis Measure set and urges the Renal Standing Committee to recommend endorsement of both measures.