NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Purple text represents the responses from measure developers.
Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 2701

Corresponding Measures:

De.2. Measure Title: Avoidance of Utilization of High Ultrafiltration Rate (>=13 ml/kg/hour)

Co.1.1. Measure Steward: Kidney Care Quality Alliance (KCQA)

De.3. Brief Description of Measure: Percentage of adult in-center hemodialysis patients in the facility whose average ultrafiltration rate (UFR) is >=13 ml/kg/hour AND who receive an average of <240 minutes per treatment during the calculation period.

1b.1. Developer Rationale: Ultrafiltration rates (UFRs) are determined by the amount of fluid that must be removed from the patient and dialysis session length. As treatment time decreases, UFR tends to increase and vice versa. Both high UFRs (>=13 ml/kg/hour) and abbreviated session duration (<240 minutes) are associated with a greater risk of all-cause and cardiovascular mortality in hemodialysis patients, and research suggests that dialysis sessions >240 minutes are independently associated with a significantly reduced relative risk of mortality.

The intent of this measure is to generally foster the use of slower, gentler dialysis sessions to reduce hemodialysis-related mortality. Success for the measure can be achieved by employing either or both of two approaches: 1) dialyzing patients at an average UFR <13 ml/kg/hour AND/OR 2) dialyzing patients for an average of >240 minutes per session during the reporting period. Adherence to these conventions will help attenuate the rapid fluctuations in fluid balance and blood pressure that contribute to cardiovascular morbidity and mortality in hemodialysis patients.

S.4. Numerator Statement: Number of patients* from the denominator whose average UFR is >=13 mg/kg/hr (NOT just >13) hour AND who receive an average of <240 minutes per treatment during the calculation period.**

*To address the fact that patients may contribute varying amounts of time to the annual denominator population, results will be reported using a "patient-month" construction.

** The calculation period is defined as the same week that the monthly Kt/V is drawn.

S.6. Denominator Statement: Number of adult in-center hemodialysis patients in an outpatient dialysis facility undergoing chronic maintenance hemodialysis during the calculation period.

S.8. Denominator Exclusions: The following patients are excluded from the denominator population:

1. Patients <18 years of age (implicit in denominator definition).

2. Home dialysis patients (implicit in denominator definition).
3. Patients in a facility <30 days.
4. Patients with >4 hemodialysis treatments during the calculation period.
5. Patients with <7 hemodialysis treatments in the facility during the reporting month.
6. Patients without a completed CMS Medical Evidence Form (Form CMS-2728) in the reporting month.
7. Kidney transplant recipients with a functioning graft.
8. Facilities treating <=25 adult in-center hemodialysis patients during the reporting month.

De.1. Measure Type: Process

S.17. Data Source: Electronic Health Records

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Oct 02, 2015 **Most Recent Endorsement Date:** Oct 02, 2015

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not applicable.

Preliminary Analysis: Maintenance of Endorsement Measure

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Evidence Summary or **Summary of prior review in 2015**

- This is an EHR-based process measure at the facility level that assesses the percentage of adult in-center hemodialysis patients in the facility whose average ultrafiltration rate (UFR) is >=13 ml/kg/hour AND who receive an average of <240 minutes per treatment during the calculation period.
- The measure is based on one Kidney Disease Outcomes Quality Initiative (KDOQI) clinical guideline and a systematic review of the evidence.
    - The KDOQI clinical practice guidelines for hemodialysis adequacy: Achievement of optimal "dry" weight (CPG 5.1) gave the evidence a grade of A (high quality of evidence).
    - The developer clarified that the measure requires either having dialyzing patients at an average UFR ≤13 ml/kg/hour and/or dialyzing patients for an average of >240 minutes per session during the reporting period.
    - Upon review of the evidence submitted, the Committee noted that none of the articles reviewed during the systemic review addressed those specific requirements and different cutoffs are listed for both the timeframe and UFR.
- Since the initial submission, the KDOQI guidelines have been updated. This submission cites the most recent recommendations along with additional evidence.

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- Developer cites updated 2015 KDOQI HD Guidelines recommendations 4.1 and 4.2
    - 4.1 We recommend that patients with low residual kidney function (<2 mL/min) undergoing thrice weekly hemodialysis be prescribed a bare minimum of 3 hours per session. (1D - Grade Level 1: "strong recommendation" for which "most patients should receive the recommended course of action" and "the recommendation can be adopted as policy for most situations." Grade D: "very low quality of evidence" for which "the estimate of effect is very uncertain and often will be far from the truth")
        - Developer notes that the KDOQI Work Group's rationale for the Grade D is that much of the supporting body of evidence for Guideline 4.1 were observational studies and that the generalizability of findings from the cited nocturnal dialysis studies to patients undergoing conventional in-center HD has not been definitively established.
        - Nevertheless, the KDOQI Work Group notes that because "the strength of a recommendation is determined not just by the quality of the evidence, but also by other, often complex judgments regarding the size of the net medical benefit, values and preferences, and costs", KDOQI maintains "most patients should receive the recommended course of action" (extended treatment times beyond a "bare minimum" of three hours) and that "the recommendation can be adopted as policy in most situations."
        - A portion of KDOQI's supporting rationale for Guideline 4.1 is copied here:
            - While there is a paucity of clinical trial data to inform recommendations to inform recommendations for optimal length of treatment time, several observational studies have associated shorter HD sessions with higher mortality.
            - Importantly, the [KDOQI] Work Group could find no evidence to suggest harm from extending treatment times.
            - In a recent observational study of 746 patients using propensity score matching to compare those treated with thrice-weekly in-center nocturnal HD

(7.85 hours) or conventional in-center HD (3.75 hours), conversion to nocturnal HD was associated with a 25% reduction in the risk for death after adjustment for age, body mass index, and dialysis vintage (HR, 0.75; 95% CI, 0.61-0.91; P50.004).

    - 4.1.1 Consider additional hemodialysis sessions or longer hemodialysis treatment times for patients with large weight gains, high ultrafiltration rates, poorly controlled blood pressure, difficulty achieving dry weight, or poor metabolic control (such as hyperphosphatemia, metabolic acidosis, and/or hyperkalemia). (Not Graded)
    - 4.2 We recommend both reducing dietary sodium intake as well as adequate sodium/water removal with hemodialysis to manage hypertension, hypervolemia, and left ventricular hypertrophy. (1B – Grade B: "moderate quality of evidence," with "the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different).
    - 4.2.1 Prescribe an ultrafiltration rate for each hemodialysis session that allows for an optimal balance among achieving euvolemia, adequate blood pressure control and solute clearance, while minimizing hemodynamic instability and intradialytic symptoms. (Not Graded)
- While 4.2 is an important consideration, the focus of the measure aligns more closely with 4.1. It is concerning that the evidence was Grade D for 4.1.
- Developer cites updated 2019 UK Renal Association Clinical Practice Guideline on Haemodialysis recommendation 4.1
    - 4.1 Fluid Assessment and Management in Adults recommends "avoiding excessive UFRs by addressing fluid gains, accepting staged achievement of target weight, or using an augmented schedule, as necessary". [1B - Grade 1: recommendation to do something where the benefits clearly outweigh the risks for most, if not all patients. Grade B: "moderate-quality evidence from randomized trials that suffer from serious flaws in conduct, inconsistency, indirectness, imprecise estimates, reporting bias, or some combination of these limitations, or from other study designs with special strength."]
    - This recommendation, wherein an absolute UFR threshold is not identified, represents a change from the 2009 iteration in which a maximum rate of 10 ml/kg/hour was recommended.
- Developer provides summaries of additional studies that assess the impact of negative outcomes from UHR.

Questions for the Committee:

- The evidence for this measure carries a strong recommendation but has limitations on the evidence provided to support the recommendation. Is the supporting evidence within the 2015 KDOQI guideline and supplemental evidence provided by the developer sufficient?

Guidance from the Evidence Algorithm

Measure assesses performance on a process (Box 3) -> Evidence presented includes graded practice guidelines based on systematic review of literature (Box 4) -> Quantity: low/high; Quality: low/moderate; Consistency: moderate/moderate (Box 5) -> Moderate (Box 5b) -> Moderate

Preliminary rating for evidence: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

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

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- Performance scores obtained during testing are as follows:

- Mean Score = 11.66% (lower = better performance); 95% CI = 11.46-11.87%; Standard Deviation = 6.92
- Minimum Score = 0%; Maximum Score = 50%*
- Median = 10.88%; Mode = 8.00%; Interquartile Range = 8.14

- Results show a significant spread between both the minimum and maximum scores, as well as the median and minimum and maximum scores, indicating that the measure identifies clinically and practically meaningful differences in performance among the measured entities.

Disparities

- Developer asserts that no disparities data is yet available for NQF 2701:
  - A structural reporting measure based on NQF 2701 is currently being implemented in the ESRD Quality Incentive Program (QIP) for PY 2020. As such, no data on disparities are yet available.
  - Likewise, SDS data were not collected during measure testing by the developer.
- Developer notes that while existing evidence on disparities in this area remains limited, a large observational study of 118,394 dialysis patients in a large dialysis organization between 2008 – 2012 demonstrated disparities in both UFR prescription and related outcomes.
  - Specifically, Assimon et al., found in their 2016 AJKD publication an association between high UFRs and all-cause mortality in blacks, non-Hispanics, and in patients with a higher BMI.
  - The authors also found that patients with an average UFR >13 were significantly more likely to be female, non-black, and Hispanic.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?
- No disparities information is provided for the measure, but studies are cited; are you aware of additional evidence that disparities exist in this area of healthcare?

Preliminary rating for opportunity for improvement: ☐ High ☒ Moderate ☐ Low ☐ Insufficient

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures–are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure.

- The CPGs used for evidence have ungraded or low quality evidence and the 23 observational studies used to support only include 4 since 2015, supportive of concepts but not necessarily with data as to specifics of this measure
- Process Measure. The intent of this measure is to foster the use of slower, gentilier dialysis sessions to reduce hemodialysis related mortality. Adherence to either one or two of the approaches will help attenuate the rapid fluctuations in fluid balance and blood pressure that contribute to cardiovascular morbidity and mortality in hemodialysis patients. Measure is derived from data collected.
- The goal of the measure is to encourage slower ultrafiltration rates by removing less fluid or by increasing treatment time. The evidence is weakly related to the goal since these are separate processes and longer treatments may not necessarily be associated with slow UFR

- Evidence is based on guidelines, which in turn is based on retrospective research. Not that fluid overload and removal isn't important but little actual prospective evidence exist to support this specific quantitative measure.
- While recommended by KDOQI, evidence in the literature is not based on randomized controlled trials. The measure does apply directly to the KDOQI guidelines, although KDOQI doesn't specify actual rate of ultrafiltration (but rather that it should "allow for optimal balance among achieving euvolemia, adequate BP control and solute clearance, while minimizing hemodynamic instability and intradialytic symptoms). I am concerned that the 2015 KDOQI guideline is based on expert opinion rather than truly evidence based medicine.
- There is a great deal of evidence associating shorter dialysis time and high UFR with increased morbidity but less consensus on maximum acceptable UFR (e.g. > 10 vs >13).
- Moderate rating
- I will not be evaluating NQF 2701 because of a conflict of interest.
- low to moderate evidence
- The structure and process relate well to the desired outcome (decreased relative risk of mortality). The outcome remains based on observational studies and comparison between potentially differing cohorts (in-center nocturnal thrice weekly patients are different from many of their regular in-center peers; they generally do not rely on medical transportation, have higher physical function, have help at home, etc.). While the desired outcome is decrease in mortality, I question whether the HD patients value this outcome if it comes at the cost of more time (or more days) at the dialysis clinic.
- The measure is supported by empirical data. There have been additional studies and update to clinical guidelines for the continued use of this measure.
- I agree
- The measure directly assesses the percentage of patients whose UFr >13 and receives less than 240 minutes per treatment.
- The evidences cited is from the 2015 KDQOI and UK guideline for hemodialysis. The evidence was given a Grade D which related to low quality due to most of the studies were observational. 4.1.1 of longer treatment times for increased weight gains was not graded nor was 4.2.1 using prescribed UFR's for achieving euvolemia. 4.2 of dietary restrictions of sodium and water was graded as B. The evidence of high UFR's and shorter dialysis treatments do correlate to SHR and SMR.

1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?

- Median 10.9% with rage 0 to 50% from combined LDO data; more limited disparity data but some provided suggesting some disparities in care
- Mean Score - 11.66% (lower= performance)' 95% cl= 11.46-11.87% standard deviation = 6.92. Minimum Score=0%; Maximum Score = 50% Median = 10.88%;Mode =8% Interquartile =8.14. Results have shown significant spread between both the minimum and maximum scores, as well as the median and minimum and maximum scores indicating that the measure identifies clinically and practically meaningful differences in performance among the measured entities. No disparities data is available.
- There is a gap between minimum and maximum scores
- Moderate: A gap does exist in the data
- Agree with moderate rating. I am not aware of any evidence that disparities exist; however, I don't know that this has been evaluated in great depth

- A performance gap exists but little data on disparities available.
- Moderate rating
- I will not be evaluating NQF 2701 because of a conflict of interest.
- no convincing disparities data , yes performance gap present
- Current performance measures suggest a meaningful gap. There are few data on disparities.
- Both UFR >=13 and tx times still have room for improvement. More studies are needed to explore ways to decrease the disparities. From a regulatory and compliance perspective, there continues to be noncompliance from the QAPI team to analyze root cause related to UFRs and tx times. Shortened treatment times and no shows are also not explored as expected.
- The data showed a performance gap of 10% of dialysis treatments do not meet the outcome measure. Also, only 30% of dialysis treatment session have a session length of >240 min. No disparities were found.
- Developer did not provide disparity data although there are publications that assert there is
- 2701 is used as a QIP measure. A retrospective study was conducted that showed significant spread in performance of measured entities. DOPPS data shows 10% of facilities have UFR's >= 13 and only 10% of facilities treatment times are >240. No disparity or SDS data is available.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Renal Committee Staff

NQF Staff Evaluation Summary

Reliability

- Testing was conducted at a total of 4,252 dialysis facilities from three dialysis providers.
- ICCs were calculated for each provider organization.
    - The KCQA measure was analyzed for within- and between-facility variance among patients' dialysis sessions that did not meet the quality standard specifications.
    - The "within" facility variation is the "error variance" or "noise" that reflects the degree of between-month variation in the measure that occurs within a facility.
    - The "between" facility variation is the explained or "systematic variance" (i.e., the "signal") that is attributable to variation in performance between facilities and represents real differences in performance.
    - The intra-class correlation coefficient (ICC) was calculated to estimate the ratio of the between- to the within-facility variance, standardized for both the level of variation and the numbers of observations examined.

Dialysis Organization	Intra Class Correlation	Ratio of Between to Within Facility Co Variance
A	.60	1.7
B	.65	2.0
C	.70	2.3

- Moderate ICCs reported suggesting good reliability at the facility score level.

Validity

- Measure developer tested score level validity using convergent validity, a common approach to score level testing.
- Validity of the measure was evaluated by correlating facility-specific scores from other NQF endorsed measures that look at mortality and readmissions.
    - Standardized Hospitalization Ratio for Admissions measure (SHR, NQF #1463)
    - Standardized Mortality Ratio* measure (SMR, NQF #0369)
- Results were statistically significant and directionally appropriate, with low, positive values (0.03-0.17).

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The staff is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The staff is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

Preliminary rating for reliability:	☐ **High**	☒ **Moderate**	☐ **Low**	☐ **Insufficient**
Preliminary rating for validity:	☐ **High**	☒ **Moderate**	☐ **Low**	☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?

- elements seem well defined and no concerns regarding implementation
- Data elements were clearly defined - patient data from health records.
- Testing shows good performance on reliability, validity, and specifications
- Moderate: for in center dialysis, the data elements are well defined. No accommodation is made for residual renal function which may or may not affect this outcome
- No concerns about reliability or validity. Don't think there is a need to discuss and/or vote.
- ICC's suggest good reliability
- Moderate rating
- I will not be evaluating NQF 2701 because of a conflict of interest.
- moderate reliability evidence
- No major concern
- None
- Data elements are clearly defined. I have no concerns about the likelihood that this measure can be implemented.
- Data elements are defined. Measure can be implemented. No concerns
- The data is collected from the EMR. The UFR rates are also a reported QIP measure. No concerns about implementation.

2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?

- data provided that showed differentiation between centers is achieved
- No
- No
- No
- no
- no
- No concerns
- I will not be evaluating NQF 2701 because of a conflict of interest.
- between and within facility data are reproducible
- No major concern
- None
- no concerns
- No concerns
- Data that was gathered was with 3 large dialysis facilities. Data showed high variability between facilities and low variability within facilities.

2b1. Validity -Testing: Do you have any concerns with the testing results?

- compared measure results to SHR and SMR with association in direction desired

- No concerns with validity. Overall rating of validity - moderate - is the highest eligibility rating if score - level testing has NOT been conducted
- No
- No
- no
- no
- No concerns with Validity. Moderate Rating
- I will not be evaluating NQF 2701 because of a conflict of interest.
- moderate (right direction and statistically significant)
- No concern
- None
- no concerns
- No concerns
- Validity was tested against other related measures of SMR and SHR. No concerns

2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment) 2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure? 2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?

- exclusion criteria seem clear and reasonable
- No risk adjustment or risk stratification,.
- The measure looks at two different and only partly related results: limited ultrafiltration rate and treatment duration. To assess how this measure impacts outcomes, it needs to measure a single factor
- It may be interested to look at CHF status as a risk adjustor.
- Exclusions seem to be appropriate. Hard to control for inter-dialytic weight gain depending on adherence of patients to their fluid restriction. Not sure if I saw an appropriate risk adjustment strategy.
- It is possible that social factors might affect risk but no data presented.
- No issues identified
- I will not be evaluating NQF 2701 because of a conflict of interest.
- no risk adjustments and no SD data available exclusions appropriate
- Exclusions are appropriate. I think there exists social risk factors that remain difficult to quantify or include in analyses. Some patients cannot afford to longer treatment times due to work schedules, child care demands, limited opportunities for transportation to and from dialysis that precludes longer times. I do not see an easy way to case adjust for these rather intangible factors.
- The exclusions were appropriate for this measure.
- 2b2: no exclusions 2b3: no risk adjustment
- Exclusions are appropriate.
- Risk factor data was not available

2b4-6. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data) 2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?

- No specific concerns
- No 5%c estimated missing data which is inconsequential.
- No
- Only if residual renal function may make a difference
- Agree that these items listed can represent threats to validity, as well as when patient goes against medical advice and asks to be taken off early.
- Correlates similar to other established measures
- No concerns with validity
- I will not be evaluating NQF 2701 because of a conflict of interest.
- exclusion criteria well justified, data readily obtainable
- Missing data may constitute a threat to this measure.
- The data demonstrates a comparability. And missing data would not be a threat to validity.
- 2b4-6: No threats to validity. 2b4: the results show meaningful differences 2b5: n/a 2b6: the missing data did not skew the total data
- No concerns
- Data shows that few facilities have longer treatment times to correct for higher UF rates. This demonstrates differences in quality

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. **Feasibility** is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

Required data elements are in electronic sources.

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Is the data collection strategy ready to be put into operational use?

Preliminary rating for feasibility: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?

- seems very feasible
- No concerns.

- None
- No issues
- No concerns re: feasibility
- no concerns
- elements are electronically captured to lessen reporting burden. High rating for feasibility
- I will not be evaluating NQF 2701 because of a conflict of interest.
- no concerns
- Presently data available and used in common healthcare delivery. No concerns.
- All the data is electronically available and currently used as part of QIP.
- The data can be captured from electronic records.
- No concerns
- All data elements are generated through the EMR. No concerns

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No** ☐ **UNCLEAR**

OR

Planned use in an accountability program? ☐ **Yes** ☐ **No**

Accountability program details

- NQF 2701 is currently being implemented as a facility-level reporting measure within CMS´s ESRD QIP for the purposes of public reporting, payment, and external quality improvement/benchmarking. The ESRD QIP is a nation-wide program encompassing all dialysis facilities receiving payment from Medicare as "a provider of services or a renal dialysis facility for renal dialysis services" under the ESRD PPS.
- The measure is also being used as a facility-level internal quality improvement metric by numerous dialysis organizations.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others [vetting]

- Feedback gathered from the participating and other dialysis organizations, dialysis facilities, patient groups, and other KCQA member organizations indicate that that measure is feasible, meaningful, and will provide an accurate reflection of quality.

Additional Feedback: Not available

Questions for the Committee:

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- How has the measure been vetted in real-world settings by those being measured or others?

Preliminary rating for Use: ☒ Pass ☐ No Pass

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results No information on improvement provided.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation Findings not yet available

Potential harms Not available

Additional Feedback: Not available

Questions for the Committee:

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- Do the benefits of the measure outweigh any potential unintended consequences?

Preliminary rating for Usability and use: ☐ High ☒ Moderate ☐ Low ☐ Insufficient

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided? 4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?

- In use by ESRD QIP and by some LDOs
- Measure is publicly reported by CMS.
- Publicly, it is reported as an ultrafiltration measure but that is not what it measures consistently.
- yes, no issues.
- Not sure

- Vetted and usable
- This measure is part of the ESRD QIP program.
- I will not be evaluating NQF 2701 because of a conflict of interest.
- main concern is whether the target number is truly correct goal (<=13) the data supporting a Specific number are weak ,it is a reasonable reporting measure, but MAY NOT perform as well as a VALID QUALITY measure , the ICC ratio ranged from .6 to .7 suggesting fair but not really robust ability to accurately discriminate quality differences between facilities
- Yes, presently in use and transparent.
- These measures are publicly reported. The 3 LDOs make up the majority of dialysis facilities in the U.S. and KCQA's committee expertise provided a robust representation for feedback.
- 4a1 - yes the measure is publicly reported and in an accountability program. 4a2: Use of the measure has been reported and opportunity for feedback provided.
- Measures are used in accountability of program and is reported.
- The measure is publicly reported via QIP. Facilities receive their results.

4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations? 4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.

- Usability seems clearly presented without significant concerns for harms or deleterious unintended consequences
- For clinical measures, the required data elements are routinely generated and used during care delivery i.e. blood pressure, lab test, diagnosis, medication order.
- It is not clear that the results can be used since two different measures are being combined
- The unintended consequence is that this standardized metric, combined with patient preference for treatment time, discourages personalized prescribing in some cases.
- Patients may be left fluid overloaded without addressing the inability to achieve EDW, thereby leading to increased BPs and other consequences of fluid overload; particularly if patient is unable or is not offered the opportunity for extra HD treatment outside of routine scheduled treatment time
- Potentially less relevant for elderly patients who may not tolerate longer dialysis
- no findings related to unintended consequences
- I will not be evaluating NQF 2701 because of a conflict of interest.
- see above it is unlikely that the measure will result in harm to patients , though it is POSSIBLE that attempts to achieve UFR limits (less than = to 13) will lead to higher EDW, volume expansion and related unwanted outcomes and events
- Potential unintended consequences:

  1) uniform HD time of 4 hours for all patients, regardless of needs or presence of residual renal function;

  2) blanket limit of UF rate to <13 cc/kg/hr, resulting in patients leaving over their ideal dry weight on days of high interdialytic weight gains;

  3) increased time on dialysis with mal-alignment of care goals between the healthcare providers and the patient when the patient views less time on dialysis as more valuable than potential lower mortality.

- The measure has yet to convert to a clinical metric but the kidney community anticipates the conversion. There has been no data to suggest harm. However, from an operational perspective, I would like to see the impact of scheduling and staffing to achieve the goal of longer tx times.
- 4b1 - improvement can be achieved by tracking this measure. 4b2: Keeping ultrafiltration rates low to prevent harm is a long-term benefit for the patients but some patients struggle with fluid management and feel they are "harmed" when the total fluid is not removed during their treatment. Many struggle with understanding the concept.
- Rationale has been provided on how performance results will improve care. Do not think there are unexpected harms.
- No unintended consequences. Benefits of longer treatment times and lower UFR's is clear.

Criterion 5: Related and Competing Measures

Related or competing measures

0249 : Delivered Dose of Hemodialysis Above Minimum

0256 : Minimizing Use of Catheters as Chronic Dialysis Access

0257 : Maximizing Placement of Arterial Venous Fistula (AVF)

0258 : Consumer Assessment of Healthcare Providers and Systems (CAHPS) In-Center Hemodialysis Survey (ICH CAHPS)

1460 : Bloodstream Infection in Hemodialysis Outpatients

2977 : Hemodialysis Vascular Access: Standardized Fistula Rate

2978 : Hemodialysis Vascular Access: Long-term Catheter Rate

Harmonization

Measures are harmonized with related and competing measures to the extent possible.

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?

- Does not seem applicable
- Yes other measures. The measure specifications are harmonized with related measures.
- No
- No
- No
- Harmonized
- harmonized with competing measures listed in the measure worksheet
- I will not be evaluating NQF 2701 because of a conflict of interest.
- no
- Delivering a minimal dose of dialysis, related perhaps, but not competing.
- Yes, but these measures are harmonized.
- 0249: Delivered Dose of Hemodialysis Above Minimum. Oft times when patients reach the minimum Kt/V they are satisfied even if they are not obtaining their estimated dry weight. Patients are

frequently motivated by the time element of dialysis - always wanting to decrease their time amount of treatment.

- Measures was harmonized with completing measures.
- There are related measures. No additional harmonization needed

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 01/15/2021

- Comment by: Kidney Care Partners

  Kidney Care Partners (KCP) appreciates the opportunity to submit early (pre-Standing Committee meeting) comments on the measures under consideration for endorsement in the National Quality Forum's Renal Project Fall 2020 Cycle. KCP is a coalition of members of the kidney care community that includes the full spectrum of stakeholders related to dialysis care—patient advocates, healthcare professionals, dialysis providers, researchers, and manufacturers and suppliers—organized to advance policies that improve the quality of care for individuals with both chronic kidney disease and end stage renal disease. We commend NQF for undertaking this important work and offer comment on both measures under review. KCP believes fluid management is a critical area to address through performance measurement and supports continued endorsement of this measure.

- Of the 1 NQF member who have submitted a support/non-support choice:
  - 1 support the measure
  - 0 do not support the measure

Combined Methods Panel Scientific Acceptability Evaluation

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 2701

Measure Title: Avoidance of Utilization of High Ultrafiltration Rate (>/=13 ml/kg/hour)

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☒ **Electronic Health Data** ☒ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented? ☒ Yes ☐ No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   ***NOTE**: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**

   - **None identified**

RELIABILITY: TESTING

Submission document: "MIF_2701" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level ☒ Measure score ☐ Data element ☐ Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure ☒ Yes ☐ No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of **patient-level data** conducted?

   ☐ Yes ☐ No

6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

   - Testing was conducted at a total of 4,252 dialysis facilities from three dialysis providers.
   - ICCs were calculated for each provider organization.
     - The KCQA measure was analyzed for within- and between-facility variance among patients' dialysis sessions that did not meet the quality standard specifications.
     - The "within" facility variation is the "error variance" or "noise" that reflects the degree of between-month variation in the measure that occurs within a facility.
     - The "between" facility variation is the explained or "systematic variance" (i.e., the "signal") that is attributable to variation in performance between facilities and represents real differences in performance.
     - The intra-class correlation coefficient (ICC) was calculated to estimate the ratio of the between- to the within-facility variance, standardized for both the level of variation and the numbers of observations examined.

Dialysis Organization	Intra Class Correlation	Ratio of Between to Within Facility Co Variance
A	.60	1.7
B	.65	2.0
C	.70	2.3

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   - Moderate ICCs reported, suggesting good reliability at the facility measure score level.

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

   ☒ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and **all** testing results):

    ☐ **High** (NOTE: Can be HIGH **only if** score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has **not** been conducted)

    ☐ **Low** (NOTE: Should rate **LOW** if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: Should rate **INSUFFICIENT** if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**
    - Moderate ICC results provided; testing was appropriately conducted using an common methodological approach.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2.
    - Exclusion appear appropriate and occurring with low to moderate frequency, with in-facility care, home dialysis, <7 HD treatments occurring most often (5-7% each).

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.
    - None identified

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**

    Submission document: Testing attachment, section 2b5.
    - None identified

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6.
    - None identified

16. **Risk Adjustment**

    16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☒ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☒ No

VALIDITY: TESTING

17. **Validity testing level: ☒ Measure score ☐ Data element ☐ Both**
18. **Method of establishing validity of the measure score:**

☒ Face validity

☒ Empirical validity testing of the measure score

☐ N/A (score-level testing not conducted)

19. **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

- Measure developer tested score level validity using convergent validity, a common approach to score level testing.
- Validity of the measure was evaluated by correlating facility-specific scores from other NQF endorsed measures that look at mortality and readmissions.
  - Standardized Hospitalization Ratio for Admissions measure (SHR, NQF #1463)
  - Standardized Mortality Ratio* measure (SMR, NQF #0369)
- Results were statistically significant and directionally appropriate, with low, positive values(0.03-0.17)

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

21. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**

NOTE that data element validation from the literature is acceptable.

Submission document: *Testing attachment, section 2b1.*

☐ **Yes**

☐ **No**

☒ **Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there **are** threats to validity and/or relevant threats to validity were **not assessed OR** if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level **is required**; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**
    - Measure developer provided a facility score-level convergent validity Pearson's correlation analysis between the measure of interest and two external measures of quality. The results established weak, positive correlations that were directionally appropriate and statistically significant. This is not an uncommon method or result.

ADDITIONAL RECOMMENDATIONS

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**
    - **None identified.**

Brief Measure Information

NQF #: 2701

Corresponding Measures:

De.2. Measure Title: Avoidance of Utilization of High Ultrafiltration Rate (>=13 ml/kg/hour)

Co.1.1. Measure Steward: Kidney Care Quality Alliance (KCQA)

De.3. Brief Description of Measure: Percentage of adult in-center hemodialysis patients in the facility whose average ultrafiltration rate (UFR) is >=13 ml/kg/hour AND who receive an average of <240 minutes per treatment during the calculation period.

1b.1. Developer Rationale: Ultrafiltration rates (UFRs) are determined by the amount of fluid that must be removed from the patient and dialysis session length. As treatment time decreases, UFR tends to increase and vice versa. Both high UFRs (>=13 ml/kg/hour) and abbreviated session duration (<240 minutes) are associated with a greater risk of all-cause and cardiovascular mortality in hemodialysis patients, and research suggests that dialysis sessions >240 minutes are independently associated with a significantly reduced relative risk of mortality.

The intent of this measure is to generally foster the use of slower, gentler dialysis sessions to reduce hemodialysis-related mortality. Success for the measure can be achieved by employing either or both of two approaches: 1) dialyzing patients at an average UFR <13 ml/kg/hour AND/OR 2) dialyzing patients for an average of >240 minutes per session during the reporting period. Adherence to these conventions will help attenuate the rapid fluctuations in fluid balance and blood pressure that contribute to cardiovascular morbidity and mortality in hemodialysis patients.

S.4. Numerator Statement: Number of patients* from the denominator whose average UFR is >=13 mg/kg/hr (NOT just >13) hour AND who receive an average of <240 minutes per treatment during the calculation period.**

*To address the fact that patients may contribute varying amounts of time to the annual denominator population, results will be reported using a "patient-month" construction.

** The calculation period is defined as the same week that the monthly Kt/V is drawn.

S.6. Denominator Statement: Number of adult in-center hemodialysis patients in an outpatient dialysis facility undergoing chronic maintenance hemodialysis during the calculation period.

S.8. Denominator Exclusions: The following patients are excluded from the denominator population:

1. Patients <18 years of age (implicit in denominator definition).
2. Home dialysis patients (implicit in denominator definition).
3. Patients in a facility <30 days.
4. Patients with >4 hemodialysis treatments during the calculation period.
5. Patients with <7 hemodialysis treatments in the facility during the reporting month.
6. Patients without a completed CMS Medical Evidence Form (Form CMS-2728) in the reporting month.
7. Kidney transplant recipients with a functioning graft.
8. Facilities treating <=25 adult in-center hemodialysis patients during the reporting month.

De.1. Measure Type: Process

S.17. Data Source: Electronic Health Records

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Oct 02, 2015 **Most Recent Endorsement Date:** Oct 02, 2015

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not applicable.

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

NQF_EvidenceAttachment_10-27-20REDLINE.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

Yes

1a. Evidence (subcriterion 1a)

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☒ Process: **Avoidance of use of ultrafiltration rate (UFR) >=13 ml/kg/hour AND/OR dialysis session time <240 minutes.** (NOTE: Success for the measure can be achieved by employing **either** or **both** of two approaches: 1) Dialyzing patients at an average UFR <13 ml/kg/hour; AND/OR 2) Dialyzing patients for an average of >=240 minutes per session during the reporting period.)

☐ Appropriate use measure:

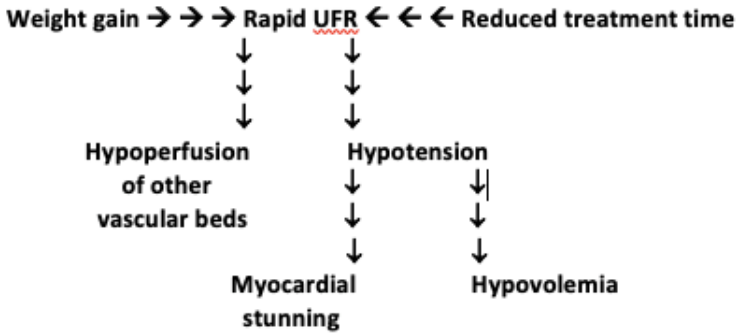
☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

Process Being Measured: Avoidance of use of ultrafiltration rate (UFR) >=13 ml/kg/hour and/or dialysis session time <240 minutes.

Logic Diagram: Potential mechanisms underlying the association between high UFR and adverse outcomes:



Rationale: Ultrafiltration rate (UFR) is determined by the amount of fluid that must be removed from the patient and dialysis session length. As treatment time decreases, UFR tends to increase and vice versa. Both high UFR (>=13 ml/kg/hour) and abbreviated session duration (<240 minutes) are associated with a greater risk of all-cause and cardiovascular mortality in hemodialysis patients, with research suggesting that dialysis session length >=240 minutes is independently associated with a significantly reduced relative risk of mortality.

The intent of this measure is thus to generally foster the use of slower, gentler dialysis sessions to reduce hemodialysis-related mortality. The measure criteria can be met by employing either or both of two approaches: 1) dialyzing patients at an average UFR <13 ml/kg/hour; and/or 2) dialyzing patients for an average of >=240 minutes per session during the reporting period. Adherence to these conventions will help attenuate the rapid fluctuations in fluid balance and blood pressure that contribute to cardiovascular morbidity and mortality in hemodialysis patients.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

Not applicable.

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

Not applicable.

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based

on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☒ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☒ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • Title: • Author: • Date: • Citation, including page number: • URL:	Title: Kidney Disease Outcomes Quality Initiative (KDOQI). Clinical Practice Guideline for Hemodialysis Adequacy Author: National Kidney Foundation Date: 2015 Citation: National Kidney Foundation. Kidney Disease Outcomes Quality Initiative (KDOQI). Clinical Practice Guideline for Hemodialysis Adequacy: 2015 Update. *Am J Kidney Dis.* 2015;66(5):884-930. Relevant Pages: 913-916. URL: https://www.ajkd.org/action/showPdf?pii=S0272-6386%2815%2901019-7 NOTE: The relevant KDOQI Guidelines have been updated since the measure was endorsed in 2015. This endorsement maintenance submission cites these most recent recommendations; the original submission relied on the 2006 Guidelines.

Systematic Review	Evidence
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	**Guideline 4: Volume and Blood Pressure Control: Treatment Time and Ultrafiltration Rate** **4.1** We recommend that patients with low residual kidney function (<2 mL/min) undergoing thrice weekly hemodialysis be prescribed a bare minimum of 3 hours per session. (1D) • **4.1.1** Consider additional hemodialysis sessions or longer hemodialysis treatment times for patients with large weight gains, high ultrafiltration rates, poorly controlled blood pressure, difficulty achieving dry weight, or poor metabolic control (such as hyperphosphatemia, metabolic acidosis, and/or hyperkalemia). (Not Graded) **4.2** We recommend both reducing dietary sodium intake as well as adequate sodium/water removal with hemodialysis to manage hypertension, hypervolemia, and left ventricular hypertrophy. (1B) • **4.2.1** Prescribe an ultrafiltration rate for each hemodialysis session that allows for an optimal balance among achieving euvolemia, adequate blood pressure control and solute clearance, while minimizing hemodynamic instability and intradialytic symptoms. (Not Graded)
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Appraisal of the quality of the evidence and the strength of recommendations followed the Grading of Recommendation Assessment, Development, and Evaluation (GRADE) approach. **Recommendation 4.1 Evidence Grade: D** • Evidence Grade D is defined as "very low quality of evidence" for which "the estimate of effect is very uncertain and often will be far from the truth." **Recommendation 4.2 Evidence Grade: B** • Evidence Grade B is defined as "moderate quality of evidence," with "the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different." **Recommendations 4.1.1. and 4.2.1: Not Graded** • "Not Graded" was used typically "to provide guidance based on common sense or where the topic does not allow adequate application of evidence. The ungraded recommendations are generally written as simple declarative statements, but are not meant to be interpreted as being stronger recommendations than Level 1 or 2 recommendations."
Provide all other grades and definitions from the evidence grading system	Other evidence grades in the grading system: • **Evidence Grade A:** High quality of evidence for which KDOQI is "confident that the true effect lies close to that of the estimate of the effect." • **Evidence Grade C:** Low quality of evidence for which "the true effect may be substantially different from the estimate of the effect."

Systematic Review	Evidence
Grade assigned to the **recommendation** with definition of the grade	**Recommendation 4.1 Grade: 1** • Recommendation Grade Level 1 is defined as a "strong recommendation" for which "most patients should receive the recommended course of action" and "the recommendation can be adopted as policy for most situations." **Recommendation 4.2 Grade: 1** • Recommendation Grade Level 1 (see above). **Recommendations 4.1.1. and 4.2.1: Not Graded** • "Not Graded" was used typically "to provide guidance based on common sense or where the topic does not allow adequate application of evidence. The ungraded recommendations are generally written as simple declarative statements, but are not meant to be interpreted as being stronger recommendations than Level 1 or 2 recommendations."
Provide all other grades and definitions from the recommendation grading system	Other recommendation grades include: • **Recommendation Level 2:** Conditional recommendation for which "different choices will be appropriate for different patients" and "the recommendation is likely to require substantial debate and involvement of stakeholders before policy can be determined."
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	**Quantity:** A total of 39 studies were cited by KDOQI as evidence supporting Guideline 4.2. **Quality:** The 2015 KDOQI update included a review of clinical trials and observational studies published between 2000 and March 2014 on topics including high-frequency hemodialysis and risks; prescription flexibility in initiation timing, frequency, duration, and ultrafiltration rate; and volume and blood pressure control.
Estimates of benefit and consistency across studies	**Recommendation 4.1:** "The estimate of effect is very uncertain and often will be far from the truth." **Recommendation 4.2:** "The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different." **Recommendations 4.1.1 and 4.2.1:** "Guidance based on common sense or where the topic does not allow adequate application of evidence." The publication did not provide an estimate of consistency across the cited studies.

Systematic Review	Evidence
What harms were identified?	No harms were cited for the recommendation; the supporting body of evidence validates the recommendations' premises—failure to prescribe an appropriate hemodialysis UFR and session duration to achieve euvolemia and minimize hemodynamic instability is associated with adverse outcomes ranging from cardiovascular events and mortality to hypotensive seizures. Nevertheless, potential harms stemming from the process of setting UFR to achieve a set target ("dry") weight post-dialysis were noted. While this has been the accepted method of maintaining a consistent volume state, the inaccuracy of estimation is widely appreciated. Both over- and underestimation are common, with the former contributing to hypertension and left ventricular hypertrophy, and the latter accelerating the loss of residual kidney function and perhaps risking myocardial stunning.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Numerous studies addressing UFR and dialysis treatment time have been published since this guideline was released in 2015 (see 1a.4.1 and 1a.4.2 below for details and citations); none were identified that contradict the KDOQI recommendation, which does not identify an absolute threshold for UFR and establishes a "bare minimum" treatment time.

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	Title: Clinical Practice Guideline on Haemodialysis Author: UK Renal Association Date: 2019 Citation: UK Renal Association. Clinical Practice Guideline on Haemodialysis. *BMC Nephrology.* 2019;20:379-415. Relevant Pages: 3-4 (382-383). URL: https://bmcnephrol.biomedcentral.com/track/pdf/10.1186/s12882-019-1527-3. NOTE: The UK Renal Association Guidelines have been updated since 2701 was endorsed in 2015. This endorsement maintenance submission cites the most recent recommendations; the original submission referenced the 2009 Guidelines.

Systematic Review	Evidence
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	**Guideline 4.1: Fluid Assessment and Management in Adults** We recommend avoiding excessive UFRs by addressing fluid gains, accepting staged achievement of target weight, or using an augmented schedule, as necessary. [1B] **NOTE:** This recommendation, wherein an absolute UFR threshold is not identified, represents a change from the 2009 iteration in which a maximum rate of 10 ml/kg/hour was recommended.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Appraisal of the quality of the evidence and the strength of recommendations followed a modified Grading of Recommendation Assessment, Development, and Evaluation (GRADE) approach. **Guideline 4.1 Evidence Grade: B** • Evidence Grade B is defined as "moderate-quality evidence from randomized trials that suffer from serious flaws in conduct, inconsistency, indirectness, imprecise estimates, reporting bias, or some combination of these limitations, or from other study designs with special strength."
Provide all other grades and definitions from the evidence grading system	Other evidence grades in the grading system include: • **Evidence Grade A:** "High-quality evidence that comes from consistent results from well-performed randomized controlled trials, or overwhelming evidence of some other sort (such as well-executed observational studies with very strong effects)." • **Evidence Grade C:** "Low-quality evidence from observational studies, or from controlled trials with several very serious limitations." • **Evidence Grade D:** "Evidence is based only on case studies or expert opinion."
Grade assigned to the **recommendation** with definition of the grade	**Guideline 4.1 Recommendation Grade: 1** • Recommendation Grade 1 is a strong recommendation to do (or not do) something, where the benefits clearly outweigh the risks (or vice versa) for most, if not all patients.
Provide all other grades and definitions from the recommendation grading system	Other recommendation grades in the grading system include: • **Recommendation Grade 2:** "A weaker recommendation, where the risks and benefits are more closely balanced or are more uncertain."
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	**Quantity:** A total of 15 studies were cited as evidence supporting Guideline 4.1. **Quality:** The guideline is an update of a previous version written in 2009 and included systematic literature searches of clinical trials and observational studies undertaken by lead authors to identify all relevant evidence published up until the end of June 2018.
Estimates of benefit and consistency across studies	The publication did not provide an estimate of benefit or consistency across the cited studies.

Systematic Review	Evidence
What harms were identified?	As with the KDOQI guidelines, the cited body of evidence support the premise of the recommendation—failure to appropriately address fluid gains through achievement of target weights and/or use of augmented schedules to avoid excessive UFR is associated with mortality and adverse cardiovascular-related outcomes. Nevertheless, it was noted that potential harms might stem from the process of setting UFR to achieve a set target ("dry") weight post-dialysis, for which the inaccuracy of estimation is widely appreciated. Both over- and underestimation are common, with the former contributing to hypertension and left ventricular hypertrophy, and the latter accelerating the loss of residual kidney function and perhaps risking myocardial stunning.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Few studies addressing UFR have been published since this guideline was released in 2019 (see 1a.4.1 and 1a.4.2 below for details and citations); none were identified that contradict the UK recommendation, which does not identify an absolute threshold for UFR.

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

UFR is recognized as an important and modifiable risk factor for morbidity and mortality among patients receiving maintenance hemodialysis, yet identifying a specific UFR target and determining the appropriate amount of fluid to remove during dialysis remains a clinical challenge. Both volume overload with excessive interdialytic weight gain (IDWG) and recurrent episodes of intradialytic hypotension associated with the use of higher UFRs are recognized as important predictors of morbidity and mortality.[1,2,3]

[1] Agarwal R, Weir MR. Dry-weight: A concept revisited in an effort to avoid medication-directed approaches for blood pressure control in hemodialysis patients. *CJASN.* 2010;5:1255–1260.

[2] Chou JA, Kalantar-Zadeh K. Volume balance and intradialytic ultrafiltration rate in the hemodialysis patient. *Curr Heart Fail Rep.* 2017;14(5):421-427.

[3] Lopot F, Kotyk P, Blaha J, Forejt J. Use of continuous blood volume monitoring to detect inadequately high dry weight. *International Journal of Artificial Organs.* 1996;19:411–414.

Numerous observational studies in recent years have assessed outcomes associated with UFR:[4,5]

- The first included 22,000 prevalent hemodialysis patients from seven countries in the Dialysis Outcomes and Practice Patterns Study (DOPPS). Saran et al. found that UFR >10 ml/kg/hour was associated with both intradialytic hypotension (OR = 1.3, p = 0.045) and all-cause mortality (adjusted HR 1.09, p = 0.02).[6]
- Another prospective cohort study of 287 prevalent hemodialysis patients in Italy demonstrated that for every 1 ml/kg/hour increase in UFR there was a 22% increase in mortality risk (p <0.01). In secondary analyses, the authors identified a UFR of 12.4 ml/kg/hour as the most discriminatory cut-point for predicting two-year mortality.[7] Notably, the study was also restricted to patients with a urine output of 150 ml/day or less, offering compelling evidence for the association absent confounding from residual kidney function, typically associated with better clinical outcomes.[8]
    - An analysis of the Hemodialysis (HEMO) Study, a randomized controlled trial of 1,846 patients followed prospectively for 7 years, found that UFRs >13 ml/kg/hour were associated with a 59% increased risk of all-cause mortality (HR 1.59, 1.29-1.96) and a 71% increased risk of cardiovascular mortality (HR 1.71, 1.23-2.38 after adjustment (p <0.001 for both). In spline analyses, risk increased sharply after 10 ml/kg/hour.[9]
    - An observational study of 118,394 hemodialysis patients in a large dialysis organization between 2008 and 2012 dichotomized mean UFR over a 30-day period as <= or >10 and <= or >13 ml/kg/hour. Here again, UFRs >10 and >13 were both associated with higher all-cause mortality compared to their respective references (adjusted HRs 1.22 [1.20-1.24] and 1.31 [1.28-1.34]). The association was more pronounced in blacks, non-Hispanics, patients with longer dialysis vintage, longer session duration, and patients with higher BMI. When UFR was treated as a continuous variable, each 1 ml/kg/hour increase was associated with 3% increased risk for mortality.[10]
    - The preceding studies included markers of health such as albumin, blood pressure, and comorbidities, but did not fully account for potential residual confounders like patient resiliency or frailty. Shorter dialysis sessions may, for instance, be prescribed to frailer patients due to dialysis intolerance or a lower body weight. Flythe et al. offered some clarity on this issue by assessing treatment time and outcomes in a national cohort of patients from a large dialysis organization undergoing thrice weekly in-center hemodialysis. Patients prescribed dialysis sessions > and <240 minutes were pair-matched on post-dialysis weight, age, gender, and vascular access type. Session length <240 minutes was significantly associated with increased all-cause mortality (adjusted HR 1.26, 1.07–1.48; p = 0.005), with a dose-response between prescribed session length

---

[4] Slinlin Y, Babu M, Ishani A. Ultrafiltration rate in conventional hemodialysis: Where are the limits and what are the consequences? *Seminars in Dialysis.* 2018;31(6):544-550.

[5] Assimon MM, Flythe JE. Rapid ultrafiltration rates and outcomes among hemodialysis patients: Re-examining the evidence base. *Curr Opin Nephrol Hypertens.* 2015;24(6):525-530.

[6] Saran R, Bragg-Gresham JL, Levin NW, et al. Longer treatment time and slower ultrafiltration in hemodialysis: Associations with reduced mortality in the DOPPS. *Kidney Int.* 2006;69:1222-1228.

[7] Movilli E, Gaggia P, Zubani R, et al. Association between high ultrafiltration rates and mortality in uraemic patients on regular haemodialysis. A 5-year prospective observational multicentre study. *Nephrol Dial Transplant.* 2007;22:3547–3552.

[8] Vilar E, Farrington K. Emerging importance of residual renal function in endstage renal failure. *Semin Dial.* 2011; 24:487–494.

[9] Flythe JE, Kimmel SE, Brunelli SM. Rapid fluid removal during dialysis is associated with cardiovascular morbidity and mortality. *Kidney Int.* 2011;79:250-257.

[10] Assimon MM, Wenger JB, Wang L, Flythe JE. Ultrafiltration rate and mortality in maintenance hemodialysis patients. *Am J Kidney Dis.* 2016;68:911-922.

and survival.[11] The study did not directly address UFR, but suggests that the slower fluid removal facilitated by longer treatment times may be advantageous, independent of body weight.[12]

The near-linear association between high UFR and adverse outcomes illustrated in these landmark studies highlights a considerable opportunity to improve care and outcomes for dialysis patients—and offers a compelling framework upon which a performance metric can be constructed to address this vitally important aspect of dialysis care. Yet while the literature provides persuasive evidence supporting the use of UFR between 10 and 13—with more recent publications[13] suggesting that even lower rates may prove beneficial—KCQA is cognizant of the fact that imposing too restrictive a limitation would not be without consequence, increasing risk for pervasive failures of target weight achievement and volume expansion over time.[14]

KCQA recognizes that any effective fluid management measure must provide for clinical judgement, allowing physicians ample room to respond appropriately to varying clinical presentations in a manner that also meets the needs and preferences of their patients. KCQA and the larger renal community thus selected the <13 ml/kg/hour parameter not only because it carries the greatest consensus among experts, but because we believe it also offers a balanced, feasible approach to prevent the deleterious consequences of excessive UFR.

1a.4.2 What process was used to identify the evidence?

The process used to identify the evidence supporting NQF 2701 consisted of an extensive literature review, the clinical experience and expert consensus of KCQA members and the KCQA Feasibility/Testing Workgroup, and a retrospective review of pertinent database data from three large dialysis organizations with a correlation to existing measures of adverse outcomes (i.e., hospitalization and mortality) over the same time period. Relevant peer-reviewed publications since the measure was endorsed have been incorporated into this maintenance review submission.

1a.4.3. Provide the citation(s) for the evidence.

1. Agarwal R, Weir MR. Dry-weight: A concept revisited in an effort to avoid medication-directed approaches for blood pressure control in hemodialysis patients. *CJASN.* 2010;5:1255–1260.
2. Chou JA, Kalantar-Zadeh K. Volume balance and intradialytic ultrafiltration rate in the hemodialysis patient. *Curr Heart Fail Rep.* 2017;14(5):421-427.
3. Lopot F, Kotyk P, Blaha J, Forejt J. Use of continuous blood volume monitoring to detect inadequately high dry weight. *International Journal of Artificial Organs.* 1996;19:411–414.

[11] Flythe JE, Curhan GC, Brunelli SM. Shorter length dialysis sessions are associated with increased mortality, independent of body weight. *Kidney Int.* 2013; 83:104–113.

[12] Assimon MM, Flythe JE. Rapid ultrafiltration rates and outcomes among hemodialysis patients: Re-examining the evidence base. *Curr Opin Nephrol Hypertens.* 2015;24(6):525-530.

[13] Lee YJ et al. Ultrafiltration rate, residual kidney function, and survival among patients treated with reduced-frequency hemodialysis. *AJKD.* 2020;75(3):342-350.

[14] Flythe JE. Ultrafiltration rate clinical performance measures: Ready for primetime? *Semin Dial.* 2016;29(6):425-434.

4. Slinlin Y, Babu M, Ishani A. Ultrafiltration rate in conventional hemodialysis: Where are the limits and what are the consequences? *Seminars in Dialysis.* 2018;31(6):544-550.
5. Assimon MM, Flythe JE. Rapid ultrafiltration rates and outcomes among hemodialysis patients: Re-examining the evidence base. *Curr Opin Nephrol Hypertens.* 2015;24(6):525-530.
6. Saran R, Bragg-Gresham JL, Levin NW, et al. Longer treatment time and slower ultrafiltration in hemodialysis: Associations with reduced mortality in the DOPPS. *Kidney Int.* 2006;69:1222-1228.
7. Movilli E, Gaggia P, Zubani R, et al. Association between high ultrafiltration rates and mortality in uraemic patients on regular haemodialysis. A 5-year prospective observational multicentre study. *Nephrol Dial Transplant.* 2007;22:3547–3552.
8. Vilar E, Farrington K. Emerging importance of residual renal function in endstage renal failure. *Semin Dial.* 2011; 24:487–494.
9. Flythe JE, Kimmel SE, Brunelli SM. Rapid fluid removal during dialysis is associated with cardiovascular morbidity and mortality. *Kidney Int.* 2011;79:250-257.
10. Assimon MM, Wenger JB, Wang L, Flythe JE. Ultrafiltration rate and mortality in maintenance hemodialysis patients. *Am J Kidney Dis.* 2016;68:911-922.
11. Flythe JE, Curhan GC, Brunelli SM. Shorter length dialysis sessions are associated with increased mortality, independent of body weight. *Kidney Int.* 2013; 83:104–113.
12. Assimon MM, Flythe JE. Rapid ultrafiltration rates and outcomes among hemodialysis patients: Re-examining the evidence base. *Curr Opin Nephrol Hypertens.* 2015;24(6):525-530.
13. Lee YJ et al. Ultrafiltration rate, residual kidney function, and survival among patients treated with reduced-frequency hemodialysis. *AJKD.* 2020;75(3):342-350.
14. Flythe JE. Ultrafiltration rate clinical performance measures: Ready for primetime? *Semin Dial.* 2016;29(6):425-434.
15. Sinha AD, Agawaral A. Opinion: The fallacy of low interdialytic weight gain and lower ultrafiltration rate: Lower is not always better. *Semin Dial.* 2014;27(1):11-13.
16. Flythe JE, Curhan GC, Brunelli SM. Disentangling the ultrafiltration rate-mortality association: The respective roles of session length and weight gain. *Clin J Am Soc Nephrol*. 2013;8(7):1151-1161.
17. Lindberg M, Pruetz KG, Lindberg P et al. Interdialytic weight gain and ultrafiltration rate in hemodialysis: Lessons about fluid adherence from a national registry of clinical practice. *Hemodial Int.* 2013;17(4):548-556.
18. Curatola G, Bolignano D, Rastelli S, et al. Ultrafiltration intensification in hemodialysis patients improves hypertension but increases AV fistula complications and cardiovascular events. *J Nephrol.* 2011;24(4):465-473.
19. Flythe JE, Brunelli SM. The risks of high ultrafiltration rate in chronic hemodialysis: Implications for patient care. *Semin Dial*. 2011;24(3):259-265.
20. London GM. Ultrafiltration intensification for achievement of dry weight and hypertension control is not always the therapeutic gold standard. *J Nephrol.* 2011;24(4)395-397.
21. Brunelli SM, Chertow GM, Ankers ED, et al. Shorter dialysis times are associated with higher mortality among incident hemodialysis patients. *Kidney Int.* 2010;77:630–636.
22. Burton JO, Jefferies HJ, Selby NM, and McIntyre CW. Hemodialysis-induced cardiac injury: Determinants and associated outcomes. *Clin J Am Soc Nephrol.* 2009;4:914–920.
23. Miller JE, Kovesdy CP, Nissenson AR, et al. Association of hemodialysis treatment time and dose with mortality and the role of race and sex. *Am J Kidney Dis.* 2010; 55:100–112.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE *(e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.*

Ultrafiltration rates (UFRs) are determined by the amount of fluid that must be removed from the patient and dialysis session length. As treatment time decreases, UFR tends to increase and vice versa. Both high UFRs (>=13 ml/kg/hour) and abbreviated session duration (<240 minutes) are associated with a greater risk of all-cause and cardiovascular mortality in hemodialysis patients, and research suggests that dialysis sessions >240 minutes are independently associated with a significantly reduced relative risk of mortality.

The intent of this measure is to generally foster the use of slower, gentler dialysis sessions to reduce hemodialysis-related mortality. Success for the measure can be achieved by employing either or both of two approaches: 1) dialyzing patients at an average UFR <13 ml/kg/hour AND/OR 2) dialyzing patients for an average of >240 minutes per session during the reporting period. Adherence to these conventions will help attenuate the rapid fluctuations in fluid balance and blood pressure that contribute to cardiovascular morbidity and mortality in hemodialysis patients.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. ***(This is required for maintenance of endorsement****. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

A structural reporting measure based on KCQA's UFR metric is being implemented in the ESRD Quality Incentive Program (QIP) for PY 2020. Facilities must report all data elements required by the measure to calculate the UFR:

- HD Kt/V Date
- Post-Dialysis Weight
- Pre-Dialysis Weight
- Delivered Minutes of Hemodialysis
- Number of dialysis sessions delivered to the patient in the reporting month

Thus while at this time we thus have no formal performance scores to report, as has been the case with other reporting measures in the QIP we anticipate the UFR measure will be converted to a clinical measure after confirmation that the data elements can be feasibly captured.

Additionally, the measure was tested using data from three KCQA member dialysis organizations, each with the capacity to provide retrospective analyses from a data warehouse/repository. All pertinent data from all eligible (i.e., adult in-center hemodialysis) patients of the participating organizations during the testing period were included in the datasets. Testing encompassed 4,252 dialysis facilities and 412,522 patients across the three organizations. The study was conducted retrospectively on data from January 1, 2013-December 31, 2013.

Performance scores obtained during testing are as follows:

- Mean Score = 11.66% (lower = better performance)
- 95% CI = 11.46-11.87%
- Standard Deviation = 6.92
- Minimum Score = 0%
- Maximum Score = 50%*
- Median = 10.88%
- Mode = 8.00%
- Interquartile Range = 8.14

Results show a significant spread between both the minimum and maximum scores, as well as the median and minimum and maximum scores, indicating that the measure identifies clinically and practically meaningful differences in performance among the measured entities.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

In addition to the testing data presented above, a recent national sample of DOPPS data indicate that hemodialysis sessions performed at UFR >=13.0 ml/kg/hour remains at approximately 10 percent as of February 2020, indicating continued room for improvement in this aspect of dialysis care. (See Graph 1 in Appendix.)

The same DOPPS data also demonstrate considerable room for improvement in achieved average dialysis session length >240 minutes, the second approach by which the measure criteria can be met. This benchmark has moved little over the past decade, remaining at only approximately 30 percent as of February 2020 for a national sample. (See Graphs 2 and 3 in Appendix.)

Reference: US DOPPS Practice Monitor, April 2020, https://www.dopps.org/DPM.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. ***(This is required for maintenance of endorsement.*** *Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

As noted, a structural reporting measure based on NQF 2701 is currently being implemented in the ESRD Quality Incentive Program (QIP) for PY 2020. As such, no data on disparities are yet available. Likewise, SDS data were not collected during measure testing by KCQA.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

DOPPS also provides little additional information in this regard. Only two SDS groups—black vs non-black—are displayed, and the data show little difference between the two in either average UFR or achieved session length as of February 2019, with blacks faring slightly better in both categories. (See Graphs 4 and 5 in Appendix.).

However, as cited in the Evidence Attachment, we have identified one large observational study of 118,394 hemodialysis patients in a large dialysis organization between 2008 and 2012 that demonstrates a more pronounced association between high UFRs and all-cause mortality in blacks, non-Hispanics, and in patients with a higher BMI. The authors also found that patients with average UFR >13 were significantly more likely (p

<0.005 for all associations) to be female (1.33 [1.29-1.37]), non-black (1.28 [1.24-1.31]), and Hispanic (1.20 [1.14-1.27]).

References:

1. US DOPPS Practice Monitor, April 2020, https://www.dopps.org/DPM.
2. Assimon MM et al. Ultrafiltration rate and mortality in maintenance hemodialysis patients. Am J Kidney Dis. 2016;68:911-922.

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1, 2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

Measure Testing (subcriteria 2a2, 2b1-2b6)

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

The reliability of the measure was assessed using a repeated-measures analysis of variance (ANOVA) test. Data were statistically analyzed using the facility and the treatment month as independent variables and the measure scores as dependent variables. Analyses were conducted separately for all the facilities in each of the three dialysis organizations. The rationale for using a repeated-measures ANOVA is that there should be relatively little within-facility variation in the monthly proportion of patients' dialysis treatment sessions that do not meet the clinical standard threshold (e.g., have a UFR >=13 ml/kg/hour). Rather, if the measure is helpful in discriminating between high and low performing dialysis facilities, the level of variation from month to month should be high between facilities.

The KCQA measure was analyzed for within- and between-facility variance among patients' dialysis sessions that did not meet the quality standard specifications. The "within" facility variation is the "error variance" or "noise" that reflects the degree of between-month variation in the measure that occurs within a facility. The "between" facility variation is the explained or "systematic variance" (i.e., the "signal") that is attributable to variation in performance between facilities and represents real differences in performance. The intra-class correlation coefficient (ICC) was calculated to estimate the ratio of the between- to the within-facility variance, standardized for both the level of variation and the numbers of observations examined. The higher the ICC, the greater the reliability of the measures. The ratio of the between- to within-facility variation was also examined as a "signal to noise" ratio. Greater between-facility variation than within-facility variation indicates that the measure is discriminating between facilities.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

The table below reports the measure's intra-class correlation coefficient for each of the participating dialysis organizations, as well as the ratio of between- to within-facility variation:

Dialysis Organization	Intra Class Correlation	Ratio of Between to Within Facility Co Variance
A	.60	1.7
B	.65	2.0
C	.70	2.3

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

As demonstrated in the above table, the intra-class correlation for all organizations is high, indicating a good level of reliability within facilities over the course of the 12 months. Additionally, the estimated between-facility variance is greater than the within-facility variance, again suggesting that the measure discriminates between the participating facilities. Across all groups, there is more variation between facilities than within facilities, which when considered in light of the relatively high intra-class correlation coefficients, suggests that the measure is reliable and differentiates between facilities.

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☒ **Empirical validity testing**

☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

Criterion predictive (correlative) validity was used, assessing the correlation of the computed measure scores against some criterion (e.g., another measure of the same construct or an outcome) deemed valid. Specifically, the validity of the measure was evaluated by correlating facility-specific scores with each facility's 2013 Standardized Hospitalization Ratio for Admissions measure (SHR, NQF #1463) and Standardized Mortality Ratio* measure (SMR, NQF #0369) scores using Pearson's Correlation Coefficient. Both the SHR and SMR are NQF-endorsed publicly available dialysis facility outcome measures that the KCQA measure could be expected to impact.

To allow for correlation with the most current SHR and SMR scores publicly available on Dialysis Facility Compare (DFC), 2013 facility data were used for testing. If available, correlation to 2013 hospitalization rates from the facilities' DFRs also were analyzed.

Additionally, between July 2014 and February 2015, KCQA conducted an iterative assessment of face validity based on a series of conferences of the KCQA Steering Committee and the KCQA Feasibility/Testing Workgroup, as well as repetitive polling of the full KCQA at various stages of the measure development process.

*** The SMR specifications are based on a 4-year rolling period.**

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

The Pearson's Correlation Coefficients are summarized as follows:

Dialysis Organization	2013 SHR	2013 SMR	2013 Hospitalization Rate (from DFR)
A	0.12	- -	0.17
B	0.11	0.11	0.07
C	0.09	0.08	0.03

- - cell intentionally left blank

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

The correlation between the quality performance measure of avoidance of high UFR and the SMR and SHR are statistically significant and in the expected direction; facilities that have fewer patients with high UFR have lower mortality and hospitalization. The correlation supports the hypothesized underlying construct of the measure—that improving fluid management in dialysis patients will reduce the ultimate adverse patient outcomes of mortality and hospitalization.

2b2. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section 2b4*

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

For each facility across the three participating dialysis organizations the overall number and percentages of patients meeting each exclusion criterion was recorded for each of the 12 months. The monthly and annual frequencies of the occurrence of each exclusion and the variability of the exclusions were then analyzed.

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

The annual counts for the individual denominator exclusions across the three organizations were:

- **Age <18 years = 32,085 patient-months (0.66%)**
- **Patients receiving care in a facility <30 days = 306,860 patients-months (7.58%)**
- **Home dialysis patients = 192,645 patient-months (5.08%)**
- **<7 hemodialysis treatments in the facility during the reporting month = 335,606 patient-months (7.64%)**
- **Patients without a completed CMS Medical Evidence Form (Form CMS-2728) in the reporting month = 32,806 patient-months (0.65%)**
- **Kidney transplant recipients with a functioning graft = Not tested (discussed further below)**

- **Patients who receive 4 or more dialysis sessions during the calculation period = 72,133 patient-months (1.58%)**

The total number of annual exclusions across the three organizations was 657,227 patient-months, with a range of 18,439 to 458,112 patient-months excluded per organization. The average monthly exclusion across the three organizations was 55,769 patients, with a range of 1 to 288 patients excluded per facility each month.

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis.* ***Note**: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

As can be seen, with one exception (transplant recipients with functioning graft), the frequency with which the exclusions were encountered during testing is sufficient to demonstrate they are necessary to prevent unfair distortion of performance results; Table 4 of KCQA's Testing Data Attachment (attached to this form) documents that the variability in their occurrence across providers also supports the need for the exclusions.

We found that the "kidney transplant recipients with a functioning graft" exclusion could not be operationalized consistently across providers during the limited time for testing. However, we have retained it in the measure specifications since it remains clinically relevant and appropriate.

Because KCQA tested a separate fluid management measure using every session, exclusions were documented at that level and are presented here at that level. We did not perform a specific examination of exclusion rates during only the Kt/V week, since there is no reason to presume that the weekly rates would vary from monthly rates—i.e., it can be inferred that the percentages are equivalent for testing purposes.

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO PM, or resource use measure, skip to section 2b5.

2b3.1. What method of controlling for differences in case mix is used?

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with risk factors**

☐ **Stratification by risk categories**

☐ **Other,**

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of*

p<0.10; correlation of x or higher; patient factors should be present at the start of care) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

2b3.4a. What were the statistical results of the analyses used to select risk factors?

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

2b3.9. Results of Risk Stratification Analysis:

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

2b3.11. Optional Additional Testing for Risk Adjustment (***not required****, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

Descriptive statistics for the annual performance measure scores for all tested entities (facilities) were constructed. These statistics include the mean, standard deviation and standard error, 95% confidence interval, median, mode, range of scores, and the interquartile range of scores across the measured entities. We defined meaningful difference as a significant spread (>20%) between minimum and maximum scores or a significant spread between median and minimum or median and maximum scores.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

Findings are summarized here:

N	Range of Scores	Mean Score	Median Score	Mode of Scores	Interquartile Range
4,251 facilities	0-50%	11.66%	10.88%	8.00%	8.14
*	*	SD =6.92	*	*	*
*	*	SE = 0.11	*	*	*
*	*	95% CI = 11.46-11.87%	*	*	*

*cell intentionally left blank

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

Results are interpreted as showing a significant spread between both the minimum and maximum scores, as well as the median and minimum and maximum scores, indicating that the measure identifies clinically and practically meaningful differences in performance among the measured entities.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped

Note*: This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

All facilities remained in the analysis, regardless of the magnitude of missing data; however, participating dialysis organizations were instructed to remove individual dialysis sessions missing any data necessary to calculate the measure scores from the analysis. For each facility across the three participating dialysis organizations, the overall number, average, and range of dialysis sessions with missing data were recorded for each of the 12 months. The monthly and annual frequencies of sessions with missing data were then analyzed.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; **if no empirical sensitivity analysis**, identify the approaches for handling missing data that were considered and pros and cons of each*)

Findings are provided in the following table. As with the exclusions, we did not perform a specific examination of missing data rates during only the Kt/V week; the data are for all sessions. Nevertheless, it can be inferred that the percentage of missing data during that week is equivalent to the rate for the entire month; it may in fact be less, since facilities are attuned to data collection during that week.

In summary, 75,188 individual dialysis sessions across all three dialysis organizations were excluded from the analysis over the testing year due to missing data. The average annual number of sessions excluded per facility was 17.68.

The average monthly number of excluded treatments across the three organizations was 6,266, with a range of 1 to 323 sessions excluded per facility each month secondary to missing data. The average monthly number of sessions excluded per facility was 1.47.

DIALYSIS SESSIONS EXCLUDED ACROSS ALL FACILITIES
(TOTAL NUMBER / AVERAGE PER FACILITY / FACILITY RANGES)

TOTAL⇓

Measures	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Missing/ incomplete data: Org A	T=1,984 A=9.36 R=(0-323)	1,679 7.92 (0-288)	1,853 8.74 (0-209)	1,701 8.02 (0-152)	3,572 16.85 (0-214)	1,437 6.78 (0-113)	1,554 7.33 (0-146)	1,638 7.73 (0-156)	1,517 7.16 (0-136)	1,580 7.45 (0-136)	1,629 7.68 (0-186)	1,518 7.16 (0-187)	**T=21,662** **A=102.18**
Missing/ incomplete data: Org B	2,969 1.49 (0-19)	2,372 1.19 (0-6)	2,790 1.40 (0-8)	2,950 1.48 (0-9)	2,870 1.44 (0-10)	2,910 1.46 (0-15)	3,169 1.59 (0-16)	3,408 1.71 (0-14)	3,149 1.58 (0-17)	2,651 1.33 (0-12)	2,571 1.29 (0-11)	3,009 1.51 (0-24)	**34,818** **17.47**
Missing/ incomplete data: Org C	1,494 0.73 (0-21)	1,392 0.68 (0-20)	1,433 0.70 (0-13)	1,494 0.73 (0-40)	1,617 0.79 (0-49)	1,576 0.77 (0-280)	1,781 0.87 (0-53)	1,617 0.79 (0-54)	1,616 0.79 (0-34)	1,556 0.76 (0-27)	1,494 0.73 (0-36)	1,638 0.80 (0-23)	**18,708** **9.14**
Sessions excluded each month across all organizations	**T=6,447** **A=1.52** **R=(0-323)**	**5,443** **1.28** **(0-288)**	**6,076** **1.43** **(0-209)**	**6,145** **1.45** **(0-152)**	**8,059** **1.90** **(0-214)**	**5,923** **1.39** **(0-280)**	**6,504** **1.53** **(0-146)**	**6,663** **1.57** **(0-156)**	**6,282** **1.48** **(0-136)**	**5,787** **1.36** **(0-136)**	**5,694** **1.34** **(0-186)**	**6,165** **1.45** **(0-187)**	**T=75,188** **A=17.68**

Monthly average per organization = 6,266 dialysis sessions with missing data per organization per month

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted;* ***if no empirical analysis****, provide rationale for the selected approach for missing data*)

As a correlation between missing data and the patient-months contributed to the testing period by specific patients was not conducted, we cannot identify the total percentage of dialysis sessions that were excluded over the course of the year due to missing data elements. Consequently, we instead estimate here what that proportion would be, were all patients to contribute a full year's worth of data.

Specifically, testing encompassed 412,522 patients. If all patients contributed the still-conventional average of 12 to 13 dialysis sessions per month (or 144 to 156 sessions per year), the total number of dialysis sessions for all patients over the course of the year would be between 59,403,168 and 64,353,432. The total of 75,188 sessions that were excluded from the analysis due to missing data would then be 0.12 to 0.13% of all dialysis treatments.

While there is no widely held cut-off regarding an acceptable percentage of missing data in a data set for valid statistical inferences, conservative current statistical literature suggests that a missing rate of 5% or less is inconsequential.[1,2,3] Thus, even were the total patient-months significantly less than estimated in the above calculations (i.e., secondary to deaths, hospitalizations, and transplants over the course of a year), the rate of missing data still would not be sufficient to bias performance results for the measure. For instance, if patient-months contributing to the denominator were halved from the assumptions above, the 75,188 sessions excluded due to missing data would not surpass 0.25% of all treatments for the year.

1. **Dong Y and Peng CJ. Principled missing data methods for researchers. *Springerplus.* 2013;2:222-241.**
2. **Schafer JL. Multiple imputation: A primer. *Stat Methods in Med.* 1999;8(1):3–15.**
3. **Bennet DA. How can I deal with missing data in my study? *Aust N Z J Public Health.* 2001;25(5):464–469.**

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, **as specified**, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

Renal, Renal : End Stage Renal Disease (ESRD)

De.6. Non-Condition Specific*(check all the areas that apply):*

Care Coordination : Transitions of Care, Safety : Complications

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

Elderly, Populations at Risk, Populations at Risk : Dual eligible beneficiaries, Populations at Risk : Individuals with multiple chronic conditions

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

https://kidneycarepartners.com/wp-content/uploads/2020/10/tbKCQA_NQFendorsedSpecs10-10-20.pdf

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

No data dictionary **Attachment:**

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

No

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

Not applicable; no changes.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

***IF an OUTCOME MEASURE**, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).*

Number of patients* from the denominator whose average UFR is >=13 mg/kg/hr (NOT just >13) hour AND who receive an average of <240 minutes per treatment during the calculation period.**

*To address the fact that patients may contribute varying amounts of time to the annual denominator population, results will be reported using a "patient-month" construction.

** The calculation period is defined as the same week that the monthly Kt/V is drawn.

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

***IF an OUTCOME MEASURE,** describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).*

Numerator Data Elements

For all patients meeting the denominator criteria in the reporting month, collect the following data elements for all dialysis sessions (including supplemental sessions) falling within the same week that the monthly Kt/V is drawn:*

- Pre-Dialysis Weight for Session
- Post-Dialysis Weight for Session
- Time Delivered Per Session, in Minutes
- Session Date
- Sessions Per Week

* If more than one Kt/V is drawn in a given month, the last draw for the month will be used to define the data collection period (i.e., these data elements will be collected during the week that the final Kt/V value of the month is drawn).

Numerator Case Identification

For each facility, for all dialysis sessions falling within the calculation period for all patients meeting the denominator criteria:

1. Calculate the UFR (in ml/kg/hour) for each dialysis session (including supplemental sessions):

Session X UFR = ([{Session X Pre-Dialysis Weight in kg – Session X Post-Dialysis Weight in kg} x 1000 ml/kg] ÷ Session X Post-Dialysis Weight in kg) ÷ (Session X Delivered Treatment Time in minutes) x 60 minutes/hour

2. Calculate each patient's average UFR for all dialysis sessions (including supplemental sessions) during the calculation period:

Average UFR = (UFR1 + UFR2 + …. + UFRX) ÷ X Treatments

3. Calculate each patient's average treatment time over all dialysis sessions (including supplemental sessions) during the calculation period:

Average Treatment Time (in minutes) = (Time1 + Time 2 + … + TimeX) ÷ X Treatments

4. Identify all patients with <4 dialysis sessions during the calculation period.
5. For each facility, include in the numerator all patients with:

- an average UFR during the calculation period (Step 2 value) >=13 ml/kg/hour; AND

- an average treatment time during the calculation period (Step 3 value) <240 minutes.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

Number of adult in-center hemodialysis patients in an outpatient dialysis facility undergoing chronic maintenance hemodialysis during the calculation period.

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

***IF an OUTCOME MEASURE**, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).*

Identify all patients in the dialysis facility during the reporting period whose:

- Primary Type Treatment/Modality = Hemodialysis.
- Primary/Current Dialysis Setting = In-center.
- Date of Birth = >18 years prior to treatment date.

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

The following patients are excluded from the denominator population:

1. Patients <18 years of age (implicit in denominator definition).
2. Home dialysis patients (implicit in denominator definition).
3. Patients in a facility <30 days.
4. Patients with >4 hemodialysis treatments during the calculation period.
5. Patients with <7 hemodialysis treatments in the facility during the reporting month.
6. Patients without a completed CMS Medical Evidence Form (Form CMS-2728) in the reporting month.
7. Kidney transplant recipients with a functioning graft.
8. Facilities treating <=25 adult in-center hemodialysis patients during the reporting month.

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

For all patients meeting the denominator criteria in the reporting month, identify all patients meeting any of the following exclusion criteria during the calculation period and remove from the denominator population:

1. Date of Birth = <18 years prior to treatment date (implicit in denominator definition).
2. Primary Type Treatment/Modality = Peritoneal dialysis or home hemodialysis (implicit in denominator definition).
3. Date Patient Started Chronic Dialysis at Current Facility = >30 days prior to treatment date.
4. Sessions Per Week = >4
5. Transient Status = Not transient OR patients with <7 hemodialysis treatments in the facility during the reporting month.
6. Patients without a completed CMS Medical Evidence Form (Form CMS-2728) in the reporting month.
7. Kidney transplant recipients with a functioning graft

Note: Facilities treating <=25 adult in-center hemodialysis patients during the reporting month are also excluded.

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

Not applicable.

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Lower score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

Data are collected and scores for each facility are calculated on a monthly basis; scores are then averaged over the 12-month reporting period to obtain the facility's annual score.

Scores are calculated using the following algorithm:

1. Build the "Month 1 Raw Denominator Population."

For the Month 1 calculation period,* identify all patients in the facility during the reporting month whose:

   a. Primary Type Treatment/Modality = Hemodialysis
   b. Primary/Current Dialysis Setting = In-center
   c. Date of Birth = >18 years prior to treatment date

* The calculation period is defined as the same week that the monthly Kt/V is drawn. If more than one Kt/V is drawn in a given month, the last draw for the month will be used to define the data collection period (i.e., these data elements will be collected during the week that the final Kt/V value of the month is drawn).

2. Remove patients with exclusions to define the "Month 1 Final Denominator Population."

For all patients meeting all of the Step 1 requirements, identify all patients meeting any of the following exclusion criteria and remove from the denominator population:

   a. Date Patient Started Chronic Dialysis at Current Facility = >30 days prior to treatment date.
   b. Transient Status = Not transient OR patients with <7 hemodialysis treatments in the facility during the month.
   c. Sessions Per Week = >4.
   d. Patients without a completed CMS Medical Evidence Form (Form CMS-2728) in the reporting month.
   e. Kidney transplant recipients with a functioning graft.

3. Identify the "Month 1 Numerator Data Elements."

For all patients remaining in the denominator after Step 2, collect each of the following data elements for each dialysis session (including supplemental sessions) delivered during the Month 1 calculation period:

a. Pre-Dialysis Weight for Session

b. Post-Dialysis Weight for Session

c. Session Date

d. Time Delivered Per Session, in Minutes

e. Sessions Per Week

4. Build the "Month 1 Numerator Population."

For each patient, for all dialysis sessions included in the final Month 1 Numerator Data Set:

a. Calculate the UFR (in ml/kg/hour) for each dialysis session (including supplemental sessions):

Session X UFR = ([{Session X Pre-Dialysis Weight in kg – Session X Post-Dialysis Weight in kg} x 1000 ml/kg] ÷ Session X Post-Dialysis Weight in kg) ÷ (Session X Delivered Treatment Time in minutes) x 60 minutes/hour

b. Calculate each patient's average UFR for all dialysis sessions (including supplemental sessions) during the calculation period:

Average UFR = (UFR1 + UFR2 + …. + UFRX) ÷ X Treatments

c. Calculate each patient's average treatment time over all dialysis sessions (including supplemental sessions) during the calculation period:

Average Treatment Time (in minutes) = (Time1 + Time 2 + … + TimeX) ÷ X Treatments

d. For each facility, include in the numerator all patients with:

i. an average UFR during the calculation period (4.b. value) >= 13 ml/kg/hour;

AND

ii. an average treatment time during the calculation period (4.c. value) <240 minutes.

5. Calculate the facility's Month 1 performance score:

Month 1 Performance Score = Month 1 Numerator Population ÷ Month 1 Denominator Population

6. Repeat Steps 1 through 5 for each of the remaining 11 months of the reporting year.

7. Calculate the facility's annual performance score:

Facility's Average Annual Performance Score = (Facility's Month 1 Score + Month 2 Score +….. + Month 12 Score) ÷ 12

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

Not applicable.

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

Not applicable.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Electronic Health Records

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

CROWNWeb Electronic Data Interchange, available at URL: https://mycrownweb.org

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Facility

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Post-Acute Care

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

Not applicable.

2. Validity – See attached Measure Testing Submission Form

NQF2701_TestingAttachment_10-27-20-637408819382156345.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

No

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

No

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*)
Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS)

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

Not applicable; all data elements defined fields in electronic clinical data.

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

The measure was readily implemented by the three large dialysis organizations that participated in testing; no difficulties were encountered regarding data collection or availability, missing data, sampling, patient confidentiality, time, frequency, cost, or other feasibility/implementation issues. No modifications were made following testing.

Additionally, CMS is currently implementing a structural reporting measure based on NQF 2701 in the ESRD QIP for PY 2020. We note, however, that unlike most "checkbox" style reporting measures, CMS is requiring that all the discrete data elements defined by 2701 to calculate UFR be reported:

1. HD Kt/V Date
2. Post-Dialysis Weight

3. Pre-Dialysis Weight

4. Delivered Minutes of Hemodialysis

5. Number of sessions of dialysis delivered to the patient in the reporting month

Thus while CMS is not yet calculating facilities' UFR per se, successful collection of these data elements will effectively and fully demonstrate the feasibility of 2701.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

Not applicable.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
Not in use	Public Reporting ESRD QIP https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP Payment Program ESRD Quality Incentive Program (QIP) https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP Quality Improvement (external benchmarking to organizations) ERSD QIP https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP Quality Improvement (Internal to the specific organization) Measure is being used by numerous dialysis organizations for IQI NA

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

NQF 2701 is currently being implemented as a facility-level reporting measure within CMS´s ESRD QIP for the purposes of public reporting, payment, and external quality improvement/benchmarking. The ESRD QIP is a nation-wide program encompassing all dialysis facilities receiving payment from Medicare as "a provider of services or a renal dialysis facility for renal dialysis services" under the ESRD PPS.
The measure is also being used as a facility-level internal quality improvement metric by numerous dialysis organizations.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)
Currently publicly reported.

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

Currently publicly reported.

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

KCQA is not the measure implementer and thus does not have access to the requested information from the current implementation of the measure specifications in the ESRD QIP. However, as previously noted, the measure was tested within three KCQA member large dialysis organizations (LDOs). Assistance was available throughout the study, and performance results and feedback were provided to each LDO upon conclusion of field-testing.

There was no sampling; all pertinent data from all eligible (i.e., adult in-center hemodialysis) patients in the participating organizations during the testing period were included in the dataset. Testing encompassed 4,252 dialysis facilities, with an average mean facility census (i.e., the number of patients receiving care at the facility) of 84.11 patients (range 1-664 patients per month).

412,522 patients across the three organizations met the measure's denominator criteria and were included, with a range of 15,184 to 215,008 patients per organization. The following is a composite description of patient demographics:

- Mean patient age: 61.66 years
- Range of patient ages: 18.01-104.00 years
- Gender: 56.26% male, 43.74% female
- Race/Ethnicity:
- 52.37% white
- 36.33% African American
- 2.82% Asian
- 1.16% American Indian/Native Alaska
- 0.67% Native Hawaiian/other Pacific Islander
- 0.57% other/missing/declined
- 15.60% Hispanic (independent of race)

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

Assistance was available to each participating LDO at all times on request throughout the study. Performance results and feedback were provided to each LDO upon conclusion of field-testing. Descriptive statistics for the annual performance measure scores for all tested entities (individual facilities within the LDOs) were constructed, including the mean, standard deviation and standard error, 95% confidence interval, median, mode, range of scores, and the interquartile range of scores across the measured entities.

The measure was readily implemented by the LDOs, with no difficulties encountered with data collection or availability, patient confidentiality, or other feasibility/implementation issues. No comprehension difficulties were encountered, with no need to additional educational/explanatory efforts.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

The measure was readily implemented by the three participating LDOs and their member dialysis facilities, with no difficulties reported regarding data collection or availability, patient confidentiality, or other feasibility/implementation issues. Participating entities reported no difficulty understanding the measure concept or scores.

Additionally, the LDOs' and other dialysis organization member representatives participated in KCQA's face validity assessment for the measure, indicating that the scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good from poor quality.

4a2.2.2. Summarize the feedback obtained from those being measured.

Representatives from KCQA's member patient organizations and advocacy groups also participated in KCQA's face validity assessment for the measure, unanimously indicating that the scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good from poor quality.

4a2.2.3. Summarize the feedback obtained from other users

Likewise, representatives from KCQA's other member groups participated in the measure's face validity assessment, again indicating that the scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good from poor quality.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

As noted, feedback gathered from the participating and other dialysis organizations, dialysis facilities, patient groups, and other KCQA member organizations indicate that that measure is feasible, meaningful, and will provide an accurate reflection of quality. No modifications were made following testing.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

A structural reporting measure based on NQF 2701 is being implemented in the ESRD Quality Incentive Program for PY 2020. While we anticipate subsequent conversion to a clinical metric, data on improvement trends are not yet available. However, we postulate that performance results will promote the use of lower UFRs and longer dialysis session lengths to help attenuate rapid fluctuations in fluid balance and blood pressure that contribute to cardiovascular morbidity and mortality in hemodialysis patients. Associated hospitalization, readmissions, and mortality will consequently be minimized.

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

The measure is currently being implemented for PY2020 in CMS´s ESRD QIP; no data on unexpected findings are yet available. No unintended consequences were identified during testing.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

The measure is currently being implemented for PY2020 in CMS´s ESRD QIP; no data on unexpected benefits are yet available.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria **and** there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

0249 : Delivered Dose of Hemodialysis Above Minimum

0256 : Minimizing Use of Catheters as Chronic Dialysis Access

0257 : Maximizing Placement of Arterial Venous Fistula (AVF)

0258 : Consumer Assessment of Healthcare Providers and Systems (CAHPS) In-Center Hemodialysis Survey (ICH CAHPS)

1460 : Bloodstream Infection in Hemodialysis Outpatients

2977 : Hemodialysis Vascular Access: Standardized Fistula Rate

2978 : Hemodialysis Vascular Access: Long-term Catheter Rate

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

NHSN Dialysis Event Reporting Measure, CDC

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):

Are the measure specifications harmonized to the extent possible?

Yes

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

Not applicable; specifications of this and other NQF-endorsed facility-level performance measures applicable to adult in-center ESRD hemodialysis patients are harmonized to extent possible.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);

OR

Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Not applicable; no competing NQF-endorsed measures.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Kidney Care Quality Alliance (KCQA)

Co.2 Point of Contact: Lisa, McGonigal, lmcgon@msn.com, 203-530-9524-

Co.3 Measure Developer if different from Measure Steward: Kidney Care Quality Alliance (KCQA)

Co.4 Point of Contact: Lisa, McGonigal, lmcgon@msn.com, 203-530-9524-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

The KCQA Steering Committee guides the measure development process and decision-making. Steering Committee members include:

- Edward Jones, MD, KCQA Co-Chair—Renal Physicians Association
- Allen Nissenson, MD, KCQA Co-Chair—DaVita
- Akhtar Ashfaq, MD—Amgen
- Donna Bednarski, RN, MSN—American Nephrology Nurses Association
- Barbara Fivush, MD—American Society of Pediatric Nephrology
- Raymond Hakim, MD, PhD—American Society of Nephrology
- Eduardo Lacson, Jr., MD, MPH—Fresenius Medical Care North America
- Chris Lovell, RN, MSN—Dialysis Clinics, Inc.
- Thomas Manley, RN, BSN—National Kidney Foundation
- Gail Wick, MHSA, BSN, RN—American Kidney Fund
- Shari M. Ling, MD, Chief Medical Officer, Centers for Medicare and Medicaid Services, Center for Clinical Standards and Quality (CCSQ)–CMS Liaison Member

The KCQA Measure Feasibility/Testing Workgroup provided technical expertise and guidance during the measure development process. Workgroup members include:

- Scott Bieber, DO—Northwest Kidney Centers
- Steven Brunelli, MD, MSCE—DaVita
- Maggie Carey—Forum of ESRD Networks
- Allan Collins, MD—NxStage Medical
- Joseph Flynn, MD—American Society of Pediatric Nephrology
- Lori Hartwell—Renal Support Network
- Jeffrey Hymes, MD—Fresenius Medical Care North America
- Mahesh Krishnan, MD, MPH, MBA, FASN—DaVita
- Eduardo Lacson, MD, MPH—Fresenius Medical Care North America
- Klemens Meyer, MD—Dialysis Clinics, Inc.

- Paul Miller, MD—Renal Physicians Association
- Donald Molony, MD—Forum of ESRD Networks
- Tom Parker, MD—Renal Ventures Management
- Glenda Payne, MS, RN, CNN—American Nephrology Nurses Association
- Daniel Weiner, MD, MS—National Kidney Foundation

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2015

Ad.3 Month and Year of most recent revision: 02, 2015

Ad.4 What is your frequency for review/update of this measure? annually and as needed with changes or additions to the evidence base.

Ad.5 When is the next scheduled review/update for this measure? 10, 2021

Ad.7 Disclaimers: Dialysis facility performance measures (Measures) and related data specifications, developed by the Kidney Care Quality Alliance (KCQA), primarily funded by Kidney Care Partners, are intended to facilitate quality improvement activities by dialysis providers.

These Measures are intended to assist dialysis facilities in enhancing quality of care. Measures are designed for use by any dialysis facility. These performance Measures are not clinical guidelines and do not establish a standard of medical care. KCQA has not tested its Measures for all potential applications. KCQA encourages the evaluation of its Measures.

Measures are subject to review and may be revised or rescinded at any time by KCQA. The Measures may not be altered without the prior written approval of KCQA. Measures developed by KCQA, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by dialysis providers in connection with their care delivery or for research. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measures require a license agreement between the user and Kidney Care Partners, on behalf of KCQA.

Neither KCQA nor its members shall be responsible for any use of these Measures.

THE MEASURES ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND

Ad.8 Additional Information/Comments: