NATIONAL QUALITY FORUM

Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

Brief Measure Information

NQF #: 1423
Corresponding Measures:
De.2. **Measure Title:** Minimum spKt/V for Pediatric Hemodialysis Patients
Co.1.1. **Measure Steward:** Centers for Medicare & Medicaid Services
De.3. **Brief Description of Measure:** Percentage of patient months for all pediatric (<18 years old) in-center hemodialysis patients in which the delivered dose of hemodialysis (calculated from the last measurement of the month using the UKM or Daugirdas II formula) was spKt/V >= 1.2.
1b.1. **Developer Rationale:** In considering target spKt/V, the pediatric population should receive at least a spKt/V of 1.2, which is the minimum requirement for the adult population in order to allow for the increased nutritional needs of children. Analysis of CPM data further support this cut-off since adolescents with spKt/V below 1.2 were found to have significantly increased risk of hospitalization as compared to those with spKt/V of 1.2-1.4.

S.4. **Numerator Statement:** Number of patient months from the denominator in which the delivered dose of hemodialysis (calculated from the last measurement of the month using the UKM or Daugirdas II formula) was spKt/V >= 1.2.
S.6. **Denominator Statement:** To be included in the denominator for particular month, a patient must be on hemodialysis for the entire month, must be <18 years old at the beginning of the month, must have had ESRD for greater than 90 days at the beginning of the month, must be on thrice weekly in-center hemodialysis during the month, and must be assigned to that facility for the entire month.
S.8. **Denominator Exclusions:** Exclusions that are implicit in the denominator definition include:
1) Patients on home hemodialysis
2) Patients on peritoneal dialysis
3) Adult patients (>=18 years old)
4) Patients on ESRD less than 91 days
5) Patients not on thrice weekly dialysis
6) Patients not assigned to the facility for the entire month

There are no additional exclusions for this measure.

De.1. **Measure Type:** Outcome: Intermediate Clinical Outcome
S.17. **Data Source:** Claims, Registry Data
S.20. **Level of Analysis:** Facility

IF Endorsement Maintenance – Original Endorsement Date: Aug 16, 2011 **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?** N/A

1. Evidence, Performance Gap, Priority – 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

1423_Evidence.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.

No

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.

In considering target spKt/V, the pediatric population should receive at least a spKt/V of 1.2, which is the minimum requirement for the adult population in order to allow for the increased nutritional needs of children. Analysis of CPM data further support this cut-off since adolescents with spKt/V below 1.2 were found to have significantly increased risk of hospitalization as compared to those with spKt/V of 1.2-1.4.

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.*

Among the 14 facilities that have at least 11 eligible patients, we generated the following statistics of their performance scores (based on the patient month) using the January – December 2017 CROWNWeb and Medicare claims data: mean (SD)=95.2% (4.6%); min=85.4%; max=100.0%; 25th percentile = 92.1%; 50th percentile =97.0%; 75th percentile = 99.0%. A description of the data is included in questions 1.1-1.7 under "Scientific Acceptability".

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.**

N/A

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.*

Given that the number of facilities included in the calculation in 1b.2 is only 13, the sample was determined to be too small to display useful disparities data.

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**

Observational pediatric data exist showing that older, larger, and African-American children are less likely to receive an spKt/V greater than 1.2 consistently [1]. Additionally, in the North American Pediatric Renal Transplant Cooperative Study (NAPRTCS), monthly hemodialysis adequacy data were analyzed from 138 children from 32 centers. Multivariate modeling indicated that after adjusting for body surface area and lack of any Kt/V center measures, the mean Kt/V dose was significantly higher among females compared to

males (ß=0.13, p<0.05) and among Nonblack patients compared to Black patients (ß=0.22, p<0.001) [2].

1. Frankenfield DL, Neu AM, Warady BA, Watkins SL, Friedman AL, Fivush BA: Adolescent hemodialysis: results of the 2000 ESRD Clinical Performance Measures Project. Pediatr Nephrol 17:10-15, 2002
2. Leonard MB, et al. Racial and center differences in hemodialysis adequacy in children treated at pediatric centers: a North American Pediatric Renal Transplant Cooperative Study (NAPRTCS) report. J Am Soc Nephrol. 2004 Nov;15(11):2923-32 -32

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):*

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

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.)*
N/A

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)*
Attachment **Attachment:** 1423_Code_List.xlsx

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.

There have been no changes to the measure specifications since the last endorsement in 2015.

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 patient months from the denominator in which the delivered dose of hemodialysis (calculated from the last measurement of the month using the UKM or Daugirdas II formula) was spKt/V >= 1.2.

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).
Months with spKt/V >=1.2 are counted in the numerator. Eligible spKt/V values are those >=1.2 during the reporting month. The last spKt/V value reported, not including missing, expired, and not performed, is selected when multiple values are reported in the month.

Missing, expired, and not performed are not counted as achieving the minimum spKt/V threshold.

S.6. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
To be included in the denominator for particular month, a patient must be on hemodialysis for the entire month, must be <18 years old at the beginning of the month, must have had ESRD for greater than 90 days at the beginning of the month, must be on thrice weekly in-center hemodialysis during the month, and must be assigned to that facility for the entire month.

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).
A treatment history file is the data source for the denominator calculation used for the analyses supporting this submission. This file provides a complete history of the status, location, and dialysis treatment modality of an ESRD patient from the date of the first ESRD service until the patient dies or the data collection cutoff date is reached. For each patient, a new record is created each time he/she changes facility or treatment modality. Each record represents a time period associated with a specific modality and dialysis facility. CROWNWeb is the primary basis for placing patients at dialysis facilities and dialysis claims are used as an additional source of information in certain situations. Information regarding first ESRD service date, death, and transplant is obtained from CROWNWeb (including the CMS Medical Evidence Form (Form CMS-2728) and the Death Notification Form (Form CMS-2746)) and Medicare claims, as well as the Organ Procurement and Transplant Network (OPTN).

S.8. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
Exclusions that are implicit in the denominator definition include:
1) Patients on home hemodialysis
2) Patients on peritoneal dialysis
3) Adult patients (>=18 years old)
4) Patients on ESRD less than 91 days
5) Patients not on thrice weekly dialysis
6) Patients not assigned to the facility for the entire month

There are no additional exclusions for this measure.

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.)*

N/A

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.)*
N/A

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 = Higher 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.)*
Denominator:
For the reporting month, patients are included in the denominator if:
Patient modality is indicated as Hemodialysis during the entire month (in-center)
Patient is dialyzing thrice weekly during the month
Patient age as of the beginning of the reporting month is less than 18 years
Patient has had ESRD for greater than 90 days at the beginning of the month
Patient is assigned to the facility for the entire month

Numerator:
For the reporting month, patient months from the denominator are also included in the numerator if they have a spKt/V >=1.2. The last spKt/V value reported, not including missing, expired, and not performed, is selected when multiple values are reported in the month.

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.
N/A

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.
N/A

S.17. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.18.
Claims, Registry Data

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.
For the analyses supporting this submission, the measure is calculated using CROWNWeb as the primary data source for the Kt/V values used to determine the numerator. If a patient's Kt/V data are missing in CROWNWeb, Kt/V values from Medicare claims are used as an additional source for obtaining that information. Please see the attached data dictionary for a list of specific data elements that are used from each data source.

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)*
Other
If other: Dialysis Facility

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.)*
N/A

2. Validity – See attached Measure Testing Submission Form
1423_testing_01072019-636824705522261718.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.
Yes

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.
Yes

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 by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition
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 a combination of electronic sources

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).

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.
Data collection is accomplished via CROWNWeb, a web-based and electronic batch submission platform maintained and operated by CMS contractors. Measures reported on DFC are reviewed on a regular basis by dialysis facility providers and rare instances of inaccurate or missing data are present based on comments reported in the DFC ticketing system.

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)*.
N/A

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)
	Public Reporting Dialysis Facility Compare http://www.medicare.gov/dialysisfacilitycompare/ Dialysis Facility Compare

	http://www.medicare.gov/dialysisfacilitycompare/ Payment Program ESRD QIP http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP/ ESRD QIP http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP/

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

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

DFC:
Purpose: Dialysis Facility Compare helps patients find detailed information about Medicare-certified dialysis facilities. They can compare the services and the quality of care that facilities provide.
Geographic area: United States
Number of accountable entities: All Medicare-certified dialysis facilities who are eligible for the measure, and have at least 11 patients (due to public reporting requirements). For the most recent update to Dialysis Facility Compare (January 2019), 14 facilities had a score reported.
Patients included: All patients who meet the requirements to be included in the measure.

QIP:
Purpose: The ESRD QIP will reduce payments to ESRD facilities that do not meet or exceed certain performance standards. The measure was added to the program for PY2015. In PY2019, the QIP began reporting a comprehensive Kt/V measure, for which this the data used in this measure is counted. For the purposes of this review, we are considering this an active implementation of this measure.
Geographic area: United States
Number of accountable entities: All Medicare-certified dialysis facilities who are eligible for the measure, and have at least 11 patients (due to public reporting requirements). For the most recent QIP report (PY 2019), this was 6835 facilities. Since the QIP reports a comprehensive Kt/V measure, the number of facilities counted here is larger than for DFC.
Patients included: All patients who meet the requirements to be included in the measure.

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?*)
N/A

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.*)
N/A

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.
Results of this measure are currently reported on Dialysis Facility Compare and in the ESRD Quality Incentive Program (via the comprehensive Kt/V measure described above). All Medicare-certified dialysis facilities are eligible for reporting in both programs (approximately 7,000 dialysis facilities). Each program has a helpdesk and supporting documentation available to assist with

interpretation of the measure results.

The measure developer (UM-KECC) produces and distributes the DFC data under contract with CMS. Other CMS contractors calculate and distribute the ESRD QIP measure results.

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.**
For DFC, the results are first reported to facilities via a closed preview period, where facilities can review their data prior to each of the quarterly updates of the public facing Dialysis Facility Compare website. These preview reports are posted on dialysisdata.org, where facilities can also find a detailed Guide to the Quarterly Dialysis Facility Compare Reports and other supporting documentation. Facilities can submit comments/questions about their results at any time, and can request patient lists for their facilities during the specified preview periods.

For the ESRD QIP, results are first reported to facilities via closed preview period on an annual basis; facilities can review their data prior to the results becoming public at the end of the calendar year. These preview reports are posted on qualitynet.org, where facilities can also find supporting documentation and can submit comments/questions about their results.

A measures manual that describes the calculations for both of these programs in detail is published on the CMS website: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP/06_MeasuringQuality.html

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.
For DFC, feedback can be provided any time through contacting the dialysisdata.org helpdesk. Preview periods allow for specific times for facilities review and comment on measure calculations, and provide an opportunity to request a patient list.

For the ESRD QIP, feedback can be provided any time through contacting the QIP helpdesk. Preview periods allow for specific times for facilities review and comment on measure calculations. Comments can also be submitted in response to the Notice of Proposed Rulemaking for each QIP payment year.

4a2.2.2. **Summarize the feedback obtained from those being measured.**
We reviewed the comments and questions submitted during the DFC preview periods that have taken place since the last maintenance (2016-present). Outside of questions about facility-specific results (such as questioning the Kt/V value on record for a particular patient), we receive a handful of questions each preview period regarding the measure specifications, such as the determination of dialysis modality.

Note that since UM-KECC is not the contractor responsible for the ESRD Quality Incentive Program, we do not have access to the detailed comments/requested that are submitted during the annual preview period for that program.

4a2.2.3. **Summarize the feedback obtained from other users**
We reviewed the public comments that were addressed in the ESRD QIP Final Rules (FRs) that have been published since the last endorsement (PY2019 – PY2022). Since PY 2019, the ESRD QIP has been reporting a combined Kt/V measure in order to allow for more reporting of data for pediatric and peritoneal dialysis patients. Most of the comments addressed in the rule have to do with that decision. In the FR for PY 2019, there were also a number of questions about how the combined measure would be specified that were along similar lines to what is often asked via the DFC preview period.

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.**
The measure specifications have not been revised since the last maintenance cycle in 2015. Feedback received during DFC preview periods has resulted in more detailed and accurate documentation available to the public, primarily via the ESRD Measures Manual and the Guide to the Quarterly Dialysis Facility Reports.

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.
The following table reports the performance scores for this measure at the yearly level for 2015 - 2017. This analysis suggests some slight increase in performance across three years for the measure as implemented on DFC.

Year 2015: N = 17 Mean = 88.0%, Std Dev =8.6%, Min = 71.3%, Max = 100.0%
Year 2016: N = 14, Mean =90.3%, Std Dev = 7.8%, Min = 74.8%, Max = 100.0%
Year 2017: N = 14, Mean = 95.2%, Std Dev = 4.6%, Min = 85.4%, Max = 100.0%

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.**
We have been encouraged by the magnitude of improvement in measure results after implementation noted in 4b1 above.
We have not been notified of documented unintended impacts on patients as a result of measure implementation.

4b2.2. **Please explain any unexpected benefits from implementation of this measure.**
None that we are aware of, other than facility improvements over the last three reporting periods as noted in 4b1 and commented on in 4b2.1

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.
No

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

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

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?

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on**

interpretability and data collection burden.

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.)

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.
No appendix **Attachment:**

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** Centers for Medicare & Medicaid Services
Co.2 **Point of Contact:** Helen, Dollar-Maples, Helen.Dollar-Maples@cms.hhs.gov, 410-786-7214-
Co.3 **Measure Developer if different from Measure Steward:** University of Michigan Kidney Epidemiology and Cost Center
Co.4 **Point of Contact:** Casey, Parrotte, parrotte@med.umich.edu

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.

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.2 **Year the measure was first released:**
Ad.3 **Month and Year of most recent revision:** 04, 2019
Ad.4 **What is your frequency for review/update of this measure?** Annually
Ad.5 **When is the next scheduled review/update for this measure?** 04, 2020

Ad.6 **Copyright statement:** N/A
Ad.7 **Disclaimers:** N/A

Ad.8 **Additional Information/Comments:**

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 1423
Measure Title: Minimum spKt/V for Pediatric Hemodialysis Patients
IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title
Date of Submission: 4/2/2019

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Outcome: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one

step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.
6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

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

☐ Outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the 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*): Kt/V

☐ Process: Click here to name what is being measured

☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

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.

The measure focus is measurement of spKt/V >= 1.2. This process leads to improvement in mortality as follows: Measure spKt/V--> Assess value-->Impact on mortality.

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.)

N/A

****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.

N/A

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)

X 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

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	Clinical Practice Guidelines for Hemodialysis Adequacy: KDOQI Guideline 8. Pediatric Hemodialysis Prescription and Adequacy: 2006. http://www2.kidney.org/professionals/KDOQI/guideline_upHD_PD_VA/hd_guide8.htm
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.	8.3.1 Children should receive at least the delivered dialysis dose as recommended for the adult population. (A)
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	N/A
Provide all other grades and definitions from the evidence grading system	N/A
Grade assigned to the **recommendation** with definition of the grade	KDOQI CPG 8.3.1 rating strength grade is 'A'. The recommendation for Grade A guidelines states 'It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.'
Provide all other grades and definitions from the recommendation grading system	The rating system defined in the KDOQI Guidelines was used to grade the strength of the Guideline recommendation. KDOQI defined grades as follows: Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

	Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes. Grade CPR: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	N/A
Estimates of benefit and consistency across studies	N/A
What harms were identified?	N/A
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	The 2009 clinical pediatric dialysis adequacy TEP conducted a literature search, where we retrieved a total of 190 articles using several sources. First, we retrieved 79 articles using a PubMed search of articles with human subjects, published in English since January 1, 2005. The search terms were: [(pediatric OR pediatrics OR children) and (dialysis OR hemodialysis OR peritoneal dialysis) and (adequacy OR "dialysis dose" OR "dose monitoring" OR "residual renal function" OR "urea clearance" OR "solute clearance" OR "phosphate clearance" OR "amino acid clearance" OR "folate clearance" OR "Kt/V" OR "peritoneal equilibration test" OR ("ultrafiltration" and peritoneal)) and NOT (cvvhd OR "continuous veno venous" OR transplant OR "kidney transplant" OR transplantation)]. Second, we reviewed 61 citations from the Kidney Disease Outcomes Quality Initiative Guidelines on pediatric peritoneal dialysis and hemodialysis. Third, we reviewed the tables of contents of the journal Pediatric Nephrology and retrieved two articles from early on-line publishing that had not yet been included in PubMed. Finally, we reviewed the citations in 14 articles previously identified; this found an additional 65 articles for review. Duplicate articles were excluded.

	A total of 124 articles were found to be relevant for measure development. Four pieces of evidence listed below [1-4] were determined to be relevant to this specific measure. An additional literature search was conducted in May 2014 and additional evidence has been added to the list of citations [5-8]. 1. Lowrie EG, et al. Effect of the hemodialysis prescription of patient morbidity: report from the National Cooperative Dialysis Study. N Engl J Med 305:1176–1181, 1981. Abstract: This report summarizes morbidity in 151 patients in a cooperative trial designed to evaluate the clinical effects of different dialysis prescriptions. Four treatment groups were divided along two dimensions: dialysis treatment time (long or short), and blood urea nitrogen (BUN) concentration averaged with respect to time (TACurea) (high or low). Dietary protein was not restricted. There was no difference in mortality between the groups. Withdrawal of patients from the high-BUN groups for medical reasons was significantly greater than withdrawal from the lowBUN groups. Hospitalization was also greater in the high-BUN groups, but dialysis treatment time had no significant effects. The data indicate that the occurrence of morbid events is affected by the dialysis prescription. Increased morbidity appears to accompany prescriptions associated with a relatively high BUN. Conversely, morbidity may be decreased by prescriptions associated with more

	efficient removal of urea if the dietary intake of protein and other nutrients is adequate. (N Engl J Med. 1981; 305:1176–81.) 2. Owen WF Jr, et al. The urea reduction ratio and serum albumin concentration as predictors of mortality in patients undergoing hemodialysis. N Engl J Med 329:1001–1006, 1993. BACKGROUND: Among patients with end-stage renal disease who are treated with hemodialysis, solute clearance during dialysis and nutritional adequacy are determinants of mortality. We determined the effects of reductions in blood urea nitrogen concentrations during dialysis and changes in serum albumin concentrations, as an indicator of nutritional status, on mortality in a large group of patients treated with hemodialysis. METHODS: We analyzed retrospectively the demographic characteristics, mortality rate, duration of hemodialysis, serum albumin concentration, and urea reduction ratio (defined as the percent reduction in blood urea nitrogen concentration during a single dialysis treatment) in 13,473 patients treated from October 1, 1990, through March 31, 1991. The risk of death was determined as a function of the urea reduction ratio and serum albumin concentration. RESULTS: As compared with patients with urea reduction ratios of 65 to 69 percent, patients with values below 60 percent

	had a higher risk of death during follow-up (odds ratio, 1.28 for urea reduction ratios of 55 to 59 percent and 1.39 for ratios below 55 percent). Fifty-five percent of the patients had urea reduction ratios below 60 percent. The duration of dialysis was not predictive of mortality. The serum albumin concentration was a more powerful (21 times greater) predictor of death than the urea reduction ratio, and 60 percent of the patients had serum albumin concentrations predictive of an increased risk of death (values below 4.0 g per deciliter). The odds ratio for death was 1.48 for serum albumin concentrations of 3.5 to 3.9 g per deciliter and 3.13 for concentrations of 3.0 to 3.4 g per deciliter. Diabetic patients had lower serum albumin concentrations and urea reduction ratios than nondiabetic patients. CONCLUSIONS: Low urea reduction ratios during dialysis are associated with increased odds ratios for death. These risks are worsened by inadequate nutrition. 3. Gorman G, et al. Clinical outcomes and dialysis adequacy in adolescent hemodialysis patients. Am Journal Kidney Dis; 47: 285-93, 2006. BACKGROUND: The National Kidney Foundation-Kidney Disease Outcomes Quality Initiative guidelines recommend that adult hemodialysis (HD) patients receive a minimum dialysis dose by single-pooled Kt/V (spKt/V) of 1.2 or greater. There are no data to support a minimum spKt/V

	dose for children on HD therapy. We aim to determine the association of spKt/V with mortality and hospitalization in adolescents. METHODS: Clinical characteristics of adolescent HD patients aged 12 to 18 years old included in the 2000/2001 End-Stage Renal Disease Clinical Performance Measures Project were linked to US Renal Data System data from October 1, 1999, to October 15, 2001. Hospitalization risks after adjustment for time on dialysis therapy, access, hemoglobin level, albumin level, and height were determined by means of Poisson regression. spKt/V was analyzed by the adult target (< versus > or = 1.2) and by intervals. RESULTS: There were 613 patients with 477 patient-years of follow-up, during which there were 14 deaths and 185 hospitalizations covering 1,108 days. After adjustment, patients with an spKt/V less than 1.2 had increased hospitalization risk (1.59; 95% confidence interval, 0.98 to 2.56; P = 0.06) compared with those with an spKt/V of 1.2 or greater. Compared with patients with an spKt/V of 1.2 to 1.4, patients with an spKt/V less than 1.2 had increased adjusted risk for hospitalization (2.46; 95% confidence interval, 1.23 to 4.94; P = 0.01). Increases in spKt/V beyond 1.4 were not associated with improved outcomes. CONCLUSION: Applying the current adequacy guideline to adolescent HD patients is justified by the increased hospitalization risk of those

	who fail to attain an spKt/V of 1.2 or greater. However, attaining an spKt/V in excess of 1.4 was not associated with greater benefit. 4. Fischbach M, et al. Intensified and daily hemodialysis in children might improve statural growth. Pediatr Nephrol 21:1746–1752, 2006. Abstract: In children conventional hemodialysis does not often improve growth. We determined linear growth in five children on in-center intensified and daily hemodialysis (IDd) regimen, with a mean age of 8 years 7 months at enrollment. Four of five were on growth hormone started for a median of 28.5 months before IDd. IDd was delivered 5 to 6 times weekly, for three hours each session. Mean follow up of IDd was 18.6 months. Dropout from IDd was kidney transplantation (n=4) or transfer to another center (n=1). IDd and free diet improved appetite, thereby protein intake, was above 2 g/kg/BW. Median weekly Kt/V(urea) was 9.1 (8.7 to 10.4). Predialysis phosphorus blood levels were higher at the start (2.04+/-0.34 mmol/L) than at end of IDd (1.39+/-0.41 mmol/L) without need for carbonate of calcium in four of five cases. During conventional dialysis ht SDS decreased from -0.8 to -1.44, which occurred predominantly before rhGH start. Conversion to IDd significantly increased growth velocity to a mean of 13 cm/year (10.3-18) with a mean change of +1.84 ht SDS/year (0.4 to 2.7). This preliminary report suggests the potential efficacy of IDd regimen in promising growth velocity, either directly

from a higher dialysis dose or indirectly through an improved nutritional status.

5. Daugirdas JT. Dialysis dosing for chronic hemodialysis: beyond Kt/V. Semin Dial. 2014 Mar;27(2):98-107.

Abstract: Current views regarding hemodialysis adequacy reach beyond indices of small solute removal such as Kt/V. Nevertheless, new Kt/V-based constructs such as the standard Kt/V, which adjusts not only for dialysis frequency, but which also represents removal of sequestered solutes rather than easily removed urea, continue to be useful. The scaling of dialysis dose to measures of size other than body water results in higher recommended doses of dialysis for children, small patients, and women, compared with the current body water-based scaling approach. Aside from small solute removal, increasing weekly time on dialysis results in slower removal of fluid with better tolerance and with increased removal of phosphorus, although both salt and water and phosphorus control often respond to efforts to reduce intake. The intermediate term benefits of removing larger middle molecules such as beta-2-microglobulin appear to be modest, and the benefits of removal of protein-bound uremic toxins remain to be proved in controlled trials.

6. Kaur A, Davenport A. Hemodialysis for infants, children, and adolescents. Hemodial Int. 2014 Apr 14. doi: 10.1111/hdi.12163. [Epub ahead of print]

	Abstract: Children with chronic kidney disease stage 5 requiring dialysis can be treated by peritoneal or hemodialysis. In the United Kingdom nearly twice as many children receive peritoneal dialysis compared with hemodialysis. Technical aspects of pediatric hemodialysis are challenging and include the relative size of extracorporeal circuit and child's blood volume, assessment of adequacy, technical and complications of vascular access. Alternatives to standard hospital-based hemodialysis are also increasingly available. Optimizing nutritional status with the support of specialist pediatric dietitians is key to the management of children receiving hemodialysis. The effects of chronic illness on growth and school achievement, as well as the psychological, emotional, and social development of the child should not be underestimated. This review focuses on the above elements and highlights common pediatric practice in the United Kingdom. 7. Dunne N, Campbell M, Fitzpatrick M, Callery P. Comparison of Kt/V and urea reduction ratio in measuring dialysis adequacy in paediatric haemodialysis in England.J Ren Care. 2014 Jun;40(2):117-24. doi: 10.1111/jorc.12059. Epub 2014 Mar 20. Abstract: Background: The National Kidney Foundation-Dialysis Outcomes Quality Initiative (KDOQI) guidelines and the Renal Association recommend the use of either Kt/V or urea reduction ratio (URR) to measure haemodialysis adequacy.

	Objectives: To determine the methods used to measure paediatric haemodialysis adequacy and to assess consistency between calculations of single pool Kt/V (spKt/V) and URR. Design: A service evaluation was conducted to establish current practices in measuring dialysis adequacy. A prospective longitudinal study was conducted to compare spKt/V and URR. Participants: Thirty-two children were recruited consisting of 13 males and 19 females in five paediatric dialysis centres. Results: Inconsistencies were reported of the method of post-urea sampling with 4 of the 10 centres using the KDOQI recommended sampling method. Five dialysis centres reported using URR and five reported using spKt/V. There were substantial differences between the two measures. Using URR suggested that up to 44% of children did not receive adequate dialysis, whereas measurement by spKt/V suggested no more than 6% of the same dialysis sessions were not adequate. Conclusion: One standard measure should be used to assess dialysis adequacy in paediatric centres in England. KDOQI guidelines were not consistently followed in obtaining a post-urea blood sample and this procedure should be standardised. 8. Cadnapaphornchai MA, Teitelbaum I. Strategies for the preservation of residual renal function in pediatric dialysis patients. Pediatr Nephrol. 2014 May;29(5):825-36; quiz 832. doi: 10.1007/s00467-013-2554-0. Epub 2013

	Jul 19. Abstract: In adults with end-stage renal disease (ESRD), the preservation of residual renal function (RRF) has been shown to be associated with decreased mortality and improved control of complications of chronic kidney disease. However, less is known on the benefits of RRF in the pediatric dialysis population. The purpose of this article is to review the clinical significance of RRF and to discuss strategies for the preservation of RRF in children with ESRD.

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.

N/A

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

N/A

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

N/A

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 1423
Measure Title: Minimum spKt/V for Pediatric Hemodialysis Patients
Date of Submission: 1/7/2019
Type of Measure:

☐ Outcome (*including PRO-PM*)	☐ Composite – ***STOP – use composite testing form***
☒ Intermediate Clinical Outcome	☐ Cost/resource
☐ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.**
- **For outcome and resource use measures**, section **2b3** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b5** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b1-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates 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. For **instrument-based measures** (including PRO-PMs) **and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b1. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **instrument-based measures (including PRO-PMs) and composite performance measures**, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; [12]

AND
If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [13]

2b3. For outcome measures and other measures when indicated (e.g., resource use):
- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR
- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;
OR
there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses 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 nonresponders) and how the specified handling of missing data minimizes bias.

Notes
10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).
11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.
12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.
13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.
14. Risk factors that influence outcomes should not be specified as exclusions.
15. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☒ claims	☒ claims
☒ registry	☒ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

CROWNWeb and Medicare Claims Data from January 2013 to December 2013

For the Spring 2019 maintenance submission, 2017 CROWNWeb and Medicare claims data were used.

1.3. What are the dates of the data used in testing? January 2013 to December 2013

For the Spring 2019 maintenance submission, January – December 2017 data were used.

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☐ other: Click here to describe	☐ other: Click here to describe

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities*

included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample)

13 facilities with at least 11 eligible pediatric patients. Public reporting of this measure on DFC or in the ESRD QIP would be restricted to facilities with at least 11 eligible patients for the measure. We have applied this restriction to all the reliability and validity testing reported here.

For the Spring 2019 maintenance submission, 14 facilities that had at least 11 eligible patients during January 2017 – December 2017 were included in the analyses. Public reporting of this measure on DFC or in the ESRD QIP would be restricted to facilities with at least 11 eligible patients for the measure to comply with restrictions on reporting of potentially patient identifiable information related to small cell size. We have applied this restriction to all the reliability and validity testing reported here.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

180 HD patients and 1,195 patient months in facilities with at least 11 eligible pediatric patients.

For the Spring 2019 maintenance submission, 225 patients (1,603 patient-months) who are from 14 facilities with at least 11 eligible patients were included in the analyses.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

N/A

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

N/A

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

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*)

We used January 2013 – December 2013 CROWNWeb and Medicare claims data to calculate the inter-unit reliability (IUR) for the overall 12 months to assess the reliability of this measure. The NQF-recommended approach for determining measure reliability is a one-way analysis of variance (ANOVA), in which the between and within facility variation in the measure is determined. The inter-unit reliability (IUR) measures the proportion of the measure variability that is attributable to the between-facility variance. The yearly based IUR was estimated using a bootstrap approach, which uses a resampling scheme to estimate the within facility variation that cannot be directly estimated by ANOVA. We note that the method for calculating the IUR was developed for measures that are approximately normally distributed across facilities. Since this measure is not normally distributed, the IUR value should be interpreted with some caution.

For the Spring 2019 maintenance submission, we followed the same methodology as described above, using January 2017 – December 2017 CROWNWeb and Medicare Claims data.

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*)

For reliability we calculated the monthly and annual IUR across the 12 reporting months. As explained above, the method for calculating the IUR was developed for measures that are approximately normally distributed across facilities. IUR=0.807 with the confidence interval being (0.623, 0.929). This suggests that 81% of variation in the measure is attributed to between facility variation.

For the Spring 2019 maintenance submission, the annual IUR=0.750 across 12 reporting months, which suggests 75% of variation in the measure is attributed to between facility variation.

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?*)

The IUR suggest this measure is moderately reliable. However, since the distribution of performance scores is skewed, the IUR value should be interpreted with some caution.

For the Spring 2019 maintenance submission, the IUR again suggests this measure is reliable. However, since the distribution of performance scores is skewed, the IUR value should be interpreted with some caution.

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 <u>performance measure score</u> 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)*

Assessment based on face validity by the 2010 HD pediatric TEP.

For the Spring 2019 maintenance submission, the measure is being maintained on the basis of face validity.

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

This measure is being maintained on the basis of face validity.

Use of small solute clearance (urea reduction ratio and more recently Kt/V) as a dialysis quality measure was initially developed and approved by the Clinical TEP in 2010 on pediatric dialysis adequacy, which agreed that this quality measure domain is important in the assessment of the quality of care for pediatric dialysis patients. Achieving target Kt/V was finalized for the ESRD QIP beginning with PY 2015, and has been reported on Dialysis Facility Compare since January 2013.

In July 2015, CMS and UM-KECC revised this measure in response to concerns from the NQF Steering Committee regarding the appropriateness of single pool Kt/V for measuring Kt/V in patients who are dialyzing 3 or 4 times per week. Members of the committee argued that single pool Kt/V is not appropriate for assessing patients on different dialysis frequencies, i.e., 3 or 4 times. Standard Kt/V would be more appropriate for assessing different dialysis frequencies, such as 3 or 4 times per week.

The 2010 TEP that recommended this measure originally specified that the measure include patients on dialysis 3 or 4 times per week. This was based in part on analyses presented at that 2010 TEP meeting showing that 4 times per week hemodialysis was observed in approximately 5.6% of pediatric patient weeks, and nearly 90% of pediatric patient weeks reflected either 3 or 4 times per week hemodialysis (based on 2007 Medicare claims data). Given that this was not an insignificant proportion of patients, as the TEP concluded that these patients should all be included in this measure. UM-KECC recently updated this analysis using 2014 Medicare claims data. This showed that 4 times per week hemodialysis is now observed in 8.06% of patient weeks and approximately 92% of pediatric patient weeks reflected either 3 or 4 times per week hemodialysis. Results were generally similar using 2014 CROWNWeb data. About 7% of pediatric patient weeks were 4 times per week hemodialysis, and 94% of pediatric patient weeks were 3 or 4 times per week.

In response to the NQF Steering Committee concerns of including patients on 3 or 4 times per week dialysis, UM-KECC contacted the members of the 2010 TEP and asked them to consider a revision to limit the measure to pediatric patients on three times a week dialysis. This revision would make this

pediatric hemodialysis adequacy measure consistent with the corresponding adult measure (#0249). A majority of the TEP members supported this revision. To date, 5 of the 2010 TEP members (including the TEP chair) voted to revise the measure and limit it to pediatric patients on 3 times per week dialysis (the 2 remaining TEP members have not yet responded to the request for their feedback on this proposed change). The revised measure retains face validity based on the results of this vote.

Here is the specific decision provided by the TEP Chair on behalf of the TEP:

Specifically, the Pediatric CMS-TEP (2010) members who participated in the Pediatric Hemodialysis Adequacy Technical Expert Panel acknowledged that the measure developed in 2010 pertaining to the use of spKt/V stipulated that the measure should include patients on 3 and 4 times per week dialysis. However, they also appreciate the concern of the NQF steering committee charged with reviewing the measure for maintenance of endorsement that "the UKM or Daugirdas formulas are designed for a fixed number of dialysis treatments a week, not 3 or 4". In addition, they recognize that there are no data justifying the inclusion of children who receive more than 3 sessions of dialysis per week in this measure. As a result, the TEP members are in favor of revising the specifications to limit the measure to pediatric patients receiving dialysis 3 times per week. They believe that this would be consistent with the adult measure, which specifies that the measure is for adult hemodialysis patients dialyzing 3 times per week (patients dialyzing <3 or >3 times per week are excluded).

For the Spring 2019 submission, the text from the previous submission above still applies. No further revisions have been made to the measure to require further TEP review.

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?*)

This measure was established on the basis of face validity. Clinical TEP members agreed that this measure will improve quality of care for pediatric HD patients.

For the Spring 2019 submission, the measure is maintained on the basis of face validity. Clinical TEP members agreed that this measure will improve quality of care for pediatric HD patients. The measure was implemented in the ESRD QIP beginning with PY 2015, and has been reported on DFC since 2013.

_____

2b2. EXCLUSIONS ANALYSIS

NA ☒ no exclusions — *skip to section 2b3*

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*)

N/A

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*)

N/A

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*)
N/A

_____

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 2b4.

2b3.1. What method of controlling for differences in case mix is used?
☒ **No risk adjustment or stratification**
☐ **Statistical risk model with** Click here to enter number of factors **risk factors**
☐ **Stratification by** Click here to enter number of categories **risk categories**
☐ **Other,** Click here to enter description

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.
N/A

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.
N/A

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?
N/A

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?
N/A

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.**

N/A

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

N/A

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

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

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

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*)

N/A

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*)

N/A

_

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)*

Given that the number of facilities included in the calculation in 1.5 is only 13, the sample was determined to be too small to display useful data on meaningful differences.

For the Spring 2019 Maintenance submission, the number of facilities included in the calculation is still small (14), therefore analyses on statistically meaningful differences were not conducted.

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*)

N/A

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?*)

N/A

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*)

N/A

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*)

N/A

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*)

N/A

_____

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 nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

Reporting months with missing values are not excluded from this measure. Missing months are used to help define the measure numerator (missing is counted as not meeting the minimum threshold), so introduction of bias from exclusion of missing values is not a consideration for the measure as specified.

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*)

N/A

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 nonresponders) 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*)

N/A