

NATIONAL QUALITY FORUM

Measure Submission and Evaluation Worksheet 5.0

This form contains the information submitted by measure developers/stewards, organized according to NQF's measure evaluation criteria and process. The evaluation criteria, evaluation guidance documents, and a blank online submission form are available on the [submitting standards web page](#).

| | |
|--|--|
| NQF #: 0320 | NQF Project: Renal Endorsement Maintenance 2011 |
| (for Endorsement Maintenance Review) | |
| Original Endorsement Date: Nov 15, 2007 Most Recent Endorsement Date: Nov 15, 2007 Last Updated Date: May 08, 2012 | |
| BRIEF MEASURE INFORMATION | |
| De.1 Measure Title: Patient Education Awareness—Physician Level | |
| Co.1.1 Measure Steward: Kidney Care Quality Alliance | |
| De.2 Brief Description of Measure: Percentage of end stage renal disease (ESRD) patients aged 18 years and older with medical record documentation of a discussion of renal replacement therapy modalities (including hemodialysis, peritoneal dialysis, home hemodialysis, transplants and identification of potential living donors, and no/cessation of renal replacement therapy) at least once during the 12-month reporting period. | |
| 2a1.1 Numerator Statement: Number of patients from the denominator with medical record documentation of a discussion of renal replacement therapy modalities (including hemodialysis, peritoneal dialysis, home hemodialysis, transplants and identification of potential living donors, and no/cessation of renal replacement therapy) at least once during the 12-month reporting period. | |
| 2a1.4 Denominator Statement: All ESRD patients aged 18 years and older receiving renal replacement therapy. | |
| 2a1.8 Denominator Exclusions: None. | |
| 1.1 Measure Type: Process | |
| 2a1. 25-26 Data Source: Other, Paper Records | |
| 2a1.33 Level of Analysis: Clinician : Individual | |
| 1.2-1.4 Is this measure paired with another measure? No | |
| De.3 If included in a composite, please identify the composite measure (title and NQF number if endorsed): Not applicable. | |

| |
|---|
| STAFF NOTES (issues or questions regarding any criteria) |
| Comments on Conditions for Consideration: |
| Is the measure untested? Yes <input checked="" type="radio"/> No <input checked="" type="radio"/> If untested, explain how it meets criteria for consideration for time-limited endorsement: |
| 1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (check De.5): |
| 5. Similar/related endorsed or submitted measures (check 5.1): |
| Other Criteria: |

Staff Reviewer Name(s):

1. IMPACT, OPPORTUNITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT

Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All three subcriteria must be met to pass this criterion. See [guidance on evidence](#).

Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)

1a. High Impact: **H O M O L O I O**

(The measure directly addresses a specific national health goal/priority identified by DHHS or NPP, or some other high impact aspect of healthcare.)

De.4 Subject/Topic Areas (Check all the areas that apply): Palliative Care and End-of-Life Care, Renal, Renal : End Stage Renal Disease (ESRD)

De.5 Non-Condition Specific (Check all the areas that apply): Access to Care, Care Coordination, Disparities Sensitive, Person-and Family-Centered Care, Safety : Complications

1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, A leading cause of morbidity/mortality, Patient/societal consequences of poor quality, Severity of illness

1a.2 If "Other," please describe:

1a.3 Summary of Evidence of High Impact (Provide epidemiologic or resource use data):

In 2008, the adjusted incident rate of end-stage renal disease (ESRD) cases in the United States was 350.8 per million population, and the adjusted rate of prevalent cases rose 1.9 percent to 1,699 per million population. This rate is nearly 20 percent greater than that seen in 2000, and the annual rate of increase has remained between 1.9 and 2.3 percent since 2003. Total Medicare costs rose nearly 11 percent in 2008—up from a 7 percent rise the previous year—to \$454 billion. ESRD costs rose 13.2 percent to \$26.8 billion, and accounted for 5.9 percent of the Medicare budget.(1) Additionally, resource utilization by ESRD patients is substantial. For instance, in 2008 nearly 92% of the 112,476 incident U.S. ESRD patients were being treated via hemodialysis. Of these, 98.5% were being dialyzed three or more times per week at three or more hours per session. Additionally, the risk of hospitalization is 1.25 times greater in ESRD patients than in patients without, and adjusted hospital admission rates for dialysis patients have fallen only 1.5 percent since 1993.

Kidney disease is a major cause of morbidity and is the ninth leading cause of death in the United States. Nearly 85,000 Americans die with kidney failure each year, with adjusted rates of all-cause mortality 6.4 to 7.8 times higher for dialysis patients than for individuals in the general population. The mortality rate is highest within the first six months of initiating dialysis at approximately 30 percent. The rate then declines over the next six months before increasing gradually again over the next four years. (1)

Education programs for chronic kidney disease (CKD) patients have been shown to delay the time to dialysis and even improve survival (2,3), and research indicates that patients with greater knowledge about dialysis at initiation are more likely to use an arteriovenous fistula or graft than a catheter.(4) Initiatives such as the RightStart Program, a patient education plan implemented in a number of Fresenius Medical Care dialysis clinics, have similarly demonstrated that focus on patient education and support can dramatically improve outcomes for chronic hemodialysis patients. In this study, a total of 918 incident hemodialysis patients were prospectively enrolled in a multicenter RightStart Program and were compared with a time-concurrent group of 1020 control patients from non-RightStart clinics. RightStart patients received three months of intervention in anemia management, dialysis dosage, nutrition, dialysis access, and a comprehensive educational program. Outcomes were tracked for 12 months. The RightStart patients had

significantly improved Mental Composite Scores and reduced hospitalization and mortality rates compared to control subjects, demonstrating that a structured program of prompt medical and educational strategies in incident hemodialysis patients results in improved morbidity and mortality that last up to one year.(5)

These findings strongly support the underlying construct of the KCQA patient education awareness measures—i.e., that all ESRD patients should be educated on all renal replacement therapy modality options on at least a yearly basis to improve patient outcomes. We also note that the measure is consistent with the regulatory imperative of CMS's new Conditions for Coverage for dialysis facilities, which mandate patient education on renal replacement therapy modalities and end-of-life. Specifically, the Conditions require that documentation in patient records demonstrate that facility staff provide unbiased education to patients/designees about transplantation and all dialysis treatment options, and that the patient has the right to receive resource information for dialysis modalities not offered by the facility.(6)

1a.4 Citations for Evidence of High Impact cited in 1a.3: 1. U.S. Renal Dialysis System, USRDS 2010 Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States. 2010.

2. Devins G, Mendelssohn D, Barre P and Binik Y. Predialysis psychoeducational intervention and coping styles influence time to dialysis in chronic kidney disease. Am J Kidney Dis. 2003;42:693–703.

3. Devins G, Mendelssohn D, Barre P, Taub K and Binik Y. Predialysis psychoeducational intervention extends survival in CKD: a 20-year follow-up. Am J Kidney Dis. 2005;46:1088–98.

4. Cavanaugh K, Wingard R, Hakin R, Elas T and Ikizler T. Patient dialysis knowledge is associated with permanent arteriovenous access use in chronic hemodialysis. Clin J Am Soc Nephrol. 2009;4:950-56.

5. Wingard R, Pupim L, Krishnan M, Shintani A, Ikizler T and Hakim R. Early intervention improves mortality and hospitalization rates in incident hemodialysis patients: RightStart Program. Clin J Am Soc Nephrol. 2007;2:1170-75.

6. Federal Register (2008). Conditions for Coverage for ESRD Facilities: 42 CFR Part 494, April 2008.

1b. Opportunity for Improvement: H● M● L● I●

(There is a demonstrated performance gap - variability or overall less than optimal performance)

1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:

The measure will improve patient outcomes by ensuring that ESRD patients are made aware of all renal replacement therapy options and are familiar with the benefits and limitations of each through the promotion of routine patient education by dialysis facilities. We again note that the measure is also consistent with the new Conditions for Coverage, which require patient education on renal replacement therapy modalities and end-of-life.

1b.2 Summary of Data Demonstrating Performance Gap *(Variation or overall less than optimal performance across providers): [For Maintenance – Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.]*

NOTE: The specifications for the facility level and individual clinician level KCQA Patient Education Awareness measures differ only in the level of analysis. Pilot testing was conducted in both environments, and collective results and conclusions are presented in the two Patient Education Awareness measure submission forms for clarity.

Facility Testing:

KCQA tested its ESRD measures through a prospective cohort study on a nationally drawn sample of 53

dialysis facilities containing a mix of for-profit and not-for-profit providers; hospital-affiliated and freestanding facilities within large, small, and independent dialysis organizations; urban, suburban, and rural settings; and facilities both with and without electronic health records (EHRs). Approximately 25 patients per facility were sought, resulting in a final sample size of 1,115 patients. Both facility and patient samples were structured to be generally representative of the national industry profile as identified by the United States Renal Data Systems (USRDS) 2007 Annual Data Report. Facility records were used as the data source, given CMS's intent to include the measure in its Phase III Clinical Performance Measures, which involve CROWNWeb electronic transmission of data from facility medical records. Because CROWNWeb was not operational at the time, standardized, paper-based data collection sheets were constructed from the endorsed specifications and were employed during data collection for the testing. Both hemodialysis and peritoneal dialysis patients are encompassed by the measure specifications. Patient education data were provided on all 1,115 patients in the sample.

Over the course of the study, 924 of the 1,115 patients received education about at least one modality option. Successful performance on the measure, however, requires documentation that the patient was educated in all the options. The following is a summary of the number of individuals receiving education on each modality:

- Hemodialysis = 850 patients (64.0%)
- Peritoneal dialysis = 623 (67.4%)
- Home hemodialysis = 572 (61.9%)
- Transplants = 850 (92.0%)
- Identification of living donors = 266 (28.8%)
- No or cessation of therapy = 237 (25.7%)

The measure specifications require documentation in the medical record that the patient was educated on ALL modalities. Medical record documentation existed for 922 of the 924 educated patients.

The data elements collected thus permit calculation of performance for the measure as follows:

Performance Rate =

$$\frac{([\text{Patients educated on all modalities}] - [\text{Patients educated but without documentation}])}{\text{Total ESRD patients}}$$

$$= (185 - 2) / 1,115 = 16.4\%$$

The performance for each individual facility in the pilot ranged from 0% to 100%, with a mean performance of 16.4%.

Physician Office Testing:

To test the measure in physician offices, KCP contracted with the Iowa Foundation for Medical Care (IFMC). IFMC was under existing contract with the American Medical Association Physician Consortium for Performance Improvement (AMA PCPI)/Renal Physicians Association (RPA) to perform on-site feasibility and implementation testing of several AMA PCPI/RPA measures, and both organizations generously permitted KCQA to "piggy-back" onto the AMA PCPI/RPA testing protocol. At the time KCQA engaged IFMC, it had already obtained consent from four nephrology practice sites that would consist of a nephrology practice alpha site local to IFMC and three sites distributed geographically across the United States (Iowa, Nevada, Texas, and Pennsylvania) of various practice sizes (5.25 to 62 physicians), and medical record types (two EHR, one paper but by the time of visit transitioning to EHR, and one hybrid). Each site was asked to pull in advance the records of the first 35 adult patients seen on or after July 1, 2007; IFMC requested what it referred to as an oversample of five patients in an effort to ensure a remaining sample of 30 patients. Additionally, following the alpha site, the following were stipulated:

- Patient had two face-to-face office visits between July 1, 2007 and June 30, 2008, or if not seen in the office twice, it was determined he/she was receiving ongoing care from the office practice by looking first at the medical reviews resulting in an annual History and Physical, then supplementing using the monthly billings until the office reached the total of 35 ESRD patients. E&M service codes included: 99201, 99202, 99203, 99204, 99205, 99212, 99213, 99214, 99215, 99241, 99242, 99243, 99244, or 99245.

- ESRD patients can be identified with an ICD-9 code of 585.6 or an ICD-10 code of N18.0 and G-codes or CPT codes descriptive of hemodialysis.

The three nephrology office sites, in addition to the alpha site, were visited by a two-person IFMC abstractor team to conduct feasibility and reliability testing. Using the KCQA data collection tool, the two abstractors individually abstracted each medical record, compared the results, and evaluated the mismatches.

Mismatch codes, previously developed by IFMC for reliability testing of project abstraction, were used to classify the reason determined for each mismatch.

Physician office performance measure results were calculated based on the data collected by IFMC, with a resultant mean facility performance rate of 97 percent.

Conclusions:

The findings from both the facility and physician office testing indicate that the majority of ESRD patients are not being educated on all renal replacement therapy modality options. Findings also demonstrate that provider performance varies significantly by modality (i.e., providers are likely to routinely discuss transplants with their patients, but will broach the topic of cessation of therapy with less regularity). The results identify an important gap in clinical performance.

1b.3 Citations for Data on Performance Gap: *[For Maintenance – Description of the data or sample for measure results reported in 1b.2 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included]*

Kidney Care Quality Alliance. KCQA Performance Measures Pilot Testing Results: Report to the National Quality Forum. Unpublished. February 2010.

1b.4 Summary of Data on Disparities by Population Group: *[For Maintenance –Descriptive statistics for performance results for this measure by population group]*

KCQA tested its ESRD measures through a prospective cohort study on a nationally drawn sample of 53 dialysis facilities. Approximately 25 patients per facility were sought, resulting in a final sample size of 1,115 patients. Both facility and patient samples were structured to be generally representative of the national industry profile as identified by the USRDS 2007 Annual Data Report. As minimal patient data were sought to protect confidentiality and the collection of race/ethnicity information was not necessary to test the measure's data elements for reliability and validity, an examination of the data for disparities trends was not conducted. However, the USRDS 2010 Annual Data Report indicates that the disease burden of ESRD disproportionately affects minority populations, in particular African American and Latino populations. The rate of ESRD in minority patients ranges from 1.5 to 4 times those of age-adjusted Caucasian patients.(1)

1b.5 Citations for Data on Disparities Cited in 1b.4: *[For Maintenance – Description of the data or sample for measure results reported in 1b.4 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included]*

U.S. Renal Dialysis System, USRDS 2010 Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States. 2010.

1c. Evidence (Measure focus is a health outcome OR meets the criteria for quantity, quality, consistency of the body of evidence.)

Is the measure focus a health outcome? Yes ☒ No ☐ If not a health outcome, rate the body of

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

Created on: 05/24/2021 at 05:30 PM

| | | | |
|--|----------------|--------------------|---|
| evidence. | | | |
| Quantity: H● M● L● I● Quality: H● M● L● I● Consistency: H● M● L● I● | | | |
| Quantity | Quality | Consistency | Does the measure pass subcriterion1c? |
| M-H | M-H | M-H | Yes● |
| L | M-H | M | Yes● IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No● |
| M-H | L | M-H | Yes● IF potential benefits to patients clearly outweigh potential harms: otherwise No● |
| L-M-H | L-M-H | L | No ● |
| Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service | | | Does the measure pass subcriterion1c? Yes● IF rationale supports relationship |
| <p>1c.1 Structure-Process-Outcome Relationship (<i>Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome</i>):</p> <p>ANTECEDENTS: Education of ESRD patients has been shown to decrease catheter use, improve survival, and improve outcomes >> PROCESS: Assessment of the proportion of a provider's ESRD patient population being educated on all renal replacement therapy modality options >> Promotion of routine modality education of ESRD patients >> OUTCOME: Improved patient awareness of renal replacement therapy modality options >> Increased active patient involvement in care decisions and processes >> Improved outcomes.</p> <p>1c.2-3 Type of Evidence (<i>Check all that apply</i>):</p> <p>Systematic review of body of evidence (other than within guideline development)</p> <p>1c.4 Directness of Evidence to the Specified Measure (<i>State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population</i>):</p> <p>Central Topic: Promotion of routine patient education on renal replacement therapy modality options.</p> <p>Population: Adult ESRD patients.</p> <p>Outcomes Addressed: Proportion of ESRD patients within a facility who have been educated on all renal replacement therapy modality options.</p> <p>Differences Between Measure Focus and Measure Target Population: None.</p> <p>1c.5 Quantity of Studies in the <u>Body of Evidence</u> (<i>Total number of studies, not articles</i>): The body of evidence presented in Section 1c.6 cites seven peer-reviewed clinical studies.</p> <p>1c.6 Quality of <u>Body of Evidence</u> (<i>Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events</i>): Summary of Body of Evidence:</p> <p>Education programs for chronic kidney disease (CKD) patients have been shown to delay the time to</p> | | | |

dialysis and even improve survival (1,2), and research indicates that patients with greater knowledge about dialysis at initiation are more likely to use an arteriovenous fistula (AVF) or graft than a catheter.(3) While several studies have demonstrated an association between patient education and improved outcomes in the ESRD population, none were identified that focused exclusively on renal replacement modality options, as is the case with the KCQA patient education measure. However, the research does exist that lends support to the underlying construct of the measure—that educating ESRD patients on the available modalities (i.e., hemodialysis, peritoneal dialysis, home hemodialysis, transplants, and no or cessation of therapy) improves clinical outcomes:

- A prospective cohort of 490 adult incident hemodialysis patients followed from May 2002 until November 2005 revealed that patients that received less education and had lower levels of dialysis knowledge (as measured using the Chronic Hemodialysis Knowledge Survey) were less likely to use an AVF for dialysis than their more knowledgeable counterparts, thereby placing them at higher risk for complications and mortality.(3)
- Educational interventions with an emphasis on empowerment have been shown to improve depression, medication adherence, treatment attendance, and choice of vascular access.(3-5)
- Research has shown that patients who take the lead in choosing their treatment modality, or work together with their medical team, are much more likely to choose home dialysis modalities-and more likely to survive and to get a transplant.(6)
- The RightStart Program, a patient education plan implemented in a number of Fresenius Medical Care dialysis clinics, demonstrated that focus on patient education and support can dramatically improve outcomes for chronic hemodialysis patients. In this study, a total of 918 incident hemodialysis patients were prospectively enrolled in a multicenter RightStart Program and were compared with a time-concurrent group of 1020 control patients from non-RightStart clinics. RightStart patients received three months of intervention in anemia management, dialysis dosage, nutrition, dialysis access, and a comprehensive educational program. Outcomes were tracked for 12 months. The RightStart patients had significantly improved Mental Composite Scores and reduced hospitalization and mortality rates compared to control subjects, demonstrating that a structured program of prompt medical and educational strategies in incident hemodialysis patients results in improved morbidity and mortality that last up to one year.(7)

These findings strongly support the underlying construct of the KCQA patient education awareness measures—i.e., that ESRD patients should routinely be educated on all renal replacement therapy modality options to improve patient outcomes.

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect):* The studies cited in Section 1c.6. consistently demonstrate that patients who have been educated on the available renal replacement therapy modalities (i.e., hemodialysis, peritoneal dialysis, home hemodialysis, transplants, and no or cessation of therapy) are more likely to use an AVF for dialysis (3), have less depression and improved medication adherence and treatment attendance (3-5), and are more likely to survive and to get a transplant.(6,7)

1c.8 Net Benefit *(Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms):*

Research consistently illustrates the net benefit of educating patients on the available renal replacement therapy modalities (i.e., hemodialysis, peritoneal dialysis, home hemodialysis, transplants, and no or cessation of therapy). Educated patients are more likely to use an AVF for dialysis (3), have less depression and improved medication adherence and treatment attendance (3-5), and are more likely to survive and to get a transplant.(6,7) We note that in recognition of these well-established benefits, the new Conditions for Coverage require patient education on renal replacement therapy modalities and end-of-life.

1c.9 Grading of Strength/Quality of the Body of Evidence. Has the body of evidence been graded? **No**

1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: **Not applicable.**

1c.11 System Used for Grading the Body of Evidence: **Other**

1c.12 If other, identify and describe the grading scale with definitions: **The body of evidence has not been formally graded.**

1c.13 Grade Assigned to the Body of Evidence: **Not applicable.**

1c.14 Summary of Controversy/Contradictory Evidence: **Research consistently illustrates the net benefit of educating patients on the available renal replacement therapy modalities (i.e., hemodialysis, peritoneal dialysis, home hemodialysis, transplants, and no or cessation of therapy). Educated patients are more likely to use an AVF for dialysis (3), have less depression and improved medication adherence and treatment attendance (3-5), and are more likely to survive and to get a transplant.(6,7) No contradictory evidence was identified.**

1c.15 Citations for Evidence other than Guidelines(*Guidelines addressed below*):

1. Devins G, Mendelssohn D, Barre P and Binik Y. Predialysis psychoeducational intervention and coping styles influence time to dialysis in chronic kidney disease. *Am J Kidney Dis.* 2003;42:693–703.

2. Devins G, Mendelssohn D, Barre P, Taub K and Binik Y. Predialysis psychoeducational intervention extends survival in CKD: a 20-year follow-up. *Am J Kidney Dis.* 2005;46:1088–98.

3. Cavanaugh K, Wingard R, Hakin R, Elas T and Ikizler T. Patient dialysis knowledge is associated with permanent arteriovenous access use in chronic hemodialysis. *Clin J Am Soc Nephrol.* 2009;4:950-56.

4. Beder J et al. Effectiveness of a social work psychoeducational program in improving adherence behavior associated with risk of CVD in ESRD patients. *J Neph Social Work.* 2003; 22:12-22.

5. Johnstone S et al. Overcoming early learning barriers in hemodialysis patients: the use of screening and educational reinforcement to improve treatment outcomes. *Adv in Chronic Kid Dis.* 2004; 11(2):210-16.

6. Stack A. Determinants of modality selection among incident U.S. dialysis patients: results from a national study. *J Am Soc Nephrol.* 2002; 13(5):1279-87.

7. Wingard R, Pupim L, Krishnan M, Shintani A, Ikizler T and Hakim R. Early intervention improves mortality and hospitalization rates in incident hemodialysis patients: RightStart Program. *Clin J Am Soc Nephrol.* 2007;2:1170-75.

1c.16 Quote verbatim, the specific guideline recommendation (Including guideline # and/or page #): **Not applicable.**

1c.17 Clinical Practice Guideline Citation: **Not applicable.**

1c.18 National Guideline Clearinghouse or other URL: **Not applicable.**

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? **No**

1c.20 If guideline recommendation graded, identify the entity that graded the evidence including

balance of representation and any disclosures regarding bias:

1c.21 System Used for Grading the Strength of Guideline Recommendation: Other

1c.22 If other, identify and describe the grading scale with definitions: There is no formal guideline addressing patient education in ESRD patients.

1c.23 Grade Assigned to the Recommendation: Not applicable.

1c.24 Rationale for Using this Guideline Over Others: Not applicable.

Based on the NQF descriptions for rating the evidence, what was the developer's assessment of the quantity, quality, and consistency of the body of evidence?

1c.25 Quantity: High **1c.26 Quality:** High **1c.27 Consistency:** High

1c.28 Attach evidence submission form:

1c.29 Attach appendix for supplemental materials:

Was the threshold criterion, *Importance to Measure and Report*, met?

(1a & 1b must be rated moderate or high and 1c yes) Yes ☐ No ☒

Provide rationale based on specific subcriteria:

For a new measure if the Committee votes NO, then STOP.

For a measure undergoing endorsement maintenance, if the Committee votes NO because of 1b. (no opportunity for improvement), it may be considered for continued endorsement and all criteria need to be evaluated.

2. RELIABILITY & 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. (**evaluation criteria**)

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate field. Supplemental materials may be referenced or attached in item 2.1. See [guidance on measure testing](#).

S.1 Measure Web Page (*In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained*). Do you have a web page where current detailed specifications for this measure can be obtained? Yes

S.2 If yes, provide web page URL: <http://www.kidneycarepartners.com>

2a. RELIABILITY. Precise Specifications and Reliability Testing: H ☒ M ☒ L ☐ I ☐

2a1. Precise Measure Specifications. (*The measure specifications precise and unambiguous.*)

2a1.1 Numerator Statement (*Brief, narrative description of the measure focus or what is being measured about the target population, e.g., cases from the target population with the target process, condition, event, or outcome*):

Number of patients from the denominator with medical record documentation of a discussion of renal replacement therapy modalities (including hemodialysis, peritoneal dialysis, home hemodialysis, transplants and identification of potential living donors, and no/cessation of renal replacement therapy) at least once during the 12-month reporting period.

2a1.2 Numerator Time Window (*The time period in which the target process, condition, event, or outcome is eligible for inclusion*):

12-month reporting period.

2a1.3 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, codes with descriptors, and/or specific data collection items/responses):*

The necessary data elements are to be collected via the Centers for Medicare and Medicaid (CMS) CROWNWeb data repository when functional, as indicated by the measure's inclusion in CMS's list of Phase III ESRD Clinical Performance Measures in effect April 1, 2008.

1. Patient modality education status (select one):

- Yes, renal replacement therapy modality options have been discussed with the patient during the 12-month reporting period >> GO TO 2.
- No, renal replacement therapy modality options have NOT been discussed with the patient during the 12-month reporting period >> END.

2. Types of modalities discussed (check all that apply): >> GO TO 3.

- Hemodialysis
- Peritoneal dialysis
- Home hemodialysis
- Transplants
- Identification of potential living donors
- No or cessation of renal replacement therapy

3. Type of documentation in the medical records (check all that apply):

- No documentation
- A note or letter prepared by the nephrologist or other healthcare professional within the nephrologist's practice
- A note prepared by facility personnel

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured):*
All ESRD patients aged 18 years and older receiving renal replacement therapy.

2a1.5 Target Population Category *(Check all the populations for which the measure is specified and tested if any):* Adult/Elderly Care

2a1.6 Denominator Time Window *(The time period in which cases are eligible for inclusion):*
12-month reporting period.

2a1.7 Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses):*

The necessary data elements are to be collected via the CMS CROWNWeb data repository when functional, as indicated by the measure's inclusion in CMS's list of Phase III ESRD Clinical Performance Measures in effect April 1, 2008.

The denominator population consists of all ESRD patients under a given physician's care. Data elements required to identify the denominator population:

1. Patient diagnosis = ESRD

AND

2. Patient's date of birth

2a1.8 Denominator Exclusions (Brief narrative description of exclusions from the target population):
None.

2a1.9 Denominator Exclusion Details (All information required to identify and calculate exclusions from the denominator such as definitions, codes with descriptors, and/or specific data collection items/responses):
Not applicable.

2a1.10 Stratification Details/Variables (All information required to stratify the measure results including the stratification variables, codes with descriptors, definitions, and/or specific data collection items/responses):
Not applicable.

2a1.11 Risk Adjustment Type (Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13): No risk adjustment or risk stratification **2a1.12 If "Other," please describe:**

2a1.13 Statistical Risk Model and Variables (Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.):
Not applicable.

2a1.14-16 Detailed Risk Model Available at Web page URL (or attachment). Include coefficients, equations, codes with descriptors, definitions, and/or specific data collection items/responses. Attach documents only if they are not available on a webpage and keep attached file to 5 MB or less. NQF strongly prefers you make documents available at a Web page URL. Please supply login/password if needed:

2a1.17-18. Type of Score: Rate/proportion

2a1.19 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

2a1.20 Calculation Algorithm/Measure Logic(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; aggregating data; risk adjustment; etc.):

DENOMINATOR

Include in the denominator all patients who meet the following criteria:

1. Diagnosis = ESRD

AND

2. Patient's age = ≥ 18 years as of the first day of the most recent month of the reporting period. (Patient's age is or shall be determined by subtracting the patient's date of birth from the first day of the most recent month of the reporting period.)

NUMERATOR

Include in the numerator all patients from the denominator who meet the following criteria:

1. Patient modality education status = Yes, renal replacement therapy modality options have been discussed with the patient during the 12-month reporting period

AND

2. Types of modalities discussed =

- Hemodialysis

AND

- Peritoneal dialysis

AND

- Home hemodialysis

AND

- Transplants

AND

- Identification of potential living donors

AND

- No or cessation of renal replacement therapy

AND

3. Type of documentation in medical records =

- A note or letter prepared by the nephrologist or other healthcare professional within the nephrologist's practice

OR

- A note prepared by facility personnel

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:

Attachment

[txKCQACalcAlgorithmPtEdPhysician06-07-11FINAL.pdf](#)

2a1.24 Sampling (Survey) Methodology. If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate):

Not applicable.

2a1.25 Data Source (Check all the sources for which the measure is specified and tested). If other, please describe:

Other, Paper Records

2a1.26 Data Source/Data Collection Instrument (Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.): All data elements for the measure can be collected using the KCQA Patient Education Data Collection Form (attached), which reflect the data elements to be included in CROWNWeb.

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment:

Attachment

[fmKCQADataFormPtEdClinicianFINAL.pdf](#)

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:

2a1.33 Level of Analysis *(Check the levels of analysis for which the measure is specified and tested):*
 Clinician : Individual

2a1.34-35 Care Setting *(Check all the settings for which the measure is specified and tested):* Ambulatory
 Care : Clinician Office, Post-Acute Care

2a2. Reliability Testing. *(Reliability testing was conducted with appropriate method, scope, and adequate demonstration of reliability.)*

2a2.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*

Facility Testing:

KCQA tested its ESRD measures through a one-year prospective cohort study on a nationally drawn sample of 53 dialysis facilities containing a mix of for-profit and not-for-profit providers; hospital-affiliated and freestanding facilities within large, small, and independent dialysis organizations; urban, suburban, and rural settings; and facilities both with and without electronic health records (EHRs). Approximately 25 patients per facility were sought, resulting in a final sample size of 1,115 patients. Both facility and patient samples were structured to be generally representative of the national industry profile as identified by the United States Renal Data Systems (USRDS) 2007 Annual Data Report. Facility records were used as the data source, given CMS's inclusion of the measure in its Phase III Clinical Performance Measures, which will involve CROWNWeb electronic transmission of data from facility medical records. Because CROWNWeb was not operational at the time, standardized, paper-based data collection sheets constructed from the endorsed specifications were employed during data collection for the testing.

Physician Office Testing:

To test the measure in physician offices, KCP contracted with IFMC, which was under an existing contract with the AMA PCPI/RPA to perform on-site feasibility and implementation testing of several AMA PCPI/RPA measures and had thus already obtained consent from four nephrology practice sites that would consist of a nephrology practice alpha site local to IFMC and three sites distributed geographically across the United States (Iowa, Nevada, Texas, and Pennsylvania) of various practice sizes (5.25 to 62 physicians), and medical record types (two EHR, one paper but by the time of visit transitioning to EHR, and one hybrid). Each site was asked to pull in advance the records of the first 35 adult hemodialysis patients seen on or after July 1, 2007; IFMC requested what it referred to as an oversample of five patients in an effort to ensure a remaining sample of 30 patients.

2a2.2 Analytic Method *(Describe method of reliability testing & rationale):*

Facility Testing:

Following the data collection period, on-site data-integrity audits were performed at 11 of the 53 facilities (21%). Audit sites were selected to provide a cross-section of facilities reflective of the sample profile. Selection criteria included geographic location, facility type (e.g., for-profit vs. not-for-profit, urban vs. rural), and EHR use. Pertinent data were reabstracted from the patients' medical records and were compared to the information submitted by the facility throughout the pilot to assess the measure's reliability.

Physician Office Testing:

The three nephrology office sites, in addition to the alpha site, were visited by a two-person IFMC abstractor

team to conduct feasibility and reliability testing. Using the KCQA data collection tool, the two abstractors individually abstracted each medical record, compared the results, and evaluated the mismatches. Mismatch codes, previously developed by IFMC for reliability testing of project abstraction, were used to classify the reason determined for each mismatch.

2a2.3 Testing Results (*Reliability statistics, assessment of adequacy in the context of norms for the test conducted*):

Facility Testing:

Inter-rater reliability was assessed during the on-site audits through a direct comparison of data submitted by the facilities throughout the pilot to data reabstracted by the auditor(s). (See Table 1 [Measure Performance, Submitted vs. Reabstracted Data] in the accompanying Attachment A.) Reliability was quantitatively summarized using Cohen's Kappa with confidence intervals. The resulting Kappa statistic for the measure was found to be -0.0026 with a 95% confidence interval of -0.1251 - 0.1199 . (See Table 2 [Measure Aggregate Reliability] in Attachment A.) Based on the literature, this negative Kappa value indicates that the auditor obtained the same results as the facility abstractor less than would be expected by chance alone.

Consistent with the aggregate reliability results, the patient education measure also had a relatively low concordance rate, again demonstrating substantial inter-abstractor disagreement. Specifically, the percent agreement between the auditor and facility abstractors (i.e., the reliability percentage) was found to be 71.2%. (See Table 3 [Measure Reliability Percentage and Error Type] in Attachment A.) The reason for the inter-rater discrepancies was further characterized by discordance codes, as noted in the table.

However, for the following reasons, we did not believe that the negative Kappa and low inter-rater concordance was due to unreliability of the measure specifications/tool per se:

- First, if the reliability issues were inherent to the measure/tool itself, one would have expected the types of errors to be evenly distributed, which they were not: There were significantly more false negatives (Y/N—see Table 2) that resulted from facilities failing to count education that was present in their records—this issue can be remedied by improving the instructions to and education of personnel as to what constitutes education for purposes of meeting the specifications.
- Second, the errors for the patient education measure were largely of one type (Table 3 [Measure Reliability Percentage and Error Type] in Attachment A): “information missed,” meaning the auditors found the information in the facility's documentation, whereas the facility personnel failed to account for it. Based on interviews with personnel during the audit, while a few cases were true instances where information was overlooked, in the majority of instances the discordance arose because the facility personnel did not realize that discussions about advance directives, for example, constituted education related to end-of-life/cessation of therapy for purposes of the measure specifications.

We thus posited that the negative Kappa was attributed to two possibilities (or a combination thereof):

- First, the negative Kappa may have resulted from a subgroup of facilities accounting for significant numbers of errors that lead to a negative Kappa that “swamped” a subgroup that reliably collected data.
- Second, the negative Kappa may have been a result of an apparent confusion on the part of the facility abstractors over what, for the purposes of this measure, constitutes education, leading them to under-report meeting the measure specifications. For instance, facility abstractors frequently overlooked advance directive discussions in which there was clear documentation that cessation of renal replacement therapy was presented as an option to the patient. Given that all modalities must be discussed to achieve credit for the measure, this omission alone led to substantial inter-abstractor disagreement in the overall measure results.

To test these hypotheses, we performed additional facility-by-facility error analyses and reliability analyses by data element. We recognize the overall results for the patient education measures are not optimal. Based on the additional test hypotheses and analyses that follow, however, we believe we demonstrate that the patient education measures can be reliably collected. Some facility staff clearly understood the requirements and reliably reported data. We note that the then new Conditions for Coverage, which require patient education on modalities and end-of-life, went into effect a month after the field period began. We posit that as these requirements are implemented, coupled with training about the measure (based on what was gleaned through the pilot), that the measure will be reliably reported when it is incorporated into CROWNWeb.

Facility Error Rates and Bimodal Distribution: To further assess the hypothesis that facilities can reliably collect the patient education data elements and that the Kappa issues were not inherent to the instrument itself, we reviewed the error rates for the 11 audited facilities. As Figure 1 in Attachment A (Percent of Errors for Patient Education Measure Data Elements) demonstrates, the distribution of facilities is somewhat bimodal—four that reliably collected the data (Facilities E, F, H, J) and seven that contributed the bulk of the errors (Facilities A, B, C, D, G, I, K). These findings support the need for more detailed instructions and educating facilities about what constitutes patient education for purposes of the measure specification. They also demonstrate that the measure is not intrinsically unreliable, since some facilities clearly can reliably report the data. The finding also supports the hypothesis that a subgroup of facilities account for significant numbers of errors—i.e., that the errors are not evenly distributed and that some facilities more reliably reported the education that was documented in their records.

Facility Kappas and Bimodal Distribution: To further examine whether a subset of facilities were largely responsible for the poor overall negative Kappa, individual Kappas were calculated for each facility. Table 4 (Kappas by Facility) in Attachment A presents this analysis. (A Kappa of 0.61 and greater is considered “substantial agreement” for purposes of assessing reliability.) As with the previous analysis, this analysis again supports the hypothesis that one subset of facilities reliably reported data, while a second set did less so.

Reliability of Individual Data Elements: To further investigate the origin of the measure’s reliability statistics, Kappas also were calculated for each individual data element contained in the specifications (see Table 5 [Measure Reliability if Individual Data Elements] in Attachment A). The first row in each category (black type) illustrates the actual results obtained during the audit. As can be seen in the table, a substantial number of false negatives (Y/N) were identified for nearly all of the data elements, again demonstrating that the facilities underreported their education sessions. The documentation was there; the facility personnel, however, failed to “count” what had been performed.

To determine the extent to which the underreporting adversely affected reliability scores, we next calculated Kappas for each data element by counting the false negatives for each data element as true positives (blue type)—as would be the case had facilities properly identified their education sessions. Note that to more easily calculate a Kappa, we arbitrarily assigned 2 false negatives instead of taking the value to 0.

As can be seen in Table 5, the results of this analysis significantly improved Kappas for all data elements and for the overall measure results. These findings support our hypothesis that the negative Kappa for the patient education measure is a function of a lack of diligence by participating facilities in identifying and reporting the patient education sessions documented in their records; the reliability problem is not intrinsic to the measure itself. We posit that better educating facilities on what, for the purposes of the measure, constitutes education will effectively address the audit findings of the measure’s negative Kappa.

Summary of Conclusions from Additional Analyses

We believe the additional analyses demonstrate the following:

- The negative Kappa for the overall patient education measure performance is not an indication that the measure specifications are inherently unreliable (Table 2).
 - o The type of error was not random (Table 3). Rather, significantly more errors were “missed information,” that lead to underreporting (false negatives). Further, the underreporting often stemmed from an apparent lack of understanding by some facilities as to what constituted education (and was documented in the facilities’ records) for purpose of the measure specifications (e.g., regarding advance directives/end-of-life/no therapy).
 - o Facilities either reliably collected data or did not. The distribution of errors among the facilities was not even (Figure 1). Error rates and facility-level Kappas showed a bimodal distribution (Figure 2).
 - o “Almost perfect,” reliability (Kappa 0.8474) resulted from testing in physician offices using two professional IFMC auditor/abstractors, as further described later in this chapter.
- Improving the instructions and educating facilities to recognize what constitutes meeting the specifications should reduce the high numbers of false negatives. When reduction scenarios of the high false-positive rate are analyzed, the Kappas indicate excellent agreement and reliability (Table 5).
- Ongoing implementation of the new Conditions for Coverage—which require the education encompassed by the KCQA measure specifications—will improve reliability by sensitizing facility personnel to systematize and organize their processes and recordkeeping so as to enable them to more reliably collect the data elements. In fact, we note that during the audit we observed some facilities had already begun more systematically documenting their patient education efforts merely in response to the Conditions for Coverage. As such efforts become more commonplace, they have the ancillary effect of supporting more reliable data collection related to this measure.
- Implementation of CROWNWeb and accountability for patient education can improve reliability by deploying more detailed instructions and training and by sensitizing facility personnel to systematize and organize their processes and recordkeeping so as to enable them to more reliably collect the data elements.

Based on the additional test hypotheses and analyses, we believe we demonstrate that the patient education measures can be reliably collected.

Physician Office Testing:

To determine whether the ESRD measure definitions and specifications, as prepared by KCQA, yield stable, consistent measurements when applied in the physician office setting, inter-rater reliability was also assessed by IFMC. In contrast to testing the same specifications in facilities, the patient education measure in this setting has a Kappa of 0.8474, indicating excellent reproducibility. The literature indicates that Kappa scores of 0.75 and above denote “excellent agreement beyond chance.” More specifically, the Kappa of 0.8474 indicates “almost perfect” agreement. From this value, we can conclude that the high inter-rater reliability rates observed for this measure are reproducible when used on these records and are fundamental to the specifications/tool itself, which was the same one used in the facility testing.

IFMC also examined the reliability percentage for the patient education measure, as summarized in Table 6 of Attachment A (Measure Reliability Percentage, Physician Office Setting). From these values, we again conclude that the high inter-rater reliability rates observed in the physician office testing are reproducible and are fundamental to the specifications and tool itself.

2b. VALIDITY. Validity, Testing, including all Threats to Validity: H● M● L● I●

2b1.1 Describe how the measure specifications (measure focus, target population, and exclusions) are consistent with the evidence cited in support of the measure focus (criterion 1c) and identify any differences from the evidence:

The measure focus and target population are consistent with the evidence cited in Section 1c; no differences have been identified. In both the body of evidence and the measure specifications, the target population is adult ESRD patients and the central topic is the promotion of routine patient education on renal replacement therapy modalities to increase patient awareness of treatment options and their involvement in care decisions to improve clinical outcomes.

2b2. Validity Testing. (*Validity testing was conducted with appropriate method, scope, and adequate demonstration of validity.*)

2b2.1 Data/Sample (*Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*):

KCQA tested its ESRD measures through a one-year prospective cohort study on a nationally drawn sample of 53 dialysis facilities. Approximately 25 patients per facility were sought, resulting in a final sample size of 1,115 patients. Both facility and patient samples were structured to be generally representative of the national industry profile as identified in the USRDS 2007 Annual Data Report.

2b2.2 Analytic Method (*Describe method of validity testing and rationale; if face validity, describe systematic assessment*):

Validity refers to the degree to which a performance measure truly measures what it was intended to measure (i.e., construct validity) and the degree to which the conclusions drawn from a test would hold for other persons, places, and times (external validity).

Construct Validity:

A test is said to have construct validity when it measures a construct (or theory) accurately. For the KCQA performance measures, the construct being tested is that the measures will accurately assess and depict a provider's practices. In claiming construct validity, we would thus be asserting that our pilot test confirmed that KCQA's patient education measure does in fact effectively portray a facility's and a physician office's patient education awareness practices. Specifically, KCQA asserts the measures meet the following types of construct validity: face validity and content validity.

- A measure is said to have face validity when it appears to be valid—i.e., on its “face” it seems like a good translation of the construct being tested. Face validity uses common-sense rules—for example, to assess a facility's vaccination practices, a measure should quantify its vaccination rate. While face validity is the weakest means of demonstrating construct validity, its strength can be improved by making the process more systematic—for instance, by utilizing a panel of experts to confirm that the measure appears to be a proper translation of the construct.

- Content validity centers on a measure's ability to include or represent all of the content of the construct in question. Content under-representation occurs when important areas are missed, and construct-irrelevant variation occurs when irrelevant factors contaminate the measure. Determination of content validity requires agreement among experts in the field in question. Thus, while face validity can be established by one person, content validity must be determined by a panel.

The KCQA measures have both face and content validity based on the following: The measures were deemed appropriate and valid by (1) expert opinion within the KCP and KCQA; (2) expert opinion within the NQF ESRD TAPs, Steering Committee, and the CSAC, all of which advanced the measures to the next stage of the CDP; and (3) broad agreement as demonstrated through the NQF review and voting processes.

External Validity:

A test is said to have external validity when results can be reliably generalized to the larger relevant population. External validity can be improved by employing appropriate methods to draw the sampling model from a population. For instance, when feasible, random selection should be utilized over a nonrandom procedure. Likewise, researchers should work to assure that respondents participate and that dropout rates are minimized.

KCQA posits that external validity has been met through the diligence with which the original sampling schema was crafted to reflect the national industry and patient vintage and access profiles. Because the sample is representative of the U.S. dialysis population, results can be generalized with confidence.

2b2.3 Testing Results *(Statistical results, assessment of adequacy in the context of norms for the test conducted; if face validity, describe results of systematic assessment):*

External Validity:

External validity of the KCQA measures was established through the meticulous construction of patient and facility samples, modeled to reflect the national industry and patient vintage and access profiles as per the 2007 USRDS Annual Data Report of Chronic Kidney Disease & End-Stage Renal Disease, the most current volume available at the time the sample was constructed. Because the sample is representative of the U.S. dialysis population, results can be generalized with confidence.

Facility Sampling: In the United States, dialysis services are provided at more than 4,800 sites (freestanding non-profit and for-profit centers, hospital-based, and government-affiliated entities—i.e., Department of Veterans Affairs or state/county/city-run). Based on the industry profile in the 2007 U.S. Renal Data System (USRDS), a recruitment list of 71 facilities that mirrored this profile was identified so as to reach a target of 60 facilities, from which we assumed additional attrition might occur during the one-year course of data collection. Department of Veterans Affairs (VA)-affiliated and other public facilities were excluded to streamline the facility recruitment process. (VA and other public facilities represent less than two percent of dialysis sites, and less than one percent of the patient population.) Based on the USRDS data, the following target facility distribution was constructed:

- 60% from for-profit large dialysis organizations (LDO),
- 15% from non-profit LDOs,
- 20% from for-profit non-LDOs, and
- 5% from non-profit non-LDOs.

Ultimately, 53 facilities participated in the pilot. The final facility sample contained a mix of both for-profit and not-for-profit providers; hospital-affiliated and freestanding facilities within large, small, and independent dialysis organizations; urban, suburban, and rural settings; and facilities both with and without electronic medical records, and was generally representative of the national industry profile. The facility distribution in the final sample was:

- 59% from for-profit LDOs,
- 8% from non-profit LDOs,
- 21% from for-profit non-LDOs, and
- 13% from non-profit non-LDOs.

Additionally, KCP members represent approximately 85% of the community; the final sample contained facilities involved with KCP members (47 facilities; 89%) and those not (6 facilities; 11%).

Patient Sampling: Twenty-five patients per facility were sought, and three primary patient-related variables were identified: dialysis type (hemodialysis, peritoneal dialysis, or home hemodialysis), vintage on dialysis, and vascular access type. Per the 2007 USDRS report, approximately 94.5 percent of patients are on in-

center hemodialysis, 5 percent on peritoneal dialysis, and 0.5 percent on home hemodialysis. The sample at the outset of the study was 92.6 percent in-center hemodialysis, 4.8 percent peritoneal dialysis, and 2.7 percent home hemodialysis. At the study's conclusion, the profile was 92.1 percent in-center hemodialysis 5.2 percent peritoneal dialysis, and 2.7 percent home hemodialysis. (The slight overrepresentation of home HD patients resulted from the participation of a facility caring exclusively for home-based hemodialysis patients. We also note that the 2007 USRDS atlas reports on data as of the end of 2005. In fact, the home hemodialysis population has been growing, and is currently estimated by community members to be 1-2%. Thus, the actual sample more accurately reflects the current situation. Regardless, nothing in the current literature indicated this small sampling difference from the national norm would have any impact on the pilot test results, and so the pilot proceeded with the original sample rather than exclude the facility with only home hemodialysis and/or attempt to replace it.)

The initial patient sample size equated to 1,325 adult patients (25 patients/53 facilities), but was reduced to 1,295 because some facilities did not have enough patients of a given type. This number was reduced to 1,115 by the study's conclusion due to patient death, transplantation, or patient transfer out of the participating facility.

Face Validity:

The KCQA measures have face validity based on the following: The measures were deemed appropriate and valid by (1) expert opinion within the KCP and KCQA; (2) expert opinion within the NQF ESRD TAPs, Steering Committee, and the CSAC, all of which advanced the measures (in some cases with recommended changes adopted by KCQA) to the next stage of the CDP; and (3) broad agreement as demonstrated through the NQF review and voting processes.

Content Validity:

The KCQA measures have content validity based on the following: The measures were deemed appropriate and valid by: (1) consensus of KCQA's expert panel; and (2) consensus of NQF's ESRD Technical Advisory Panels and Steering Committee.

POTENTIAL THREATS TO VALIDITY. *(All potential threats to validity were appropriately tested with adequate results.)*

2b3. Measure Exclusions. *(Exclusions were supported by the clinical evidence in 1c or appropriately tested with results demonstrating the need to specify them.)*

2b3.1 Data/Sample for analysis of exclusions *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*
Not applicable.

2b3.2 Analytic Method *(Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference):*
Not applicable.

2b3.3 Results *(Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):*
Not applicable.

2b4. Risk Adjustment Strategy. *(For outcome measures, adjustment for differences in case mix (severity) across measured entities was appropriately tested with adequate results.)*

2b4.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*
Not applicable.

2b4.2 Analytic Method (*Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables*):

Not applicable.

2b4.3 Testing Results (*Statistical risk model: Provide quantitative assessment of relative contribution of model risk factors; risk model performance metrics including cross-validation discrimination and calibration statistics, calibration curve and risk decile plot, and assessment of adequacy in the context of norms for risk models. Risk stratification: Provide quantitative assessment of relationship of risk factors to the outcome and differences in outcomes among the strata*):

Not applicable.

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment: Not applicable. (Not an outcome or resource measure.)

2b5. Identification of Meaningful Differences in Performance. (*The performance measure scores were appropriately analyzed and discriminated meaningful differences in quality.*)

2b5.1 Data/Sample (*Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*):

As previously described, KCQA tested its ESRD measures through a year-long prospective cohort study on a nationally drawn sample of 53 dialysis facilities containing a mix of for-profit and not-for-profit providers; hospital-affiliated and freestanding facilities within large, small, and independent dialysis organizations; urban, suburban, and rural settings; and facilities both with and without electronic health records (EHRs). Approximately 25 patients per facility were sought, resulting in a final sample size of 1,115 patients. Both facility and patient samples were structured to be generally representative of the national industry profile as identified by the USRDS 2007 Annual Data Report. Facility records were used as the data source, given CMS inclusion of the measure in its Phase III Clinical Performance Measures, for which it intends to involve CROWNWeb electronic transmission of data from facility medical records. Because CROWNWeb was not operational at the time, standardized, paper-based data collection sheets were constructed from the endorsed specifications and were employed during testing.

2b5.2 Analytic Method (*Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance*):

The data elements necessary for measure calculation were collected from the 53 participating facilities on the 1,115 ESRD patients in the study sample. Over the course of the study, 924 of the 1,115 patients received education about at least one modality option. Successful performance on the measure, however, requires documentation that the patient was educated in all the options. The data elements collected thus permit calculation of performance for the measure as follows:

Performance Rate =

$$\frac{[\text{Patients educated on all modalities}] - [\text{Patients educated but without documentation}]}{\text{Total ESRD patients}}$$

2b5.3 Results (*Provide measure performance results/scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance*):

Over the course of the study, 924 of the 1,115 patients received education about at least one modality option. Successful performance on the measure, however, requires documentation that the patient was educated in all the options. The following is a summary of the number of individuals receiving education on each modality:

- Hemodialysis = 850 patients (64.0%)
- Peritoneal Dialysis = 623 (67.4%)
- Home Hemodialysis = 572 (61.9%)

- Transplants = 850 (92.0%)
- Identification of Living Donors = 266 (28.8%)
- No or Cessation of Therapy = 237 (25.7%)

The measure specifications require documentation in the medical record that the patient was educated on ALL modalities. Medical record documentation existed for 922 of the 924 educated patients.

The data elements collected thus permit calculation of performance for the measure as follows:

Performance Rate =
 ([Patients educated on all modalities] – [Patients educated but without documentation]) ÷ Total ESRD patients
 = (185-2) / 1,115 = 16.4%

The performance for each individual facility in the pilot ranged from 0% to 100%, with a mean performance of 16.4%. These findings indicate that while the majority of dialysis providers are educating their patients on some renal replacement therapy modality options, few are discussing all options with their patients. The results identify an important gap and meaningful differences in patient care.

2b6. Comparability of Multiple Data Sources/Methods. (If specified for more than one data source, the various approaches result in comparable scores.)

2b6.1 Data/Sample (Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):
 Not applicable.

2b6.2 Analytic Method (Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure):
 Not applicable.

2b6.3 Testing Results (Provide statistical results, e.g., correlation statistics, comparison of rankings; assessment of adequacy in the context of norms for the test conducted):
 Not applicable.

2c. Disparities in Care: H ☐ M ☐ L ☐ I ☐ NA ☐ (If applicable, the measure specifications allow identification of disparities.)

2c.1 If measure is stratified for disparities, provide stratified results (Scores by stratified categories/cohorts): Not applicable.

2c.2 If disparities have been reported/identified (e.g., in 1b), but measure is not specified to detect disparities, please explain:
 The measure could be reported in a stratified manner to monitor the disparities in patient education by race/ethnicity, socioeconomic status, and gender.

2.1-2.3 Supplemental Testing Methodology Information:
 Attachment
[tbKCQAAttachmentAPtEdClinician06-07-11FINAL.pdf](#)

Steering Committee: Overall, was the criterion, Scientific Acceptability of Measure Properties, met? (Reliability and Validity must be rated moderate or high) Yes ☐ No ☐
 Provide rationale based on specific subcriteria:

If the Committee votes No, STOP

3. USABILITY

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (**evaluation criteria**)

C.1 Intended Actual/Planned Use (Check all the planned uses for which the measure is intended): [Public Reporting](#), [Quality Improvement \(external benchmarking to organizations\)](#), [Quality Improvement \(Internal to the specific organization\)](#), [Regulatory and Accreditation Programs](#)

3.1 Current Use (Check all that apply; for any that are checked, provide the specific program information in the following questions): [Quality Improvement \(Internal to the specific organization\)](#)

3a. Usefulness for Public Reporting: **H● M● L● I●**

(The measure is meaningful, understandable and useful for public reporting.)

3a.1. Use in Public Reporting - disclosure of performance results to the public at large (If used in a public reporting program, provide name of program(s), locations, Web page URL(s)). If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: **[For Maintenance** – If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be considered.]

The measure is intended to be used by CMS for its public reporting and payment initiatives once CMS brings CROWNWeb fully online.

3a.2. Provide a rationale for why the measure performance results are meaningful, understandable, and useful for public reporting. If usefulness was demonstrated (e.g., focus group, cognitive testing), describe the data, method, and results: While measure results have not been tested for interpretability in public reporting, the Kidney Care Partners' dialysis patient group members support the measure and concur that the availability of performance data on this measure is an important indicator of quality of care and that the measure is readily interpreted by dialysis patients.

3.2 Use for other Accountability Functions (payment, certification, accreditation). If used in a public accountability program, provide name of program(s), locations, Web page URL(s): The measure is intended to be used by CMS for its public reporting and payment initiatives.

3b. Usefulness for Quality Improvement: **H● M● L● I●**

(The measure is meaningful, understandable and useful for quality improvement.)

3b.1. Use in QI. If used in quality improvement program, provide name of program(s), locations, Web page URL(s):

[For Maintenance – If not used for QI, indicate the reasons and describe progress toward using performance results for improvement].

The measure is intended to be used by CMS for its public reporting and payment initiatives, and data will be collected via the CROWNWeb data repository. The ESRD Conditions for Coverage (section §494.180 [h]) state that data collected through CROWNWeb are to be used in a national ESRD information system and in compilations relevant to performance assessment and quality improvement.

3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement. If usefulness was demonstrated (e.g., QI initiative), describe the data, method and results:

On-site interviews of participating facility personnel were conducted during the data integrity audits. Both facility management and the staff responsible for collecting and entering the necessary data elements agreed that the measure is an important indicator of quality that will be useful for quality improvement.

Overall, to what extent was the criterion, *Usability*, met? H● M● L● I●
Provide rationale based on specific subcriteria:

4. FEASIBILITY

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (**evaluation criteria**)

4a. Data Generated as a Byproduct of Care Processes: H● M● L● I●

4a.1-2 How are the data elements needed to compute measure scores generated? (*Check all that apply*).

Data used in the measure are:

generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition, Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

4b. Electronic Sources: H● M● L● I●

4b.1 Are the data elements needed for the measure as specified available electronically (*Elements that are needed to compute measure scores are in defined, computer-readable fields*): Some data elements are in electronic sources

4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources:

4c. Susceptibility to Inaccuracies, Errors, or Unintended Consequences: H● M● L● I●

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:

Percent agreement between the auditor and facility abstractors (i.e., the reliability percentage) was assessed during the on-site audits through a direct comparison of data submitted by the facilities throughout the pilot to data reabstracted by the auditor(s). (See Table 3 [Measure Reliability Percentage and Error Type] in Attachment A.) As previously noted, this marker of accuracy was found to be 71.2%. However, we did not believe that the low inter-rater concordance was due to unreliability of the measure specifications/tool per se for the following reasons:

- First, if the reliability issues were inherent to the measure/tool itself, one would have expected the types of errors to be evenly distributed, which they were not: There were significantly more false negatives (Y/N—see Table 2) that resulted from facilities failing to count education that was present in their records—this issue can be remedied by improving the instructions to and education of personnel as to what constitutes education for purposes of meeting the specifications.

- Second, the errors for the patient education measure were largely of one type (Table 3 [Measure Reliability Percentage and Error Type] in Attachment A): “information missed,” meaning the auditors found the information in the facility’s documentation, whereas the facility personnel failed to account for it. Based on interviews with personnel during the audit, while a few cases were true instances where information was overlooked, in the majority of instances the discordance arose because the facility personnel did not realize that discussions about advance directives, for example, constituted education related to end-of-life/cessation of therapy for purposes of the measure specifications.

- Finally, using the same data collection tool, “almost perfect,” reliability (Kappa 0.8474) was observed during testing in physician offices, indicating that the high inter-rater reliability rates observed for this

measure are reproducible when used on these records and are fundamental to the specifications/tool itself.

4d. Data Collection Strategy/Implementation: H ☐ M ☒ L ☐ I ☐

A.2 Please check if either of the following apply (regarding proprietary measures):

4d.1 Describe what you have learned/modified 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 (e.g., fees for use of proprietary measures):

Despite the identified inter-rater discordance that has been discussed, we believe that the measure can be reliably collected:

- Improving the instructions and educating facilities to recognize what constitutes meeting the specifications should reduce the high numbers of false negatives. When reduction scenarios of the high false-positive rate are analyzed, the Kappas indicate excellent agreement and reliability (Table 5).
- Ongoing implementation of the new, more Comprehensive Conditions for Coverage—which require the education encompassed by the KCQA measure specifications—will improve reliability by sensitizing facility personnel to systematize and organize their processes and recordkeeping so as to enable them to more reliably collect the data elements.
- Implementation of CROWNWeb and accountability for patient education can improve reliability by deploying detailed instructions and training and by sensitizing facility personnel to systematize and organize their processes and recordkeeping so as to enable them to more reliably collect the data elements.
- Excellent reliability was observed during testing of the same data collection instrument in physician offices, indicating that the high inter-rater reliability rates observed for this measure are reproducible when used on these records and are fundamental to the specifications/tool itself.

Overall, to what extent was the criterion, *Feasibility*, met? H ☐ M ☒ L ☐ I ☐

Provide rationale based on specific subcriteria:

OVERALL SUITABILITY FOR ENDORSEMENT

Does the measure meet all the NQF criteria for endorsement? Yes ☒ No ☐

Rationale:

If the Committee votes No, STOP.

If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.

5. COMPARISON TO RELATED AND 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 before a final recommendation is made.

5.1 If there are related measures (either same measure focus or target population) or competing measures (both the same measure focus and same target population), list the NQF # and title of all related and/or competing measures:

5a. Harmonization

5a.1 If this measure has EITHER the same measure focus OR the same target population as [NQF-endorsed measure\(s\)](#): Are the measure specifications completely harmonized?

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

5b. Competing Measure(s)

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

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

[Not applicable.](#)

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner): [Kidney Care Quality Alliance, 2550 M Street, NW, Washington, District Of Columbia, 20037](#)

Co.2 Point of Contact: [Lisa, McGonigal, MD, MPPH, lmcgon@msn.com, 203-298-0567-](#)

Co.3 Measure Developer if different from Measure Steward: [Kidney Care Quality Alliance, 2550 M Street, NW, Washington, District Of Columbia, 20037](#)

Co.4 Point of Contact: [Lisa, McGonigal, MD, MPPH, lmcgon@msn.com, 203-298-0567-](#)

Co.5 Submitter: [Lisa, McGonigal, MD, MPPH, lmcgon@msn.com, 203-298-0567-, Kidney Care Quality Alliance](#)

Co.6 Additional organizations that sponsored/participated in measure development:
[Not applicable.](#)

Co.7 Public Contact: [Lisa, McGonigal, MD, MPPH, lmcgon@msn.com, 203-298-0567-, Kidney Care Quality Alliance](#)

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development

Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

[KCQA Clinical Measures Work Group Members \(developed measures\):](#)

- [1. William Haley, MD — Mayo Clinic](#)
- [2. John Burkart, MD — GatesMcDonald Health Plus](#)
- [3. Al Collins, MD — University of Minnesota](#)
- [4. Charlie McAllister, MD — DaVita, Inc.](#)
- [5. Jerry Yee, MD — Henry Ford Hospital](#)

[KCQA Clinical Measures Task Group Members \(approved measures\):](#)

- [1. Charlie McAllister, MD—DaVita, Inc.](#)
- [2. Raymond M. Hakim, MD, PhD — Fresenius Medical Care](#)
- [3. Alan Kliger — Yale University](#)
- [4. Ed Jones — Renal Physicians Association](#)

5. Allen Nissenson — DaVita, Inc.
6. William Goodman — Amgen, Inc.
7. William Haley, MD — Mayo Clinic
8. Robert Provenzano — DaVita, Inc.
9. Gail Wick — American Nephrology Nurses Association
10. Rulan Parekh — American Kidney Fund

Kidney Care Quality Alliance Steering Committee Members (oversaw testing):

- Raymond M. Hakim, MD, PhD (Co-Chair) — Fresenius Medical Care
- Gail S. Wick, BSN, RN, CNN (Co-Chair) — American Nephrology Nurses Association
- Dolph Chianchiano, JD — National Kidney Foundation
- Richard S. Goldman, MD — Renal Physicians Association
- Barbara Fivush, MD — American Society of Pediatric Nephrology
- Maureen Michael, BSN, MBA — National Renal Administrators Association
- Allen Nissenson, MD — DaVita
- Barry M. Straube, MD — Centers for Medicare and Medicaid Services (Liaison Member)

Ad.2 If adapted, provide title of original measure, NQF # if endorsed, and measure steward. Briefly describe the reasons for adapting the original measure and any work with the original measure steward: Not applicable.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.3 Year the measure was first released: 2007

Ad.4 Month and Year of most recent revision: 02, 2010

Ad.5 What is your frequency for review/update of this measure? As needed with changes or additions to the evidence base, but at minimum every three years.

Ad.6 When is the next scheduled review/update for this measure? 02, 2013

Ad.7 Copyright statement: © 2010 Kidney Care Quality Alliance. All Rights Reserved.

Ad.8 Disclaimers:

Ad.9 Additional Information/Comments: <http://www.kidneycarepartners.com>

Date of Submission (MM/DD/YY): 10/05/2011