NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized): Complete all <mark>yellow highlighted</mark> areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: If there is no TAP or workgroup, the SC also evaluates the subcriteria (<mark>yellow</mark> highlighted areas).

Steering Committee: Complete all <mark>pink</mark> highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 1430   **NQF Project:** End Stage Renal Disease
MEASURE DESCRIPTIVE INFORMATION
De.1 Measure Title: Lower Limit of Hemoglobin for Pediatric Patients
De.2 Brief description of measure: Percentage of pediatric (<18 years old) hemodialysis and peritoneal dialysis patients, with ESRD >=3 months, who have a mean hemoglobin <10 g/dL for a 3 month reporting period, irrespective of ESA use. The hemoglobin value reported at the end of each reporting month (end-of-month hemoglobin) is used for the calculation.
1.1-2 Type of Measure: Outcome **De.3** If included in a composite or paired with another measure, please identify composite or paired measure This measure is paired with the proposed Pediatric Anemia Hemoglobin Process Measure.
De.4 National Priority Partners Priority Area: Population health **De.5** IOM Quality Domain: Effectiveness **De.6** Consumer Care Need: Living with illness

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* **A.1** Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)? Yes **A.2 Indicate if** Proprietary Measure (*as defined in measure steward agreement*): **A.3** Measure Steward Agreement: Government entity and in the public domain - no agreement necessary **A.4** Measure Steward Agreement attached:	A Y ☐ N ☐

B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	**B** Y ☐ N ☐
C. The intended use of the measure includes <u>both</u> public reporting <u>and</u> quality improvement. ▶**Purpose:** Public reporting, Internal quality improvement	**C** Y ☐ N ☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. **D.1Testing:** Yes, fully developed and tested **D.2** Have NQF-endorsed measures been reviewed to identify if there are similar or related measures? Yes	**D** Y ☐ N ☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*):	**Met** Y ☐ N ☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*):	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. *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**	**Eval Ratin g**
(for NQF staff use) Specific NPP goal:	
1a.1 Demonstrated High Impact Aspect of Healthcare: Frequently performed procedure, High resource use, Severity of illness, Patient/societal consequences of poor quality **1a.2** **1a.3 Summary of Evidence of High Impact:** Kidney disease results in a deficiency of erythropoietin, a hormone which stimulates the production of red blood cells, leading to the development of anemia. Recent studies suggest that among Chronic Kidney Disease (CKD) pediatric patients, anemia is associated with adverse outcomes including increased mortality risk and hospitalizations [1-3]. Staples et al analyzed stage II-V predialysis CKD patients and found that anemic children, defined as hematocrit<33%, were 55% more likely to be hospitalized compared to non-anemic children. Warady and Ho studied pediatric hemodialysis and peritoneal dialysis patients at the initiation of dialysis and showed that 68% of patients were anemic (hematocrit<33%), and that anemia was associated with a 55% increase in mortality risk. Mortality and hospitalization rates among adolescent hemodialysis patients were assessed in the Amaral et al study, and an increased risk of mortality with lower hemoglobin levels was observed. The mortality risk among adolescent hemodialysis patients with Hb 11-12 g/dL was 70% lower compared to patients with Hb<10 g/dL. These studies therefore suggest that a hematocrit<33% (approximately equal to Hb<10 g/dl) is associated with adverse outcomes. Furthermore, because the normal range of hemoglobin levels varies in the pediatric population according to age group and gender, the definition of anemia ideally should be age- and gender-dependent, with	**1a** C ☐ P ☐ M ☐ N ☐

Comment [KP1]: 1a. The measure focus addresses:
●a specific national health goal/priority identified by NQF's National Priorities Partners; OR
●a demonstrated high impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).

anemia being defined as below 5th percentile for hemoglobin levels in each age and gender category from the Third National Health and Nutrition Examination Survey (NHANES III) study [4]. This is consistent with the definition of anemia in the pediatric population based on the Kidney Disease Outcomes Quality Initiative (KDOQI) guidelines [5]. However, having an age and gender-specific hemoglobin target will be difficult to implement and identification of a single lower bound limit is appropriate. Across the NHANES III categories, cut-off levels for the definition of anemia are all above 10 g/dl. Thus, using a cut-off of 10 g/dl is a feasible and achievable target regardless of the pediatric age and gender category.

1a.4 Citations for Evidence of High Impact: 1. Warady B Ho M. Morbidity and mortality in children with anemia at initiation of dialysis. Pediatr Nephrol 18:1055-1062, 2003.

2. Amaral S, Hwang W, Fivush B, Neu A, Frankenfield D, Furth S. Association of mortality and hospitalization with achievement of adult hemoglobin targets in adolescents maintained on hemodialysis. J Am Soc Nephrol 17:2878-85, 2006.

3. Staples AO, Wong CS, Smith JM, Gipson DS, Filler G, Warady BA, Martz K, Greenbaum LA. Anemia and risk of hospitalization in pediatric chronic kidney disease. Clin J Am Soc Nephrol. Jan; 4(1):48-56, 2009.

4. Hollowell JG, van Assendelft OW, Gunter EW, Lewis BG, Najjar M, Pfeiffer C. Hematological and iron-related analytes—Reference data for persons aged 1 year and over: United States, 1988-94. Vital Health Stat 11:1-156, 2005

5. Clinical Practice Recommendation for Pediatrics 1.1: Identifying Patients and Initiating Evaluation; KDOQI Clinical Practice Guideline and Clinical Practice Recommendations for Anemia in Chronic Kidney Disease: 2007.

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: Recent studies have shown that pediatric patients with hemoglobin<10 g/dL had higher hospitalization and mortality rates. Furthermore, given that hospitalization is associated with increased costs of care, it is possible that efforts to reduce the prevalence of anemia in the pediatric ESRD population, which may potentially reduce hospitalization rates, may be cost-effective.

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
Analyses of ESRD Clinical Performance Measures (CPM) data from October through December 2007 showed that 19% of pediatric patients on hemodialysis had a mean hemoglobin<10.0 g/dL, and 14% of pediatric peritoneal dialysis patients had mean hemoglobin<10 g/dL.

1b.3 Citations for data on performance gap:
Internal analysis of CPM data.

1b.4 Summary of Data on disparities by population group:
In an analysis of CPM data examining anemia management in the pediatric ESRD, gender, race and ethnicity were not found to be predictive of anemia [1]. However, the United States Renal Data System (USRDS) 2008 Annual Data Report (ADR) demonstrated that hemoglobin levels tended to vary by age group, and that hemoglobin levels are 0.5g/dL higher in white children as compared to African American children and other races [2].

1b.5 Citations for data on Disparities:
1. Frankenfield DL, Neu AM, Warady BA, et al. Anemia in pediatric hemodialysis patients: results from the 2001 Clinical Performance Measures Project. Kidney International 64:1120-4, 2003.
2. U.S. Renal Data System, USRDS 2008 Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2008.

1b	
C	☐
P	☐
M	☐
N	☐

1c. Outcome or Evidence to Support Measure Focus

1c	
C	☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

3

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*): Although large scale clinical trials have not been conducted in the pediatric population, smaller scale observational and cohort studies have shown an association between anemia and poor outcomes including poor quality of life, cardiovascular disease, morbidity, and mortality. Additionally, the 2007 KDOQI Clinical Practice Recommendation (CPR) hemoglobin target for pediatrics suggests that hemoglobin be in the range of 11 to 12 g/dl (CPR 2.1.2). In considering the lower bound for hemoglobin levels (10 vs 11gm/dl), the following were taken into account: 1) age-specific hemoglobin cut-off levels for the general pediatric population and 2) the National Quality Forum (NQF) endorsed adult measure for hemoglobin targets. Thus, although the KDOQI guidelines recommend a lower bound of 11g/dl, the proposed measure cut-off is 10gm/dl. This is described further below.

P ☐
M ☐
N ☐

1c.2-3. Type of Evidence: Cohort study, Observational study, Evidence-based guideline

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*):
The presence of anemia in the pediatric population has been associated with increased morbidity and mortality [1,2]. Lower hemoglobin levels have also been associated with cardiovascular disease [3] and quality of life [4]. Staples et al analyzed stage II-V predialysis CKD patients and found that anemic children, defined as hematocrit <33%, were 55% more likely to be hospitalized compared to non-anemic children [5]. Warady and Ho studied pediatric hemodialysis and peritoneal dialysis patients at the initiation of dialysis and showed that 68% of patients were anemic (hematocrit <33%), and that anemia was associated with a 55% increase in mortality risk. Mortality and hospitalization rates among adolescent hemodialysis patients were assessed in the Amaral et al study, and an increased risk of mortality with lower hemoglobin levels was observed. The mortality risk among adolescent hemodialysis patients with Hgb 11-12 g/dL was 70% lower compared to patients with Hgb<10 g/dL. These studies therefore suggest that a hematocrit<33% (approximately equal to Hgb<10 g/dl) is associated with adverse outcomes.
The 2007 KDOQI Update of Hemoglobin Target CPR 2.1.2 recommends that in pediatric dialysis and nondialysis patients with CKD receiving ESA therapy, the selected hemoglobin target should be in the range of 11.0 to 12.0 g/dL. The NQF endorsed adult anemia measure a uses a hemoglobin lower bound of 10 g/dl. There is no evidence to approach anemia differently in the pediatric population, thus the lower bound was set at 10. Furthermore, because the normal range of hemoglobin levels varies in the pediatric population according to age group and gender, the definition of anemia ideally should be age- and gender-dependent, with anemia being defined as below 5th percentile for hemoglobin levels in each age and gender category from the NHANES III study (Hollowell, 2005). This is consistent with the definition of anemia in the pediatric population based on the KDOQI guidelines (CPR for Pediatrics 1.1; 2007). However, having an age and gender-specific hemoglobin target will be difficult to implement and identification of a single lower bound limit is appropriate. Across the NHANES III categories, cut-off levels for the definition of anemia are all above 10 g/dl. Thus, using a cut-off of 10 g/dl is a feasible and achievable target regardless of the pediatric age and gender category.

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom*):
The pediatric adequacy clinical Technical Expert Panel (TEP) rated the strength of this measure as high.

1c.6 Method for rating evidence: The clinical TEP followed similar methods of evidence assessment as that used by the KDOQI clinical practice guidelines.

1c.7 Summary of Controversy/Contradictory Evidence: The clinical TEP discussed using a cut-off of 11gm/dl vs 10 gm/dl. For the arguments cited above, the clinical TEP concluded that a cut-off of 10gm/dl was appropriate.
The panel discussed whether an upper limit for hemoglobin targets should be developed as a measure. However, there is insufficient evidence to set an upper bound for hemoglobin targets in the pediatric population and a measure was therefore not developed.

1c.8 Citations for Evidence (*other than guidelines*): 1. Warady B Ho M. Morbidity and mortality in children with anemia at initiation of dialysis. Pediatr Nephrol 18:1055-1062, 2003.

Comment [k5]: 4 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 multi-step process, the step with the greatest effect on the desired outcome should be selected as the focus of measurement. For example, although assessment of immunization status and recommending immunization are necessary steps, they are not sufficient to achieve the desired impact on health status – patients must be vaccinated to achieve immunity. This does not preclude consideration of measures of preventive screening interventions where there is a strong link with desired outcomes (e.g., mammography) or measures for multiple care processes that affect a single outcome.

Comment [k6]: 3 The strength of the body of evidence for the specific measure focus should be systematically assessed and rated (e.g., USPSTF grading system http://www.ahrq.gov/clinic/uspstf07/methods/benefit.htm). If the USPSTF grading system was not used, the grading system is explained including how it relates to the USPSTF grades or why it does not. However, evidence is not limited to quantitative studies and the best type of evidence depends upon the question being studied (e.g., randomized controlled trials appropriate for studying drug efficacy are not well suited for complex system changes). When qualitative studies are used, appropriate qualitative research criteria are used to judge the strength of the evidence.

2. Amaral S, Hwang W, Fivush B, Neu A, Frankenfield D, Furth S. Association of mortality and hospitalization with achievement of adult hemoglobin targets in adolescents maintained on hemodialysis. J Am Soc Nephrol 17:2878-85, 2006.

3. Mitsnefes MM, Daniels SR, Schwartz SM, Meyer RA, Khoury P, Strife CF. Severe left ventricular hypertrophy in pediatric dialysis: Prevalence and predictors. Pediatr Nephrol 14:898-902, 2000.

4. Gerson A, et al. Anemia and health-related quality of life in adolescents with chronic kidney disease. Am J Kidney Dis 44:1017-1023, 2004.

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number*):
KDOQI Clinical Practice Guideline and Clinical Practice Recommendations for Anemia in Chronic Kidney Disease: 2007 Update of Hemoglobin Target
CPR 2.1.2 (FULLY APPLICABLE TO CHILDREN) In the opinion of the Work Group, in pediatric dialysis and nondialysis patients with CKD receiving ESA therapy, the selected Hb target should generally be in the range of 11.0 to 12.0 g/dL. (Clinical Practice RECOMMENDATION)
Please note that these are clinical practice recommendations and are therefore defined in the KDOQI document as "expert opinion" based recommendations.

1c.10 Clinical Practice Guideline Citation: Clinical Practice Recommendation for Pediatrics 2.1: Hemoglobin Target; KDOQI Clinical Practice Guideline and Clinical Practice Recommendations for Anemia in Chronic Kidney Disease: 2007.
1c.11 National Guideline Clearinghouse or other URL: N/A

1c.12 Rating of strength of recommendation (*also provide narrative description of the rating and by whom*):
The strength of this recommendation was not graded, but was based on Work Group consensus.

1c.13 Method for rating strength of recommendation (*If different from* USPSTF system, *also describe rating and how it relates to USPSTF*):
The method used is the same as was used in developing the 2006 KDOQI guidelines, in which experts decided which recommendations were supported by evidence and which were supported by consensus of Work Group opinion. Evidence-based guideline recommendations were graded as strong or moderate or weak. This approach is consistent with the United States Preventive Services Task Force (USPSTF) grading method.

1c.14 Rationale for using this guideline over others:
There are no other known pediatric anemia guidelines. The KDOQI clinical practice guidelines and recommendations are widely utilized by the nephrology community.

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	1
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? Rationale:	1 Y☐ N☐

2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

Extent to which the measure, <u>as specified</u>, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	Eval Rating

2a. MEASURE SPECIFICATIONS

S.1 Do you have a web page where current detailed measure specifications can be obtained? **S.2** If yes, provide web page URL: 2a. Precisely Specified	2a-specs C☐ P☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable   5

Comment [k7]: USPSTF grading system http://www.ahrq.gov/clinic/uspstf/grades.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

Comment [KP8]: 2a. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NQF's Health Information Technology Expert Panel (HITEP) .

2a.1 Numerator Statement (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome*):
Number of pediatric (<18 years old) hemodialysis and peritoneal dialysis patients, with End Stage Renal Disease (ESRD) >= 3 months, who have a mean hemoglobin <10.0 g/dL for a 3 month reporting period, irrespective of erythropoiesis-stimulating agent (ESA) use. The hemoglobin value reported for the end of each reporting month (end-of-month hemoglobin) is used for the calculation.

2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator*):
Three months from the start of the reporting period.

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):
The numerator will be determined by counting all patients in the denominator who have mean Hemoglobin values less than 10 g/dl.

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*):
All pediatric (<18 years old) hemodialysis and peritoneal dialysis patients with ESRD >= 3 months

2a.5 Target population gender: Female, Male
2a.6 Target population age range: Pediatric patients less than 18 years old.

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator*):
Three months from the start of the reporting period.

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):
Patients are counted as being in the facility for the entire calendar month if "Admit Date" to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged ("Discharge Date" is null or blank), OR "Discharge Date" from the facility is greater than or equal to the last day of the study period. The patient's age will be determined by subtracting the patient's date of birth from the first day of the reporting month, and the length of dialysis will be calculated as the difference between the first "Hemoglobin Collection Date" and "Date Regular Chronic Dialysis Began". Patients will be included in the denominator if they are in the facility for the entire calendar month and their age is less than 18, length of dialysis is greater than or equal to 90 days, AND both "Hemoglobin" and "Hemoglobin Collection Date" are recorded in each of the 3 reporting months.

2a.9 Denominator Exclusions (*Brief text description of exclusions from the target population*): Patients on dialysis <3 months at the start of the reporting period, patients who are not in the facility for the entire one-month study period.

Comment [k9]: 11 Risk factors that influence outcomes should not be specified as exclusions.
12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

2a.10 Denominator Exclusion Details (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions*):
See denominator exclusions.

2a.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions*):
No stratification is required for this measure.

2a.12-13 Risk Adjustment Type: No risk adjustment necessary

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*):
N/A

2a.15-17 Detailed risk model available Web page URL or attachment:

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Lower score

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps*):
Patients are counted as being in the facility for the entire calendar month if "Admit Date" to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged ("Discharge Date" is null or blank), OR "Discharge Date" from the facility is greater than or equal to the last day of the study period. The patient's age will be determined by subtracting the patient's date of birth from the first day of the reporting month, and the length of dialysis will be calculated as the difference between the first "Hemoglobin Collection Date" and "Date Regular Chronic Dialysis Began". Patients will be included in the denominator if they are in the facility for the entire calendar month and their age is less than 18, length of dialysis is greater than or equal to 90 days, AND both "Hemoglobin" and "Hemoglobin Collection Date" are recorded in each of the 3 reporting months. The numerator will be determined by counting all patients in the denominator who have mean Hemoglobin values less than 10 g/dl.

2a.22 Describe the method for discriminating performance (*e.g., significance testing*):
The performance of the facility will be compared to state, Network and national performance.

2a.23 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)*:
N/A

2a.24 Data Source (*Check the source(s) for which the measure is specified and tested*)
Electronic clinical data

2a.25 Data source/data collection instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*):
CROWNWeb (Consolidated Renal Operations in a Web Enabled Network)

2a.26-28 Data source/data collection instrument reference web page URL or attachment: URL
http://www.projectcrownweb.org/crown/index.php

2a.29-31 Data dictionary/code table web page URL or attachment: URL
http://www.projectcrownweb.org/crown/index.php?page=Public_Documents&subPage=Release_Document s

2a.32-35 Level of Measurement/Analysis (*Check the level(s) for which the measure is specified and tested*)
Facility/Agency

2a.36-37 Care Settings (*Check the setting(s) for which the measure is specified and tested*)
Dialysis Facility

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)
Dialysis

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample (*description of data/sample and size*): For the 2008 ESRD CPM project, inter-rater reliability was assessed using facility abstracted and Network re-abstracted data. A total of 301 randomly selected medical records from both adult and pediatric patients were included in the analysis.

2b.2 Analytic Method (*type of reliability & rationale, method for testing*):
To analyze the inter-rater reliability of the ESRD CPM data agreement rates, levels of concurrence, and kappa statistics were computed. Agreement rates were calculated for continuous data, and kappa statistics and levels of concurrence were jointly used to analyze categorical data.

2b.3 Testing Results (*reliability statistics, assessment of adequacy in the context of norms for the test conducted*):
For Hemoglobin>=9, the average kappa statistic (of October, November, and December) for non-missing data was 0.97 and for missing vs. non-missing was 0.80. The average level of concurrence (LOC) for missing

2b
C☐
P☐
M☐
N☐

Comment [KP10]: 2b. Reliability testing demonstrates the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.

Comment [k11]: 8 Examples of reliability testing 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 may address the data items or final measure score.

vs. non-missing was 98%, and for non-missing data only was 100%. Generally, acceptable agreement rates are 0.80 or higher and concurrence targets are 90% or higher.

2c. Validity testing

2c.1 Data/sample *(description of data/sample and size)*: This measure was established on the basis of face validity. All clinical TEP members agreed that this measure will improve quality of care for pediatric hemodialysis and peritoneal dialysis patients.

2c.2 Analytic Method *(type of validity & rationale, method for testing)*:
Face validity is the only validity assessed, as there is no gold standard for defining minimum target hemoglobin values in pediatric dialysis patients.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
N/A

2c
C ☐
P ☐
M ☐
N ☐

2d. Exclusions Justified

2d.1 Summary of Evidence supporting exclusion(s):
Exclusions are not supported by evidence. However, they are limited to those with a compelling clinical rationale and are precisely defined.

2d.2 Citations for Evidence:
N/A

2d.3 Data/sample *(description of data/sample and size)*: N/A

2d.4 Analytic Method *(type analysis & rationale)*:
N/A

2d.5 Testing Results *(e.g., frequency, variability, sensitivity analyses)*:
N/A

2d
C ☐
P ☐
M ☐
N ☐
NA ☐

2e. Risk Adjustment for Outcomes/ Resource Use Measures

2e.1 Data/sample *(description of data/sample and size)*: No risk adjustments are necessary for this measure.

2e.2 Analytic Method *(type of risk adjustment, analysis, & rationale)*:
N/A

2e.3 Testing Results *(risk model performance metrics)*:
N/A

2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: N/A

2e
C ☐
P ☐
M ☐
N ☐
NA ☐

2f. Identification of Meaningful Differences in Performance

2f.1 Data/sample from Testing or Current Use *(description of data/sample and size)*: Data from the ESRD CPM Project were used to perform analyses on determining differences in performance in facilities with pediatric hemodialysis and peritoneal dialysis patients. In the 2008 study, CPM data were collected on all pediatric hemodialysis patients from October 2007 through December 2007 from a total of 317 facilities.

2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance *(type of analysis & rationale)*:
Facility level performance was evaluated by the calculation of facility percentages.

2f.3 Provide Measure Scores from Testing or Current Use *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in*

2f
C ☐
P ☐
M ☐
N ☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

Comment [KP12]: 2c. Validity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.

Comment [k13]: 9 Examples of validity testing include, but are not limited to: determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores on some other related valid measure; content validity for multi-item scales/tests. Face validity is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with BP < 140/90 is a marker of quality). If face validity is the only validity addressed, it is systematically assessed (e.g., ratings by relevant stakeholders) and the measure is judged to represent quality care for the specific topic and that the measure focus is the most important aspect of quality for the specific topic.

Comment [KP14]: 2d. Clinically necessary measure exclusions are identified and must be:
•supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;
AND ... [1]

Comment [k15]: 10 Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.

Comment [KP16]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:
•an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outc... [2]

Comment [k17]: 13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and wo... [3]

Comment [KP18]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

Comment [k19]: 14 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 cessatio... [4]

performance):
 CPM facility level data indicated that 29 (9%) of the 317 facilities reported that all of their pediatric dialysis patients were below the minimum hemoglobin level of 10 g/dL.

2g. Comparability of Multiple Data Sources/Methods **2g.1 Data/sample** *(description of data/sample and size)*: CROWNWeb. Phase 1 and 2 CROWNWeb Beta Testing Data: Data are based on the 18 facilities participating in Phase 1 and the 180 facilities participating in Phase 2 plus about 3000 additional batch-submission facilities in CROWNWeb. These data include about 60% of dialysis facilities and 75% of dialysis patients and are heavily weighted towards large dialysis organization facilities. **2g.2 Analytic Method** *(type of analysis & rationale)*: Multiple data sources are not allowed for this measure, and therefore testing is not needed. **2g.3 Testing Results** *(e.g., correlation statistics, comparison of rankings)*: N/A	2g C ☐ P ☐ M ☐ N ☐ NA ☐
2h. Disparities in Care **2h.1 If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: This measure is not stratified. **2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** Stratification by age group, race, and gender may be considered since NHANES suggests that age-specific hemoglobin targets vary in the general non-ESRD pediatric population.	2h C ☐ P ☐ M ☐ N ☐ NA ☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties*?	2
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, **met?** **Rationale:**	*2* C ☐ P ☐ M ☐ N ☐

Comment [KP20]: 2g. If multiple data sources/methods are allowed, there is demonstration they produce comparable results.

Comment [KP21]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender);OR rationale/data justifies why stratification is not necessary or not feasible.

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)	Eval Rating
3a. Meaningful, Understandable, and Useful Information **3a.1 Current Use:** Not in use but testing completed **3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*): The 2008 ESRD Clinical Performance Measures Project presents data on the adult and pediatric ESRD population. URL: www.cms.hhs.gov/CPMProject. **3a.3 If used in other programs/initiatives** (*If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years*): The North American Pediatric Renal Transplant Cooperative Study (NAPRTCS) is a voluntary initiative of pediatric ESRD treatment facilities to report transplantation and dialysis outcomes. An analysis of patients <18 years of age in the NAPRTCS found that 68% of patients were anemic. Additionally, results showed anemia was associated with a 52% higher risk of death (Warady B Ho M. Morbidity and mortality in children with anemia at initiation of dialysis. Pediatr Nephrol 18:1055-1062, 2003).	3a C ☐ P ☐ M ☐ N ☐

Comment [KP22]: 3a. Demonstration that information produced by the measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives). An important outcome that may not have an identified improvement strategy still can be useful for informing quality improvement by identifying the need for and stimulating new approaches to improvement.

Testing of Interpretability (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*)
3a.4 Data/sample (*description of data/sample and size*): Testing of interpretability has not been performed.

3a.5 Methods (*e.g., focus group, survey, QI project*):
N/A

3a.6 Results (*qualitative and/or quantitative results and conclusions*):
N/A

3b/3c. Relation to other NQF-endorsed measures

3b.1 NQF # and Title of similar or related measures:
0370 ESRD- Anemia Management CPM Ib- Monitoring Hemoglobin Levels Below Target Minimum

(for NQF staff use) Notes on similar/related endorsed or submitted measures:

3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): **3b.2** Are the measure specifications harmonized? If not, why? The measure specifications are harmonized with the adult NQF endorsed measure.	**3b** C ☐ P ☐ M ☐ N ☐ NA ☐
3c. Distinctive or Additive Value **3c.1** Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: This measure is for pediatric (<18 years) patients only. The NQF endorsed measure is for patients >=18 years. **5.1** If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality: This measure is for pediatric (<18 years) patients only. The NQF endorsed measure is for patients >=18 years.	**3c** C ☐ P ☐ M ☐ N ☐ NA ☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability*?	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? Rationale:	**3** C ☐ P ☐ M ☐ N ☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes **4a.1-2** How are the data elements that are needed to compute measure scores generated? Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition)	**4a** C ☐ P ☐ M ☐ N ☐
4b. Electronic Sources **4b.1** Are all the data elements available electronically? (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) Yes	**4b** C ☐ P ☐ M ☐ N ☐

Comment [KP23]: 3b. The measure specifications are harmonized with other measures, and are applicable to multiple levels and settings.

Comment [k24]: 16 Measure harmonization refers to the standardization of specifications for similar measures on the same topic (e.g., *influenza immunization* of patients in hospitals or nursing homes), or related measures for the same target population (e.g., eye exam and HbA1c for *patients with diabetes*), or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are dictated by the evidence. The dimensions of harmonization can include numerator, denominator, exclusions, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.

Comment [KP25]: 3c. Review of existing endorsed measures and measure sets demonstrates that the measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare, is a more valid or efficient way to measure).

Comment [KP26]: 4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery. (e.g., BP recorded in the electronic record, not abstracted from the record later by other personnel; patient self-assessment tools, e.g., depression scale; lab values, meds, etc.)

Comment [KP27]: 4b. The required data elements are available in electronic sources. If the required data are not in existing electronic sources, a credible, near-term path to electronic collection by most providers is specified and clinical data elements are specified for transition to the electronic health record.

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

4b.2 If not, specify the near-term path to achieve electronic capture by most providers.

4c. Exclusions

4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?
No

4c.2 If yes, provide justification.

4c	
C	☐
P	☐
M	☐
N	☐
NA	☐

> **Comment [KP28]:** 4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.

4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences

4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.
Since the data elements required for this measure are already being collected as part of the ESRD CPM project, facilities are familiar with data required for this measure. This reduces the likelihood of errors in the data collection process.

4d	
C	☐
P	☐
M	☐
N	☐

> **Comment [KP29]:** 4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.

4e. Data Collection Strategy/Implementation

4e.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/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:
Since the data elements required for this measure are already being collected as part of the ESRD CPM project, facilities are familiar with data required for this measure. This reduces the likelihood of errors in the data collection process.

4e.2 Costs to implement the measure (*costs of data collection, fees associated with proprietary measures*):
The estimated data collection burden and associated cost estimate is presented in Tables 1-3 in the Federal Register. Vol. 73, No. 73 page 20469.
URL:http://www.cms.gov/CFCsAndCoPs/downloads/ESRDfinalrule0415.pdf.

4e.3 Evidence for costs:
See above reference to Federal Register.

4e.4 Business case documentation: No formal studies evaluating the cost-effectiveness of maintaining a hemoglobin level above 10gm/dl in the pediatric ESRD population have been published. However, as described above, anemia has been associated with increased hospitalization rates in this population. Given that hospitalization is associated with increased costs of care, it is possible that efforts to reduce the prevalence of anemia in the pediatric ESRD population, which may potentially reduce hospitalization rates, may be cost-effective.

4e	
C	☐
P	☐
M	☐
N	☐

> **Comment [KP30]:** 4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility*?	4

Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? Rationale:	4
	C ☐
	P ☐
	M ☐
	N ☐

RECOMMENDATION

(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.	Time-limited ☐

Steering Committee: Do you recommend for endorsement?	Y ☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable   11

Comments:	N☐ A☐

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner)
Co.1 <u>Organization</u>
Centers for Medicare & Medicaid Services, 7500 Security Boulevard, Baltimore, Maryland, 21244

Co.2 <u>Point of Contact</u>
Thomas, Dudley, Thomas.Dudley@cms.hhs.gov, 410-786-1442-

Measure Developer If different from Measure Steward
Co.3 <u>Organization</u>
Arbor Research/UM-KECC, 315 W. Huron Street, Ann Arbor, Michigan, 48103

Co.4 <u>Point of Contact</u>
Adrienne, Janney, adrienne.janney@arborresearch.org, 734-665-4108-

Co.5 Submitter If different from Measure Steward POC
Thomas, Dudley, Thomas.Dudley@cms.hhs.gov, 410-786-1442-, Centers for Medicare & Medicaid Services

Co.6 Additional organizations that sponsored/participated in measure development

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.
Dr. Bradley Warady, panel chair (University of Missouri, Kansas City School of Medicine, Kansas City, MO)
Dr. Carolyn Abitbol (University of Miami, Holtz Children's Hospital, Miami, FL)
Dr. Eileen Brewer (Baylor College of Medicine/Texas Children's Hospital, Houston, TX)
Dr. Stuart Goldstein (Baylor College of Medicine/Texas Children's Hospital, Houston, TX)
Dr. Alicia Neu (Johns Hopkins Medical Institution, Baltimore, MD)
Dr. Irene Restaino (Children's Hospital of The King Daughters, Norfolk, VA)
Dr. Douglas Silverstein (Children´s National Medical Center, Washington, D.C.)
Dr. Sylvia Ramirez, Moderator (Arbor Research Collaborative for Health)
Alissa Kapke, Analyst (Arbor Research Collaborative for Health)
Jeffrey Pearson, Analytic Manager, (Arbor Research Collaborative for Health)

Ad.2 If adapted, provide name of original measure:
Ad.3-5 If adapted, provide original specifications URL or attachment

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 Year the measure was first released:
Ad.7 Month and Year of most recent revision:
Ad.8 What is your frequency for review/update of this measure? Three years
Ad.9 When is the next scheduled review/update for this measure? 2013

Ad.10 Copyright statement/disclaimers:

Ad.11 -13 Additional Information web page URL or attachment:

Date of Submission (*MM/DD/YY*): 09/28/2010

2d. Clinically necessary measure exclusions are identified and must be:
• supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; AND
• a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus;
 AND
• precisely defined and specified:
  − if there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);
if patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that it strongly impacts performance on the measure and 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).

2e. For outcome measures and other measures (e.g., resource use) when indicated:
• an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care;[Error! Bookmark not defined.] OR
rationale/data support no risk adjustment.

13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and women).   It is preferable to stratify measures by race and socioeconomic status rather than adjusting out differences.

14 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% v. 75%) 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 poor performance may not demonstrate much variability across providers.