NATIONAL QUALITY FORUM

Measure Information

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

Brief Measure Information

NQF #: 1454
Corresponding Measures:
De.2. **Measure Title:** Proportion of patients with hypercalcemia
Co.1.1. **Measure Steward:** Centers for Medicare & Medicaid Services
De.3. **Brief Description of Measure:** Percentage of adult dialysis patients with a 3-month rolling average of total uncorrected calcium (serum or plasma) greater than 10.2 mg/dL (hypercalcemia)
1b.1. **Developer Rationale:** The hypercalcemia measure was developed in 2010 based on the recommendations of a clinical TEP's consideration of the multiple large, risk-adjusted observational studies demonstrating a consistent relationship between presence of hypercalcemia and patient mortality. TEP members felt that while small, the population of patients with hypercalcemia was at increased risk of cardiovascular events and therefore the condition needs to be identified and appropriately treated. The TEP agreed that therapy should be focused on preventing the development of a sustained serum calcium greater than 10.2 mg/dL.

The measure was re-evaluated by a second clinical TEP in 2013. The 2013 TEP identified additional observational studies supporting the measure and affirmed their agreement with the measure's focus as a safety measure, emphasizing avoidance of hypercalcemia to prevent adverse clinical consequences. Given both the 2010 TEP and 2013 TEP recommendations, and the additional evidence cited in the current NQF submission, we maintain its importance as a clinical intermediate outcome and patient safety measure.

We acknowledge the lack of interventional trials supporting a specific threshold. However, the number of large, risk-adjusted observational studies with consistent direction of association between hypercalcemia and mortality cannot be ignored. Given this, at the time of the last review of this measure in 2015, several Renal committee reviewers agreed with the prior TEPs'opinions that the measure represented an appropriate safety-net.

As an additional concern, the Protecting Access to Medicare Act of 2014 mandated the implementation of conditions treated through oral-only medications in the ESRD Quality Incentive Program (QIP) as a safety measure against over-use of oral-only medications following changes to the ESRD PPS Bundle payment. We believe Congress recognized the need for more safety measures in the ESRD program, particularly in the area of drug overuse, following similar concerns for the use of ESAs in treating anemia in the same population. This hypercalcemia measure is the only measure of which we are aware that meets these requirements and the NQF criteria. Other relevant measures have been presented to NQF in the past, but have not received endorsement due to a lack of evidence.

S.4. **Numerator Statement:** Number of patient-months in the denominator with 3-month rolling average of total uncorrected serum (or plasma) calcium greater than 10.2 mg/dL or missing.
S.6. **Denominator Statement:** Number of patient reporting months (see S.5. above) among adult (greater than or equal to 18 years old) in-center hemodialysis, home hemodialysis, or peritoneal dialysis patients under the care of the dialysis facility for the entire reporting month who have had ESRD for greater than 90 days.
S.8. **Denominator Exclusions:** Exclusions that are implicit in the denominator definition include all patients who are <18 years old, who have not been in the facility the entire reporting month (transient patients), and patients who have had ESRD for <91 days. There are no additional exclusions for this measure.

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

IF Endorsement Maintenance – Original Endorsement Date: Aug 16, 2011 **Most Recent Endorsement Date:** Oct 02, 2015

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

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

De.4. **IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?** N/A

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

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

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form
1454_Evidence.docx
1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?
Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.
Yes

1b. Performance Gap

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

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

1b.1. **Briefly explain the rationale for this measure** *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*
If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.
The hypercalcemia measure was developed in 2010 based on the recommendations of a clinical TEP's consideration of the multiple large, risk-adjusted observational studies demonstrating a consistent relationship between presence of hypercalcemia and patient mortality. TEP members felt that while small, the population of patients with hypercalcemia was at increased risk of cardiovascular events and therefore the condition needs to be identified and appropriately treated. The TEP agreed that therapy should be focused on preventing the development of a sustained serum calcium greater than 10.2 mg/dL.

The measure was re-evaluated by a second clinical TEP in 2013. The 2013 TEP identified additional observational studies supporting the measure and affirmed their agreement with the measure's focus as a safety measure, emphasizing avoidance of hypercalcemia to prevent adverse clinical consequences. Given both the 2010 TEP and 2013 TEP recommendations, and the additional evidence cited in the current NQF submission, we maintain its importance as a clinical intermediate outcome and patient safety measure.

We acknowledge the lack of interventional trials supporting a specific threshold. However, the number of large, risk-adjusted observational studies with consistent direction of association between hypercalcemia and mortality cannot be ignored. Given this, at the time of the last review of this measure in 2015, several Renal committee reviewers agreed with the prior TEPs'opinions that the measure represented an appropriate safety-net.

As an additional concern, the Protecting Access to Medicare Act of 2014 mandated the implementation of conditions treated through oral-only medications in the ESRD Quality Incentive Program (QIP) as a safety measure against over-use of oral-only medications following changes to the ESRD PPS Bundle payment. We believe Congress recognized the need for more safety measures in the ESRD program, particularly in the area of drug overuse, following similar concerns for the use of ESAs in treating anemia in the same population. This hypercalcemia measure is the only measure of which we are aware that meets these requirements and the NQF criteria. Other relevant measures have been presented to NQF in the past, but have not received endorsement due to a lack of evidence.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*
Among the 7,097 facilities that have at least one eligible patient, we generated the following statistics of their performance scores using the January – December 2017 CROWNWeb clinical data. Percentages reflect patient-months: mean (SD)=3.3% (9.0%); min=0%; max=100 %; 25th percentile=0.5%; 50th percentile=1.5%; 75th percentile=2.9%.

During the recent DFC performance period, 1738 facilities had 0% of patients with hypercalcemia, 1825 facilities had 1% of patients with hypercalcemia, 1335 facilities had 2%, 822 had 3%, and 1377 facilities had 4% or more of their patients with hypercalcemia.

The distribution demonstrates the success of many facilities in their ability to achieve extremely low rates of hypercalcemia, as over 3000 facilities have 1% or less patients with hypercalcemia. However, when one looks at the average national performance of 2%, they may interpret that statistic as demonstrating the absence of a performance gap for this safety measure. That interpretation ignores the highly skewed distribution of facility performance for this safety measure. For this safety measure, the performance gap is clearly demonstrated by comparing the 1377 US dialysis facilities (19% of the total reported facilities) with 4% or greater patients with hypercalcemia to the majority of dialysis facilities that achieve extremely low hypercalcemia rates. We maintain that the measure is important for safety monitoring, as nearly one-fourth of US dialysis facilities are relatively poor at preventing hypercalcemia, an intermediate outcome consistently associated with poorer patient survival and clearly influenced by providers' bone and mineral disease management practices.

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

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*
Disparity analyses were performed among the entire eligible adult population (n=592,121) to examine the difference in performance scores by sex, race, ethnicity, age, insurance status and nursing home status.

In particular, for each facility, the percent of patient-months by demographic group (sex, race, ethnicity, age, insurance status and nursing home status) was calculated. Then, the facilities were divided into quintiles (Q1-Q5) based on the percentage of patient-months in the particular demographic category (i.e., a facility with percentage of females similar to the national median will be included in quintile 3). The top 20% of facilities in terms of rank, based on the percentages of females, were classified as Q5, while the bottom 20% of facilities were classified as Q1. Average (mean) performance for the measure was calculated for each quintile, and the means were examined for trend across quintiles (Q1-Q5). A test for trend was performed to assess disparities in performance scores. There were no increasing (or decreasing) linear trends for each group across quintiles. All the test results imply that we do not have enough evidence to prove that the performance scores will increase (or decrease) as the respective percentage of demographic group increase.
The mean performance scores for percentage of patients with hypercalcemia in each quintile, by demographic group, are presented below.

Facility Level Quintiles by Population Group (Quintile 1-5):

Females (Q1=3.0%, Q2=2.2%, Q3=2.2%, Q4=2.2%, Q5=3.0%)
Black (Q1=2.3%, Q2=2.2%, Q3= 2.8%, Q4=2.8% %, Q5=2.6%)
White (Q1=2.6%, Q2=2.8%, Q3=2.7%, Q4=2.4%, Q5=2.2%)
Hispanic (Q1=2.3%, Q2=1.8%, Q3=2.6%, Q4=3.1%, Q5=2.8%)

Age 65+ (Q1=3.7%, Q2=2.1%, Q3=2.5%, Q4=2.3%, Q5=2.4%)
Dual Eligible (Q1=2.8%, Q2=2.4%, Q3=2.3%, Q4=2.4%, Q5=2.8%)
Nursing Home (Q1=3.3%, Q2=2.3%, Q3=2.2%, Q4=2.2%, Q5=2.7%)

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

2. Reliability and Validity—Scientific Acceptability of Measure Properties

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

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

De.5. **Subject/Topic Area** *(check all the areas that apply):*
Renal, Renal : End Stage Renal Disease (ESRD)

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

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

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

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

S.2b. **Data Dictionary, Code Table, or Value Sets** *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*
No data dictionary **Attachment:**

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.
No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.
Not an instrument-based measure

S.3.1. **For maintenance of endorsement:** Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.
No

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

measure update and explain the reasons.
There have been no changes to the measure specifications since the annual update submitted in late 2016 and approved in early 2017, which clarified that missing data is included in both the numerator and the denominator.

S.4. **Numerator Statement** *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).
Number of patient-months in the denominator with 3-month rolling average of total uncorrected serum (or plasma) calcium greater than 10.2 mg/dL or missing.

S.5. **Numerator Details** *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).
For this measure, the patient reporting month is the last month of the three month reporting period; for example, for the March 2017 reporting month, the hypercalcemia value is the average of the reporting month + the past two months (January – March 2017). If there are multiple calcium measurements during the month, the last non-missing value will be used for the calculation. Calcium measurements can be based on either serum or plasma calcium.

Missing is defined as no value in the current or preceding two months. A 3-month average that is missing is included in the numerator (i.e. all three months are missing).

S.6. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
Number of patient reporting months (see S.5. above) among adult (greater than or equal to 18 years old) in-center hemodialysis, home hemodialysis, or peritoneal dialysis patients under the care of the dialysis facility for the entire reporting month who have had ESRD for greater than 90 days.

S.7. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*
IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).
For this measure, the patient reporting month is the last month of the three month reporting period; for example, for the March 2017 reporting month, the hypercalcemia value is the average of the reporting month + the past two months (January – March 2017).

The patient's age will be determined by subtracting the patient's date of birth from the first day of the most recent month of the reporting period. The patient's time on dialysis will be determined by subtracting the patient's date regular Chronic Dialysis Began from the first day of the most recent month of the reporting period. Patients on dialysis are determined as follows: Primary Type of Dialysis is Hemodialysis, Home Hemodialysis, CAPD or CCPD in the most recent month of the reporting period. Patients under the care of the facility for at least 30 days are determined as follows: if the discharge date from the specified facility is missing/null or is after the last day of the most recent month of the reporting period, then the patient's time under the care of the facility is calculated from the admit date to the last day of the most recent month of the reporting period; if the discharge date is prior to the last day of the most recent month of the reporting period, the patient is excluded from the calculation.

A treatment history file is the data source for the denominator calculation used for the analyses supporting this submission. This file provides a complete history of the status, location, and dialysis treatment modality of an ESRD patient from the date of the first ESRD service until the patient dies or the data collection cutoff date is reached. For each patient, a new record is created each time he/she changes facility or treatment modality. Each record represents a time period associated with a specific modality and dialysis facility. CROWNWeb is the primary basis for placing patients at dialysis facilities and dialysis claims are used as an additional source

of information in certain situations. Information regarding first ESRD service date, death, and transplant is obtained from CROWNWeb (including the CMS Medical Evidence Form (Form CMS-2728) and the Death Notification Form (Form CMS-2746)) and Medicare claims, as well as the Organ Procurement and Transplant Network (OPTN).

S.8. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
Exclusions that are implicit in the denominator definition include all patients who are <18 years old, who have not been in the facility the entire reporting month (transient patients), and patients who have had ESRD for <91 days. There are no additional exclusions for this measure.

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

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

S.11. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in measure testing attachment)
No risk adjustment or risk stratification
If other:

S.12. **Type of score:**
Rate/proportion
If other:

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

S.14. **Calculation Algorithm/Measure Logic** (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*
Patients are included in the denominator if they are >= 18 years old as of the first day of the three month study period, are ESRD for more than 90 days as of the first day of the most recent month of the study period, and are under the care of the facility for at least 30 days as of the last day of the most recent month 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 most recent month of the study period. The patient's time on dialysis will be determined by subtracting the patient's date regular Chronic Dialysis Began from the first day of the most recent month of the study period. Patients on dialysis are determined as follows: Primary Type of Dialysis is Hemodialysis, Home Hemodialysis, CAPD or CCPD in the most recent month of the study period. Patients under the care of the facility for at least 30 days are determined as follows: if the discharge date from the specified facility is missing/null or is after the last day of the most recent month of the study period, then the patient's time under the care of the facility is calculated from the admit date to the last day of the most recent month of the study period; if the discharge date is prior to the last day of the most recent month of the study period, the patient is excluded from the calculation.

The numerator will be determined by counting the patient months in the denominator that meet the following criteria: the average total serum or plasma calcium over the 3-month study period is greater than 10.2 mg/dL or missing. If there is more than one serum or plasma calcium measurement within each month of the study period, the last value for the month shall be used for the calculation of the average. Missing is defined as no value in the current or preceding two months.

S.15. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.
N/A

S.16. **Survey/Patient-reported data** *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*
Specify calculation of response rates to be reported with performance measure results.
N/A

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

S.18. **Data Source or Collection Instrument** *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*
IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.
CROWNWeb and Medicare claims.

S.19. **Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
No data collection instrument provided

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

S.21. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Other
If other: Dialysis Facility

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

2. Validity – See attached Measure Testing Submission Form
1454_Testing_01072019.docx

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

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

2.3 For maintenance of endorsement
Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.
No - This measure is not risk-adjusted

3. Feasibility

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

3a. Byproduct of Care Processes

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

3a.1. **Data Elements Generated as Byproduct of Care Processes.**
generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition
If other:

3b. Electronic Sources

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

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

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

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

3c. Data Collection Strategy

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

3c.1. **Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**
IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.
Data collection is accomplished via CROWNWeb, a web-based and electronic batch submission platform maintained and operated by CMS contractors. Measures reported on DFC are reviewed on a regular basis by dialysis facility providers and rare instances of inaccurate or missing data are present based on comments reported in the DFC ticketing system.

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

4. Usability and Use

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

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are

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

4.1. Current and Planned Use

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

Specific Plan for Use	Current Use (for current use provide URL)
	Public Reporting Dialysis Facility Compare http://www.medicare.gov/dialysisfacilitycompare/ Dialysis Facility Compare http://www.medicare.gov/dialysisfacilitycompare/ Payment Program ESRD QIP http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP/

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

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

DFC:
Purpose: Dialysis Facility Compare helps patients find detailed information about Medicare-certified dialysis facilities. They can compare the services and the quality of care that facilities provide.
Geographic area: United States
Number of accountable entities: All Medicare-certified dialysis facilities who are eligible for the measure, and have at least 11 patients (due to public reporting requirements). For the most recent update to Dialysis Facility Compare (January 2019), 6415 facilities had a score reported.
Patients included: All patients who meet the requirements to be included in the measure.

QIP:
Purpose: The ESRD QIP will reduce payments to ESRD facilities that do not meet or exceed certain performance standards. The measure was added to the program for PY2016.
Geographic area: United States
Number of accountable entities: All Medicare-certified dialysis facilities who are eligible for the measure, and have at least 11 patients (due to public reporting requirements). For the most recent QIP report (PY 2019), this was 6753 facilities.
Patients included: All patients who meet the requirements to be included in the measure.

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

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

4a2.1.1. **Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.**
How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.
Results of this measure are currently reported on Dialysis Facility Compare and in the ESRD Quality Incentive Program. All Medicare-certified dialysis facilities are eligible for reporting in both programs (approximately 7,000 dialysis facilities). Each program has a helpdesk and supporting documentation available to assist with interpretation of the measure results.

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

4a2.1.2. **Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.**
For DFC, the results are first reported to facilities via a closed preview period, where facilities can review their data prior to each of the quarterly updates of the public facing Dialysis Facility Compare website. These preview reports are posted on dialysisdata.org, where facilities can also find a detailed Guide to the Quarterly Dialysis Facility Compare Reports and other supporting documentation. Facilities can submit comments/questions about their results at any time, and can request patient lists for their facilities during the specified preview periods.

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

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

4a2.2.1. **Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.**
Describe how feedback was obtained.
For DFC, feedback can be provided any time through contacting the dialysisdata.org helpdesk. Preview periods allow for specific times for facilities review and comment on measure calculations, and provide an opportunity to request a patient list.

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

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

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

4a2.2.3. **Summarize the feedback obtained from other users**
We reviewed the public comments that were addressed in the ESRD QIP Final Rules (FRs) that have been published since the last endorsement (PY2019 – PY2022). The bulk of the questions addressed in the FRs during this time period had to do with the use of the measure to meet the Protecting Access to Medicare Act (PAMA) requirements.

4a2.3. **Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.**
Feedback received during DFC preview periods has resulted in more detailed and accurate documentation available to the public, primarily via the ESRD Measures Manual and the Guide to the Quarterly Dialysis Facility Reports.

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

4b1. **Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)**
If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.
The following reports the performance scores for this measure for eligible US dialysis facilities at the yearly level over three years, 2015 - 2017. This analysis demonstrates incremental improvement in performance across three years for the measure as implemented on DFC.

2015: 5.2%
2016: 2.4%
2017: 1.5%

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

4b2.1. **Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.**
This measure was endorsed with reserve status, given the high national performance on the measure when NQF evaluated in 2015. We were encouraged by the magnitude of improvement in measure results after implementation noted in 4b1 above. We have not been notified of documented unintended impacts on patients as a result of measure implementation.

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

5. Comparison to Related or Competing Measures

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

5. Relation to Other NQF-endorsed Measures
Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.
No

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

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

5a. Harmonization of Related Measures
The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. **If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):**
Are the measure specifications harmonized to the extent possible?

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

5b. Competing Measures

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

5b.1. **If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
N/A

Appendix

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

Contact Information

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

Additional Information

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

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

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

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

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 1454
Measure Title: Proportion of patients with hypercalcemia
IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title
Date of Submission: 4/2/2019

Instructions

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

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

1a. Evidence to Support the Measure Focus

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

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

Notes

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

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

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

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

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

☐ Outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the PRO

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

☒ Intermediate clinical outcome (*e.g., lab value*): Serum or Plasma calcium >10.2 (3-mo. rolling average)

☐ Process: Click here to name what is being measured

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

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

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

Numerous studies have demonstrated the association of prolonged calcium and phosphorus dysregulation on patient morbidity and mortality (KDOQI 2003; KDIGO 2009; KDIGO Update 2017). Observational cohort studies show a consistent adverse association of hypercalcemia with cardiovascular events and all-cause mortality [1-9]. Clinical data demonstrate the association of increased serum calcium with vascular [10,11] and valvular calcifications [12]. The basic science also supports a pathological role of high calcium in promoting soft tissue and vascular calcification [13-15]. Although there are no interventional studies demonstrating the benefit of correcting hypercalcemia, the current available evidence indicates that serum calcium concentrations >10.2 mg/dL place the patient at increased risk of poor outcomes.

(Citations included at the end of this document, with newly added citations in red).

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

N/A

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

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

N/A

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

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

X Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

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

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	1) Kidney Disease: Improving Global Outcomes (KDIGO) CKDMBD Work Group: KDIGO Clinical Practice Guideline for the Diagnosis, Evaluation, Prevention, and Treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD). Kidney International 2009 76 (Suppl 113): S1-S130. 2) National Kidney Foundation: K/DOQI Clinical Practice Guidelines for Bone Metabolism and Disease in Chronic Kidney Disease. American Journal of Kidney Disease 2003 42:S1-S202 (suppl 3). 3) KDIGO 2017 Clinical Practice Guidelines Update for the Diagnosis, Evaluation, Prevention, and Treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD). Kidney International Supplements 2017 7(1):1-59.
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	1) KDIGO Clinical Practice Guideline for the Diagnosis, Evaluation, Prevention, and Treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD): 4.1.2. In patients with CKD stages 3-5D, we suggest maintaining serum calcium in the normal range (2D).

	2) KDOQI Clinical Practice Guidelines for Bone Metabolism and Disease in Chronic Kidney Disease: Guideline 6.2.In CKD Patients With Kidney Failure (Stage 5): Serum levels of corrected total calcium should be maintained within the normal range for the laboratory used, preferably toward the lower end (8.4 to 9.5 mg/dL [2.10 to 2.37 mmol/L]). (OPINION) 3) KDIGO 2017 Clinical Practice Guidelines Update for the Diagnosis, Evaluation, Prevention, and Treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD): 4.1.3 In adult patients with CKD G3a-G5D, we suggest avoiding hypercalcemia (2c). In children with CKD G3a-G5D, we suggest maintaining serum calcium in the age-appropriate normal range (2c).
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	The 2009 KDIGO guideline recommendation was graded 2D. The KDOQI recommendation was not graded. The 2017 KDIGO guideline recommendation was graded 2C. Grade C: Low quality of evidence. The true effect may be substantially different from the estimate of the effect.
Provide all other grades and definitions from the evidence grading system	The rating system defined in the KDOQI Guidelines was used to grade the strength of the Guideline recommendation. KDOQI defined grades as follows: A: High quality of evidence. We are confident that the true effect lies close to that of the estimate of the effect. B: Moderate quality of evidence. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. C: Low quality of evidence. The true effect may be substantially different from the estimate of the effect. D: Very low quality of evidence. The estimate of effect is very uncertain and often will be far from the truth.
Grade assigned to the **recommendation** with definition of the grade	The 2009 KDIGO guideline recommendation was graded 2D. The KDOQI recommendation was not graded. The 2017 KDIGO guideline recommendation was graded 2C. Level 2 (conditional recommendation/suggestion): "We Suggest". Implications:

	• Patients: The majority of people in your situation would want the recommended course of action, but many would not. • Clinicians: Different choices will be appropriate for different patients. Each patient needs help to arrive at a management decision consistent with her or his values and preferences. Policy: The recommendation is likely to require substantial debate and involvement of stakeholders before policy can be determined.
Provide all other grades and definitions from the recommendation grading system	Level 1 (strong recommendation): "We Recommend". Implications • Patients: Most people in your situation would want the recommended course of action and only a small proportion would not. • Clinicians: Most patients should receive the recommended course of action. Policy: The recommendation can be adopted as policy in most situations.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	N/A
Estimates of benefit and consistency across studies	N/A
What harms were identified?	N/A
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	See below for a list of new studies (and their abstracts) that have been published in recent years. These studies support the SR's cited above.

1a.4 OTHER SOURCE OF EVIDENCE

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

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

N/A

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

N/A

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

N/A

1) Kim Y, Yoo KD, Kim HJ et al. Association of serum mineral parameters with mortality in hemodialysis patients: Data from the Korean end-stage renal disease registry. Kidney Res Clin Pract. 2018 Sep;37(3):266-276. doi: 10.23876/j.krcp.2018.37.3.266. Epub 2018 Sep 30.

BACKGROUND:
We investigated the associations between mineral metabolism parameters and mortality to identify optimal targets in Korean hemodialysis patients.
METHODS:
Among hemodialysis patients registered in the end-stage renal disease registry of the Korean Society of Nephrology between March 2012 and June 2017, those with serum calcium, phosphorus, and intact parathyroid hormone (iPTH) measured at enrollment were included. Association of serum levels of calcium, phosphorus, and iPTH with all-cause mortality was analyzed.
RESULTS:
Among 21,433 enrolled patients, 3,135 (14.6%) died during 24.8 ± 14.5 months of follow-up. After multivariable adjustment, patients in the first quintile of corrected calcium were associated with lower mortality (hazard ratio [HR], 0.84; 95% confidence interval [95% CI], 0.71-0.99; P = 0.003), while those in the fifth quintile were associated with higher mortality (HR, 1.39; 95% CI, 1.20-1.61; P < 0.001) compared with those in the third quintile. For phosphorus, only the lowest quintile was significantly associated with increased mortality (HR, 1.24; 95% CI, 1.08-1.43; P = 0.003). The lowest (HR, 1.18; 95% CI, 1.02-1.36; P = 0.026) and highest quintiles of iPTH (HR, 1.24; 95% CI, 1.05-1.46; P = 0.013) were associated with increased mortality. For target counts achieved according to the Kidney Disease Outcomes Quality Initiative guideline, patients who did not achieve any mineral parameter targets hadhigher mortality than those who achieved all three targets (HR, 1.37; 95% CI, 1.12-1.67; P = 0.003).

CONCLUSION:
In Korean hemodialysis patients, high serum calcium, low phosphorus, and high and low iPTH levels were associated with increased all-cause mortality.

2) Wang M, Obi Y, Streja E et al. Association of Parameters of Mineral Bone Disorder with Mortality in Patients on Hemodialysis according to Level of Residual Kidney Function. Clin J Am Soc Nephrol. 2017 Jul 7;12(7):1118-1127. doi: 10.2215/CJN.11931116. Epub 2017 May 9.

BACKGROUND AND OBJECTIVES:
The relationship between mineral and bone disorders and survival according to residual kidney function status has not been previously studied in patients on hemodialysis. We hypothesized that residual kidney function, defined by renal urea clearance, modifies the association between mineral and bone disorder parameters and mortality.
DESIGN, SETTING, PARTICIPANTS, & MEASUREMENTS:
The associations of serum phosphorus, albumin-corrected calcium, intact parathyroid hormone, and alkaline phosphatase with all-cause mortality were examined across three strata (<1.5, 1.5 to <3.0, and ≥3.0 ml/min per 1.73 m2) of baseline residual renal urea clearance using Cox models adjusted for clinical characteristics and laboratory measurements in 35,114 incident hemodialysis patients from a large United States dialysis organization over the period of 2007-2011.
RESULTS:
A total of 8102 (23%) patients died during the median follow-up of 1.3 years (interquartile

range, 0.6-2.3 years). There was an incremental mortality risk across higher serum phosphorus concentrations, which was pronounced among patients with higher residual renal urea clearance (Pinteraction=0.001). Lower concentrations of serum intact parathyroid hormone were associated with higher mortality among patients with low residual renal urea clearance (i.e., <1.5 ml/min per 1.73 m2), whereas higher concentrations showed a higher mortality risk among patients with greater residual renal urea clearance (i.e., ≥1.5 ml/min per 1.73 m2; Pinteraction<0.001). Higher serum corrected total calcium and higher alkaline phosphatase concentrations consistently showed higher mortality risk (Ptrend<0.001 for both) irrespective of residual renal urea clearance strata (Pinteraction=0.34 and Pinteraction=0.53, respectively).
CONCLUSIONS:
Residual kidney function modified the mortality risk associated with serum phosphorus and intact parathyroid hormone among incident hemodialysis patients. Future studies are needed to examine whether taking account for residual kidney function into the assessment of mortality risk associated with serum phosphorus and intact parathyroid hormone improves patient management and clinical outcomes in the hemodialysis population.

3) Liu CT, Lin YC, Lin YC et al. Roles of Serum Calcium, Phosphorus, PTH and ALP on Mortality in Peritoneal Dialysis Patients: A Nationwide, Population-based Longitudinal Study Using TWRDS 2005-2012. Sci Rep. 2017 Feb 24;7(1):33. doi: 10.1038/s41598-017-00080-4.

Biomarkers of chronic kidney disease-mineral and bone disorder (CKD-MBD) correlate with morbidity and mortality in dialysis patients. However, the comparative roles of each CKD-MBD biomarker remained undetermined on long-term peritoneal dialysis (PD) patients. This retrospective study, employing a population-based database, aimed to evaluate the performance and provide the best evidence of each biomarker of CKD-MBD as predictor of all-cause mortality. Throughout the 8-year study period, total 12,116 PD patients were included in this study. Cox proportional regression and Kaplan-Meier method were used for survival analysis. For Cox regression model, baseline measurements and time-varying covariates were used for analysis. In Cox regression model using time-dependent covariates, serum calcium level of ≧9.5 mg/dL was associated with increased mortality. For phosphorus, serum levels of either ≧6.5 mg/dL or <3.5 mg/dL were associated with increased mortality. For parathyroid hormone (PTH), higher serum levels were not associated increased mortality. For alkaline phosphatase (ALP), mortality increased at levels ≧100 IU/L. Our findings suggested that the detrimental effect of ALP on survival was more consistent, while serum calcium, phosphorus and PTH may have a less prominent effect on mortality. This study provided additional information for manipulating CKD-MBD biomarkers in PD patients.

4) Soohoo M, Feng M, Obi Y et al. Changes in Markers of Mineral and Bone Disorders and Mortality in Incident Hemodialysis Patients.Am J Nephrol. 2016;43(2):85-96. doi: 10.1159/000444890. Epub 2016 Mar 8.

BACKGROUND:
Abnormalities in mineral and bone disorder (MBD) markers are common in patients with chronic kidney disease. However, previous studies have not accounted for their changes over time, and it is unclear whether these changes are associated with survival.
METHODS:
We examined the association of change in MBD markers (serum phosphorus (Phos), albumin-corrected calcium (Ca(Alb)), intact parathyroid hormone (iPTH) and alkaline phosphatase (ALP)) during the first 6 months of hemodialysis (HD) with all-cause mortality across baseline MBD

strata using survival models adjusted for clinical characteristics and laboratory measurements in 102,754 incident HD patients treated in a large dialysis organization between 2007 and 2011.
RESULTS:
Across all MBD markers (Phos, Ca(Alb), iPTH and ALP), among patients whose baseline MBD levels were higher than the reference range, increase in MBD levels was associated with higher mortality (reference group: MBD level within reference range at baseline and no change at 6 months follow-up). Conversely, decrease in Phos and iPTH, among baseline Phos and iPTH levels lower than the reference range, respectively, were associated with higher mortality. An increase in ALP was associated with higher mortality across baseline strata of ALP ≥80 U/l. However, patients with baseline ALP <80 U/l trended towards a lower risk of mortality irrespective of the direction of change at 6 months follow-up.
CONCLUSIONS:
There is a differential association between changes in MBD markers with mortality across varying baseline levels in HD patients. Further study is needed to determine if consideration of both baseline and longitudinal changes in the management of MBD derangements improves outcomes in this population.

5) Fernandez-Martin JL, Martinez-Camblor P, Dionisi MP et al. Improvement of mineral and bone metabolism markers is associated with better survival in haemodialysis patients: the COSMOS study. Nephrol Dial Transplant. 2015 Sep;30(9):1542-51. doi: 10.1093/ndt/gfv099. Epub 2015 Apr 28.
BACKGROUND:
Abnormalities in serum phosphorus, calcium and parathyroid hormone (PTH) have been associated with poor survival in haemodialysis patients. This COSMOS (Current management Of Secondary hyperparathyroidism: a Multicentre Observational Study) analysis assesses the association of high and low serum phosphorus, calcium and PTH with a relative risk of mortality. Furthermore, the impact of changes in these parameters on the relative risk of mortality throughout the 3-year follow-up has been investigated.
METHODS:
COSMOS is a 3-year, multicentre, open-cohort, prospective study carried out in 6797 adult chronic haemodialysis patients randomly selected from 20 European countries.
RESULTS:
Using Cox proportional hazard regression models and penalized splines analysis, it was found that both high and low serum phosphorus, calcium and PTH were associated with a higher risk of mortality. The serum values associated with the minimum relative risk of mortality were 4.4 mg/dL for serum phosphorus, 8.8 mg/dL for serum calcium and 398 pg/mL for serum PTH. The lowest mortality risk ranges obtained using as base the previous values were 3.6-5.2 mg/dL for serum phosphorus, 7.9-9.5 mg/dL for serum calcium and 168-674 pg/mL for serum PTH. Decreases in serum phosphorus and calcium and increases in serum PTH in patients with baseline values of >5.2 mg/dL (phosphorus), >9.5 mg/dL (calcium) and <168 pg/mL (PTH), respectively, were associated with improved survival.
CONCLUSIONS:
COSMOS provides evidence of the association of serum phosphorus, calcium and PTH and mortality, and suggests survival benefits of controlling chronic kidney disease-mineral and bone disorder biochemical parameters in CKD5D patients.

6) Rivara MB, Ravel V, Kalantar-Zadeh K et al. Uncorrected and Albumin-Corrected Calcium, Phosphorus, and Mortality in Patients Undergoing Maintenance Dialysis. J Am Soc Nephrol. 2015 Jul;26(7):1671-81. doi: 10.1681/ASN.2014050472. Epub 2015 Jan 22.

Uncorrected serum calcium concentration is the first mineral metabolism metric planned for use as a quality measure in the United States ESRD population. Few studies in patients undergoing either peritoneal dialysis (PD) or hemodialysis (HD) have assessed the association of uncorrected serum calcium concentration with clinical outcomes. We obtained data from 129,076 patients on dialysis (PD, 10,066; HD, 119,010) treated in DaVita, Inc. facilities between July 1, 2001, and June 30, 2006. After adjustment for potential confounders, uncorrected serum calcium <8.5 and ≥10.2 mg/dl were associated with excess mortality in patients on PD or HD (comparison group uncorrected calcium 9.0 to <9.5 mg/dl). Additional adjustment for serum albumin concentration substantially attenuated the all-cause mortality hazard ratios (HRs) associated with uncorrected calcium <8.5 mg/dl (HR, 1.29; 95% confidence interval [95% CI], 1.16 to 1.44 for PD; HR, 1.17; 95% CI, 1.13 to 1.20 for HD) and amplified the HRs associated with calcium ≥10.2 mg/dl (HR, 1.65; 95% CI, 1.42 to 1.91 for PD; HR, 1.59; 95% CI, 1.53 to 1.65 for HD). Albumin-corrected calcium ≥10.2 mg/dl and serum phosphorus ≥6.4 mg/dl were also associated with increased risk for death, irrespective of dialysis modality. In summary, in a large nationally representative cohort of patients on dialysis, abnormalities in markers of mineral metabolism, particularly high concentrations of serum calcium and phosphorus, were associated with increased mortality risk. Additional studies are needed to investigate whether control of hypercalcemia and hyperphosphatemia in patients undergoing dialysis results in improved clinical outcomes.

7) Block GA, Klassen PS, Lazarus JM, et al. Mineral metabolism, mortality, and morbidity in maintenance hemodialysis. Journal of the American Society of Nephrology : JASN 2004 15:2208-18.

Mortality rates in ESRD are unacceptably high. Disorders of mineral metabolism (hyperphosphatemia, hypercalcemia, and secondary hyperparathyroidism) are potentially modifiable. For determining associations among disorders of mineral metabolism, mortality, and morbidity in hemodialysis patients, data on 40,538 hemodialysis patients with at least one determination of serum phosphorus and calcium during the last 3 mo of 1997 were analyzed. Unadjusted, case mix-adjusted, and multivariable-adjusted relative risks of death were calculated for categories of serum phosphorus, calcium, calcium x phosphorus product, and intact parathyroid hormone (PTH) using proportional hazards regression. Also determined was whether disorders of mineral metabolism were associated with all-cause, cardiovascular, infection-related, fracture-related, and vascular access-related hospitalization. After adjustment for case mix and laboratory variables, serum phosphorus concentrations >5.0 mg/dl were associated with an increased relative risk of death (1.07, 1.25, 1.43, 1.67, and 2.02 for serum phosphorus 5.0 to 6.0, 6.0 to 7.0, 7.0 to 8.0, 8.0 to 9.0, and >/=9.0 mg/dl). Higher adjusted serum calcium concentrations were also associated with an increased risk of death, even when examined within narrow ranges of serum phosphorus. Moderate to severe hyperparathyroidism (PTH concentrations >/=600 pg/ml) was associated with an increase in the relative risk of death, whereas more modest increases in PTH were not. When examined collectively, the population attributable risk percentage for disorders of mineral metabolism was 17.5%, owing largely to the high prevalence of hyperphosphatemia. Hyperphosphatemia and hyperparathyroidism were significantly associated with all-cause, cardiovascular, and fracture-related hospitalization. Disorders of mineral metabolism are independently associated with mortality and morbidity

associated with cardiovascular disease and fracture in hemodialysis patients.

8) Young EW, Albert JM, Satayathum S, et al. Predictors and consequences of altered mineral metabolism: the Dialysis Outcomes and Practice Patterns Study. Kidney international 2005 67:1179-87.

BACKGROUND:
Altered mineral metabolism contributes to bone disease, cardiovascular disease, and other clinical problems in patients with end-stage renal disease.
METHODS:
This study describes the recent status, significant predictors, and potential consequences of abnormal mineral metabolism in representative groups of hemodialysis facilities (N= 307) and patients (N= 17,236) participating in the Dialysis Outcomes and Practice Patterns Study (DOPPS) in the United States, Europe, and Japan from 1996 to 2001.
RESULTS:
Many patients fell out of the recommended guideline range for serum concentrations of phosphorus (8% of patients below lower target range, 52% of patients above upper target range), albumin-corrected calcium (9% below, 50% above), calcium-phosphorus product (44% above), and intact PTH (51% below, 27% above). All-cause mortality was significantly and independently associated with serum concentrations of phosphorus (RR 1.04 per 1 mg/dL, P= 0.0003), calcium (RR 1.10 per 1 mg/dL, P < 0.0001), calcium-phosphorus product (RR 1.02 per 5 mg(2)/dL(2), P= 0.0001), PTH (1.01 per 100 pg/dL, P= 0.04), and dialysate calcium (RR 1.13 per 1 mEq/L, P= 0.01). Cardiovascular mortality was significantly associated with the serum concentrations of phosphorus (RR 1.09, P < 0.0001), calcium (RR 1.14, P < 0.0001), calcium-phosphorus product (RR 1.05, P < 0.0001), and PTH (RR 1.02, P= 0.03). The adjusted rate of parathyroidectomy varied 4-fold across the DOPPS countries, and was significantly associated with baseline concentrations of phosphorus (RR 1.17, P < 0.0001), calcium (RR 1.58, P < 0.0001), calcium-phosphorus product (RR 1.11, P < 0.0001), PTH (RR 1.07, P < 0.0001), and dialysate calcium concentration (RR 0.57, P= 0.03). Overall, 52% of patients received some form of vitamin D therapy, with parenteral forms almost exclusively restricted to the United States. Vitamin D was potentially underused in up to 34% of patients with high PTH, and overused in up to 46% of patients with low PTH. Phosphorus binders (mostly calcium salts during the study period) were used by 81% of patients, with potential overuse in up to 77% patients with low serum phosphorus concentration, and potential underuse in up to 18% of patients with a high serum phosphorus concentration.
CONCLUSION:
This study expands our understanding of the relationship between altered mineral metabolism and outcomes and identifies several potential opportunities for improved practice in this area.

9) Kalantar-Zadeh K, Kuwae N, Regidor DL, et al. Survival predictability of time-varying indicators of bone disease in maintenance hemodialysis patients. Kidney international 2006 70:771-80.

Although renal osteodystrophy and vitamin D analogs may be related to survival in maintenance hemodialysis (MHD) patients, most studies have examined associations between baseline values and survival without accounting for variations in clinical and laboratory measures over time. We examined associations between survival and quarterly laboratory values and administered paricalcitol in a 2-year (July 2001-June 2003) cohort of 58,058 MHD patients from all DaVita dialysis clinics in USA using both time-dependent Cox models with repeated measures and fixed-covariate Cox models with only baseline values. Whereas hypercalcemia and

hyperphosphatemia were robust predictors of higher death risk in all models, the association between serum calcium and mortality was different in time-varying models. Changes in baseline calcium and phosphorus values beyond the Kidney Disease Outcome Quality Initiative recommended targets were associated with increased mortality. Associations between high serum parathyroid hormone and increased death risk were masked by case-mix characteristics of MHD patients. Time-varying serum alkaline phosphatase had an incremental association with mortality. Administration of any dose of paricalcitol was associated with improved survival in time-varying models. Controlling for nutritional markers may introduce overadjustment bias owing to their strong collinearity with osteodystrophy surrogates. Whereas both time-dependent and fixed-covariate Cox models result in similar associations between osteodystrophy indicators and survival, subtle but potentially clinically relevant differences between the two models exist, probably because fixed models do not account for variations of osteodystrophy indices and changes in medication dose over time.

10) Kimata N, Albert JM, Akiba T, et al. Association of mineral metabolism factors with all-cause and cardiovascular mortality in hemodialysis patients: the Japan dialysis outcomes and practice patterns study. Hemodialysis international. International Symposium on Home Hemodialysis 2007 11:340-8.

Abnormalities in mineral metabolism have been linked to mortality in hemodialysis (HD) patients. We postulated that these abnormalities would have a particularly large deleterious impact on deaths due to cardiovascular causes in Japan. This study describes the recent status of abnormal mineral metabolism, significant predictors, and potential consequences in the Dialysis Outcomes and Practice Patterns Study (DOPPS), Phases 1 and 2, in Japan. Major predictor variables were patient demographics, comorbidities, and laboratory markers of mineral metabolism such as albumin-adjusted serum calcium (calciumAlb), phosphorus, and intact PTH (iPTH). In a cross section of 3973 Japanese HD patients in DOPPS I and II, a large faction had laboratory values outside of the recommended Kidney Disease Outcomes Quality Initiative (K/DOQI) guideline range for serum concentrations of phosphorus (51% of patients above upper target range), calciumAlb (43.7% above), calcium-phosphorus (Ca x P) product (41.1% above), and iPTH (18.6% above). All-cause mortality was significantly and independently associated with calciumAlb (relative risk [RR]=1.22 per 1 mg/dL, p=0.0005) and iPTH (RR=1.04 per 100 pg/mL, p=0.04). Cardiovascular mortality was significantly associated with calciumAlb (RR=1.28, p=0.02), phosphorus (RR=1.13 per 1 mg/dL, p=0.008), Ca x P product (RR=1.07 per 2 mg(2)/dL(2), p=0.002), and PTH (RR=1.08, p=0.0001). This study expands our understanding of the relationship between altered mineral metabolism and mortality outcomes, showing slightly stronger associations with cardiovascular causes than observed for all-cause mortality. These findings have important therapeutic implications for Japanese HD patients.

11) Tentori F, Blayney MJ, Albert JM, et al. Mortality risk for dialysis patients with different levels of serum calcium, phosphorus, and PTH: the Dialysis Outcomes and Practice Patterns Study (DOPPS). American journal of kidney diseases : the official journal of the National Kidney Foundation 2008 52:519-30.

BACKGROUND:

Abnormalities in serum calcium, phosphorus, and parathyroid hormone (PTH) concentrations are common in patients with chronic kidney disease and have been associated with increased morbidity and mortality. No clinical trials have been conducted to clearly identify categories of calcium, phosphorus, and PTH levels associated with the lowest mortality risk. Current clinical practice guidelines are based largely on expert opinions, and clinically relevant differences exist among guidelines across countries. We sought to describe international trends in calcium, phosphorus, and PTH levels during 10 years and identify mortality risk categories in the Dialysis Outcomes and Practice Patterns Study (DOPPS), an international study of hemodialysis practices and associated outcomes.
STUDY DESIGN:
Prospective cohort study.
PARTICIPANTS:
25,588 patients with end-stage renal disease on hemodialysis therapy for longer than 180 days at 925 facilities in DOPPS I (1996-2001), DOPPS II (2002-2004), or DOPPS III (2005-2007).
PREDICTORS:
Serum calcium, albumin-corrected calcium (Ca(Alb)), phosphorus, and PTH levels.
OUTCOMES:
Adjusted hazard ratios for all-cause and cardiovascular mortality calculated using Cox models.
RESULTS:
Distributions of mineral metabolism markers differed across DOPPS countries and phases, with lower calcium and phosphorus levels observed in the most recent phase of DOPPS. Survival models identified categories with the lowest mortality risk for calcium (8.6 to 10.0 mg/dL), Ca(Alb) (7.6 to 9.5 mg/dL), phosphorus (3.6 to 5.0 mg/dL), and PTH (101 to 300 pg/mL). The greatest risk of mortality was found for calcium or Ca(Alb) levels greater than 10.0 mg/dL, phosphorus levels greater than 7.0 mg/dL, and PTH levels greater than 600 pg/mL and in patients with combinations of high-risk categories of calcium, phosphorus, and PTH.
LIMITATIONS:
Because of the observational nature of DOPPS, this study can only indicate an association between mineral metabolism categories and mortality.
CONCLUSIONS:
Our results provide important information about mineral metabolism trends in hemodialysis patients in 12 countries during a decade. The risk categories identified in the DOPPS cohort may be relevant to efforts at international harmonization of existing clinical guidelines for mineral metabolism.

12) Chertow G.M., Raggi P., Chasan-Taber S., Bommer J., Holzer H., Burke S.K. Determinants of progressive vascular calcification in haemodialysis patients. Nephrology Dialysis Transplantation 2004 19 (6), pp. 1489- 1496.

BACKGROUND:
We determined recently that targeted treatment with calcium-based phosphate binders (calcium acetate and carbonate) led to progressive coronary artery and aortic calcification by electron beam tomography (EBT), while treatment with the non-calcium-containing phosphate binder, sevelamer, did not. Aside from the provision of calcium, we hypothesized that other factors might be related to the likelihood of progressive calcification in both or either treatment groups.
METHODS:

We explored potential determinants of progressive vascular calcification in 150 randomized study subjects who underwent EBT at baseline and at least once during follow-up (week 26 or 52).
RESULTS:
Among calcium-treated subjects, higher time-averaged concentrations of calcium, phosphorus and the calcium-phosphorus product were associated with more pronounced increases in EBT scores; no such associations were demonstrated in sevelamer-treated subjects. The relation between parathyroid hormone (PTH) and the progression of calcification was more complex. Lower PTH was associated with more extensive calcification in calcium-treated subjects, whereas higher PTH was associated with calcification in sevelamer-treated subjects. Serum albumin was inversely correlated with progression in aortic calcification. Sevelamer was associated with favourable effects on lipids, although the link between these effects and the observed attenuation in vascular calcification remains to be elucidated.
CONCLUSION:
Calcium-based phosphate binders are associated with progressive coronary artery and aortic calcification, especially when mineral metabolism is not well controlled. Calcium may directly or indirectly (via PTH) adversely influence the balance of skeletal and extraskeletal calcification in haemodialysis patients.

13) Dhingra R, Sullivan LM, Fox CS, Wang TJ, D´Agostino RB Sr, Gaziano JM, Vasan RS: Relations of serum phosphorus and calcium levels to the incidence of cardiovascular disease in the community. Arch Intern Med 2007 167: 879–885.

BACKGROUND:
Higher levels of serum phosphorus and the calcium-phosphorus product are associated with increased mortality from cardiovascular disease (CVD) in patients with chronic kidney disease (CKD) or prior CVD. However, it is unknown if serum phosphorus levels influence vascular risk in individuals without CKD or CVD.
METHODS:
We prospectively evaluated 3368 Framingham Offspring study participants (mean age, 44 years; 51% were women) free of CVD and CKD. We used multivariable Cox models to relate serum phosphorus and calcium levels to CVD incidence.
RESULTS:
On follow-up (mean duration, 16.1 years), there were 524 incident CVD events (159 in women). In multivariable analyses and adjusting for established risk factors and additionally for glomerular filtration rate and for hemoglobin, serum albumin, proteinuria, and C-reactive protein levels, a higher level of serum phosphorus was associated with an increased CVD risk in a continuous fashion (adjusted hazard ratio per increment of milligrams per deciliter, 1.31; 95% confidence interval, 1.05-1.63; P=.02; P value for trend across quartiles = .004). Individuals in the highest serum phosphorus quartile experienced a multivariable-adjusted 1.55-fold CVD risk (95% confidence interval, 1.16%-2.07%; P=.004) compared with those in the lowest quartile. These findings remained robust in time-dependent models that updated CVD risk factors every 4 years and in analyses restricted to individuals without proteinuria and an estimated glomerular filtration rate greater than 90 mL/min per 1.73 m(2). Serum calcium was not related to CVD risk.
CONCLUSION:
Higher serum phosphorus levels are associated with an increased CVD risk in individuals free of CKD and CVD in the community. These observations emphasize the need for additional research

to elucidate the potential link between phosphorus homeostasis and vascular risk.

14) Wang AY, Lam CW, Wang M, Chan IH, Lui SF, Sanderson JE. Is valvular calcification a part of the missing link between residual kidney function and cardiac hypertrophy in peritoneal dialysis patients? Clinical journal of the American Society of Nephrology 2009 4:1629-36.

BACKGROUND AND OBJECTIVES:

Residual renal function (RRF) predicts survival and shows an important inverse relation with cardiac hypertrophy in peritoneal dialysis (PD) patients. We hypothesized that valvular calcification and the calcification milieu may be part of the process linking loss of RRF and cardiac hypertrophy.

DESIGN, SETTING, PARTICIPANTS, & MEASUREMENTS:
A cross-sectional study was conducted by performing two-dimensional echocardiography on 230 PD patients to assess valvular calcification and left ventricular (LV) mass and collecting 24-h urine for estimation of RRF.
RESULTS:
Patients having valvular calcification had lower RRF than those without. Patients with no RRF showed higher calcium-phosphorus product (Ca x P) and C-reactive protein (CRP). Using multiple logistic regression analysis, every 1-ml/min per 1.73 m(2) increase in residual GFR was associated with a 28% reduction in the risk for valvular calcification. The association was lost after additional adjustment for Ca x P and CRP. Using multiple linear regression analysis, loss of RRF showed significant association with increased LV mass index, but this association was lost after additional adjustment for CRP, Ca x P, and valvular calcification. Patients with all three calcification risk factors, namely inflammation, high CaxP, and no RRF, showed the highest prevalence of valvular calcification and had the most severe cardiac hypertrophy.
CONCLUSIONS:
The association among loss of RRF, valvular calcification, and cardiac hypertrophy was closely linked to increased inflammation and high Ca x P in PD patients. These data suggest that valvular calcification and the calcification milieu are part of the processes linking loss of RRF and worsening cardiac hypertrophy in PD.

15) Ketteler M, Schlieper G, Floege J. Calcification and cardiovascular health: new insights into an old phenomenon. Hypertension 2006 47:1027–1034.
Abstract:

Uremic cardiovascular disease is characterized by accelerated calcifying atherosclerosis and valvular heart disease. Vascular calcification develops at different sites within the vessel wall. Although intimal plaque calcification is a feature of genuine atherosclerosis, medial calcification is restricted to the smooth muscle cell layer and especially to the elastic laminae of arterial vessels (Figure 1). Both entities can be frequently observed in chronic kidney disease (CKD) patients. Dialysis patients with intimal calcifications are elderly and characterized by a history of “traditional” risk factors (eg, smoking and dyslipidemia) before the start of dialysis, whereas those with medial calcifications are, on average, 20 years younger and characterized by a longer time on dialysis treatment and a higher incidence of derangements in their calcium (Ca) phosphate (P) balance. Another recent study in incident dialysis patients showed that those with rapid arterial calcification progress already had calcified coronary arteries before reaching the

dialysis stage. This emphasizes that diagnostic, preventive, and therapeutic measures need to be initiated in early CKD stages. The clinical importance of this notion is stressed by a number of reports demonstrating that coronary artery and valvular calcifications occur prematurely and are very prevalent in dialysis patients and that they are independent risk factors of cardiovascular death in this patient group. Such calcifications can, therefore, serve to at least partially explain why cardiovascular mortality is dramatically increased in the uremic as compared with a normal population and why it is not appropriately explained by the traditional Framingham risk factors. One of the mechanisms by which medial vascular calcification feeds into cardiovascular mortality may be via the associated increase in aortic pulse wave velocity. Calcified arteries become stiffer, causing quicker return of the systolic pulse wave from the periphery, thereby increasing left ventricular afterload. Through this mechanism, a high aortic pulse wave velocity is associated with increased left ventricular mass.

16) Giachelli CM. Vascular calcification mechanisms. Journal of the American Society of Nephrology : JASN 2004 15:2959–2964.
Abstract:

Vascular calcification is highly correlated with cardiovascular disease mortality, especially in patients with ESRD or diabetes. In addition to the devastating effects of inappropriate biomineralization seen in cardiac valvulopathies, calciphylaxis, and idiopathic arterial calcification, vascular calcification is now recognized as a marker of atherosclerotic plaque burden as well as a major contributor to loss of arterial compliance and increased pulse pressure seen with age, diabetes, and renal insufficiency. In recent years, several mechanisms to explain vascular calcification have been identified including (1) loss of inhibition, (2) induction of bone formation, (3) circulating nucleational complexes, and (4) cell death. Alterations in calcium (Ca) and phosphorus (P) balance as seen in patients with ESRD promotes vascular calcification via multiple mechanisms and may explain the alarmingly high levels of cardiovascular disease deaths in these patients. Strategies to control Ca and P levels in patients with ESRD have met with early success in preventing progression of vascular calcification. Whether or not vascular calcification can be reversed is not yet known, but exciting new studies suggest that this may be possible in the future.

17) Yang H, Curinga G, Giachelli CM. Elevated extracellular calcium levels induce smooth muscle cell matrix mineralization in vitro. Kidney Int. 2004;66(6):2293–2299.

Hyperphosphatemia, elevated calcium x phosphorus product (Ca x P), and calcium burden, major causes of vascular calcification, are correlated with increased cardiovascular morbidity and mortality in dialysis patients.

METHODS:
To address the underlying mechanisms responsible for these findings, we have utilized an in vitro human smooth muscle cell (HSMC) model of vascular calcification. Previous studies using this system demonstrated enhanced calcification of HSMC cultures treated with phosphorus levels in the hyperphosphatemic range, and implicated a sodium-dependent phosphate cotransport-dependent mechanism in this effect. In the present study, we examine the effect of increasing calcium concentrations on HSMC calcification in vitro.

RESULTS:
Increasing calcium to levels observed in hypercalcemic individuals increased mineralization of

HSMC cultures under normal phosphorus conditions. Importantly, at these total calcium concentrations, ionized calcium levels increased from 1.2 mmol/L to 1.7 mmol/L, consistent with levels observed physiologically in normocalcemic and hypercalcemic individuals, respectively. Furthermore, increasing both calcium and phosphorus levels led to accelerated and increased mineralization in the cultures. Calcium-induced mineralization was dependent on the function of a sodium-dependent phosphate cotransporter, since it was inhibited by phosphonoformic acid (PFA). While elevated calcium did not affect short-term phosphorus transport kinetics, long-term elevated calcium treatment of HSMCs induced expression of the sodium-dependent phosphate cotransporter, Pit-1.

CONCLUSION:
These studies suggest that elevated calcium may stimulate HSMC mineralization by elevating Ca x P product and enhancing the sodium-dependent phosphate cotransporter-dependent mineralization pathway previously observed in HSMCs.

18) Foley RN, Parfrey PS, Harnett JD, et al. Hypocalcemia, morbidity, and mortality in end-stage renal disease. American journal of nephrology 1996 16:386-93.

BACKGROUND:
Hypocalcemia and hyperphosphatemia with secondary hyper-parathyroidism are characteristic of end-stage renal disease (ESRD). Although calcium levels critically affect almost all cellular processes, the impact of chronic hypocalcemia and other abnormalities of calcium-phosphate homeostasis on the prognosis of ESRD patients is unknown.
METHODS:
An inception cohort of 433 patients starting ESRD therapy was followed prospectively for an average of 41 months. Serum calcium and other parameters were measured monthly. The mean calcium levels were 9.4 +/- 0.7 mg/dl. 23% of the patients had mean calcium levels < 8.8 mg/dl. After adjusting for baseline age, diabetes mellitus, ischemic heart disease, smoking and cholesterol levels, as well as serial albumin, hemoglobin, mean arterial blood pressure, phosphate and alkaline phosphatase levels, chronic hypocalcemia was strongly associated with mortality (RR 2.10, p = 0.006 for a mean calcium level < 8.8 mg/dl). The association with mortality was similar in hemodialysis (RR 2.10, p = 0.006) and peritoneal dialysis patients (2.67, p = 0.034). Using similar covariate adjustment, chronic hypocalcemia was associated with de novo ischemic heart disease (RR 5.23, p < 0.001), recurrent ischemic heart disease (RR 2.46, p = 0.006), de novo cardiac failure (RR 2.64, p < 0.001), and recurrent cardiac failure (RR 3.30, p < 0.001). Hypocalcemia retained its independent impact on morbidity and mortality when analyzed as a time-dependent covariate.
CONCLUSIONS:
Chronic hypocalcemia, a very common, reversible feature of chronic uremia, is independently associated with morbidity and mortality in ESRD patients.

19) Koch M, Lund R, Oldemeyer B, Meares AJ, Dunlay R. Refeeding hypophosphatemia in a chronically hyperphosphatemic hemodialysis patient. Nephron 2000;86(4):552.
Abstract:

Hyperphosphatemia is a common problem in dialysis patients and is usually due to inadequate adherence to dietary phosphate restrictions and noncompliance with pre- scribed phosphate binders. We report an unusual case of refeeding hypophosphatemia in a dialysis patient with an eating disorder and chronic hyperphosphatemia.

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

Measure Number (*if previously endorsed*)**:** 1454
Measure Title: Proportion of patients with hypercalcemia
Date of Submission: 1/7/2019
Type of Measure:

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

Instructions

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

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

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **instrument-based measures** (including PRO-PMs) **and composite performance measures**, reliability should be demonstrated for the computed performance score.

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

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

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

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

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

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

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

2b6. Analyses identify the extent and distribution of **missing data** (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

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

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

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

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

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

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

We analyzed national CROWNWeb data from January 201 – December 2013.

For the Spring 2019 Maintenance submission, we analyzed CROWNWeb data from January 2017 – December 2017.

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

For the Spring 2019 Maintenance submission, we used data from January – December 2017.

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

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

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

5,880 facilities that had at least 11 eligible adult patients during January 2013 – December 2013 were included in the analyses. Public reporting of this measure on DFC or in the ESRD QIP would be restricted to facilities with at least 11 eligible patients for the measure. We have applied this restriction to all the reliability and validity testing reported here.

For the Spring 2019 Maintenance submissions, 6,824 facilities that had at least 11 eligible adult patients during January 2017 – December 2017 were included in the analyses. Public reporting of this measure on DFC or in the ESRD QIP would be restricted to facilities with at least 11 eligible patients for the measure to comply with restrictions on reporting of potentially patient identifiable information related to small sample size. We have applied this restriction to all the reliability and validity testing reported here.

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

There were a total of 482,979 eligible patients. Among those patients, the average age was 62 years, 56% of patients were male, 57% were white, 35% were black, 5% were Asian/Pacific Islander, 18% were Hispanic, and 45% had type II diabetes as the primary cause of ESRD.

A subset of 482,444 eligible patients who belonged to the facilities that had at least 11 eligible patients were included in the testing and analyses.

For the Spring 2019 Maintenance submission, a total of 592,121 eligible patients who belonged to the facilities that had at least 11 eligible patients were included in the testing and analyses. Among those patients, the average age was 63 years, 57% of patients were male, 59% were white, 33% were black, 6% were Asian/Pacific Islander, 18% were Hispanic, and 46% had type II diabetes as the primary cause of ESRD.

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

N/A

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

N/A

_____

2a2. RELIABILITY TESTING

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

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

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

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

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

We used January 2013 – December 2013 CROWNWeb data to calculate facility level monthly and annual performance scores. 5,880 facilities that had at least 11 eligible patients were included in the testing. This included a total of 482,444 patients.

We assessed reliability by calculating facility-level Pearson correlation coefficients between the current performance month and the preceding month for reporting months during January 2013 – December 2013.

In addition, we calculated inter-unit reliability (IUR) for each reporting month and the overall 12 months. The monthly based measure was a simple average across individuals in the facility. The NQF-recommended approach for determining measure reliability is a one-way analysis of variance (ANOVA), in which the between and within facility variation in the measure is determined. The inter-unit reliability (IUR) measures the proportion of the measure variability that is attributable to the between-facility variance. The yearly based measure, however, is not a simple average and we instead estimate the IUR using a bootstrap approach, which uses a resampling scheme to estimate the within facility variation that cannot be directly estimated by ANOVA.

For the Spring 2019 Maintenance submission, we followed the methodology described above to reproduce the inter-unit reliability calculation and Pearson correlation coefficients.

We used January 2017 – December 2017 CROWNWeb data to calculate facility level monthly and annual performance scores. 6,824 facilities that had at least 11 eligible patients were included in the testing. This included a total of 592,121 patients.

We assessed reliability by calculating facility-level Pearson correlation coefficients between the current performance month and the preceding month for reporting months during January 2017 – December 2017.

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

The Pearson correlation coefficients of each pair of the current and the preceding month ranged from 0.78 to 0.84. All were statistically significant (p<0.0001), indicating this measure is reliable over time.

The monthly IURs ranged from 0.61 to 0.66, which indicates that more than half of the variation in the monthly based measure can be attributed to the between facility differences and less than half to within facility variation. The annual IUR across the 12 reporting months was 0.86, which indicates that 86% of the variation in the yearly based measure can be attributed to the between facility variation.

For the Spring 2019 Maintenance submission, the Pearson correlation coefficients of each pair of the current and the preceding month ranged from 0.78 to 0.88. All were statistically significant (p<0.0001), indicating this measure is reliable over time.

The monthly IURs ranged from 0.86 to 0.90, which indicates that more than 80% of the variation in the monthly based measure can be attributed to the between facility differences and less than half to within facility variation. The annual IUR across the 12 reporting months was 0.87, which indicates that 87% of the variation in the yearly based measure can be attributed to the between facility variation.

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

The IURs provide evidence of reliability in that most of the variation can be attributed to the between facility variation. The monthly and overall IURs suggest this measure is reliable. However, since the distribution of performance scores is skewed, the IUR value should be interpreted with some caution. The moderate to strong statistically significant results of the Pearson correlations indicates this measure is reliable over time.

For the Spring 2019 Maintenance submission, the interpretation of the IUR results is similar to the previous submission; the monthly and overall IURs suggest this measure is reliable. However, since the distribution of performance scores is skewed, the IUR value should be interpreted with some caution.

2b1. VALIDITY TESTING

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

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

☒ **Performance measure score**

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

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

We used January 2013 – December 2013 CROWNWeb data to calculate facility level monthly and annual performance scores. 5,880 facilities that had at least 11 eligible patients were included in the testing. This included a total of 482,444 patients.

We assessed validity using Poisson regression models to identify the predictive strength of facility level performance scores for the measure, on mortality, using the 2017 SMR.

In 2010, the measure was unanimously ratified by the Clinical TEP as a valid measure.

For the Spring 2019 Maintenance submission, we used January 2017 – December 2017 CROWNWeb data to calculate facility level monthly and annual performance scores. 6,824 facilities that had at least 11 eligible patients were included in the testing. This included a total of 592,121 patients.

We assessed validity using Poisson regression models to identify the predictive strength of facility level performance scores for the measure, on mortality, using the 2017 SMR. We anticipate a positive correlation with the SMR, since hypercalcemia is a marker of poor overall health.

In 2010, the measure was unanimously ratified by the Clinical TEP as a valid measure. In addition, the measure was presented to another Clinical TEP in 2013, who reaffirmed their support of hypercalcemia as a quality measure.

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

Poisson regression modeling was used to assess the predictive strength of facility level performance scores for hypercalcemia on mortality, using the 2013 NQF endorsed SMR. The results suggest the measure performance scores were predictive of mortality as measured by the SMR. For instance, the facility-level relative risk of mortality for a 10% increase in percent of patients with hypercalcemia, is 1.07 ($p<0.0001$).

For the Spring 2019 Maintenance submission, the results again suggest the measure performance scores were predictive of mortality as measured by the 2017 SMR. The facility-level relative risk of mortality for a 10% increase in percent of patients with hypercalcemia, is 1.08 ($p<0.0001$).

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

The results of the Poisson regression suggest that facilities with a higher percentage of patient-months with hypercalcemia experience a higher standardized mortality rate relative to facilities with a lower percentage of patients with hypercalcemia. The direction of the relationship is as expected.

For the Spring 2019 Maintenance submission, the results of the Poisson regression again suggests that facilities with a higher percentage of patient-months with hypercalcemia experience a higher standardized mortality rate relative to facilities with a lower percentage of patients with hypercalcemia. The direction of the relationship is as expected.

We also maintain the measure on the basis of face validity as determined by the 2010 and 2013 TEPs.

2b2. EXCLUSIONS ANALYSIS

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

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

N/A

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

N/A

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

N/A

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b4.

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

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

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

N/A

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case

mix) is not needed to achieve fair comparisons across measured entities.
Risk adjustment is not necessary for this measure. Disparities were examined at the facility level in section 1b.4, and no disparities were observed.

No risk adjustment or risk stratification is required for this measure. The body of evidence and data on disparities presented in the Importance section do not support the need for risk adjustment (see response to 1b.4 and 1b.5).

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)
Also discuss any "ordering" of risk factor inclusion; for example, are social risk factors added after all clinical factors?

N/A

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

- ☐ **Published literature**
- ☐ **Internal data analysis**
- ☐ **Other (please describe)**

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

N/A

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

N/A

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

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

N/A

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

N/A

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

N/A

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

N/A

2b3.9. Results of Risk Stratification Analysis:

N/A

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

N/A

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

N/A

_____

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

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

Differences in measure performance were evaluated separately for each facility using patient level analyses. The proportion of patients with three-month (current month and the two prior months) rolling average of serum or plasma calcium greater than 10.2 (hypercalcemia), calculated at the year-level, was compared between one facility and the overall national distribution, and repeated for each individual facility.

Note that the monthly based measure is a simple average of binary outcomes across individuals in the facility, for which the binary outcome equals to 0 (failure=hypercalcemia or if the value of calcium for the reporting month is missing). The differences in proportions can be compared using Fisher's Exact tests or its normal approximation. The yearly based measure, however, is not a simple average of binary outcomes and we instead used a re-sampling based exact test, with re-sampling generated from the population distribution of the patient level outcomes. Due to non-symmetric structure of the measure distributions, a one-sided test with significance level 0.025 is used (corresponding to cutoff=0.05 in a

two-sided test). To calculate the p-value, we assess the probability that the facility would experience a number of events (i.e., percentage with hypercalcemia) more extreme than that observed if the null hypothesis were true, the null hypothesis being that the facility's distribution of hypercalcemia will follow the overall national distribution.

For the Spring 2019 submission, we reproduced the significance analysis using data from January – December 2017. We have revised the description of the analysis to be clearer:

Testing was performed on the yearly based performance score. We used a re-sampling based exact test, with re-sampling generated from the population distribution of the patient level outcomes. Note that a one-sided test with significance level 0.025 is used (corresponding to cutoff=0.05 in a two-sided test) due to non-symmetric structure of the measure's distribution. To calculate the p-value, we compute the probability that the facility would experience a number of events (i.e., percentage with hypercalcemia) more extreme than that observed if the null hypothesis were true, with the null hypothesis being that the facility's distribution of hypercalcemia will follow the overall national distribution.

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

Proportion of facilities with significant p-values (significance level ≤0.025) is shown as follows:

Category	Number of facilities	Percent of facilities
As expected	5004	85.1%
Worse than expected	876	14.9%

For the Spring 2019 Maintenance submission, the results were as follows:

Proportion of facilities with significant p-values (significance level ≤0.025) is shown as follows:

Category	Number of facilities	Percent of facilities
As expected	6,329	92.7%
Worse than expected	495	7.3%

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

Using monthly hypercalcemia, calculated at the year-level, as the performance measure, 5004 (85%) facilities have achieved expected performance, and 876 facilities (15%) performed worse than expected.

In general, higher rates of hypercalcemia represent worse quality of care. This analysis demonstrates the ability to identify both practical and statistically significant differences in performance across facilities based on their proportion of hypercalcemia.

For the Spring 2019 Maintenance submission, using monthly hypercalcemia, calculated at the year-level, as the performance measure, 6,329 (93%) facilities have achieved expected performance, and 495 facilities (7%) performed worse than expected.

In general, higher rates of hypercalcemia represent worse quality of care. This analysis demonstrates the ability to identify both practical and statistically significant differences in performance across facilities based on their proportion of hypercalcemia.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
***If only one set of specifications, this section can be skipped*.**

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

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

N/A

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

N/A

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

N/A

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

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

Reporting months with missing values are not excluded from this measure. Missing months are used as a component of the measure numerator, so introduction of bias from exclusion of missing values is not a consideration for the measure as specified.

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

N/A

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

N/A