

Dear Measure Steward/Developer:

Thank you for submitting measures to the Renal project for consideration of endorsement – initial or endorsement maintenance. Your submission has been forwarded to the Steering Committee for review in preparation for its meeting on August 16-17.

In the initial review of your measures, we have identified some questions for clarification or further information needed for the Committee to complete its evaluation. We ask that you send your responses **by Monday, August 8th** so we can provide to the Steering Committee in advance of the meeting. If that's not possible, please let us know.

We realize that the detailed questions on Evidence are new and NQF reviewed the new guidance on measure evaluation and measure submission items on a measure developer webinar in May 2011. The guidance documents as well as the updated slides from the webinar are posted on the [NQF Submitting Standards web page](#).

If you have any questions, please let us know.

KCQA

- **0320** Patient Education Awareness—Physician Level
- **0324** Patient Education Awareness—Facility Level
- **0251** Vascular Access—Functional AVF or Evaluation by Vascular Surgeon for Placement
- **0262** Vascular Access—Catheter Vascular Access and Evaluation by Vascular Surgeon for Permanent Access.

1c. Evidence

We identified some common issues across many of the submissions noted below. Please review your submissions and send us any clarifications. Please keep in mind that NQF does not expect the developer to conduct a primary evidence review; rather the developer is asked to report on the review/grading of body of evidence that was conducted by others – specifically on the quantity, quality, and consistency of the body of evidence.

- **1c6. Quality of the Body of the Evidence.** Most of the submissions did not address the quality of the body of evidence. What did the systematic review of the body of evidence determine about the overall quality of the body of evidence? We ask for a grade in item 1c13, but for this item, we want substantive information about the quality. If the review did not address the overall quality of the body of evidence, please state that and indicate what was identified about the quality of individual studies.

KCQA RESPONSE

Vascular Access: As we noted in our submission, applying the recommendations set forth by the NQF Evidence Task Force, the quality of the studies composing the body of evidence upon which both the KCQA vascular access measures and the 2006 KDOQI Clinical Practice Guidelines for Vascular Access are based can be judged as “moderate” when using a scale of High, Moderate, or Low. The evidence is of moderate, rather than high, quality due to the fact that

there are no randomized clinical trials (RCTs) comparing the three available vascular access types or demonstrating the superiority of AVFs, as a treatment/placebo RCT in that regard would be unethical given the known risks of catheters. Despite this, we have provided evidence from seven studies and two review articles, together encompassing more than 300,000 hemodialysis patients in the United States, the United Kingdom, and Canada, that consistently demonstrate that AVFs have superior longevity, fewer complications, and are associated with the lowest mortality and costs. The studies included in the body of evidence control for confounders that could account for other plausible explanations for these findings and provide a large, precise estimate of effect.

Patient Education: As noted in our submission, applying the recommendations set forth by the NQF Evidence Task Force, the quality of the studies composing the body of evidence upon which the KCQA patient education awareness measures are based can be judged as “high” when using a scale of High, Moderate, or Low, indicating that there are relevant RCTs of direct evidence with adequate size to obtain precise estimates of effect and without serious flaws that introduce bias.

- **1c7. Consistency of Results across Studies.** Most of the submissions did not provide information on consistency of the magnitude and direction of effect across the studies in the body of evidence. For the outcomes studied, what was the magnitude and direction of effect? If a meta-analysis was conducted, the results would be important evidence. Information from evidence tables can be used to provide substantive information on effect size.

KCQA RESPONSE

Vascular Access: The studies cited in the vascular access body of evidence consistently demonstrate that AVFs have the lowest complications rate, the lowest costs of implantation and maintenance, require the fewest interventions, provide longer survival of the access, and are associated with increased survival and lower hospitalization rates than either AV grafts or catheters. As noted in the measure submissions, the number of access-related events is three to seven-fold greater in prosthetic bridge grafts than in native AVFs (1,2,4), costs of implantation and access maintenance are the lowest for AVFs (4-6), and AVFs have been demonstrated to have lower rates of infection than grafts, which, in turn, are less prone to infection than catheters (8). Consequently, AVFs are associated with increased survival and lower hospitalization rates than either AV grafts or catheters (9). Research indicates that patients dialyzed via catheters and grafts have a greater mortality risk (relative risk = 2.3 and 1.47, respectively) than patients dialyzed with AVFs (9). After controlling for age and comorbidities, studies included in the body of evidence indicate that mortality rates are approximately 25 percent for patients being dialyzed via AVFs, 28 percent for patients with AV grafts, and greater than 41 percent for patients with catheters (9-12). Research has likewise indicated that patients with AVFs have the lowest ($P < 0.0001$) likelihood of death compared with those with AV grafts (hazard ratio [HR], 1.160; 95% confidence interval [CI], 1.084 to 1.241) or catheters (HR, 1.696; 95% CI, 1.593 to 1.806) (11). In diabetes mellitus (DM) patients with ESRD, the associated relative mortality risk was higher for those with AV grafts ($RR = 1.41$, $P < 0.003$) and catheters ($RR = 1.54$, $P < 0.002$) as compared with AVFs. In non-DM patients, those with catheters had a higher associated mortality ($RR = 1.70$, $P < 0.001$), as did, to a lesser degree, those with AV grafts ($RR = 1.08$, $P = 0.35$) when compared with AVF. Cause-specific analyses found higher infection-related deaths for patients with catheters ($RR = 2.30$, $P < 0.06$) and AV grafts ($RR = 2.47$, $P < 0.02$) compared with DM patients with AVFs; in non-DM patients, risk was higher also for catheters

(RR = 1.83, $P < 0.04$) and AV grafts (RR = 1.27, $P < 0.33$). Deaths caused by cardiac causes were higher in catheters than AVFs for both DM (RR = 1.47, $P < 0.05$) and non-DM (RR = 1.34, $P < 0.05$) patients (10).

Patient Education: The studies we cited in the patient education body of evidence on the submission form consistently demonstrate that patients who have been educated on the available renal replacement therapy modalities (i.e., hemodialysis, peritoneal dialysis, home hemodialysis, transplants, and no or cessation of therapy) are more likely to use an AVF for dialysis (15), have less depression and improved medication adherence and treatment attendance (15-17), and are more likely to survive and to get a transplant (18,19). For example, multivariable analyses at three and six months after dialysis initiation indicate that patients who score 20 percentage points higher on the Chronic Hemodialysis Knowledge Survey (CHeKS) at baseline are more likely to use an AVF or AV graft for dialysis access than a catheter (OR [95% CI] = 1.49[1.16, 1.93]; $P = 0.002$ and 1.33[1.03, 1.72]; $P = 0.03$, respectively) (15). Likewise, an intensive patient education program (i.e., the RightStart program) coupled with interventions focusing on specific areas of clinical care (e.g., anemia management, adequate dialysis dose, nutrition) was shown to significantly reduce mortality rates among incident hemodialysis patients. Specifically, mortality rates at three, six, and twelve months were 20, 18, and 17 percent for RightStart patients versus 39, 33, and 30 percent for control subjects, respectively ($P < 0.001$ for all comparisons). After adjustment for the random effects of facility difference, the decrease in hazard ratio (HR) of death for the RightStart versus the control group at 365 days remained significant (HR 0.59; 95% CI 0.45 to 0.79; $P < 0.001$). There was a 41 percent reduction in 365-day risk for death in patients who participated in the RightStart program (19).

Additionally, a study published after the KCQA Patient Education Awareness measures were submitted to NQF for endorsement maintenance review demonstrates that attendees of a predialysis national treatment options program (TOPs) who initiated long-term dialysis therapy (median, 3.4 months) at Fresenius Medical Care, North America facilities throughout 2008 more frequently selected home dialysis and had lower catheter rates and mortality risk during the first 90 days of dialysis when compared with period-prevalent incident patients receiving standard care. Specifically, the unadjusted OR for the 3,165 TOPs attendees (10.5 percent of 30,217 incident patients admitted between January 1 and December 31, 2008) for selecting PD therapy was 8.45 (95% CI, 7.63-9.37) with a case-mix plus laboratory-adjusted OR of 5.13 (95% CI, 3.58-7.35). For TOPs patients who opted for in-center HD therapy, the OR was 2.14 (95% CI, 1.96-2.33) and adjusted OR was 2.06 (95% CI, 1.88-2.26) for starting with an AVF or AV graft. The unadjusted early mortality HR was 0.51 (95% CI, 0.43-0.60) and case-mix plus laboratory-adjusted HR was 0.61 (95% CI, 0.50-0.74) for TOPs attendees (all outcomes, $P < 0.001$) (20).

- **1c.11/1c.21 System used for grading the body of evidence and grading the recommendation.** You can select USPSTF, GRADE, or other. Other is acceptable as long as you describe it. Please note, that grading the body of evidence is different from the strength of the recommendation. Some submissions indicated that the GRADE system was used, but the actual grades provided in items 1c.13/1c.23 were different than those described in GRADE documents. In that case, please clarify the differences (was the GRADE system used, modified, just different labels for the GRADE scale) and provide the rating scale with definitions.

KCQA RESPONSE

Vascular Access: Again, as we noted in the submission form, the KDOQI guidelines are based on a systematic review of the literature available at the time of publication. In developing the KDOQI clinical practice guidelines, the National Kidney Foundation utilizes experts to decide which recommendations are supported by evidence and which are supported by consensus of the Work Group opinion. Evidence-based guideline recommendations are graded as strong, moderate, or weak in an approach consistent with—but not identical to—the USPSTF and GRADE grading methods. Specifically, KDOQI Guidelines are assigned a grade of “A”, “B”, or “C” depending whether the recommendation is based on strong, moderately strong, or weak evidence that the practice improves health outcomes:

- Grade A = “It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.”
- Grade B = “It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.”
- Grade C = “It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.”

KDOQI grades the strength of its recommendation to use AVFs over other forms of vascular access in chronic hemodialysis patients as “B”, indicating that the recommendation is based on moderately strong evidence that the practice improves health outcomes. Clinicians are recommended to routinely follow the guideline for eligible patients.

Patient Education: There are no existing clinical practice guidelines addressing patient education on renal replacement modality options and the related body of evidence has not been formally graded. However, as previously noted on the submission form, application of the recommendations set forth by the NKF Evidence Task Force indicates that because there are RCTs of direct evidence with adequate size to obtain precise estimates of effect and without serious flaws that introduce bias, the quality of the studies composing the body of evidence upon which the KCQA patient education awareness measures are based can be judged as “high” when using a scale of High, Moderate, or Low.

2a. Reliability

Measure specifications. 2a1.3, 2a1.7, 2a1.9 Measures were originally specified for medical record abstraction or for physician measures CPT-II codes; however, you indicated that data will come from CROWNWeb. Does CROWNWeb include the data needed for your measures (e.g., patient education, seen by a vascular surgeon)? The detailed specifications should include the specific CROWNWeb data fields/numbers (and any standard response options if relevant).

KCQA RESPONSE:

Existing CROWNWeb data fields/numbers that could be applied for the collection of the KCQA vascular access measures include the following:

- CROWNWeb ID 225, Date Regular Chronic Dialysis Began
- CROWNWeb ID 226, Access Type for Dialysis
- CROWNWeb ID 227, Date Access Type for Dialysis Changed
- CROWNWeb ID 228, AVF Maturing

- CROWNWeb ID 229, AVF Creation Date
- CROWNWeb ID 230, AVF Usable Date
- CROWNWeb ID 231, AVF Fistula State
- CROWNWeb ID 232, AV Grate State

Only CMS is in a position to comment on their specific plans and timeframe to incorporate the “evaluated by a vascular surgeon” or “discussion of renal replacement modality options” data elements into CROWNWeb. We do note, however, that CMS has made clear in its communications with us and through publication in the 2008 list of CPMs its intent to include both the KCQA vascular access and patient education measures in CROWNWeb.

Reliability Testing. 2a2 – Reliability testing was based on chart abstraction. How does that apply to data from CROWNWeb? How does facility chart abstraction relate to physician measures collected using CROWNWeb data?

KCQA RESPONSE: As CROWNWeb is not yet functional, it was necessary to field test the KCQA measures based on chart abstraction. CMS has indicated that the necessary data for these physician-level measures will be collected at the facility level using the CROWNWeb interface. Performance can be linked to the attending nephrologist through use of the CROWNWeb Attending Practitioner data elements (CROWNWeb ID numbers 243-246). Additionally, we note that other than batch submitters, the reliability testing based on chart abstraction is congruent with reliability of data collection for CROWNWeb because facilities will be abstracting their charts and manually entering the information into the CROWNWeb interface—just as they manually abstracted their charts and entered it into the testing spreadsheets or paper collection forms.

3. Usability

3a1 – What is progress toward public reporting and any specific plans and timeline? Is CMS planning to report on patient education or physician performance on any topic?

KCQA RESPONSE: We noted in our submissions that in its list of Phase III CPMs released in April 2008, CMS indicated its intent to include both KCQA clinician-level vascular access measures and the KCQA clinician- and facility-level patient education measures in the CROWNWeb CPMs. We also note that CMS has informed us that it remains interested in implementing the KCQA measures. However, CMS, as the implementer, is the only organization in a position to comment on specific plans and the timeline for CROWNWeb (the data collection vehicle for public reporting of these measures).

VASCULAR ACCESS CITATIONS:

1. Perera G, Mueller M, Kubaska S, Wilson S, Lawrence P, Fujitani R. Superiority of autogenous arteriovenous hemodialysis access: maintenance of function with fewer secondary interventions. *Ann Vasc Surg.* 2004;18:66–73.
2. Huber T, Carter J, Carter R, Seeger J. Patency of autogenous and polytetrafluoroethylene upper extremity arteriovenous hemodialysis accesses: a systematic review. *J Vasc Surg.* 2003;38:1005–11.
3. Pisoni R, Young E, Dykstra D, et al. Vascular access use in Europe and the United States: results from the DOPPS. *Kidney Int.* 2002;61:305–16.
4. Mehta S. Statistical summary of clinical results of vascular access procedures for haemodialysis, in

Sommer BG, Henry ML (eds): Vascular Access for Hemodialysis-II (ed 2). Chicago, IL, Gore, 1991, pp 145–57.

5. The Cost Effectiveness of Alternative Types of Vascular access and the Economic Cost of ESRD. Bethesda, MD, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 1995, pp 139–57.
6. Eggers P, Milam R. Trends in vascular access procedures and expenditures in Medicare's ESRD program, in Henry ML (ed): Vascular Access for Hemodialysis-VII. Chicago, IL, Gore, 2001, pp 133–43.
7. Gulati S, Sahu K, Avula S, Sharma R, Ayyagiri A, Pandey C. Role of vascular access as a risk factor for infections in hemodialysis. *Ren Fail*. 2003;25:967–73.
8. Nassar G and Ayus J. Infectious complications of the hemodialysis access. *Kidney Int*. 2001;60:1–13.
9. Dhingra R, Young E, Hulbert-Shearon T, Leavey S, and Port F. Type of vascular access and mortality in U.S. hemodialysis patients. *Kidney Int*. 2001;60:1443–51.
10. Woods J and Port F. The impact of vascular access for haemodialysis on patient morbidity and mortality. *Nephrol Dial Transplant*. 1997;12:657–9.
11. Xue J, Dahl D, Ebben J, and Collins A. The association of initial hemodialysis access type with mortality outcomes in elderly Medicare ESRD patients. *Am J Kidney Dis*. 2003;42:1013–19.
12. Polkinghorne K, McDonald S, Atkins R, and Kerr P. Vascular access and all-cause mortality: a propensity score analysis. *J Am Soc Nephrol*. 2004;15:477–86.

PATIENT EDUCATION CITATIONS:

13. Devins G, Mendelssohn D, Barre P and Binik Y. Predialysis psychoeducational intervention and coping styles influence time to dialysis in chronic kidney disease. *Am J Kidney Dis*. 2003;42:693–703.
14. Devins G, Mendelssohn D, Barre P, Taub K and Binik Y. Predialysis psychoeducational intervention extends survival in CKD: a 20-year follow-up. *Am J Kidney Dis*. 2005;46:1088–98.
15. Cavanaugh K, Wingard R, Hakin R, Elas T and Ikizler T. Patient dialysis knowledge is associated with permanent arteriovenous access use in chronic hemodialysis. *Clin J Am Soc Nephrol*. 2009;4:950-56.
16. Beder J et al. Effectiveness of a social work psychoeducational program in improving adherence behavior associated with risk of CVD in ESRD patients. *J Neph Social Work*. 2003;22:12-22.
17. Johnstone S et al. Overcoming early learning barriers in hemodialysis patients: the use of screening and educational reinforcement to improve treatment outcomes. *Adv in Chronic Kid Dis*. 2004;11(2):210-16.
18. Stack A. Determinants of modality selection among incident U.S. dialysis patients: results from a national study. *J Am Soc Nephrol*. 2002;13(5):1279-87.
19. Wingard R, Pupim L, Krishnan M, Shintani A, Ikizler T and Hakim R. Early intervention improves mortality and hospitalization rates in incident hemodialysis patients: RightStart Program. *Clin J Am Soc Nephrol*. 2007;2:1170-75.
20. Lacson E, Wang W, DeVries C et al. Effects of a nationwide predialysis educational program on modality choice, vascular access, and patient outcomes. *Am J Kidney Dis*. 2011;58(2):235-42.

IMIS HEALTH

- **0570** Chronic Kidney Disease (CKD): monitoring phosphorus
- **0571** Chronic Kidney Disease: monitoring parathyroid hormone (pth)
- **0574** Chronic Kidney Disease (CKD): monitoring calcium

1c. Evidence

We identified some common issues across many of the submissions noted below. Please review your submissions and send us any clarifications. Please keep in mind that NQF does not expect the developer to conduct a primary evidence review; rather the developer is asked to report on the review/grading of body of evidence that was conducted by others – specifically on the quantity, quality, and consistency of the body of evidence.

- **1c6. Quality of the Body of the Evidence.** Most of the submissions did not address the quality of the body of evidence. What did the systematic review of the body of evidence determine about the overall quality of the body of evidence? We ask for a grade in item 1c13, but for this item, we want substantive information about the quality. If the review did not address the overall quality of the body of evidence, please state that and indicate what was identified about the quality of individual studies.

The Quality of the Body of the Evidence is low and is based on expert opinion. Despite lack of evidence from clinical trials, this recommendation is rated as a “strong” recommendation from KDIGO. The majority of studies demonstrating benefit come from trials involving populations of hemodialysis patients. However, a cross sectional population study done by Levin et al. consisting of 1800 patients found abnormalities in PTH, Calcium, and Phosphorous in the majority of patients with eGFR \leq 60. No existing trials exist to support early monitoring / intervention in CKD stage III. However, based on the clear benefits exist for treatment found in the ESRD subpopulation, it is strongly recommended by NKF and KDIGO to begin monitoring Calcium, Phosphorous, and PTH at CKD stage III, as the risk of this intervention is low and the benefit is high.

- **1c7. Consistency of Results across Studies.** Most of the submissions did not provide information on consistency of the magnitude and direction of effect across the studies in the body of evidence. For the outcomes studied, what was the magnitude and direction of effect? If a meta-analysis was conducted, the results would be important evidence. Information from evidence tables can be used to provide substantive information on effect size.

There are no clinical trials demonstrating efficacy of evaluation or treatment of PTH, Calcium, and Phosphorous in CKD stage III.

- **1c.11/1c.21 System used for grading the body of evidence and grading the recommendation.** You can select USPSTF, GRADE, or other. Other is acceptable as long as you describe it. Please note, that grading the body of evidence is different from the strength of the recommendation. Some submissions indicated that the GRADE system was used, but the actual grades provided in items 1c.13/1c.23 were different than those described in GRADE documents. In that case, please clarify the differences (was the GRADE system used, modified, just different labels for the GRADE scale) and provide the rating scale with definitions.

KDIGO rates this measure as Level 1 C. (Strong recommendation, low quality of evidence)

KDIGO distinguishes their recommendations as Level 1 or Level 2 grades.

A Level 1 grade is defined as a strong recommendation (“We recommend”). KDIGO intends this level of recommendation to be viewed from three points of view: From the patient’s standpoint, most people in this situation would want the recommended course of action and only a small proportion would not.

From the clinician’s standpoint, most patients should receive the recommended course of action. From a policy standpoint, this recommendation can be adopted as policy in most situations.

Level 2 grade is defined as a suggestion. From the patient’s standpoint, the majority of people in this situation would want the recommended course of action, but many would not. From the clinician’s standpoint, different choices will be appropriate for different patients and each patient needs help to arrive at a decision consistent with his or her values and preferences. From a policy standpoint, this recommendation is likely to require further discussion and debate before policy can be determined. KDIGO rates their level of evidence between A to D.

Grade A evidence is high. The true effect lies close to the estimate of its effect.

Grade B evidence is moderate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Grade C evidence is low. The true effect may be substantially different than the estimate of the effect.

Grade D evidence is very low. The estimate of the effect is very uncertain, and often will be far from the truth.

Reliability 2a

2a2 Reliability testing— What is the rationale for considering correlation of facility scores in two time periods as a demonstration of reliability of the data? NQF is asking for evidence of the reliability of either: 1) the data used in the measure (e.g., lab date, lab method, dialysis dose); or 2) the measure score (amount of variation due to true differences vs. error/noise – signal- to- noise analysis).

The reliability of administrative data is dependent on the incentive of the respondents to give the correct information. In general, administrative claims data is more reliable in the PPO setting than the HMO setting, because in the PPO setting claims are tied to reimbursement. HMOs that use administrative data measure can take steps to incentivize the submission of appropriate data fields.

The numerator of these measures assess whether laboratory testing for Ca, PO₄, or PTH were done. In the PPO setting, for patients not on dialysis, Ca and PO₄ testing will be done by a routine laboratory, which would be highly incentivized to give the information, since reimbursement is dependent upon reporting. For patients on dialysis, Ca and PO₄ may be checked during dialysis and not be reported in administrative data as individual tests but bundled with the dialysis codes. This is why in the construction of these two measures we excluded dialysis patients. However, in the case of PTH, this testing would be done by a routine laboratory, since dialysis centers do not have the capacity to this test. Thus, we did not exclude dialysis patients in the PTH measure.

Although it is possible for providers to commit fraud and bill for test or procedure when it was not administered, we believe this is less likely to occur with laboratory testing.

In addition, to investigate the test-retest reliability of the data, we investigated the correlation of providers’ scores (who met minimal denominator threshold of 25) between two years for two plans in the PPO setting. We found that the test-retest reliability is high based on Pearson’s correlation coefficient ($c > 0.7$) [Table 1].

Table 1: Test-retest reliability

	Pearson's correlation coefficient for plan A	Pearson's correlation coefficient for plan B
CKD calcium	0.814	0.706
CKD PTH	0.825	0.897
CKD PO4	0.816	0.963

Validity 2b

2b.2 Validity testing –What is the rationale for comparing plan level scores to a range found in the literature as a demonstration of validity (accuracy) of the data elements? NQF is asking for evidence of the validity of either: 1) the data used in the measure (e.g., accuracy of the lab date, lab method, dialysis dose); or 2) the measure score (correctness of conclusion about quality). Can you provide evidence of accuracy of patient-level data?

To assess for the validity of our measures, we first assess whether the construction of the measures can accurately capture the denominator, exclusion, and numerator.

Accuracy of Denominator

The IMS denominator algorithm to identify patient with chronic renal disease (stage III or greater, $GFR < 60\text{ml/min/1.73 m}^2$) is consistent with algorithms in the literature which were able to identify patients with CKD with > 97% specificity when compared to data obtained by other sources (i.e, creatinine values).[1-3]

Accuracy of Exclusion

For the calcium and phosphorus measures, we excluded patients who were on dialysis and hospitalized during the numerator time period. We cannot accurately capture the receipt of Ca and PO4 laboratory tests because of bundled coding. The accuracy of capturing patients who were dialyzed or hospitalized is high because hospitalizations and dialysis are costly and reimbursement for these events would be completely dependent upon the providers submitting the claims. It would be difficult to submit fraudulent claims for hospitalization or dialysis because of the high cost of these services. Health plans scrutinize these patients by conducting utilization reviews and enrolling these patients in disease management programs.

Accuracy of Numerator

The accuracy of assessing whether Ca, PO4, and PTH tests were done by routine laboratory would also be high because submitting these claims would be tied to reimbursement, especially in the PPO setting.

Although it is possible for providers to submit fraudulent claims for factitious tests or procedures, we believe this is less likely to occur in the laboratory testing.

Second, we compared the rates obtained using our algorithm to those reported by the rates found in the literature to assess the validity of these measures. We found that the measure rates we obtained with our data sets were in line with the rates found in the literature (see original submission).

3. Usability

3.1 Endorsed measure (though only 18 months), but not publicly reported. 3a.1 What is plan and timeline for public reporting?

These measures can be used both in public reporting and pay-for-performance programs. As we do not own data ourselves, we cannot publicly report results, however, we encourage our clients (who do own the data) to utilize these measures for public reporting.

1. Kern, E.F., et al., *Failure of ICD-9-CM codes to identify patients with comorbid chronic kidney disease in diabetes*. Health Serv Res, 2006. **41**(2): p. 564-80.
2. Rector, T.S., et al., *Specificity and sensitivity of claims-based algorithms for identifying members of Medicare+Choice health plans that have chronic medical conditions*. Health Serv Res, 2004. **39**(6 Pt 1): p. 1839-57.
3. Birman-Deych, E., et al., *Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors*. Medical Care, 2005. **43**(5): p. 480-485.

AMA/PCPI

- **1662** Angiotensin Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy
- **1660** ESRD Patients Receiving Dialysis: Hemoglobin Level <10g/dL
- **1666** Patients on Erythropoiesis Stimulating Agent (ESA)--Hgb Level > or = 12g/dL
- **1667** (Pediatric) ESRD Patients Receiving Dialysis: Hemoglobin Level < 10g/dL
- **1633** Blood Pressure Management
- **1668** Laboratory Testing (Lipid Profile)
- **0323** Hemodialysis Adequacy: Solute
- **0321** Peritoneal Dialysis Adequacy: Solute

1c. Evidence

We identified some common issues across many of the submissions noted below. Please review your submissions and send us any clarifications. Please keep in mind that NQF does not expect the developer to conduct a primary evidence review; rather the developer is asked to report on the review/grading of body of evidence that was conducted by others – specifically on the quantity, quality, and consistency of the body of evidence.

- **1c6. Quality of the Body of the Evidence.** Most of the submissions did not address the quality of the body of evidence. What did the systematic review of the body of evidence determine about the overall quality of the body of evidence? We ask for a grade in item 1c13, but for this item, we want substantive information about the quality. If the review did not address the overall quality of the body of evidence, please state that and indicate what was identified about the quality of individual studies.

1660 ESRD Patients Receiving Dialysis: Hemoglobin Level <10g/dL

The description of the evidence review within the guideline, did not address the overall quality of the body of evidence related to this measure nor was any grade provided for the quality of the body of evidence. Therefore, the text that was included in the form, was added to describe the study design/flaws, directness of the evidence to the measure, and any imprecision within the studies.

1666 Patients on Erythropoiesis Stimulating Agent (ESA)--Hgb Level > or = 12g/dL

The guideline refers to the quality of the body of evidence related to this measure by including the GRADE scale and also by stating the following:

The quality of a body of evidence pertaining to a particular outcome of interest was initially categorized based on study design. For questions of interventions, the initial quality grade is “high” if the body of evidence consists of RCTs, “low” if it consists of observational studies, or “very low” if it consists of studies of other study designs. The grade for the quality of evidence for each intervention/outcome pair was then decreased if there were limitations to the method quality of the aggregate of studies, if there were inconsistencies in the results across studies, if there was uncertainty about the directness of evidence including limited applicability of the findings to the population of interest, if the data were imprecise or sparse, or if there was thought to be a high likelihood of reporting bias. The final grade for the quality of the evidence for an intervention/outcome pair could be one of the following 4 grades: “high,” “moderate,” “low,” or “very low.”

The text provided with respect to the grading of the body of evidence for this particular measure was provided in the form and is also provided below.

In appraising the overall evidence, the Work Group considered mortality, cardiovascular events, and HRQoL as outcomes of high importance. The Work Group rated the evidence showing a trend toward greater cardiovascular events in dialysis and nondialysis patients assigned to Hb targets greater than 13.0 g/dL to be of moderately high quality for showing harm and of high quality for showing lack of benefit. The Work Group considered the HRQoL benefits in patients assigned to higher Hb targets as low quality evidence based on the limitations of reported HRQoL evidence (see the following section, Limitations of Evidence). The conclusion that in dialysis and nondialysis patients with CKD receiving ESA therapy, the Hb target should not be greater than 13.0 g/dL reflects the Work Group's judgment that the possibility to cause harm weighs more heavily than the potential to improve quality of life and to decrease transfusions.

There is no other mention of grading of the overall quality of the body of evidence.

1667 (Pediatric) ESRD Patients Receiving Dialysis: Hemoglobin Level < 10g/dL

The description of the evidence review within the guideline was provided within the form. The additional text that was included in the form, was added to describe the study design/flaws, directness of the evidence to the measure, and any imprecision within the studies.

0323 Hemodialysis Adequacy: Solute

The description of the evidence review within the guideline, did not address the overall quality of the body of evidence related to this measure nor was any grade provided for the quality of the body of evidence. Therefore, the text that was included in the form, was added to describe the study design/flaws, directness of the evidence to the measure, and any imprecision within the studies.

0321 Peritoneal Dialysis Adequacy: Solute

The guideline provided a grade for the strength/quality of the body of evidence related to this measure. The grade, moderately strong, was provided, as well as details of the scale used by the guideline developer. Therefore, the text that was included in the form, was added to describe the study design/flaws, directness of the evidence to the measure, and any imprecision within the studies.

1633 Blood Pressure Management

The guideline did not provide a summary of the overall quality of the evidence. However, the information provided in the guideline about individual studies reviewed resulted in the Guideline Development and Evidence Review team rating the evidence as strong (for patients with chronic kidney disease [CKD] being considered in the "highest-risk" group for cardiovascular disease [CVD] for implementing recommendations for pharmacological therapy) and moderately strong (for the Blood Pressure target of < 130/80 mm Hg and Plan of Care components). Regarding the individual studies, very detailed information is provided on the NQF submission form.

1662 Angiotensin Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy

The guideline did not provide a summary of the overall quality of the evidence. However, the information provided in the guideline about individual studies reviewed resulted in the Guideline Development and Evidence Review team rating the evidence as "strong," indicating that the evidence includes results from well-designed, well-conducted study/studies in the target population

that directly assess effects on health outcomes. Regarding the individual studies, very detailed information is provided on the NQF submission form.

1668 Laboratory Testing (Lipid Profile)

The guideline did not provide a summary of the overall quality of the evidence. However, the information provided in the guideline about individual studies reviewed resulted in the Guideline Development and Evidence Review team rating the evidence as “moderately strong.” Regarding the individual studies, very detailed information is provided on the NQF submission form.

- **1c7. Consistency of Results across Studies.** Most of the submissions did not provide information on consistency of the magnitude and direction of effect across the studies in the body of evidence. For the outcomes studied, what was the magnitude and direction of effect? If a meta-analysis was conducted, the results would be important evidence. Information from evidence tables can be used to provide substantive information on effect size.

1660 ESRD Patients Receiving Dialysis: Hemoglobin Level <10g/dL

The description of the evidence review within the guideline, did not address the overall consistency of results across studies. The text provided in the form is taken from the guideline and describes the consistency and difference among studies reviewed to form the guideline recommendation.

1666 Patients on Erythropoiesis Stimulating Agent (ESA)--Hgb Level > or = 12g/dL

The description of the evidence review within the guideline, did not address the overall consistency of results across studies. The text provided in the form is taken from the guideline and describes the consistency and difference among studies reviewed to form the guideline recommendation.

1667 (Pediatric) ESRD Patients Receiving Dialysis: Hemoglobin Level < 10g/dL

The description of the evidence review, within the guideline, did not address the overall consistency of results across studies. The text provided in the form is taken from the guideline and describes the consistency and difference among studies reviewed to form the guideline recommendation.

0323 Hemodialysis Adequacy: Solute

The description of the evidence review, within the guideline, did not address the overall consistency of results across studies. The text provided in the form is taken from the guideline and describes the consistency and differences among studies reviewed to form the guideline recommendation.

0321 Peritoneal Dialysis Adequacy: Solute

The description of the evidence review, within the guideline, did not address the overall consistency of results across studies. The text provided in the form is taken from the guideline and describes the consistency and differences among studies reviewed to form the guideline recommendation.

1633 Blood Pressure Management

Our analysis of the Guideline Development and Evidence Review Team’s review of the evidence is that the results across studies are consistent, and, based on their detailed explanation (provided on the NQF submission form), it seems logical to extrapolate recommendations for patients with CKD from clinical studies performed in the general population.

1662 Angiotensin Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy

Our analysis of the Guideline Development and Evidence Review Team's review of the evidence is that the results across studies are consistent. Furthermore:

- An individual patient meta-analysis of 646 patients in 10 randomized clinical trials confirmed the results of prior studies. Consequently, the Guideline Development Work Group concluded that ACE inhibitors and ARBs are preferred agents for diabetic kidney disease with microalbuminuria and should be prescribed for patients with or without hypertension.
- Two meta-analyses have demonstrated a greater effect of ACE inhibitors compared to other classes of antihypertensive agents on reducing proteinuria in diabetic and nondiabetic kidney disease. Other studies show a larger effect of ARBs compared to other classes.
- Supporting these conclusions, a meta-analysis by the ACE Inhibition of Progressive Renal Disease (AIPRD) Study Group of patient-level data on 1,860 nondiabetic patients enrolled in 11 RCTs of various ACE inhibitors found that the ACE inhibitor group had better blood pressure control, lower urine protein excretion, and an approximately 30% reduction in the risk of kidney failure, as well as a reduction in the combined endpoint of doubling of serum creatinine or onset of kidney failure.

1668 Laboratory Testing (Lipid Profile)

Our analysis of the Guideline Development and Evidence Review Team's review of the evidence is that the results across studies are consistent, and, based on their detailed explanation (provided on the NQF submission form), it seems logical to extrapolate recommendations for patients with CKD from clinical studies performed in the general population. In addition:

- Unfortunately, there are no large, adequately powered, randomized, controlled trials testing the hypothesis that treatment of dyslipidemia preserves kidney function. However, there have been several small studies, and a meta-analysis of these studies. This meta-analysis included prospective, controlled trials published before July 1, 1999.
- **1c.11/1c.21 System used for grading the body of evidence and grading the recommendation.** You can select USPSTF, GRADE, or other. Other is acceptable as long as you describe it. Please note, that grading the body of evidence is different from the strength of the recommendation. Some submissions indicated that the GRADE system was used, but the actual grades provided in items 1c.13/1c.23 were different than those described in GRADE documents. In that case, please clarify the differences (was the GRADE system used, modified, just different labels for the GRADE scale) and provide the rating scale with definitions.

1660 ESRD Patients Receiving Dialysis: Hemoglobin Level <10g/dL

The guideline did not provide a grade for the body of evidence related to this measure.

1666 Patients on Erythropoiesis Stimulating Agent (ESA)--Hgb Level > or = 12g/dL

The text provided with respect to the grading of the body of evidence for this particular measure was provided in the form and is also provided below.

In appraising the overall evidence, the Work Group considered mortality, cardiovascular events, and HRQoL as outcomes of high importance. The Work Group rated the evidence showing a trend toward greater cardiovascular events in dialysis and nondialysis patients assigned to Hb targets greater than 13.0 g/dL to be of moderately high quality for showing harm and of high quality for showing lack of benefit. The Work Group considered the HRQoL benefits in patients assigned to higher Hb targets as low quality evidence based on the limitations of reported HRQoL evidence (see the following section, Limitations of Evidence). The conclusion that in dialysis and nondialysis

patients with CKD receiving ESA therapy, the Hb target should not be greater than 13.0 g/dL reflects the Work Group's judgment that the possibility to cause harm weighs more heavily than the potential to improve quality of life and to decrease transfusions.

There is no other mention of grading of the overall quality of the body of evidence related to this measure.

1667 (Pediatric) ESRD Patients Receiving Dialysis: Hemoglobin Level < 10g/dL

The guideline did not provide a grade for the body of evidence related to this measure.

0323 Hemodialysis Adequacy: Solute

The guideline did not provide a grade for the body of evidence related to this measure.

0321 Peritoneal Dialysis Adequacy: Solute

The body of evidence related to this measure was graded, and the grade provided was moderately strong, as indicated in the form. The system used to grade the evidence was described in the form, where indicated.

1633 Blood Pressure Management

For both the system for grading the body of evidence and grading the recommendations, new labels and definitions were provided, other than USPSTF or the GRADE system. The systems used by K/DOQI are both described on the NQF submission form.

1662 Angiotensin Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy

For both the system for grading the body of evidence and grading the recommendations, new labels and definitions were provided, other than USPSTF or the GRADE system. The systems used by K/DOQI are both described on the NQF submission form.

1668 Laboratory Testing (Lipid Profile)

For both the system for grading the body of evidence and grading the recommendations, new labels and definitions were provided, other than USPSTF or the GRADE system. The systems used by K/DOQI are both described on the NQF submission form.

Measures of specific hemoglobin value (1660, 1666, 1667).

How does the recent FDA announcement regarding ESAs and hemoglobin values influence whether there is sufficient evidence to support the measure focus (see [FDA Safety Communication](#) 6/24/11 and 6/27/11)?

After considering the announcement and discussing the new recommendations amongst our AMA-PCPI/RPA team, we are in agreement that our measures do not need to be changed or withdrawn from NQF consideration. Our Hemoglobin < 10 measures are irrespective of ESA use and seem to be in line with the ESA announcement. We believe that our Hemoglobin > 12 measure is still consistent with the new recommendations of:

For patients with the anemia of chronic kidney disease NOT on dialysis

- Consider starting ESA treatment only when the hemoglobin level is less than 10 g/dL and when certain other considerations apply.
- If the hemoglobin level exceeds 10 g/dL, reduce or interrupt the dose of ESA.

For patients with the anemia of chronic kidney disease on dialysis

- Initiate ESA treatment when the hemoglobin level is less than 10 g/dL.
- If the hemoglobin level approaches or exceeds 11 g/dL, reduce or interrupt the dose of ESA.

We believe that our measures are still consistent with the FDA, in that a patient on dialysis having a Hgb level > 11 should trigger a change in ESA treatment, but if the patient's Hgb got to a level > 12, then it would be indicative of poor care (a truly bad outcome). It seems that this same rationale would apply for patients NOT on dialysis, although a change in ESA treatment should be triggered when the patient's Hgb got to a level > 10 (according to the FDA). Therefore, our Hgb > 12 measure's value seems to be in catching egregious overuse of ESA therapy.

The upper value measure (Hgb>12) is clearly a safety measure. Given the current financial reimbursement system, there will be multiple forces leading to inadequate treatment of anemia, so some measure looking at the lower value (Hgb<10) is needed. We could have lengthy discussions as to whether the lower threshold should be 9, 9.5 or 10. However, AMA-PCPI/RPA does not anticipate and is not implying that all patients should have a serum Hgb >10.

As measures of intermediate clinical outcome there is a potential need for risk adjustment? What analysis and/or rationale supports that there is no need for risk adjustment or that exclusions are relevant to the question of risk adjustment?

The PCPI partially accounts for risk adjustment through the inclusion of exceptions. Risk factors such as those identified in 2a1.8 (reasons for exceptions) would be among those variables included in a more fully specified risk adjustment model. The PCPI recommends that clinicians document the specific reasons for exclusion in patients' medical records for purposes of optimal patient management and audit-readiness.

2a. Reliability

Measure specifications.

Why are draft PCPI eSpecifications being submitted? Some are labeled draft, they are not in the format HL7 HQMF format, and the measure has not been tested as an electronic measure.

The eSpecifications are stamped "Draft" because at the time the measures were submitted to NQF for endorsement consideration, the measures were not yet approved by the PCPI membership. The voting period concludes on August 16. PCPI can submit an updated version once the measures have been approved.

The eSpecifications are not in the HL7 HQMF eMeasure format, because currently there is no tool available to facilitate the development of the HQMF eMeasure by measure developers. Our understanding is that a Measure Authoring Tool is under development, although not yet available. The eSpecifications, however, do include the same content that would be included in the human-readable

version of the XML specification. In particular, the data elements required for the measure are mapped to the Quality Data Model version 2.1; there is a human readable representation of the measure, including the text description, and graphical representation of the measure logic, including relative timing components. The eSpecification documents the requirements to collect the measure information from an electronic health record as the data source. The testing data included on the NQF submission forms includes the testing of the measures using EHRs as a data source.

2a1.8 Exclusions (exceptions) for medical or patient reason are not precisely specified. Are the examples specified exclusions?

The appropriate method and level of granularity for exception coding is an ongoing debate, internal and external to the AMA-PCPI. Historically, the AMA-PCPI approach for representing exceptions is to include the CPT® Category II code modifier approach developed by the AMA-PCPI and CPT Performance Measures Advisory Group (PMAG) for all measure specifications. Over the past year, the PCPI—in anticipation of quality reporting via EHRs—has begun to code the specific examples listed within each of the three broad categories of measure exceptions, and these coded examples are included in the eSpecifications for the Renal measures. In these cases, we provide the relevant codes from the various clinical terminologies as part of our eSpecifications. There continues to be work done to determine how to capture the “other medical, patient, system reasons,” however the most likely reasons for the exceptions have been precisely specified, and we do consider the examples to be specified exceptions for the measure.

The measure exception categories are not available uniformly across all measures; for each measure, there must be a clear rationale to permit an exception for a medical, patient, or system reason. Where possible, examples have been provided in the measure exception language of instances that would constitute an exception. Examples are intended to guide clinicians and are not all-inclusive lists of all possible reasons why a patient could be excluded from a measure.

Reliability Testing. 2a2 What is the rationale for testing reliability of chart abstraction when the measure was implemented using CPT-II codes on claim forms? What do the results indicate about the reliability of the data used in the measures as implemented?

The purpose of using inter rater reliability testing is to evaluate whether the CKD/ESRD measure definitions and specifications yield stable, consistent measurements. The measures have also been specified for use in EHR. The 2 abstractors were able to pull the appropriate data elements from the electronic medical record and they each reviewed the same sample of charts. A percent-agreement between abstractors and the Kappa statistic (to adjust for chance agreement) are derived from the results of these chart reviews. The testing results show that the measures are highly reliable with excellent agreement beyond chance.

Validity 2b

2b3. What analysis demonstrates validity of exclusions across physicians who identify patients who should be excluded from measurement?

Each of the 4 testing sites in the AMA PCPI testing project (details on sampling methods in original submission) were studied to determine clinical appropriateness of reported exceptions. The exceptions were validated upon manual review of the medical record, against an a priori list generated by expert

opinion. Exceptions were determined to be clinically appropriate. Our original NQF submissions included the verbatim documentation for exceptions for each of the measures.

2b5. What were the performance scores for testing sites? What was the number of physicians and patients in the PQR data?

We do not assess the performance rates or generate performance scores for the individual testing sites as an element of our testing projects because this assessment is out of the scope of the project. Instead, performance results are calculated for each measure in the set and measure rates are calculated both with exceptions and without exceptions so that we can assess its ability to measure what is intended. The information below shows the performance results from both, the CKD and ESRD projects.

MEASURE PERFORMANCE RESULTS					
END STAGE RENAL DISEASE					
	Measure rate with exceptions		Measure rate without exceptions		Exception rate
<i>Influenza Immunization</i>	68 of 169	40%	68 of 169	40%	0%
<i>Vascular Access-Patients Receiving HD</i>	84 of 96	88%	84 of 107	79%	10%
<i>Plan of Care for Anemia</i>	NA	NA	1481 of 2012	74%	NA
<i>Plan of Care for Inadequate HD</i>	NA	NA	755 of 1109	68%	NA
<i>Plan of Care for Inadequate PD</i>	NA	NA	43 of 62	69%	NA
CHRONIC KIDNEY DISEASE					
	Measure rate with exceptions		Measure rate without exceptions		Exception rate
<i>Blood Pressure Management</i>	NA	NA	573 of 674	85%	NA
<i>ACE Inhibitor or ARB Therapy</i>	50 of 58	86%	50 of 71	70%	18%
<i>Laboratory Testing (Ca++, P, iPTH and Lipid Profile)</i>	75 of 112	67%	75 of 112	67%	0%
<i>Plan of Care for Anemia</i>	NA	NA	463 of 578	80%	NA
<i>Influenza Immunization</i>	26 of 110	24%	26 of 112	23%	2%
<i>Referral for Evaluation for AV Fistula</i>	35 of 98	36%	35 of 112	31%	13%

NA = no applicable exceptions

The data from the Confidential CMS PQRI 2008 Performance Information by Measure Jan-Sep TAP file shows the number of physicians reporting in the denominator for each of the measures. The data does not show information about the number of patients.

Performance Measure	Number of Physicians Reporting in the 2008 PQRI Data
END STAGE RENAL DISEASE	
#79 Influenza Immunization	24,684
#78 Vascular Access-Patients Receiving HD	38,127
#80 Plan of Care for Anemia	179,197
#81 Plan of Care for Inadequate HD	160,065

#82 Plan of Care for Inadequate PD	6,312
CHRONIC KIDNEY DISEASE	
#120 ACE Inhibitor or ARB Therapy in Patients with CKD	713
#121 Laboratory Testing (Calcium, Phosphorous, iPTH and Lipid Profile)	5,829
#122 Blood Pressure Management	45,814
#123 Plan of Care Anemia	11,979

3. Usability

3.1 Indicated measures are currently used in public reporting and professional certification. 3a1 Is individual physician performance data available to the public? If not, what are plans and timeline for public reporting? 3.2 What certification programs are using this measure?

Some of the measures are used in PQRS (public reporting). Individual physician performance is not available at this time, but CMS is working on a "Physician Compare" website. Section 10331 (a)(2) of the Affordable Care Act also requires that, no later than January 1, 2013, and with respect to reporting periods that begin no earlier than January 1, 2012, we implement a plan for making information on physician performance publicly available through the Physician Compare website.

The PCPI is currently working towards having the measures incorporated into a Maintenance of Certification program.

Amgen

- **1655** ESRD patients with PTH >400pg/mL and not treated with a calcimimetic or vitamin D analog.
- **1658** ESRD patients with PTH <130pg/mL and continued treatment with a calcimimetic or vitamin D analog

1c. Evidence

We identified some common issues across many of the submissions noted below. Please review your submissions and send us any clarifications. Please keep in mind that NQF does not expect the developer to conduct a primary evidence review; rather the developer is asked to report on the review/grading of body of evidence that was conducted by others – specifically on the quantity, quality, and consistency of the body of evidence.

- **1c6. Quality of the Body of the Evidence.** Most of the submissions did not address the quality of the body of evidence. What did the systematic review of the body of evidence determine about the overall quality of the body of evidence? We ask for a grade in item 1c13, but for this item, we want substantive information about the quality. If the review did not address the overall quality of the body of evidence, please state that and indicate what was identified about the quality of individual studies.

The measures proposed by Amgen utilize laboratory determinations of the concentration of parathyroid hormone (PTH) in serum or plasma. Disturbances in the regulation of the function of the parathyroid glands, the principal source of PTH in the circulation, are highly prevalent among patients with end-stage renal disease (ESRD). These disturbances often result in very high or very low concentrations of PTH in serum or plasma, one of several laboratory abnormalities that represent an integral component of the recently defined syndrome of chronic kidney disease-mineral and bone disorder, or CKD-MBD. Results from PTH testing are the most widely recognized and most extensively employed biochemical index of the disease of secondary hyperparathyroidism (SHPT) among those with ESRD, and they are used also to exclude the diagnosis of SHPT and to identify patients with relative hypoparathyroidism, or adynamic renal osteodystrophy. Both disorders figure prominently as discrete aspects of CKD-MBD.

Measurements of PTH are available to nearly all clinicians who provide medical care to patients with ESRD. They are essential to inform clinical decisions about the diagnosis and management of SHPT and CKD-MBD. Serial measurements over time provide important information about disease progression, the response to treatment, and expected disease-specific outcomes that affect the overall health of patients with ESRD.

The most extensive systemic assessment of evidence pertaining to PTH measurements and their value in the diagnosis and management of SHPT in the context of CKD-MBD among patients with ESRD was presented in 2009. The evidence review was included in the clinical practice guidelines developed by the Working Group of Kidney Disease: Improving Global Outcomes (KDIGO®), an initiative of the National Kidney Foundation. Evidence from published studies was evaluated using the GRADE system. This set of global recommendations and the evidence review provided in the KDIGO® document were subsequently endorsed by the Kidney Dialysis Outcomes Quality Initiative (KDOQI™) in the United States. The report from each group indicated that there was an overall paucity of evidence from randomized clinical trials to provide specific recommendations about a specific therapeutic target range for PTH in the management of patients with ESRD. Many of the studies evaluated during the review of

evidence were noted to have important methodological shortcomings. The randomized clinical trials available for review did not provide clear evidence of risk modification and/or therapeutic benefit for major patient-level outcomes such as death, hospitalization, cardiovascular events, or skeletal fracture. Fairly strict criteria were used, however, by the Working Group with respect to study duration and the numbers of subjects enrolled when selecting studies for consideration.

Definitive conclusions about the relative risk of important clinical outcomes also could not be reached after an assessment of results from observational research using values of 2.0 and 0.5 to define clinically meaningful thresholds for increases or decreases in relative risk. Despite these shortcomings, both the KDIGO® and the KDOQI™ clinical practice guidelines readily acknowledge that PTH measurements are required to identify patients with SHPT in the setting of CKD-MBD and that serial measurements are needed to inform ongoing clinical management. Moreover, results from observational studies demonstrate a consistent increase in mortality risk among patients with ESRD when PTH concentrations are markedly elevated or substantially reduced. These observations served as the basis for defining upper and lower threshold values for PTH to define levels of extreme risk.

Amgen agrees generally with the statements from KDIGO® and KDOQI™ about the overall quality of the published evidence about CKD-MBD and its limitations as a basis for developing robust clinical practice guidelines. It is our view, however, that the body of evidence indicating that SHPT is a progressive disorder, which increases in severity over time, is much stronger. For this specific disease-state issue, the information available is more compelling.

- **1c7. Consistency of Results across Studies.** Most of the submissions did not provide information on consistency of the magnitude and direction of effect across the studies in the body of evidence. For the outcomes studied, what was the magnitude and direction of effect? If a meta-analysis was conducted, the results would be important evidence. Information from evidence tables can be used to provide substantive information on effect size.

In observational studies, estimates of the relative risk for mortality range from 1.2 to 1.8 among patients with ESRD and elevated PTH concentrations compared with those with PTH values in a range considered appropriate for this population. Mortality risk increases progressively at incrementally higher PTH concentrations among patients with elevated values. Among patients with low PTH concentrations, estimates of relative mortality risk range from 1.4 to more than 2.0, and they are highest among patients with the lowest PTH concentrations. Meta-analyses that have examined the relationship between PTH concentrations and mortality risk among patients with chronic kidney disease (CKD) are likely to be flawed on the basis of an underlying assumption that the risk-relationship is linear across the entire population and because the groups of patients with CKD included in the analysis were quite heterogeneous.

When considering patients with ESRD and SHPT, both the duration of CKD and the number of years of treatment with dialysis, or dialysis vintage, have been identified consistently as predictors of disease severity as judged by the concentration of PTH in serum or plasma. Additionally, both the duration of CKD and dialysis vintage are associated with higher annual rates of surgical parathyroidectomy. This observation has been made repeatedly in the dialysis population in the United States, providing evidence of the progressive nature of the disease of SHPT.

- **1c.11/1c.21 System used for grading the body of evidence and grading the recommendation.** You can select USPSTF, GRADE, or other. Other is acceptable as long as you describe it. Please

note, that grading the body of evidence is different from the strength of the recommendation. Some submissions indicated that the GRADE system was used, but the actual grades provided in items 1c.13/1c.23 were different than those described in GRADE documents. In that case, please clarify the differences (was the GRADE system used, modified, just different labels for the GRADE scale) and provide the rating scale with definitions.

Tom, we may need help from Robyn and/or Lisa on this topic.

Reliability 2a

Specifications CROWNWeb is listed as data source (2a1.27-29) but no specific CROWNWeb fields or definitions are provided for numerator, denominator (2a1.3/2a1.7).

2a2 Reliability testing was not conducted, which is acceptable according to NQF guidance IF validity testing at data element was conducted (see below). If enough data available, can you assess the reliability of performance scores (signal-to-noise)?

In reviewing the plans for CROWNWeb, Amgen noticed that the PTH level data element is slated for collection through CROWNWeb along with several other items related to prescription etc for "Vitamin D analog." We sent a request to Tom Dudley in CMS in April 2011 requesting clarification from CMS that this will be implemented and/or re-labeled as "Vitamin D analog and/or calcimimetic," since this then appropriately covers the full range of therapeutic options.

Validity 2b

2b2 Validity testing at the data element level for EHR data generally involves analyzing agreement between data extracted electronically to visual review/abstraction from medical record or simulated testing with known values. Identifying aggregate statistics for LDOs and DOPPS data does not provide evidence of data element validity. Can you provide any other evidence of validity?

The data used to develop the proposed PTH measures were obtained directly from electronic medical records maintained by a provider of dialysis services. No physical chart abstraction was required, and no additional validation has been performed because the information represents a portion of the actual clinical medical record. It is the same information used to inform clinical decisions, and it accurately reflects the information sent to CROWNWeb by dialysis providers.

3a. Usability

3a.1. Indicated measure currently used for QI only but intended for public use by CMS; what are the projected plans and dates for public reporting?

The recommendation is to include the proposed PTH measures for reporting no later than January 1, 2012.

Updates to the NQF Submissions for the CMS ESRD Quality Measures

August 10, 2011

Contents

Measure Specifications (2a1)	2
Measure #0252: Assessment of Iron Stores.....	2
Measure #0253: Peritoneal Dialysis Adequacy- Measurement of Total Solute Clearance At Regular Intervals	4
Measure #0254: Peritoneal Dialysis Adequacy- Calculate Weekly KT/V_{urea} in the Standard Way	4
Measure #0318: Peritoneal Dialysis Adequacy- Delivered Dose of Peritoneal Dialysis Above Minimum	6
Measure #0256: Hemodialysis Vascular Access- Minimizing Use of Catheters as Chronic Dialysis Access	7
Measure #0257: Hemodialysis Vascular Access- Maximizing Placement of Arterial Venous Fistula (AVF)	8
Evidence (1c)	9
Measure #0252: Assessment of Iron Stores.....	10
Measure #0247: Hemodialysis Adequacy Clinical Performance Measure I: Hemodialysis Adequacy- Monthly Measurement of Delivered Dose	13
Measure #0248: Hemodialysis Adequacy Clinical Performance Measure II: Method of Measurement of Delivered Hemodialysis Dose.....	14
Measure #0249: Hemodialysis Adequacy Clinical Performance Measure III: Method of Measurement of Delivered Hemodialysis Dose.....	16
Measure #0253: Peritoneal Dialysis Adequacy- Measurement of Total Solute Clearance at Regular Intervals.....	17
Measure #0254: Peritoneal Dialysis Adequacy- Calculate Weekly Kt/V_{urea} in the Standard Way.....	18
Measure #0318: Peritoneal Dialysis Adequacy- Delivered Dose of Peritoneal Dialysis above Minimum.....	19
Measure #0255: Measurement of Serum Phosphorus Concentration	21
Measure #0261: Measurement of Serum Calcium Concentration	21
Measure #0256: Hemodialysis Vascular Access- Minimizing Use of Catheters as Chronic Dialysis Access	22
Measure #0257: Hemodialysis Vascular Access- Maximizing Placement of Arterial Venous Fistula (AVF)	23
Reliability Testing (2a2.2)	25
Validity Testing (2b2.3)	26
Measure #0256: Hemodialysis Vascular Access- Minimizing Use of Catheters as Chronic Dialysis Access	26
Measure #0253: Peritoneal Dialysis Adequacy- Measurement of Total Solute Clearance At Regular Intervals	27
Measure #0254: Peritoneal Dialysis Adequacy- Calculate Weekly KT/V_{urea} in the Standard Way	27

Measure #0318: Peritoneal Dialysis Adequacy- Delivered Dose of Peritoneal Dialysis above Minimum.....	28
Usability - Current Use (3.1).....	29
Measure #0369: Dialysis Facility Risk-Adjusted Standardized Mortality Ratio Level	29
Measure Title (De.1).....	29
Risk Adjustment Strategy and Testing (2b4)	29
Risk Adjustment Testing Results (2b4.3)	32

Measure Specifications (2a1)

***Question:** Be sure to have the specific CROWNWeb data fields/numbers identified (and any standard response options if relevant). This was done in most but not all the measures. If data from other sources such as claims, please be specific.*

Measure #0252: Assessment of Iron Stores

2a1.2. Numerator Time Window

Data collected for this ESRD CPM are for a three-month time period.

2a1.3. Numerator Details

Number of patients in the denominator for whom “Serum Ferritin” and “Serum Ferritin Collection Date” are populated AND EITHER “Iron Saturation (TSAT) Percentage” and “Iron Saturation (TSAT) Percentage Collection Date” are populated OR “Reticulocyte Hemoglobin (CHr)” and “Reticulocyte Hemoglobin (CHr) Collection Date” are populated at least once during a three-month reporting period for in-center hemodialysis patients, peritoneal dialysis patients, and home hemodialysis patients.

2a1.6. Denominator Time Window

Data collected for this ESRD CPM are for a three-month time period.

2a1.7. Denominator Details

The patient’s age will be determined by subtracting the patient’s date of birth from the first day of the reporting month. Hemodialysis patients are defined as follows: “Admit Date” to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged (“Discharge Date” is null or blank), OR “Discharge Date” from the facility is greater than or equal to the last day of the study period AND “Treatment Dialysis Broad Start Date” is prior or equal to the first day of the study period, AND “Dialysis Broad Type of Treatment” = ‘HD’, AND “Primary Dialysis Setting” = ‘Dialysis Facility/Center’ or ‘Home’ on the last day of the study period, AND “Date Regular Chronic Dialysis Began” is prior to the first day of the study period.

Peritoneal dialysis patients are defined as follows: “Admit Date” to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged (“Discharge Date” is null or blank), OR “Discharge Date” from the facility is greater than or equal to the last day of the study period AND “Treatment Dialysis Broad Start Date” is prior or equal to the first day of the study period, AND

“Dialysis Broad Type of Treatment” = ‘PD’ on the last day of the study period, AND “Date Regular Chronic Dialysis Began” is prior to the first day of the study period.

The denominator will include all patients greater than or equal to 18 years old who are determined to be in-center hemodialysis, home hemodialysis, or peritoneal dialysis patients for whom “ESA Prescribed” = “YES” at any time during the three-month study period OR “Hemoglobin” < 11 in at least one month of the study period. The hemoglobin value reported for the end of each study month is used for this calculation.

2a1.9. Denominator Exclusion Details

No denominator exclusions

2a1.20. Calculation Algorithm/Measure Logic

For this measure calculation, the numerator will be divided by the denominator.

Calculation of the numerator and denominator is described below.

The patient’s age will be determined by subtracting the patient’s date of birth from the first day of the reporting month. Hemodialysis patients are defined as follows: “Admit Date” to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged (“Discharge Date” is null or blank), OR “Discharge Date” from the facility is greater than or equal to the last day of the study period AND “Treatment Dialysis Broad Start Date” is prior or equal to the first day of the study period, AND “Dialysis Broad Type of Treatment” = ‘HD’, AND “Primary Dialysis Setting” = ‘Dialysis Facility/Center’ or ‘Home’ on the last day of the study period, AND “Date Regular Chronic Dialysis Began” is prior to the first day of the study period.

Peritoneal dialysis patients are defined as follows: “Admit Date” to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged (“Discharge Date” is null or blank), OR “Discharge Date” from the facility is greater than or equal to the last day of the study period AND “Treatment Dialysis Broad Start Date” is prior or equal to the first day of the study period, AND “Dialysis Broad Type of Treatment” = ‘PD’ on the last day of the study period, AND “Date Regular Chronic Dialysis Began” is prior to the first day of the study period.

The denominator will include all patients greater than or equal to 18 years old who are determined to be in-center hemodialysis, home hemodialysis, or peritoneal dialysis patients for whom “ESA Prescribed” = “YES” at any time during the three-month study period OR “Hemoglobin” < 11 in at least one month of the study period. The hemoglobin value reported for the end of each study month is used for this calculation.

Number of patients in the denominator for whom “Serum Ferritin” and “Serum Ferritin Collection Date” are populated AND EITHER “Iron Saturation (TSAT) Percentage” and “Iron Saturation (TSAT) Percentage Collection Date” are populated OR “Reticulocyte Hemoglobin (CHr)” and “Reticulocyte Hemoglobin (CHr) Collection Date” are populated at least once during a three-month reporting period for in-center hemodialysis patients, peritoneal dialysis patients, and home hemodialysis patients.

Measure #0253: Peritoneal Dialysis Adequacy- Measurement of Total Solute Clearance At Regular Intervals

2a1.3. Numerator Details

The numerator will be determined by counting the patients in the denominator who had a total solute clearance for urea (endogenous residual renal urea clearance & dialytic) measured in any of the four months in the study period. The algorithm counts the number of non-missing values for the "Weekly Kt/V value: Peritoneal Dialysis, reporting month" (pd_weekly_ktv) variable. If the number of non-missing values is ≥ 1 and the patient is counted in the denominator, then the patient is included in the numerator.

2a1.7. Denominator Details

The patient's age will be determined by subtracting the patient's date of birth from the first day of the reporting month.

Peritoneal dialysis patients are defined as follows: (1) "Admit Date" to the specified facility is prior or equal to the first day of the study period; (2) the patient has not been discharged (i.e., the discharge date is null or blank or the discharge date is greater than or equal to the last day of the study period); (3) the treatment start date is less than or equal to the date of the study period; (4) the type of dialysis treatment = 'PD'; (5) the primary dialysis setting is "Home" or "Dialysis Facility/Center"; (6) the "Date Regular Chronic Dialysis Began" is prior to the first day of the study period.

The denominator will include all patients ≥ 18 years old who are PD patients assigned to a single facility for a four month period.

2a1.9. Denominator Exclusion Details

None.

2a1.20. Calculation Algorithm/Measure Logic

This measure is calculated by dividing the number of patients in the numerator by the number of patients in the denominator.

Measure #0254: Peritoneal Dialysis Adequacy- Calculate Weekly KT/V_{urea} in the Standard Way

2a1.3. Numerator Details

The numerator will be determined by counting the patients in the denominator who have:

- (1) Weekly Kt/V_{urea} used to measure delivered peritoneal dialysis dose and endogenous renal urea clearance;
- (2) Residual renal function (unless negligible [$< 100\text{mL}$ urine in 24 hours]) assessed by measuring the renal component of Kt/V_{urea} and estimating the patient's glomerular filtration rate (GFR) by calculating the mean of urea and creatinine clearance;

(3) Total body water (V) estimated by either the Watson or Hume method using actual body weight, and BSA estimated by either the Dubois and Dubois method, the Gehan and George method, or the Haycock method of using actual body weight; during the four month study period.

Specifically, the algorithm first counts the number of non-missing values for the “Weekly Kt/V value: Peritoneal Dialysis, reporting month” (pd_weekly_ktv) variable. If the number of non-missing values is ≥ 1 and the patient is counted in the denominator, then the patient is counted in the numerator if the most recent month with data meets these conditions:

(1) “Kt/V method” either “Watson” OR “Hume”

AND

(2) “BSA method” either “Dubois & Dubois” OR “Gehan & George” OR “Haycock”

AND

(3) (“Urine Volume” < 100) OR (“Urine Volume” ≥ 100 AND “PD Residual Renal Function” = “Yes”)

2a1.7. Denominator Details

The patient’s age will be determined by subtracting the patient’s date of birth from the first day of the reporting month.

Peritoneal dialysis patients are defined as follows: (1) “Admit Date” to the specified facility is prior or equal to the first day of the study period; (2) the patient has not been discharged (i.e., the discharge date is null or blank or the discharge date is greater than or equal to the last day of the study period); (3) the treatment start date is less than or equal to the date of the study period; (4) the type of dialysis treatment = ‘PD’; (5) the primary dialysis setting is “Home” or “Dialysis Facility/Center”; (6) the “Date Regular Chronic Dialysis Began” is prior to the first day of the study period.

The denominator will include all patients ≥ 18 years old who are PD patients assigned to a single facility for a four month period.

2a1.9. Denominator Exclusion Details

None.

2a1.20. Calculation Algorithm/Measure Logic

This measure is calculated by dividing the number of patients in the numerator by the number of patients in the denominator.

Measure #0318: Peritoneal Dialysis Adequacy- Delivered Dose of Peritoneal Dialysis Above Minimum

2a1.3. Numerator Details

The numerator will be determined by counting the patients in the denominator who had a delivered peritoneal dialysis weekly Kt/Vurea of at least 1.7 (dialytic + residual) during the four month study period.

Specifically, the algorithm first counts the number of non-missing values for the “Weekly Kt/V value: Peritoneal Dialysis, reporting month” (pd_weekly_ktv) variable. If the number of non-missing values is ≥ 1 and the patient is counted in the denominator, then the patient is counted in the numerator if the most recent month with data meets these conditions:

(1) “Kt/V method” either “Watson” OR “Hume”

AND

(2) “BSA method” either “Dubois & Dubois” OR “Gehan & George” OR “Haycock”

AND

(3) (“Urine Volume” < 100) OR (“Urine Volume” ≥ 100 AND “PD Residual Renal Function” = “Yes”)

AND

(4) “Weekly Kt/V” ≥ 1.7

2a1.7. Denominator Details

The patient’s age will be determined by subtracting the patient’s date of birth from the first day of the reporting month.

Peritoneal dialysis patients are defined as follows: (1) “Admit Date” to the specified facility is prior or equal to the first day of the study period; (2) the patient has not been discharged (i.e., the discharge date is null or blank or the discharge date is greater than or equal to the last day of the study period); (3) the treatment start date is less than or equal to the date of the study period; (4) the type of dialysis treatment = ‘PD’; (5) the primary dialysis setting is “Home” or “Dialysis Facility/Center”; (6) the “Date Regular Chronic Dialysis Began” is prior to the first day of the study period.

The denominator will include all patients ≥ 18 years old who are PD patients assigned to a single facility for a four month period.

2a1.9. Denominator Exclusion Details

None.

2a1.20. Calculation Algorithm/Measure Logic

This measure is calculated by dividing the number of patients in the numerator by the number of patients in the denominator.

Measure #0256: Hemodialysis Vascular Access- Minimizing Use of Catheters as Chronic Dialysis Access

2a1.3. Numerator Details

The numerator will be determined by counting the patients in the denominator who were on maintenance hemodialysis with a chronic catheter continuously for 90 days or longer prior to the last hemodialysis session of the month (“Access Type for Dialysis” = “Catheter” AND “Date Access Type for Dialysis Changed” is blank or, if populated, is more than 90 days prior to the last hemodialysis session of the month).

2a1.7. Denominator Details

The patient’s age will be determined by subtracting the patient’s date of birth from the first day of the reporting month.

Hemodialysis patients are defined as follows: “Admit Date” to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged (“Discharge Date” is null or blank), OR “Discharge Date” from the facility is greater than or equal to the last day of the study period AND “Treatment Dialysis Broad Start Date” is prior or equal to the first day of the study period, AND “Dialysis Broad Type of Treatment” = ‘HD’, AND “Primary Dialysis Setting” = ‘Dialysis Facility/Center’ or ‘Home’ on the last day of the study period, AND “Date Regular Chronic Dialysis Began” is prior to the first day of the study period.

The denominator will include all patients at least 18 years old who are determined to be in-center hemodialysis or home hemodialysis patients.

2a1.9. Denominator Exclusion Details

See above denominator details.

2a1.20. Calculation Algorithm/Measure Logic

For this measure calculation, the numerator will be divided by the denominator.

Calculation of the numerator and denominator is described below.

The patient’s age will be determined by subtracting the patient’s date of birth from the first day of the reporting month.

Hemodialysis patients are defined as follows: “Admit Date” to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged (“Discharge Date” is null or blank), OR “Discharge Date” from the facility is greater than or equal to the last day of the study period AND “Treatment Dialysis Broad Start Date” is prior or equal to the first day of the study period, AND “Dialysis Broad Type of Treatment” = ‘HD’, AND “Primary Dialysis Setting” = ‘Dialysis Facility/Center’ or ‘Home’ on the last day of the study period, AND “Date Regular Chronic Dialysis Began” is prior to the first day of the study period.

The denominator will include all patients at least 18 years old who are determined to be in-center hemodialysis or home hemodialysis patients.

The numerator will be determined by counting the patients in the denominator who were on maintenance hemodialysis with a chronic catheter continuously for 90 days or longer prior to the last hemodialysis session of the month (“Access Type for Dialysis” = “Catheter” AND “Date Access Type for Dialysis Changed” is blank or, if populated, is more than 90 days prior to the last hemodialysis session of the month).

Measure #0257: Hemodialysis Vascular Access- Maximizing Placement of Arterial Venous Fistula (AVF)

2a1.3. Numerator Details

The numerator will be determined by counting the patients in the denominator for whom “Access Type for Dialysis” = “autogenous AV fistula with two needles” at the last treatment of the month.

2a1.7. Denominator Details

The patient’s age will be determined by subtracting the patient’s date of birth from the first day of the reporting month.

Hemodialysis patients are defined as follows: “Admit Date” to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged (“Discharge Date” is null or blank), OR “Discharge Date” from the facility is greater than or equal to the last day of the study period AND “Treatment Dialysis Broad Start Date” is prior or equal to the first day of the study period, AND “Dialysis Broad Type of Treatment” = ‘HD’, AND “Primary Dialysis Setting” = ‘Dialysis Facility/Center’ or ‘Home’ on the last day of the study period, AND “Date Regular Chronic Dialysis Began” is prior to the first day of the study period. The denominator will include all patients at least 18 years old who are determined to be in-center hemodialysis or home hemodialysis patients.

The denominator will include all patients greater than or equal to 18 years old who are determined to be in-center hemodialysis, or home hemodialysis patients.

2a1.9. Denominator Exclusion Details

See above denominator details.

2a1.20. Calculation Algorithm/Measure Logic

For this measure calculation, the numerator will be divided by the denominator.

Calculation of the numerator and denominator is described below.

The patient’s age will be determined by subtracting the patient’s date of birth from the first day of the reporting month.

Hemodialysis patients are defined as follows: “Admit Date” to the specified facility is prior or equal to the first day of the study period, AND the patient has not been discharged (“Discharge Date” is null or

blank), OR “Discharge Date” from the facility is greater than or equal to the last day of the study period AND “Treatment Dialysis Broad Start Date” is prior or equal to the first day of the study period, AND “Dialysis Broad Type of Treatment” = ‘HD’, AND “Primary Dialysis Setting” = ‘Dialysis Facility/Center’ or ‘Home’ on the last day of the study period, AND “Date Regular Chronic Dialysis Began” is prior to the first day of the study period. The denominator will include all patients at least 18 years old who are determined to be in-center hemodialysis or home hemodialysis patients. The denominator will include all patients greater than or equal to 18 years old who are determined to be in-center hemodialysis, or home hemodialysis patients.

The numerator will be determined by counting the patients in the denominator for whom “Access Type for Dialysis” = “autogenous AV fistula with two needles” at the last treatment of the month.

Evidence (1c)

***Question:** We identified some common issues across many of the submissions noted below. Please review your submissions and send us any clarifications. Please keep in mind that NQF does not expect the developer to conduct a primary evidence review; rather the developer is asked to report on the review/grading of body of evidence that was conducted by others – specifically on the quantity, quality, and consistency of the body of evidence.*

1c.6. Quality of the Body of the Evidence. *Most of the submissions did not address the quality of the body of evidence. What did the systematic review of the body of evidence determine about the overall quality of the body of evidence? We ask for a grade in item 1c13, but for this item, we want substantive information about the quality. If the review did not address the overall quality of the body of evidence, please state that and indicate what was identified about the quality of individual studies.*

1c.7. Consistency of Results across Studies. *Most of the submissions did not provide information on consistency of the magnitude and direction of effect across the studies in the body of evidence. For the outcomes studied, what was the magnitude and direction of effect? If a meta-analysis was conducted, the results would be important evidence. Information from evidence tables can be used to provide substantive information on effect size.*

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation. *You can select USPSTF, GRADE, or other. Other is acceptable as long as you describe it. Please note, that grading the body of evidence is different from the strength of the recommendation. Some submissions indicated that the GRADE system was used, but the actual grades provided in items 1c.13/1c.23 were different than those described in GRADE documents. In that case, please clarify the differences (was the GRADE system used, modified, just different labels for the GRADE scale) and provide the rating scale with definitions.*

Below we provide text in response to these questions and other updates.

Measure #0252: Assessment of Iron Stores

1c.4. Directness of evidence to the specified measure

The body of evidence shows the role of the assessment of iron stores in anemia management. ESA and iron therapies correct anemia in dialysis patients. Routine assessment of iron stores allows for prudent use of IV iron which can lower the dose of ESAs. Routine assessment can also be useful in monitoring for iron overload.

1c.6. Quality of the Body of the Evidence

The body of evidence evaluated by the KDOQI guidelines was in support of iron targets and consists of two RCTs and various other studies (tissue iron studies, iron challenge tests, and nonrandomized trials). The two RCTs provide comparative information on ferritin and TSAT targets for IV iron therapy [Besarab A, 2000; DeVita MV, 2003].

An overall grade was not assigned to the body of evidence. However, individual studies were graded in the KDOQI Guidelines based on applicability and methodological quality. Applicability was graded according to the population of interest. Three grades were defined including: (1) sample is representative of target population, or results are definitely applicable to the target population irrespective of study sample; (2) sample is representative of a relevant subgroup of the target population; and (3) sample is representative of a narrow subgroup of patients only, and may not be generalizable to other subgroups. Methodological quality, or internal validity, referred to the design, conduct and reporting of the clinical study. A 3-level classification of study quality was devised: (1) least bias; results are valid; (2) susceptible to some bias, but not sufficient to invalidate the results; and (3) significant bias that may invalidate the results.

Of the two studies referenced in the KDOQI guidelines, both received the middle grade (2) for applicability. For methodological quality, one study received the highest grade [Besarab A, 2000], while the other study one received the middle grade [DeVita MV, 2003].

Recent clinical trials provide evidence that targeting higher Hgb levels when treating anemia in patients with chronic kidney disease (CKD) may increase the risk of adverse outcomes. The Trial to Reduce Cardiovascular Endpoints with Aranesp Therapy (TREAT) study found higher rates of stroke, thromboembolism, and cancer-related deaths in patients with CKD and diabetes who were treated to the higher Hgb target. The Correction of Hemoglobin and Outcomes in Renal Insufficiency (CHOIR) study (CKD patients) [Singh AK, 2006] and the Normal Hematocrit study (dialysis patients at high cardiovascular risk) [Besarab A, 1998] both found higher rates of death and cardiovascular complications among patients treated to higher Hgb targets. Two meta-analyses, which included both dialysis and non-dialysis CKD studies, also supported these findings [Phrommintikul A, 2007; KDOQI, 2006]. Although the cause of higher event rates among patients randomized to higher Hgb targets remains incompletely understood, higher ESA doses have been implicated as a possible explanation, and recent opinion in the nephrology community has coalesced around strategies to limit ESA dose when possible. To this end, alternate methods to facilitate ESA-mediated erythropoiesis, and support Hgb levels with lower ESA

doses, are increasingly recommended, and the judicious use of IV iron therapy remains central to this strategy [Kapoian T, 2008; Pizzi LT, 2008; Singh AK, 2010].

In addition to monitoring for iron deficiency, there is utility in monitoring for iron overload as there are evidence limitations with respect to long-term safety of IV iron therapy. As standard practice, IV iron dosing decisions are based on clinical measures of iron stores including ferritin and transferrin saturation (TSAT) levels. The proposed CPMs leave treatment decisions about IV iron dosing to the judgment of the practitioner, but attention to iron stores is an important factor in anemia treatment in dialysis patients.

1c.7. Consistency of Results across Studies

The direction of effect (reduced need for ESA of with higher iron dose) was consistent across two studies reviewed by KDOQI. The magnitude ranged from 28% to 40% reduction in ESA dose. More recent studies have also found similar effects.

1c.8. Net Benefit

ESA and iron therapies correct anemia in dialysis patients. Routine assessment of iron stores allows for prudent use of intravenous (IV) iron which can lower the dose of ESAs. Routine assessment of iron stores also allows for monitoring for iron overload.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

1c.11. Remove GRADE selection. The body of evidence was not graded.

1c.21. Other

1c.22. A structured approach, facilitated by the use of evidence profiles and modeled after the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) approach, was used to grade the quality of the overall evidence and the strength of recommendations. The strength of each guideline recommendation was rated as either “strong” or “moderately strong.” A “strong” rating indicates “it is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is high quality evidence that the practice results in net medical benefit to the patient.” The “moderately strong” rating indicates “it is recommended that clinicians routinely follow the guideline for eligible patients. There is at least moderately high quality evidence that the practice results in net medical benefit to the patient.” In the absence of strong or moderately strong quality evidence or when additional considerations did not support strong or moderately strong evidence-based guideline recommendations, the Work Group could elect to issue CPRs based on consensus of expert opinions.

1c.23. Consensus of expert opinions

1c.15 Citations for Evidence

Add the following citations:

4) Besarab A, Bolton WK, Browne JK, Egrie JC, Nissenson AR, Okamoto DM, Schwab SJ, Goodkin DA: The effects of normal as compared with low hematocrit values in patients with cardiac disease who are receiving hemodialysis and epoetin. N Engl J Med 339: 584–590, 1998.

- 5) Phrommintikul A, Haas SJ, Elisk M, Klum H: Mortality and target haemoglobin concentrations in anemia patients with chronic kidney disease treated with erythropoietin: A meta-analysis. *Lancet* 369: 381–388, 2007.
- 6) Kapoian T, O'Mara NB, Singh AK, Moran J, Rizkala AR, Geronemus R, Kopelman RC, Dahl NV, Coyne DW: Ferric gluconate reduces epoetin requirements in hemodialysis patients with elevated ferritin. *J Am Soc Nephrol* 19: 372–379, 2008.
- 7) Pizzi LT, Bunz TJ, Coyne DW, Goldfarb DS, Singh AK: Ferric gluconate treatment provides cost savings in patients with high ferritin and low transferrin saturation. *Kidney Int* 74: 1588–1595, 2008.
- 8) Pfeffer MA, Burdmann EA, Chen CY, et al. A Trial of Darbepoetin Alfa in Type 2 Diabetes and Chronic Kidney Disease. *New England Journal of Medicine* 361: 2019-2032, 2009.
- 9) Phrommintikul A, Haas SJ, Elisk M, Klum H. Mortality and target haemoglobin concentrations in anemia patients with chronic kidney disease treated with erythropoietin: A meta-analysis. *Lancet* 369: 381-388, 2007.
- 10) Palmer SC, Navaneethan SD, Craig JC, et al. Meta-analysis: erythropoiesis-stimulating agents in patients with chronic kidney disease. *Annals of Internal Medicine* 153: 23-33, 2010.
- 11) Suetonia C. Palmer, *Annals of Internal Medicine* 153: 23-33, 2010; published ahead of print May 3, 2010.
- 12) Besarab A, Amin N, Ahsan M, et al: Optimization of epoetin therapy with intravenous iron therapy in hemodialysis patients. *J Am Soc Nephrol* 11:530-538, 2000
- 13) DeVita MV, Frumkin D, Mittal S, Kamran A, Fishbane S, Michelis MF: Targeting higher ferritin concentrations with intravenous iron dextran lowers erythropoietin requirement in hemodialysis patients. *Clin Nephrol* 60:335-340, 2003

1c.16 Quote verbatim, the specific guideline recommendation

3.2.1 Frequency of iron status tests:

In the opinion of the Work Group, iron status tests should be performed as follows:

3.2.1.1 Every month during initial ESA treatment.

3.2.1.2 At least every 3 months during stable ESA treatment or in patients with HD-CKD not treated with an ESA.

1c.17 Clinical Practice Guideline Citation

KDOQI; National Kidney Foundation. II. Clinical practice guidelines and clinical practice recommendations for anemia in chronic kidney disease in adults. *Am J Kidney Dis* 47[Suppl 3]: S16–S85, 2006.

1c.24 Rationale for Using this Guideline over Others

Other international guidelines are available and are generally consistent with the KDOQI recommendations, although not as comprehensive.

Measure #0247: Hemodialysis Adequacy Clinical Performance Measure I: Hemodialysis Adequacy- Monthly Measurement of Delivered Dose

1c.6. Quality of the Body of the Evidence

The evidence supporting the monthly measurement of delivered dose is indirect. The body of evidence supports the dose of dialysis and clinical outcomes, and in these studies, dose of dialysis was measured monthly. Studies supporting the relationship between delivered dose of dialysis and clinical outcomes include a clinical trial and prospective and retrospective cohort studies. There was a well-designed randomized controlled clinical trial (the HEMO study), two were prospective studies and the remaining were retrospective cohort studies. The HEMO study with its randomized design and measurement of hard outcomes was given significant importance in defining the Hemodialysis Adequacy Guidelines for the KDOQI, thereby suggesting that the quality of at least this study was high. One of the prospective studies (ref 4) is based on the Dialysis Outcomes and Practice Patterns Study, which is an international prospective study of dialysis practices on patient outcomes. Although not a clinical trial, findings from the DOPPS have generally informed the formation of KDOQI clinical guidelines because the study population is large, nationally representative by design, with adequate longitudinal follow-up. Another study cited is based on the USRDS (ref 7) which included a national US random sample of prevalent hemodialysis patients. Finally, another study cited (ref 3) compares the association between dialysis adequacy and clinical outcomes with both the CMS and DOPPS datasets. Altogether, these suggest that the body of evidence for this measure is of generally acceptable quality.

An overall grade was not assigned to the body of evidence. However, individual studies were graded in the KDOQI Guidelines based on applicability and methodological quality. Applicability was graded according to the population of interest. Three grades were defined including: (1) sample is representative of target population, or results are definitely applicable to the target population irrespective of study sample; (2) sample is representative of a relevant subgroup of the target population; and (3) sample is representative of a narrow subgroup of patients only, and may not be generalizable to other subgroups.

Methodological quality, or internal validity, referred to the design, conduct and reporting of the clinical study. A 3-level classification of study quality was devised: (1) least bias; results are valid; (2) susceptible to some bias, but not sufficient to invalidate the results; and (3) significant bias that may invalidate the results.

Of the studies referenced in 1c.15, five received the highest grade (1) for applicability, and one received a grade of 2. For methodological quality, three studies received the highest grade (1), one received a 2, and the remaining studies received a 3.

1c.7. Consistency of Results across Studies

The direction of effect (benefit of higher dialysis dose) was consistent across studies. The magnitude of effect differed by study, partly because of differences in methodology for measurement of dialysis adequacy. For instance, the NECOSAD study (ref 2) reported a mortality risk of 1.66 for patients in the lowest weekly Kt/V category, whereas another study reported a relative risk of 0.85 for higher dialysis doses only in women (ref 3). Another study reported a 17% reduction in mortality risk for 5% higher urea reduction ratio (4) and an analysis looking at body size and mortality showed a mortality reduction of 7.5% for a 0.1 point increase in eKt/V.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

1c.11. Remove GRADE selection. The body of evidence was not graded.

1c.21. Change response to 'Other'.

1c.22. The rating system defined in the KDOQI Guidelines was used to grade the strength of the Guideline recommendation. KDOQI defined grades as follows:

Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.

Grade CPR: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.

Measure #0248: Hemodialysis Adequacy Clinical Performance Measure II: Method of Measurement of Delivered Hemodialysis Dose

1c.6. Quality of the Body of the Evidence

Studies supporting the relationship between delivered dose of dialysis using spKt/V and clinical outcomes include a clinical trial and prospective and retrospective cohort studies. The HEMO study (1) with its randomized design and measurement of hard outcomes was given significant importance in defining the Hemodialysis Adequacy Guidelines for the KDOQI, thereby suggesting that the quality of at least this study was high. One of the prospective studies is based on the Dialysis Outcomes and Practice Patterns Study, which is an international prospective study of dialysis practices on patient outcomes. Although not a clinical trial, findings from the DOPPS have generally informed the formation of KDOQI clinical guidelines because the study population is large, nationally representative by design, with adequate longitudinal follow-up. Another study cited is based on the USRDS which included a national US random sample of prevalent hemodialysis patients. Finally, another study cited compares the association between dialysis adequacy and clinical outcomes with both the CMS and DOPPS datasets. Altogether, these suggest that the body of evidence for this measure is of acceptable quality.

An overall grade was not assigned to the body of evidence. However, individual studies were graded in the KDOQI Guidelines based on applicability and methodological quality. Applicability was graded according to the population of interest. Three grades were defined including: (1) sample is representative of target population, or results are definitely applicable to the target population irrespective of study sample; (2) sample is representative of a relevant subgroup of the target population; and (3) sample is representative of a narrow subgroup of patients only, and may not be generalizable to other subgroups.

Methodological quality, or internal validity, referred to the design, conduct and reporting of the clinical study. A 3-level classification of study quality was devised: (1) least bias; results are valid; (2) susceptible to some bias, but not sufficient to invalidate the results; and (3) significant bias that may invalidate the results.

Of the studies referenced in 1c.15, five received the highest grade (1) for applicability, four were graded a 2, and 1 received a 3. For methodological quality, three studies received the highest grade (1), one received a 2, and the remaining studies received a 3.

1c.7. Consistency of Results across Studies

Dialysis adequacy as calculated by spKt/V was utilized in several studies, and direction of effect (benefit of higher dialysis dose) tended to be consistent across studies. Comparison of magnitude of effect was not done because of differences in analytical methods. In the DOPPS study, the relative risk of mortality for $\text{spKt/V} < 1.2$ was 1.16 (Port FK, Pisoni RL, Bragg-Gresham JL, et al: DOPPS estimates of patient life years attributable to modifiable hemodialysis treatment practices in the United States. *Blood Purif* 22:175-180, 2004). Another analyses revealed that a 0.1 unit higher value of spKt/V had a 7.5% lower risk of mortality (8). In another cohort, patients who received an SpKt/V of 0.7 had a 2.8 increased mortality risk RR as compared to patients who received an spKt/V of between 1.2-1.3 (9). Finally, in the HEMO study (1), patient mortality did not improve at higher target dialysis above that recommended by clinical guidelines. The HEMO study also used Urea Reduction Ratio, spKt/V and eKt/V in the design of the study.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

1c.11. Remove GRADE selection. The body of evidence was not graded.

1c.21. Change response to 'Other'.

1c.22 The rating system defined in the KDOQI Guidelines was used to grade the strength of the Guideline recommendation. KDOQI defined grades as follows:

Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.

Grade CPR: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.

Measure #0249: Hemodialysis Adequacy Clinical Performance Measure III: Method of Measurement of Delivered Hemodialysis Dose

1c.6. Quality of the Body of the Evidence

Studies supporting the relationship between delivered dose of dialysis and clinical outcomes include a clinical trial and prospective and retrospective cohort studies.. There was a well-designed randomized controlled clinical trial (the HEMO study), two were prospective studies and the remaining were retrospective cohort studies. The HEMO study with its randomized design and measurement of hard outcomes was given significant importance in defining the Hemodialysis Adequacy Guidelines for the KDOQI, thereby suggesting that the quality of at least this study was high. One of the prospective studies (ref 4) is based on the Dialysis Outcomes and Practice Patterns Study, which is an international prospective study of dialysis practices on patient outcomes. Although not a clinical trial, findings from the DOPPS have generally informed the formation of KDOQI clinical guidelines because the study population is large, nationally representative by design, with adequate longitudinal follow-up. Another study cited is based on the USRDS (ref 7) which included a national US random sample of prevalent hemodialysis patients. Finally, another study cited (ref 3) compares the association between dialysis adequacy and clinical outcomes with both the CMS and DOPPS datasets. Altogether, these suggest that the body of evidence for this measure is of generally acceptable quality.

An overall grade was not assigned to the body of evidence. However, individual studies were graded in the KDOQI Guidelines based on applicability and methodological quality. Applicability was graded according to the population of interest. Three grades were defined including, (1) sample is representative of target population, or results are definitely applicable to the target population irrespective of study sample; (2) sample is representative of a relevant subgroup of the target population; and (3) sample is representative of a narrow subgroup of patients only, and may not be generalizable to other subgroups.

Methodological quality, or internal validity, referred to the design, conduct and reporting of the clinical study. A 3-level classification of study quality was devised: (1) least bias; results are valid; (2) susceptible to some bias, but not sufficient to invalidate the results; and (3) significant bias that may invalidate the results.

Of the studies referenced in 1c.15, six received the highest grade (1) for applicability, four received a grade of 2, and one received a 3. For methodological quality, three studies received the highest grade (1), one received a 2, and the remaining studies received a 3.

1c.7. Consistency of Results across Studies

Dialysis adequacy as calculated by spKt/V was utilized in several studies, and direction of effect (benefit of higher dialysis dose) tended to be consistent across studies. Comparison of magnitude of effect was not done because of differences in analytical methods. In the DOPPS study, the relative risk of mortality

for $\text{spKt/V} < 1.2$ was 1.16 (Port FK, Pisoni RL, Bragg-Gresham JL, et al: DOPPS estimates of patient life years attributable to modifiable hemodialysis treatment practices in the United States. *Blood Purif* 22:175-180, 2004). Another analyses revealed that a 0.1 unit higher value of spKt/V had a 7.5% lower risk of mortality (8). In another cohort, patients who received an spKt/V of 0.7 had a 2.8 increased mortality risk RR as compared to patients who received an spKt/V of between 1.2-1.3 (9). Finally, in the HEMO study (1), patient mortality did not improve at higher target dialysis above that recommended by clinical guidelines. The HEMO study also used Urea Reduction Ratio, spKt/V and eKt/V in the design of the study.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

1c.11. Remove GRADE selection. The body of evidence was not graded.

1c.21. Change response to 'Other'.

1c.22. The rating system defined in the KDOQI Guidelines was used to grade the strength of the Guideline recommendation. KDOQI defined grades as follows:

Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.

Grade CPR: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.

Measure #0253: Peritoneal Dialysis Adequacy- Measurement of Total Solute Clearance at Regular Intervals

1c.6. Quality of the Body of the Evidence

The KDOQI panel noted that the body of evidence shows a correlation between total solute clearance for urea and patient mortality and morbidity. Thus, this evidence supports the present measure in that that the delivered dose of dialysis should be measured frequently for assessment of adequate treatment.

An overall grade was not assigned to the body of evidence. However, individual studies were graded in the KDOQI Guidelines based on applicability and methodological quality. Applicability was graded according to the population of interest. Three grades were defined including, (1) sample is representative of target population, or results are definitely applicable to the target population irrespective of study sample; (2) sample is representative of a relevant subgroup of the target population; and (3) sample is representative of a narrow subgroup of patients only, and may not be generalizable to other subgroups.

Methodological quality, or internal validity, referred to the design, conduct and reporting of the clinical study. A 3-level classification of study quality was devised: (1) least bias; results are valid; (2) susceptible

to some bias, but not sufficient to invalidate the results; and (3) significant bias that may invalidate the results.

In particular, of the 20 studies considered in the body of evidence, the results from two randomized clinical trials were used to justify the KDOQI guidelines (Paniagua 2002; Lo 2003). These two studies were graded the highest quality (Level 1) and the highest applicability (Level 1). The results from additional observational studies also supported the KDOQI recommendations (see, e.g., Bargman 2001; Rocco 2000; Churchill 1998), and were graded level (1) applicability and level (2) quality.

1c.7. Consistency of Results across Studies

Results were consistent across studies in direction, although magnitude varied. Of the 20 studies reviewed by the KDOQI panel, 17 examined either peritoneal clearance or total urea Kt/V. Of these 17 studies, 12 found lower mortality for higher PD clearance; on the other hand, only two noted higher mortality with higher PD clearance, neither of which achieved statistical significance.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

1c.11. Change from GRADE to Other.

1c.21. Change from GRADE to Other.

The rating system defined for grading the KDOQI Guidelines was defined as follows:

Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.

Grade C: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.

Measure #0254: Peritoneal Dialysis Adequacy- Calculate Weekly Kt/V_{urea} in the Standard Way

1c.6. Quality of the Body of the Evidence

The KDOQI panel noted that the body of evidence shows a correlation between total solute clearance for urea and patient mortality and morbidity. Thus, this evidence supports the present measure in that that the delivered dose of dialysis should be measured frequently and in a standard way for assessment of adequate treatment.

An overall grade was not assigned to the body of evidence. However, individual studies were graded in the KDOQI Guidelines based on applicability and methodological quality. Applicability was graded according to the population of interest. Three grades were defined including, (1) sample is representative of target population, or results are definitely applicable to the target population

irrespective of study sample; (2) sample is representative of a relevant subgroup of the target population; and (3) sample is representative of a narrow subgroup of patients only, and may not be generalizable to other subgroups.

Methodological quality, or internal validity, referred to the design, conduct and reporting of the clinical study. A 3-level classification of study quality was devised: (1) least bias; results are valid; (2) susceptible to some bias, but not sufficient to invalidate the results; and (3) significant bias that may invalidate the results.

In particular, of the 20 studies considered in the body of evidence, the results from two randomized clinical trials were used to justify the KDOQI guidelines (Paniagua 2002; Lo 2003). These two studies were graded the highest quality (Level 1) and the highest applicability (Level 1). The results from additional observational studies also supported the KDOQI recommendations (see, e.g., Bargman 2001; Rocco 2000; Churchill 1998), and were graded level (1) applicability and level (2) quality.

1c.7. Consistency of Results across Studies

Results were consistent across studies in direction, although magnitude varied. Of the 20 studies reviewed by the KDOQI panel, 17 examined either peritoneal clearance or total urea Kt/V. Of these 17 studies, 12 found lower mortality for higher PD clearance; on the other hand, only two noted higher mortality with higher PD clearance, neither of which achieved statistical significance.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

1c11. Change from GRADE to Other.

1c21. Change from GRADE to Other.

The rating system defined for grading the KDOQI Guidelines was defined as follows:

Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.

Grade C: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.

Measure #0318: Peritoneal Dialysis Adequacy- Delivered Dose of Peritoneal Dialysis above Minimum

1c.6. Quality of the Body of the Evidence

The KDOQI panel noted that the body of evidence shows a correlation between total solute clearance for urea and patient mortality and morbidity. Thus, this evidence supports the present measure in that

that the delivered dose of dialysis should be measured frequently for assessment of adequate treatment, and treatment should be set accordingly.

An overall grade was not assigned to the body of evidence. However, individual studies were graded in the KDOQI Guidelines based on applicability and methodological quality. Applicability was graded according to the population of interest. Three grades were defined including, (1) sample is representative of target population, or results are definitely applicable to the target population irrespective of study sample; (2) sample is representative of a relevant subgroup of the target population; and (3) sample is representative of a narrow subgroup of patients only, and may not be generalizable to other subgroups.

Methodological quality, or internal validity, referred to the design, conduct and reporting of the clinical study. A 3-level classification of study quality was devised: (1) least bias; results are valid; (2) susceptible to some bias, but not sufficient to invalidate the results; and (3) significant bias that may invalidate the results.

In particular, of the 20 studies considered in the body of evidence, the results from two randomized clinical trials were used to justify the KDOQI guidelines (Paniagua 2002; Lo 2003). These two studies were graded the highest quality (Level 1) and the highest applicability (Level 1). The results from additional observational studies also supported the KDOQI recommendations (see, e.g., Bargman 2001; Rocco 2000; Churchill 1998), and were graded level (1) applicability and level (2) quality.

1c.7. Consistency of Results across Studies

Results were consistent across studies in direction, although magnitude varied. Of the 20 studies reviewed by the KDOQI panel, 17 examined either peritoneal clearance or total urea Kt/V. Of these 17 studies, 12 found lower mortality for higher PD clearance; on the other hand, only two noted higher mortality with higher PD clearance, neither of which achieved statistical significance.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

1c11. Change from GRADE to Other.

1c21. Change from GRADE to Other.

The rating system defined for grading the KDOQI Guidelines was defined as follows:

Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.

Grade C: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.

Measure #0255: Measurement of Serum Phosphorus Concentration

1c.6. Quality of the Body of the Evidence

The submitting organization recognizes the opinion-based level of evidence supporting the KDIGO Clinical Practice Guidelines for measurement of serum concentration of phosphorus. As such, the overall quality of the body of evidence or the quality of individual studies is not rated in the KDIGO guidelines. Notwithstanding, research in many studies have observed that abnormalities of serum phosphorus concentration are common in the CKD population and that failure to monitor and correct such abnormalities are strongly associated with morbidity and mortality. Observational studies have shown a consistent adverse association of low serum phosphorus with all-cause mortality. Furthermore, the basic science supports a pathological role of low serum phosphorus and intracellular phosphorus depletion in disturbed cellular function.

1c.7. Consistency of Results across Studies

Serum phosphorus is consistently demonstrated to be an important biomarker, strongly associated with adverse cardiovascular outcomes. In addition, the data from in-vitro and in-vivo animal studies establish the biologic plausibility of the adverse effects of inappropriate levels of serum phosphorus on cardiovascular outcomes. Observational data consistently report an increased level of cardiovascular events and mortality when serum phosphorus rises above the normal range in patients with Stage 5 CKD.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

The body of evidence (1c.11) and the guideline recommendations (1c.21) were not graded.

Measure #0261: Measurement of Serum Calcium Concentration

1c.6. Quality of the Body of the Evidence

The submitting organization recognizes the opinion-based level of evidence in support of the KDIGO Clinical Practice Guidelines for measurement of concentration of serum calcium. As such, the overall quality of the body of evidence or the quality of individual studies is not rated in the KDIGO guidelines. Notwithstanding, researchers in many studies have observed that abnormalities of serum calcium concentration are common in this population and that failure to monitor and correct such abnormalities are strongly associated with morbidity and mortality. The basic science also supports a pathological role of high calcium in promoting soft tissue and vascular calcification. At this time, there are no interventional studies demonstrating the benefit of correcting hypercalcemia. The overall quality of the body of evidence or the quality of individual studies was not addressed in the NKF-KDOQI guidelines.

1c.7. Consistency of Results across Studies

Observational cohort studies show a consistent adverse association of hypercalcemia with cardiovascular events and all-cause mortality. There is also clinical data demonstrating the association of increased serum calcium with vascular and valvular calcifications.

1c.11/1c.21 System used for grading the body of evidence and grading the recommendation

The body of evidence (1c.11) and the guideline recommendations (1c.21) were not graded.

Measure #0256: Hemodialysis Vascular Access- Minimizing Use of Catheters as Chronic Dialysis Access

1c.6. Quality of Body of Evidence

A large body of literature exists showing strong associations between central venous catheter use in hemodialysis patients with poorer survival and greater morbidity [1-40,44]. The prevalence of numerous patient comorbidity indicators was similar in facilities with high versus low catheter use. Lower mortality has been observed with reduction in catheter use and an increase in fistula use in facility- and patient-level access use studies [7, 10, 13, 40, 41]. Furthermore, much of the 30-40% higher case-mix adjusted mortality rate for US hemodialysis patients compared to those in several European countries appears to be explained by differences in vascular access use between these two regions [2].

The overall quality of the body of evidence for this measure has not been evaluated. Furthermore, the quality of the individual studies referenced in 1c.15 has not been graded.

1c.7. Consistency of Results across Studies

Results were consistent across all studies listed in body of evidence. For example, across 8 large observational studies of prevalent hemodialysis patients, the hazard ratio of case-mix adjusted all-cause mortality ranged from 1.32 to 1.75 (median HR~1.5, $p<0.05$ in all studies) for patients dialyzing with a catheter versus a native arteriovenous (AV) fistula [1-4,8,23,26,37,44]. Catheter use in incident patients at the time of commencing hemodialysis was associated with a 1.5-2.5 fold higher HR of mortality. Furthermore, a 20% higher mortality rate was observed for every 20% greater facility percent catheter use (compared with AV fistula use)[2]; conversion from a catheter to an AV access was associated with a 31% lower mortality rate whereas conversion from an AV access to a catheter was associated with a 80-138% higher mortality rate, in incident and prevalent patients respectively [7,10,13].

1c.11. System used for grading the body of evidence

Not graded

1c.15 Citations for Evidence, other than guidelines which are addressed below

Please add the following references:

42. USRDS, 2009 Annual Report, vol 2, Chapter 11, pg. 341.
43. National Kidney Foundation: KDOQI Clinical Practice Guidelines for Vascular Access. Update 2006.
44. Foley RN, Chen SC, Collins AJ: Hemodialysis access at initiation in the United States, 2005 to 2007: still "catheter first". *Hemodial Int* **13**:533–542, 2009.

1c.21. System used for grading the strength of guideline recommendation

Other

1c.22. If other, identify and describe the grading scale with definitions

The rating system defined in the KDOQI Guidelines was used to grade the strength of the Guideline recommendation. KDOQI defined grades as follows:

Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.

Grade CPR: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.

1c.23. Grade assigned to the recommendation

KDOQI Guideline 8.1.2.2 was graded B.

Measure #0257: Hemodialysis Vascular Access- Maximizing Placement of Arterial Venous Fistula (AVF)

1c.6. Quality of Body of Evidence

A large body of literature exists showing strong associations between central venous catheter use in hemodialysis patients with poorer survival and greater morbidity [1-40, 44]. The prevalence of numerous patient comorbidity indicators was similar in facilities with high versus low catheter use. Lower mortality has been observed with reduction in catheter use and an increase in fistula use in facility- and patient-level access use studies [7, 10, 13, 40, 41]. Furthermore, much of the 30-40% higher case-mix adjusted mortality rate for US hemodialysis patients compared to those in several European countries appears to be explained by differences in vascular access use between these two regions [2]. In addition, per person per year “access event” costs were greatest for patients with a catheter or arteriovenous graft, at \$5,960 and \$7,451, respectively, in 2007 as indicated in the 2009 USRDS Annual Report. In contrast, among patients with an arteriovenous fistula these costs averaged \$3,194 — 57 percent lower than the costs incurred by patients with an AV graft [42].

The overall quality of the body of evidence for this measure has not been evaluated. Furthermore, the quality of the individual studies referenced in 1c.15 has not been graded.

1c.7. Consistency of Results across Studies

Results were consistent across all studies listed in body of evidence. For example, across 8 large observational studies of prevalent hemodialysis patients, the hazard ratio of case-mix adjusted all-cause mortality ranged from 1.32 to 1.75 (median HR~1.5, $p<0.05$ in all studies) for patients dialyzing with a catheter versus a native arteriovenous (AV) fistula [1-4, 8,23, 26, 37, 44]. Catheter use in incident patients at the time of commencing hemodialysis was associated with a 1.5-2.5 fold higher HR of mortality. Furthermore, a 20% higher mortality rate was observed for every 20% greater facility percent catheter use (compared with AV fistula use[2]); conversion from a catheter to an AV access was associated with a 31% lower mortality rate whereas conversion from an AV access to a catheter was associated with a 80-138% higher mortality rate, in incident and prevalent patients respectively [7,10,13]. Native AV fistula use has also shown to be associated with longer survival in comparison to AV graft use. Case-mix adjusted mortality rates were 39% higher in 2 large cohort studies of incident

patients initiating HD with an AV graft versus an AV fistula [3,44]. In three large cohort studies of HD patients dialyzing with an AV graft displayed a 5-20% higher mortality rate compared with patients dialyzing with an AV fistula, while in a fourth such study, patients dialyzing with AV graft versus an AV fistula displayed an 8% higher mortality rate among non-diabetic HD patients and a 41% higher mortality rate among diabetic HD patients. Furthermore, in a facility practice-based analysis, a 9% higher mortality rate was observed for every 20% greater facility percent AV graft use (compared with AV fistula use[2]). In addition, the number of access events is 3- to 7-fold greater in prosthetic grafts than in native AV fistulae and is an important factor in the higher annual access event costs observed for AV grafts compared with AV fistulae [43].

1c.11. System used for grading the body of evidence

The body of evidence was not graded.

1c.15 Citations for Evidence, other than guidelines which are addressed below

Please add the following references:

42. USRDS, 2009 Annual Report, vol 2, Chapter 11, pg. 341.

43. National Kidney Foundation: KDOQI Clinical Practice Guidelines for Vascular Access. Update 2006.

44. Foley RN, Chen SC, Collins AJ: Hemodialysis access at initiation in the United States, 2005 to 2007: still "catheter first". *Hemodial Int* **13**:533–542, 2009.

1c.21 System used for grading the recommendation.

Other

1c.22. If other, identify and describe the grading scale with definitions

The rating system defined in the KDOQI Guidelines was used to grade the strength of the Guideline recommendation. KDOQI defined grades as follows:

Grade A: It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves health outcomes.

Grade B: It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderately strong evidence that the practice improves health outcomes.

Grade CPR: It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence or on the opinions of the Work Group and reviewers that the practice might improve health outcomes.

1c.23. Grade assigned to the recommendation

KDOQI Guideline 8.1.2.1 was graded B.

Reliability Testing (2a2.2)

Question: What is the rationale for considering correlation of facility scores in two time periods as a demonstration of reliability of the data? NQF is asking for evidence of the reliability of either: 1) the data used in the measure (e.g., lab date, lab method, dialysis dose) or 2) the measure score (amount of variation due to true differences among providers vs. error/noise – signal- to-noise analysis). With the data available from CROWNWeb, is it possible to conduct signal-to-noise analysis?

The use of correlation across different time periods is a measure of reliability that applies to any type of measure (based on normal or binary data, for example) and we find it useful from this perspective. In the case of a mixed normal model with between and within variances, σ_b^2 and σ_w^2 respectively, it can be seen that this correlation would estimate the same quantity as $1-1/F$ in the one way analysis of variance applied to a single wave. That is, it would estimate $n' \sigma_b^2 / [\sigma_w^2 + n' \sigma_b^2]$ where n' is essentially an average facility size. So, it is again measuring the relative size of the between and within variation.

The NQF document “Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties” from January 2011 suggests using ANOVA to perform a signal-to-noise analysis. ANOVA was performed on patient level data from October 2010 using each measure as the independent variable and facility as the dependent variable. The intraclass correlation coefficients ranged from 0.02 to 0.34, the interunit reliability ranged from 0.57 to 0.97, and all measures had statistically significant F tests.

ANOVA signal-to-noise results by measure using CROWNWeb 2010 data						
Measure	Intraclass Correlation (ICC)	R ²	F	Interunit Reliability (IUR=1-(1/F))	P-value for F test	
#0247 Monthly Measurement of Delivered HD Dose	0.17	0.18	16.33	0.94	<.0001	
#0248 Method of Measurement of Delivered Dose	0.26	0.27	28.25	0.96	<.0001	
#0250 Minimum Delivered Hemodialysis Dose	0.34	0.35	28.66	0.97	<.0001	
#0253 Measurement of Total Solute Clearance	0.19	0.26	4.47	0.78	<.0001	
#0254 Standard calculation of weekly Kt/Vurea	0.10	0.17	2.75	0.64	<.0001	
#0318 Delivered Dose of Peritoneal Dialysis	0.09	0.16	2.35	0.57	<.0001	
#0257 Maximizing Use of AV Fistula (AVF)	0.07	0.08	6.40	0.84	<.0001	
#0256 Minimizing Use of Catheters	0.08	0.10	6.36	0.84	<.0001	

ANOVA signal-to-noise results by measure using CROWNWeb 2010 data

Measure	Intraclass Correlation (ICC)	R ²	F	Interunit Reliability (IUR=1-(1/F))	P-value for F test
#0252 Assessment of Iron Stores	0.25	0.27	20.30	0.95	<.0001
#0255 Measurement of Serum Phosphorus	0.17	0.18	17.34	0.94	<.0001
#0261 Measurement of Serum Calcium	0.17	0.18	17.49	0.94	<.0001

Validity Testing (2b2.3)

Question: The validity testing approach was sound. Sometimes the results provided didn't support the conclusion that better performance is associated with lower mortality by quintiles (e.g., confidence intervals include 1.0). How does that change your conclusion or what are possible explanations?

We are providing revised text for several measures showing that better performance is associated with better outcomes.

Measure #0256: Hemodialysis Vascular Access- Minimizing Use of Catheters as Chronic Dialysis Access

2b2.1. Data Sample

2009 CROWNWeb data (August - December) were used to calculate monthly performance scores, and the SHR was calculated using 2009 Medicare-paid dialysis claims and the Medical Evidence Form (Form CMS-2728).

2b2.2. Analytic Method

Validity was assessed using Poisson regression models to measure the association between facility level quintiles of performance scores and the 2009 SHR. Facility-level performance scores were divided into quintiles and the relative risk (RR) of hospitalization was calculated for each quintile. The highest quintile was used as the reference group. Thus, a RR>1.0 for the lower performance score quintiles would indicate a higher relative risk of hospitalization.

Association with the 2009 SMR was also examined; while an increase in mortality was not found to be statistically significant, an association with hospital admission rates shows that catheter usage is associated with significant patient morbidity.

2b2.3. Testing Results

Quintiles of the performance scores were defined as follows:

Q1 & Q2: 0-0%, RR(CI)= 0.97 (0.95, 1.00), p = 0.02

Q3: 0%–<6%, RR(CI)= 0.89 (0.86, 0.91), $p < 0.001$

Q4: 6%–<11%, RR(CI)= 0.96 (0.93, 0.98), $p = 0.001$

Q5: 11%–<58%, RR(CI)= 1.00(ref)

Overall results from the Poisson model indicated that the percent of patients dialyzing with a catheter ≥ 90 days was significantly associated with SHR ($p < 0.001$). These findings confirm the association between low catheter use and lower hospital admissions.

Measure #0253: Peritoneal Dialysis Adequacy- Measurement of Total Solute Clearance At Regular Intervals

2b2.2. Analytic Method

Validity was assessed using Poisson regression models to measure the association between facility level measure performance and facility level mortality as indicated by the standardized mortality ratio (SMR; methodology on SMR calculations is attached). Facility-level performance scores were divided into tertiles and the relative risk (RR) of mortality was calculated for each tertile. The highest performance tertile was used as the reference group. Thus, a $RR > 1.0$ for the lower performance score tertiles would indicate a higher relative risk of mortality is associated with lower performance on the measure.

2b2.3. Testing Results

Tertiles of facility performance scores were defined as follows:

T1: 0% - 57%

T2: 57% - 72%

T3: 72% - 100%

Results from the Poisson model indicated lower performance scores were associated with SMR ($p = 0.05$). For T1, $RR = 1.06$, 95% CI: (1.01, 1.12), $p = 0.03$. For T2, $RR = 1.06$, 95% CI: (1.00, 1.11), $p = 0.04$.

Measure #0254: Peritoneal Dialysis Adequacy- Calculate Weekly KT/V_{urea} in the Standard Way

2b2.2. Analytic Method

Validity was assessed using Poisson regression models to measure the association between facility level measure performance and facility level mortality as indicated by the standardized mortality ratio (SMR; methodology on SMR calculations is attached). Facility-level performance scores were divided into tertiles and the relative risk (RR) of mortality was calculated for each tertile. The highest performance tertile was used as the reference group. Thus, a $RR > 1.0$ for the lower performance score tertiles would indicate a higher relative risk of mortality is associated with lower performance on the measure.

2b2.3. Testing Results

Tertiles of facility performance scores were defined as follows:

T1: 0% - 33%

T2: 34% - 50%

T3: 50% - 100%

Results from the Poisson model indicated lower performance scores were marginally associated with SMR ($p=0.10$ for overall test). However, both T1 and T2 had estimated RR values higher than the reference group. For T1, $RR = 1.04$, 95% CI: (0.99, 1.12), $p=0.15$. For T2, $RR = 1.06$, 95% CI: (1.00, 1.11), $p=0.04$.

Measure #0318: Peritoneal Dialysis Adequacy- Delivered Dose of Peritoneal Dialysis above Minimum

2b2.2. Analytic Method

Validity was assessed using Poisson regression models to measure the association between facility level measure performance and facility level mortality as indicated by the standardized mortality ratio (SMR; methodology on SMR calculations is attached). Facility-level performance scores were divided into tertiles and the relative risk (RR) of mortality was calculated for each tertile. The highest performance tertile was used as the reference group. Thus, a $RR>1.0$ for the lower performance score tertiles would indicate a higher relative risk of mortality is associated with lower performance on the measure.

Validity was also assessed using data from the 2008 Clinical Performance Measures (CPM) project. Patient-level PD Kt/V values were divided into quintiles and related to patient mortality using Cox proportional hazards regression, adjusted for demographic characteristics. The quintile with the highest Kt/V values was used as the reference group, therefore a $HR>1.0$ for the lower groups would indicate a higher rate of mortality is associated with lower values of Kt/V.

2b2.3. Testing Results

Tertiles of facility performance scores were defined as follows: T1: 0% - 30%; T2: 30% - 47%; T3: 47% - 100%

Results from the Poisson model indicated higher estimated SMR for facilities with lower scores on the measure, although this did not reach statistical significance ($p=0.50$ for overall test). Both T1 and T2 had estimated RR values slightly higher than the reference group. For T1, $RR = 1.03$, 95% CI: (0.97, 1.08), $p=0.34$. For T2, $RR = 1.03$, 95% CI: (0.98, 1.08), $p=0.29$.

The lack of statistical significance of the above analysis may be affected by the combination of the relative size of the PD population and limitations in the PD data collected by CROWNWeb during the testing phase. During the testing phase, reporting of PD data was not as widespread as reporting of HD data. To provide further information given these limitations, results from the 2008 CPM Project are described below. These analyses were performed at the patient level and were adjusted for patient demographics.

Kt/V quintiles were defined as follows: Q1: 0.1 – 1.85; Q2: 1.85 – 2.11; Q3: 2.11 – 2.34; Q4: 2.35 – 2.73; Q5: 2.74 – 4.90.

Results from the model indicated higher mortality for patients with lower Kt/V values (overall $p < 0.01$). For Q1: HR=2.381, $p < 0.001$; for Q2: HR=1.984, $p < 0.001$; for Q3: HR=1.834, $p < 0.001$; for Q4: HR=1.480, $p = 0.02$.

Usability - Current Use (3.1)

Question: If indicated that a measure is currently used in public reporting (3.1), why in 3a.1 is it being evaluated for public reporting and what are the specific plans and timeline? Is the measure not currently reported? If individual facility performance is not identified (e.g., in CPM reports) then it is not considered public reporting.

SMR is the only measure for which 'Public Reporting' should be checked. Additionally, the following boxes should be checked for all measures:

- Public Health/Disease Surveillance
- Regulatory and Accreditation Programs
- Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
- Quality Improvement (Internal to the specific organization)

Measure #0369: Dialysis Facility Risk-Adjusted Standardized Mortality Ratio Level

Measure Title (De.1)

Question: Why Is "(32)" in the Title?

The "32" should be removed from the title.

Risk Adjustment Strategy and Testing (2b4)

Question: How were the risk factors selected? Race and ethnicity have substantial role in risk model, which is not consistent with NQF criteria. Generally, inclusion of race and ethnicity in risk models can obscure disparities in care. What is the rationale and analyses justifying inclusion of race/ethnicity?

To assess the adequacy of the risk model, we calculated the stratified Concordance Index (C-Index) in the survival model, which measures how well the risk model discriminates between different responses, in other words, is the predicted response low for low observed responses and high for high observed responses. In this model, C-Index=0.68 which suggests relatively good predictive ability of the risk model.

We also examined the functional form of all continuous variables (Comorbidity Index, log(BMI), and age) in the model through risk decile plots.

Risk factor selection: The methods for development of the risk factor models have been published and documented (Wolfe RA et al. Using USRDS generated mortality tables to compare local ESRD mortality rates to national rates. *Kidney Int* 1992; 42: 991-96; Wolfe RA et al: New dialysis facility mortality statistics (SMRs) adjust for more patient characteristics. *J Am Soc Nephrol* 2001; 12; A1802). The

adjustments included in the model are measured at baseline and are all statistically significant in the model. Two-way interactions were examined and selected for the final model based on both the magnitude and statistical significance of the estimates.

Inclusion of race/ethnicity: While adjustment for race and ethnicity tends to obscure disparities in care for many medical conditions, such adjustment may clarify such disparities in the ESRD setting and failure to adjust for race may obscure those disparities. In a series of analyses spanning several decades, it has been consistently seen that African-American patients have lower death rates than Caucasian patients on dialysis (RR=0.82), with a variety of adjustment models intended to make “all else equal”. This difference contributes to the downward trend in mortality seen in Figure 1 for facilities with higher % Black case mix, when the mortality is unadjusted for race. In the unadjusted analysis, facilities with higher % Black have lower mortality, in part because Black patients have lower mortality. That is, facilities with more Black patients have lower death rates, in part due to unadjusted case mix differences, just as facilities treating younger patients would have lower death rates, if age were not adjusted for. The unadjusted analysis does not, and cannot, separate the effect of case mix due to race from the effect of quality of care at facilities that treat a higher percentage of Black patients. Consequently, it is unknown whether the lower mortality at facilities with greater percentages of Black patients is because Black ESRD patients have lower mortality than non-Black patients, or if it is because such facilities provide better care.

Figure 1 also shows a race-adjusted analysis of facility-level mortality with %Black. The adjusted analysis shows that when mortality at a facility is compared to the mortality that would be expected for the race mix of patients, those facilities treating higher percentages of Black patients have higher mortality, on average. Figure 2 shows that the elevated mortality at such facilities is seen among both Black and non-Black patients at those facilities. The range of disparity in mortality exceeds 10%.

In the ESRD setting, the unadjusted analysis suggests that facilities treating larger percentages of Black patients have lower mortality, but cannot answer the question of how much of that trend is due to lower mortality among black patients. The adjusted analysis, simultaneously accounting for both facility differences and for patient-covariates, suggests very strongly that, in fact, facilities treating a higher % of Black patients tend to have higher mortality, when compared to the mortality that would be expected given their case mix. In the ESRD dialysis setting, mortality analyses that are adjusted for race appear to provide the clearest evaluation of the quality of care that is provided by facilities and these adjusted analyses do not obscure disparities in access to health care, but instead, appear to clarify those disparities.

Figure 1

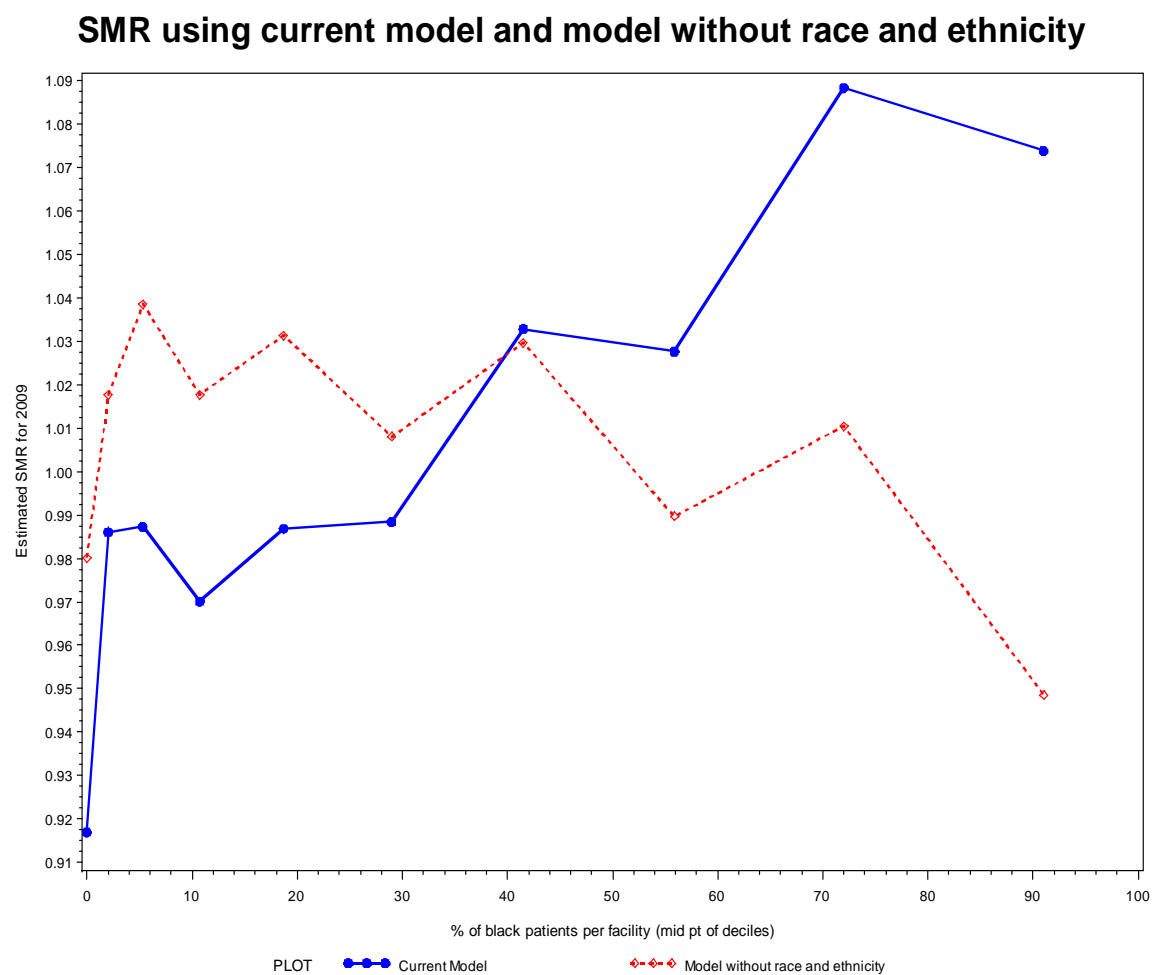
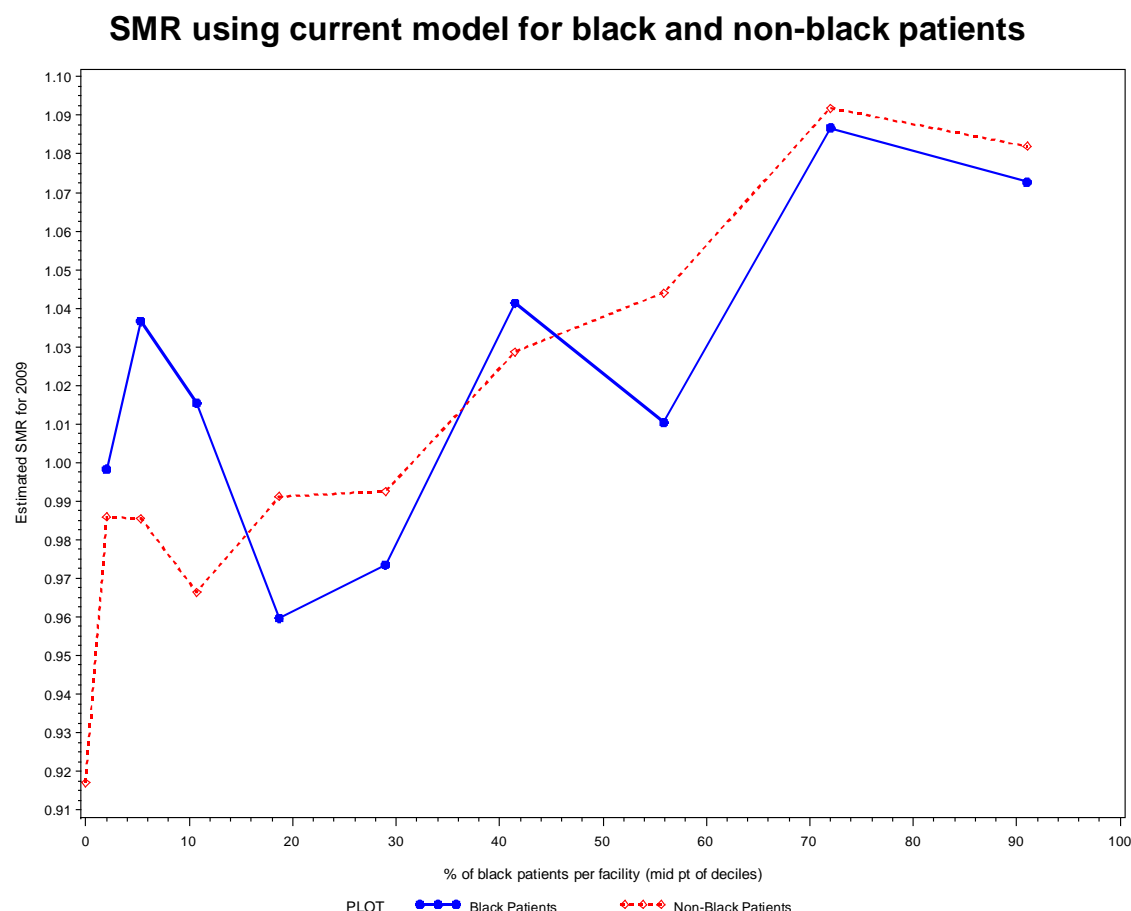
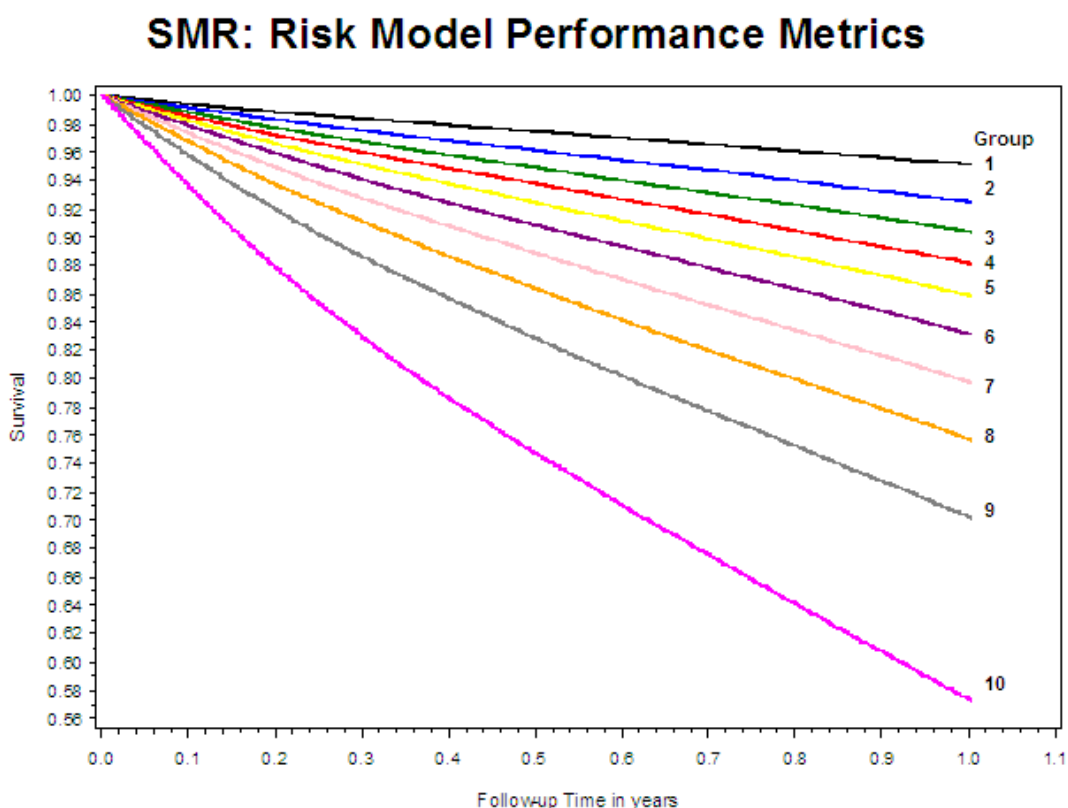


Figure 2

Risk Adjustment Testing Results (2b4.3)

Question: Is There A Risk Decile Plot For The Entire Model?

Decile plots showing estimates of the cumulative rates by years of follow up are plotted in Figure 3. The plot shows that the risk factors in the model are discriminating well between patients. There is good separation among all 10 groups and the ordering is as predicted by the model (patients predicted to be at lower risk have the best survival rates). The absolute differences between the groups is also large with survival at one year ranging from 95% for those patients predicted to have the lowest mortality rates (group 1 in black) down to 55% for those predicted to have the lowest rates of survival (group 10 in pink).

Figure 3

The Comorbidity Index is appropriately treated in the SMR model. This is illustrated by comparing a Comorbidity Index decile plot that treats the index as a continuous variable (Figure 4) with a decile plot with the Comorbidity Index = 0 indicator (Figure 6). In Figure 5, we can see a non-linearity clustered at 0 with a significant portion >20%. However, with the indicator there is a reduction of 1050 in the -2LogLikelihood , which is supported by the decile plot (Figure 5).

Similarly, comparing decile plots of $\log(\text{BMI})$ without (Figure 6) and with a linear spline (Figure 7) supports modeling $\log(\text{BMI})$ using a linear spline with a single knot at 3.5. Furthermore, the model with a linear spline had a reduction of 636 in the -2LogLikelihood compared to the model without the linear spline.

Since age has an interaction with race (black versus non-black) they are plotted separately in two trajectories in the decile plot (Figure 8). This plot shows that the knot at age 14 included in our model works well for both races.

Figure 4: Comorbidity Index Risk Decile Plot

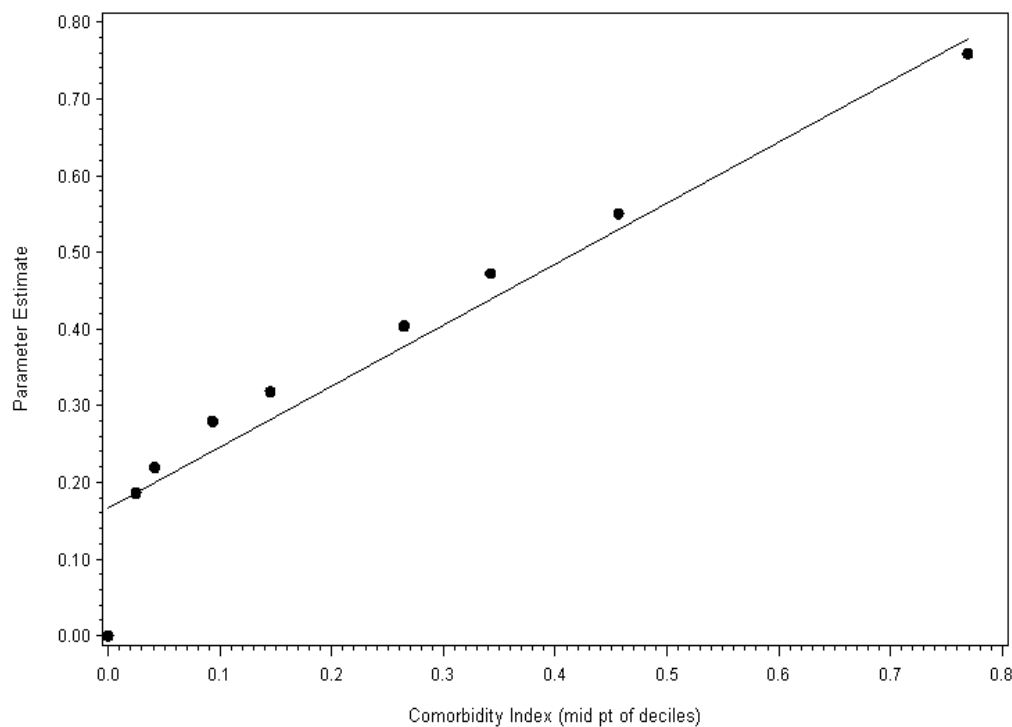


Figure 5: Risk Decile Plot for Comorbidity Index with Comorbidity Index = 0 Indicator

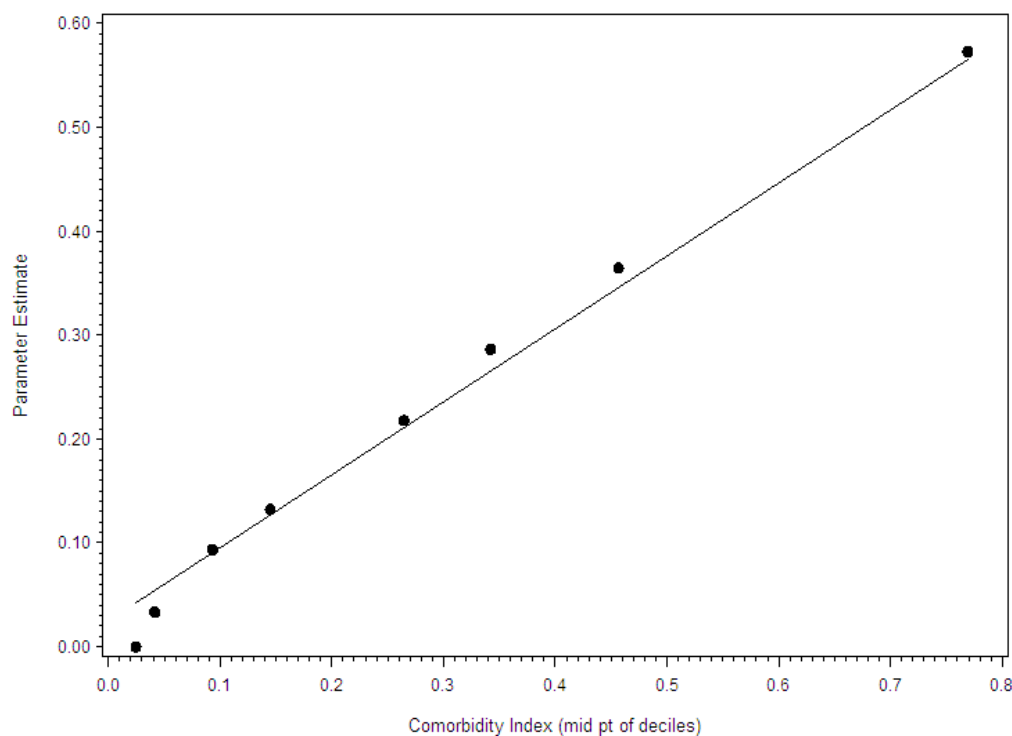


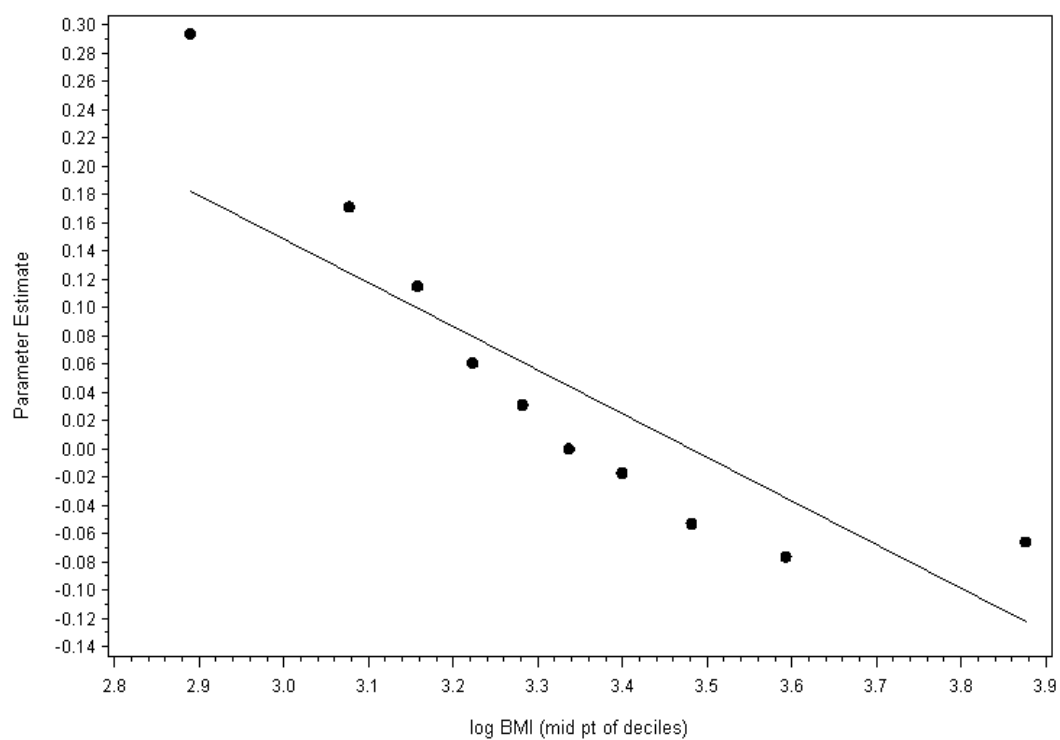
Figure 6: Risk Decile Plot for Log(BMI)

Figure 7: Risk Decile Plot for Log(BMI) with Knot at Log(BMI) = 3.5

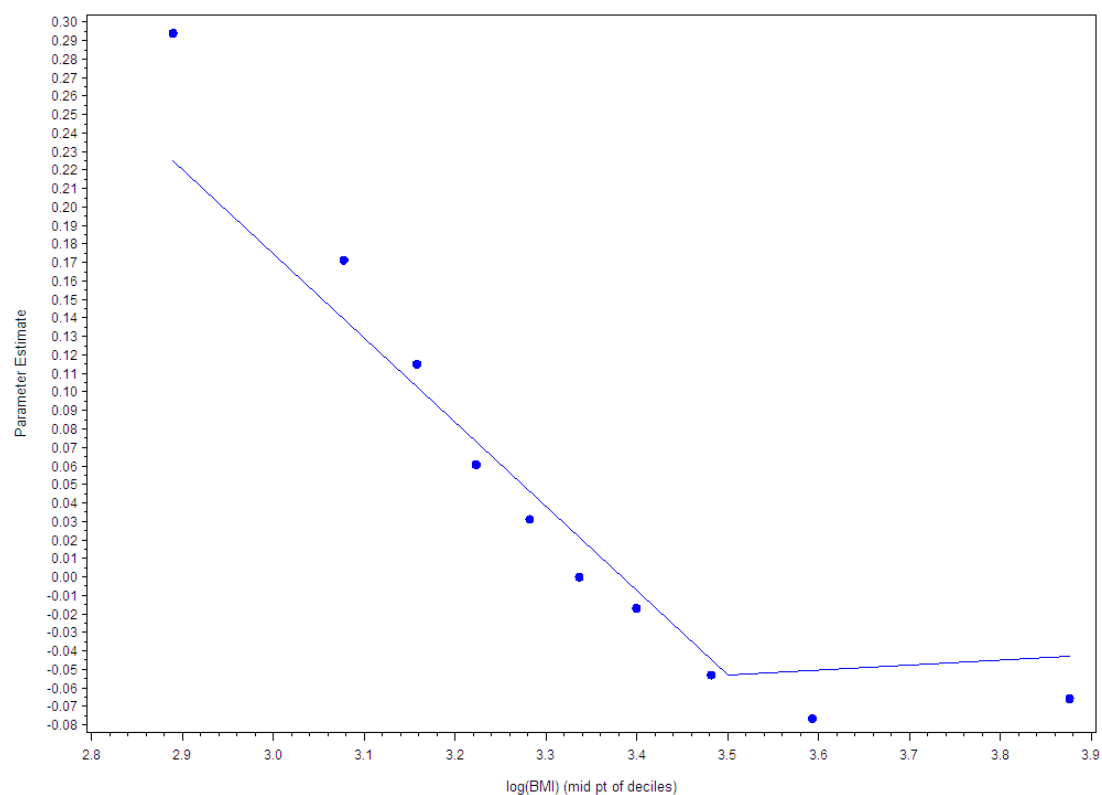
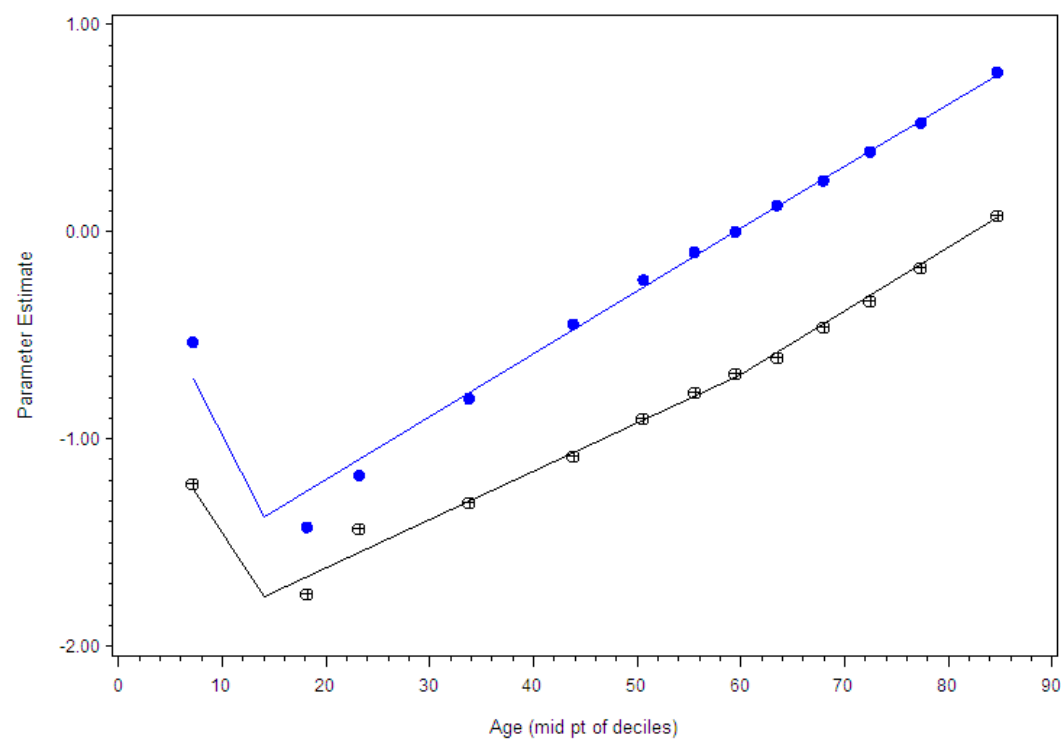


Figure 8: Risk Decile Plot for Age



ACTIVE HEALTH MANAGEMENT

- **0626** Chronic Kidney Disease - Lipid Profile Monitoring
- **0627** Chronic Kidney Disease with LDL Greater than or equal to 130 – Use of Lipid Lowering Agent

1c. Evidence

We identified some common issues across many of the submissions noted below. Please review your submissions and send us any clarifications. Please keep in mind that NQF does not expect the developer to conduct a primary evidence review; rather the developer is asked to report on the review/grading of body of evidence that was conducted by others – specifically on the quantity, quality, and consistency of the body of evidence.

- **1c6. Quality of the Body of the Evidence.** Most of the submissions did not address the quality of the body of evidence. What did the systematic review of the body of evidence determine about the overall quality of the body of evidence? We ask for a grade in item 1c13, but for this item, we want substantive information about the quality. If the review did not address the overall quality of the body of evidence, please state that and indicate what was identified about the quality of individual studies.

ActiveHealth Management Response

Citation:

- a. Chronic kidney disease: effects on the cardiovascular system. Schiffrin EL, Lipman ML, Mann JF. *Circulation*. 2007 Jul 3;116(1):85-97.
- b. K/DOQI Clinical Practice Guidelines for Managing Dyslipidemias in Chronic Kidney Disease. National Kidney Foundation American Journal of Kidney Diseases, Vol 41, No 4, Suppl 3 (April), 2003: pp S8-S9.

Taken from the K/DOQI Guidelines:

Assessment of Dyslipidemias

Evidence supporting guideline statements regarding the assessment of dyslipidemias was sought in published studies on (1) the prevalence of dyslipidemias in CKD; (2) the association between dyslipidemias and ACVD; and (3) the association between dyslipidemias and CKD progression. To ascertain the prevalence of dyslipidemias in CKD, the Work Group and Evidence Review Team examined retrospective and prospective cohort studies. To ascertain the association between dyslipidemias and ACVD or CKD progression, the Work Group and Evidence Review Team examined retrospective and prospective cohort studies, as well as case-control studies.

Treatment of Dyslipidemias

Evidence supporting guideline statements regarding the *efficacy* of treatment of dyslipidemias was sought only in randomized controlled trials of patients with CKD. Direct and indirect evidence on the *safety* of treatment of dyslipidemias in CKD was sought in controlled and uncontrolled studies of (1) the pharmacokinetics of lipid-lowering medications in CKD; (2) possible drug interactions in CKD; and (3) possible adverse reactions to lipid-lowering therapies in CKD (including small series and case reports).

Associations Between Dyslipidemias and ACVD in CKD

The incidence of ACVD is very high in patients with CKD (Fig 5). Therefore, the NKF Task Force on CVD and the K/DOQI Work Group on CKD both concluded that, in the management of risk factors such as

dyslipidemia, patients with CKD should be considered to be in the highest risk category, i.e., equivalent to that of patients with known CHD.^{2,4} There is very strong evidence from the general population that dyslipidemias cause ACVD, and this evidence has led to the ATP III guidelines for evaluation and treatment.³ It is conceivable that the pathogenesis of ACVD is different in patients with CKD, and that dyslipidemias do not contribute to ACVD in CKD. However, the relationship between dyslipidemias and ACVD in the general population is robust, i.e., it is valid in men and women^{3,46,47}; old and middle-aged^{3,46,47}; smokers and nonsmokers^{3,47}; hypertensive and non-hypertensive patients⁴⁷; diabetics and nondiabetics^{3,48}; and individuals with higher or lower LDL,^{3,47} higher or lower total cholesterol,^{3,47} higher or lower triglycerides, ^{3,47} and higher or lower HDL(Fig 4).^{3,47,49,50} There are no compelling reasons to assume that dyslipidemias do not contribute to ACVD in patients with CKD as well.

There are no randomized, controlled, intervention trials testing the hypothesis that dyslipidemias cause ACVD in patients with CKD. However, in an observational study of 3,716 patients initiating treatment for Stage 5 CKD in 1996, the use of statins in 362 (9.7%) was independently associated with lower all-cause mortality and a reduction in CVD deaths during follow-up.⁸⁴ Unfortunately, it is likely that the patients using statins had other favorable characteristics that were not accounted for in the adjusted analysis, but may have explained their reduced risk for CVD independent of their use of statins. Therefore, these study results are consistent with, but do not prove, the hypothesis that dyslipidemias contribute to ACVD in patients with CKD.

Associations Between Dyslipidemias and ACVD in Kidney Transplant Recipients

Several studies have reported a positive association between total cholesterol and ACVD in kidney transplant recipients (Table 12). Unfortunately, few of these studies examined the relationship between LDL and ACVD. Lower levels of HDL were associated with ACVD in 3 of 4 studies. In 3 of 6 studies, higher levels of triglycerides were associated with ACVD. Altogether these studies suggest that the relationship between ACVD and dyslipidemias in kidney transplant recipients is similar to that observed in the general population. However, each of these studies had design limitations; in particular, none was truly prospective. Kidney transplant recipients may also have nontraditional lipoprotein abnormalities that could theoretically contribute to ACVD.¹²⁵⁻¹²⁷ However, the role of these lipoprotein abnormalities in the pathogenesis of ACVD in CKD, as in the general population, is unclear.

Rationale for Treating High LDL Cholesterol

The ATP III Guidelines were developed using rigorous, evidence-based methods. In the absence of data from randomized trials conducted in patients with CKD, it is reasonable to assume that the interventions recommended by the ATP III will similarly reduce ACVD in patients with CKD. However, randomized trials proving that treatment of dyslipidemias reduce the incidence of ACVD ultimately need to be conducted.

The risk of CHD events is markedly increased in patients with CKD.^{2,4} Therefore, patients with CKD should be considered to have a risk equivalent to that of CHD. This risk category in the ATP III Guidelines includes patients with known ACVD, patients with diabetes, and patients with an expected 10-year risk of CHD >20%. Evidence suggests that patients with CKD have an expected 10-year CHD risk >20%,^{2,4} thereby justifying their inclusion in this highest risk category.

Treating High LDL With a Statin

There is strong evidence from studies in the general population that statins reduce CHD events

and all-cause mortality. The reduction in mortality and in CHD events is proportional to the reduction in LDL. The literature search identified only 2 small, controlled trials of simvastatin in hemodialysis patients (Table 27), and only 2 randomized trials demonstrating the efficacy of statins in peritoneal dialysis patients (Table 28). There is substantial evidence that statins are safe and effective in reducing LDL in kidney transplant recipients (Table 29). In the absence of strong evidence to the contrary, it is reasonable to assume that statins will reduce LDL and thereby ACVD in most patients with CKD. Statins are clearly the most effective class of antilipemic agents for reducing LDL.

- **1c7. Consistency of Results across Studies.** Most of the submissions did not provide information on consistency of the magnitude and direction of effect across the studies in the body of evidence. For the outcomes studied, what was the magnitude and direction of effect? If a meta-analysis was conducted, the results would be important evidence. Information from evidence tables can be used to provide substantive information on effect size.

ActiveHealth Management Response

Citation: K/DOQI Clinical Practice Guidelines for Managing Dyslipidemias in Chronic Kidney Disease. National Kidney Foundation American Journal of Kidney Diseases, Vol 41, No 4, Suppl 3 (April), 2003: pp S8-S9.

Evidence tables show that there is some association between dyslipidemia and cardiovascular disease and patients who are on peritoneal dialysis, or had a transplant; there is a possible association of dyslipidemia and hemodialysis patients.

There a number of trials showing the benefits of treatment of dyslipidemia in patients with chronic kidney disease.

The evidence tables are below for the prevalence and association of dyslipidemia in patients with chronic kidney disease.

Table 10. Associations between Dyslipidemias and Cardiovascular Disease in Hemodialysis Patients.

Study	N	Quality ^a	Applica- bility ^b	Ad- justed ^c	Cardiovascular Disease Risk with Worsening Dyslipidemia ^d			
					CHOL	LDL	HDL	TG
Cheung, ⁸⁹ 2000	936	●	†††	Yes	↔			
Kronenberg, ⁹⁰ 1999	440	●	†††	Yes	↔	↔	↓	↔
Zimmerman, ¹⁸ 1998	280	●	†††	Yes	↔	↔	↓	↔
Cressman, ⁹¹ 1992	129	●	†††	Yes	↔	↔	↔	↔
Stack, ⁹² 2001	3,925	○	†††	Yes	↔	↔	↔	↔
Degoulet, ⁹⁴ 1982	1,453	○	†††	Yes	↔			↓
Iseki, ⁹⁵ 1996	1,491	○	††	Yes	↔			↔
Goldwasser, ⁹⁶ 1993	125	○	††	No	↑		↔	
Kimura, ⁹⁷ 1996	195	○	††	Yes	↔			
Fujisawa, ⁹⁸ 2000	51	○	††	No	↔	↔		↔
Yeun, ²⁰ 2000	91	○	††	Yes	↔			

^aStudy quality was graded: ● least bias, results are valid; ○ susceptible to some bias, but not sufficient to invalidate the results; or ○ significant bias that may invalidate the results.

^bApplicability was rated: ††† representative of a wide spectrum of patients; †† representative of a relevant subgroup; or † representative of a narrow subgroup.

^cIndicates whether results were statistically adjusted for covariates.

^d↔ indicates no association between dyslipidemia and cardiovascular disease; ↓ indicates that dyslipidemia was associated with less cardiovascular disease or there was a trend that was not statistically significant (↘); and ↑ indicates that dyslipidemia was associated with more cardiovascular disease or there was a trend that was not statistically significant (↗).

Abbreviations: N, number of subjects in the study; CHOL, cholesterol; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; TRIG, triglycerides.

Table 11. Associations between Dyslipidemias and Cardiovascular Disease in Peritoneal Dialysis Patients.

Study	N	Quality ^a	Applica- bility ^b	Ad- justed ^c	Cardiovascular Disease Risk with Worsening Dyslipidemia ^d			
					CHOL	LDL	HDL	TG
Webb, ⁹⁹ 1993	75	○	††	Yes	↑	↑	↔	↔
Olivares, ¹²⁴ 1992	102	○	††	No	↔	↔	↔	↔

^aStudy quality was graded: ● least bias, results are valid; ○ susceptible to some bias, but not sufficient to invalidate the results; or ○ significant bias that may invalidate the results.

^bApplicability was rated: ††† representative of a wide spectrum of patients; †† representative of a relevant subgroup; or † representative of a narrow subgroup.

^cIndicates whether results were statistically adjusted for covariates.

^d↔ indicates no association between dyslipidemia and cardiovascular disease; ↓ indicates that dyslipidemia was associated with less cardiovascular disease or there was a trend that was not statistically significant (↘); and ↑ indicates that dyslipidemia was associated with more cardiovascular disease or there was a trend that was not statistically significant (↗).

Abbreviations: N, number of subjects in the study; CHOL, cholesterol; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; TRIG, triglycerides.

Table 12. Associations between Dyslipidemias and Cardiovascular Disease in Kidney Transplant Recipients.

Study	N	Quality ^a	Applica- bility ^b	Ad- justed ^c	Cardiovascular Disease Risk with Worsening Dyslipidemia ^d			
					CHOL	LDL	HDL	TG
Kasiske, ¹²⁹ 1996	675	●	↑↑↑	Yes	↔	↔	↓	↑
Aker, ¹²⁹ 1998	427	●	↑↑	Yes	↑	↑	↓	↔
Aakus, ¹³⁰ 1999	406	●	↑	No	↑		↔	
Kasiske, ¹³¹ 2000	1,124	○	↑↑	Yes	↑	↑	↔	↑
Ong, ¹³² 1994	192	○	↑↑	No	↑		↔	↔
Barbagallo, ¹³³ 1999	57	○	↑↑	Yes	↔	↓	↑	↔
Roodnat, ¹³⁴ 2000	676	○	↑↑	Yes	↑			
Massy, ¹³⁵ 1998	79	○	↑↑	Yes	↑		↑	↔
Biesenbach, ¹³⁶ 2000	21	○	↑	No	↔			↓

^aStudy quality was graded: ● least bias, results are valid; ○ susceptible to some bias, but not sufficient to invalidate the results; or ○ significant bias that may invalidate the results.

^bApplicability was rated: ↑↑↑ representative of a wide spectrum of patients; ↑↑ representative of a relevant subgroup; or ↑ representative of a narrow subgroup.

^cIndicates whether results were statistically adjusted for covariates.

^d↔ indicates no association between dyslipidemia and cardiovascular disease; ↓ indicates that dyslipidemia was associated with less cardiovascular disease or there was a trend that was not statistically significant (↘); and ↑ indicates that dyslipidemia was associated with more cardiovascular disease or there was a trend that was not statistically significant (↗).

Abbreviations: N, number of subjects in the study; CHOL, cholesterol; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; TRIG, triglycerides.

Table 27. Randomized Trials Evaluating the Treatment of Dyslipidemias in Hemodialysis Patients.

Study	N	Quality ^a	Applicability ^b	Treatment	Change Compared To Baseline (%) ^c			
					CHOL	LDL	HDL	TG
Soroka, ²⁹² 1998	15	●	†	Animal-based low-protein diet vs. Soy-based vegetarian diet	-5 -5	-6 -3	-10 -16	+8 -4
Khajehdehi, ²⁹³ 2000	84	○	†	Vitamin E vs.	—	-5	+15	—
				Vitamin D ₃ vs.	-4	-23	+3	-10
				Vitamin C vs.	-13*	-16*	—	-2
				Placebo	—	+5	+4	-1
Khajehdehi, ²⁹⁴ 2000	60	○	†	Fish oil (1.5 g/d) vs.	-10	-16	+138*	-11
				Corn oil (4.5 g/d) vs.	-2	-18	+42*	-10
				Sesame oil (4.5 g/d) vs.	-5	-1	+28	-2
				Placebo	+4	-13	+5	-1
Seri, ²⁹⁵ 1993	28	○	†	Fish oil vs.	-1	-24	+35	-4
				Blackcurrant seed oil vs.	—	-22	+10	-9
				Mixed oils vs.	—	+15	+39	-11
				Placebo	—	+5	+2	+3
Schrader, ²⁹⁶ 1988	70	●	††	LMWH vs.	+5	+18	—	—
				Standard heparin	+2	+38	+5	+36
Kronenberg, ²⁹⁷ 1995	48	○	††	LMWH vs.	+6*	+13*	+6*	—*
				Standard heparin	-3	-4	+22	+18
Saltissi, ²⁹⁸ 1999	36	○	††	LMWH vs.	—	+1	—	-9
				Unfractionated heparin	—	—	-3	+4
Blankenstijn, ²⁹⁹ 1995	28	●	†	High-flux vs.	-9	-3	+8	-28
				Low-flux dialysis membrane	+2	+5	+3	+2
Golper, ³⁰⁰ 1990	82	●	††	L-carnitine vs.	—	-7	-3	—
				Placebo	-3	-2	-3	—
Weschler, ³⁰¹ 1984	10	●	††	L-carnitine vs.	+17	—	—	+22
				Placebo	+15	—	—	—
Nilsson-Ehle, ³⁰² 1985	28	●	†	L-carnitine vs.	—	—	+5	+5
				Placebo	+1	+2	—	-1
Giorcelli, ³⁰³ 1980	84	○	††	L-carnitine vs.	—	—	—	-41
				Thiadenol	—	—	—	-42
Casciani, ³⁰⁴ 1982	15	○	†	L-carnitine vs.	—	—	+23	—
				Placebo	—	—	+7	—
Saltissi, ³⁰⁵ 2002	33	●	†††	Simvastatin (n=22) vs.	-21*	-25*	-6	-18
				Placebo (n=11)	-12	-14	-3	-14

Table 27. Randomized Trials Evaluating the Treatment of Dyslipidemias in Hemodialysis Patients, continued.

Chang, ³⁰⁶ 2002	62	●	†††	Simvastatin (n=31) vs. No treatment (n=31)	-16* +2	-41* +3	+3 -3	-17* +2
Chertow, ³⁰⁷ 1997	36	●	†††	Sevelamer hydrochloride vs. Placebo	-10* +4	* +2	+5 +2	—
Chertow, ³⁰⁸ 2002	200	●	†††	Sevelamer hydrochloride vs. Placebo	-22* -1	-36* +1	-2 —	-7 +8
Park, ³⁰⁹ 2000	65	●	††	Estrogen/medroxyprogesterone vs. No treatment	-1 -0	-9 +1	+12 -2	+20 +2
				Exercise vs. Control	-3 -3	NC NC	+16 -7	-33 -3

^aStudy quality was graded: ● least bias, results are valid; ○ susceptible to some bias, but not sufficient to invalidate the results; or ○ significant bias that may invalidate the results.

^bApplicability was rated: ††† representative of a wide spectrum of patients; †† representative of a relevant subgroup; or † representative of a narrow subgroup.

^cThe percent change within each group is indicated. *** indicates a statistically significant difference between treatment and control groups.

Abbreviations: LMWH, low molecular weight heparin; NC, no change (as stated by the authors, without reporting data).

Table 28. Randomized Trials Evaluating the Treatment of Dyslipidemias in Peritoneal Dialysis Patients.

Study	N	Quality ^a	Applicability ^b	Treatment	Change Compared To Baseline (%) ^c			
					CHOL	LDL	HDL	TG
Saltissi, ³⁰⁵ 2002	23	●	†††	Simvastatin (n=16) vs.	-22*	-25*	0	-2
				Placebo (n=7)	-2	-4	+5	+4
Harris, ³¹⁰ 2002	130	●	†††	Atorvastatin vs.	-29*	-40*	+7*	-14*
				Placebo	-6	-9	-4	+11

^aStudy quality was graded: ● least bias, results are valid; ○ susceptible to some bias, but not sufficient to invalidate the results; or ○ significant bias that may invalidate the results.

^bApplicability was rated: ††† representative of a wide spectrum of patients; †† representative of a relevant subgroup; or † representative of a narrow subgroup.

^cThe percent change within each group is indicated. *** indicates a statistically significant difference between treatment and control.

Table 29. Randomized Trials Evaluating the Treatment of Dyslipidemia in Kidney Transplant Recipients.

Study	N	Quality ^a	Applicability ^b	Treatment	Change Compared To Baseline (%) ^c			
					CHOL	LDL	HDL	TG
Bennett, ³¹¹ 1995	90	●	↑	High-dose EPA vs. Low-dose EPA		+6 +6*	-10 -5	
Urakaze, ³¹² 1989	30	●	↑	Fish oil vs. Placebo			-25* +4	
Maachi, ³¹³ 1995	80	○	↑	Fish oil vs. Placebo	—	—	—	—
Holdaas, ³¹⁴ 2001	364	●	↑↑↑	Fluvastatin vs. Placebo	-18* +28	-41* +26	+5 +29	-25* +44
Kasiske, ³¹⁵ 2001	141	●	↑↑↑	Simvastatin vs. Gemfibrozil vs. Placebo	— —	— —	— —	— —
Arnadottir, ³¹⁶ 1994	37	●	↑↑	Simvastatin vs. Placebo	-22* -3	-32* -4	+13* —	-12 +4
Martinez-Hernandez, ³¹⁷ 1993	21	●	↑↑	Simvastatin vs. Placebo	-16* +5	-20* +1	+7 +4	-15 +8
Castro, ³¹⁸ 1997	36	●	↑↑	Simvastatin vs. Fish oil	-16 -10	-18 -2	-3 -16	-26 -23
Katznelson, ³¹⁹ 1996	48	●	↑↑	Pravastatin vs. Control	-15* +30	* —		* —
Kliem, ³²⁰ 1996	44	●	↑↑	Pravastatin vs. Lovastatin	-16 -26	-21 -31	— -6	-6 -25
Sahu, ³²¹ 2001	65	●	↑↑	Lovastatin vs. Placebo	* —	* —		* —
Kasiske, ³²² 1990	11	●	↑↑	Lovastatin vs. Diet	-20	-27	+4	-17
Renders, ³²³ 2001	10	●	↑↑	Atorvastatin vs. Cerivastatin	-30 -32	-43 -35	+16 +6	-23 -23
Santos, ³²⁴ 2001	67	○	↑↑	Simvastatin vs. Placebo	-23* -2	-35* -2	+7 -8	-10 +2

^aStudy quality was graded: ● least bias, results are valid; ○ susceptible to some bias, but not sufficient to invalidate the results; or ○ significant bias that may invalidate the results.

^bApplicability was rated: ↑↑↑ representative of a wide spectrum of patients; ↑↑ representative of a relevant subgroup; or ↑ representative of a narrow subgroup.

^cThe percent change within each group is indicated. "*" indicates a statistically significant difference between treatment and control.

^dAbbreviation: EPA, eicosapentaenoic acid (20:3).

- 1c.11/1c.21 System used for grading the body of evidence and grading the recommendation.**
 You can select USPSTF, GRADE, or other. Other is acceptable as long as you describe it. Please note, that grading the body of evidence is different from the strength of the recommendation. Some submissions indicated that the GRADE system was used, but the actual grades provided in items 1c.13/1c.23 were different than those described in GRADE documents. In that case, please clarify the differences (was the GRADE system used, modified, just different labels for the GRADE scale) and provide the rating scale with definitions.

ActiveHealth Management Response

1c.12: Grading of the body of evidence used by K/DOQI:

Table 8. Rating the Strength of Evidence.

Outcome(s)	Population	Methodological Quality		
		Well designed and analyzed (little, if any, potential bias)	Some problems in design and/or analysis (some potential bias)	Poorly designed and/or analyzed (large potential bias)
Health outcome(s)	Target population	Strong ^a	Moderate ^b	Weak ^h
Health outcome(s)	Other than the target population	Moderate ^c	Moderate ^d	Weak ^h
Surrogate measure for health outcome(s)	Target population	Moderate ^e	Weak ^f	Weak ^h
Surrogate measure for health outcome(s)	Other than the target population	Weak ^g	Weak ^g	Weak ^{g,h}

Strong- ^aEvidence includes results from well-designed, well-conducted study/studies in the target population that directly assess effects on net health outcomes.

Moderate- ^bEvidence is sufficient to determine effects on net health outcomes in the target population, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies; ^cOR ^eevidence is from a population other than the target population, but from well-designed, well-conducted studies; ^dOR ^fevidence is from studies with some problems in design and/or analysis; ^gOR ^hevidence is from well-designed, well-conducted studies on surrogate endpoints for efficacy and/or safety in the target population.

Weak- Evidence is insufficient to assess the effects on net health outcomes because it is from studies with some problems in design and/or analysis on surrogate endpoints for efficacy and/or safety in the target population; ^gOR ^hthe evidence is only for surrogate measures in a population other than the target population; ^hthe evidence is from studies that are poorly designed and/or analyzed.

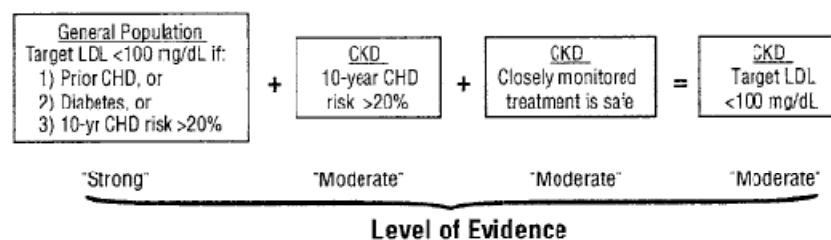


Fig 3. The chain of logic for evidence supporting the treatment of low-density lipoprotein cholesterol in patients with chronic kidney disease. See text for details. Abbreviations: LDL, low-density lipoprotein; CHD, coronary heart disease; CKD, chronic kidney disease.

1c.13: Moderate Evidence

Grading of the Recommendation

Rating the Strength of Recommendations.

Grade	Recommendation
A	It is strongly recommended that clinicians routinely follow the guideline for eligible patients. There is strong evidence that the practice improves net health outcomes.
B	It is recommended that clinicians routinely follow the guideline for eligible patients. There is moderate evidence that the practice improves net health outcomes.
C	It is recommended that clinicians consider following the guideline for eligible patients. This recommendation is based on either weak evidence, poor evidence or on the opinions of the Work Group and reviewers, that the practice might improve net health outcomes.

Health outcomes are health-related events, conditions, or symptoms that can be perceived by individuals to have an important effect on their lives. Improving net health outcomes implies that benefits outweigh any adverse effects.

1b2. Opportunity for Improvement – what is the performance data for the measure as specified?

ActiveHealth Management Response

Citation:

- a. Primary Care Management of Chronic Kidney Disease. Allen et al. J Gen Intern Med 26(4):386–92.
- b. Association of preventive health care with atherosclerotic heart disease and mortality in CKD. Snyder JJ, Collins AJ. J Am Soc Nephrol. 2009 Jul;20(7):1614-22.
- c. KDOQI hypertension, dyslipidemia, and diabetes care guidelines and current care patterns in the United States CKD population: National Health and Nutrition Examination Survey 1999-2004. Snyder JJ, Collins AJ. Am J Nephrol. 2009;30:44-54.

Here is our original response, which directly speaks to the performance gap for the measures:

“Despite the evidence for the importance of appropriate management of hyperlipidemia in CKD, and national guidelines recommending annual cholesterol screening in all patients with CKD, percent of patients with CKD, who had LDL levels, shows room for improvement. This was demonstrated in a study, which included 166 primary care physicians caring for over 300,000 adult patients, including 11,774 patients with CKD. Overall, only 75% of patients with CKD had annual cholesterol screening. There was statistically significant variability in screening in different groups based on gender (male, 76.6% vs. female, 74.5%) ethnicity (black, 72.4% vs. white, 75.5%), and type of insurance (commercial, 77.2% vs. uninsured, 45.1%). In addition, patient with morbid conditions (diabetes, hypertension or CVD) were significantly more likely to receive LDL screening than those without comorbidities.”

There are a number of other studies that show that there is a gap in performance with respect to adhering to guidelines. The Snyder and Collins study identified 14,213 patients with CKD. They found that high cholesterol was more likely in patient with stages 3 and 4 and awareness and treatment was less likely than those without CKD.

Reliability 2a

2a2. Repeatability of electronic data is ensured by testing algorithms and programming – assures will get the same result (not necessarily the correct result – that’s the validity of data question). With so much data available, can you assess the reliability of performance scores (signal- to-noise)?

ActiveHealth Management Response

I don’t understand the question.

For over 10 years, ActiveHealth has used clinically-enriched claims data to identify patients for specific interventions. As a part of the process we have built rules that err on the side of specificity, i.e., we use condition validation rules that attempt to ensure that we identify conditions appropriately and minimize false positives. We use the condition validation rules to support a number of programs including identifying patients for disease management, for our clinical decision support rules, and for our performance measurement. For example, we use the same chronic kidney disease condition validation rule to identify patients with kidney disease for disease management, the same validation rules for clinical disease management and for the denominator of the renal performance measures.

Generally, whenever if we receive feedback from a provider indicating that we did not identify a patient correctly, we adjust our rules to minimize false positives.

2b. Validity

2b2. Validity testing – Essentially repeated what is under reliability testing. Validity at the data element level should assess accuracy of the data used in the performance measure (not the accuracy of programming). Generally that entails assessing agreement with an authoritative source (or citing prior studies that addressed accuracy such as CKD diagnosis in claims, etc.) What is meant by “ensure that we receive valid codes”? In 4c1 you mention “corroborate the data” - what does that mean? What are quantitative results of these processes?

ActiveHealth’s Response

Citation: Identification of individuals with CKD from Medicare claims data: a validation study.
Winkelmayer WC, Schneeweiss S, Mogun H, Patrick AR, Avorn J, Solomon DH. Am J Kidney Dis. 2005 Aug;46(2):225-32.

Winkelmayer et al. studied the specificity and sensitivity of claims data for identifying patients with chronic kidney disease. They identified all patients who were admitted to a regional hospital with myocardial infarction and who had a creatinine on admission (n = 1,852). They compared the claims data for the patients before admission and calculated the creatinine clearance. They also completed a chart abstraction for the study. They found that claims data was specific for identification of patients with CKD

Table 3. Test Characteristics of Several Algorithms for CKD Compared With a Gold Standard of eGFR Less Than 60 mL/min/1.73 m² (<1.00 mL/s/1.73 m²) for a 6-Month Ascertainment

Diagnosis	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
DN	1.7 (1.0-2.4)	99.0 (98.3-99.8)	79.3 (64.6-94.1)	32.2 (30.1-34.3)
ARF	3.9 (2.9-5.0)	99.7 (99.2-100.0)	96.3 (91.2-100.0)	32.8 (30.7-35.0)
HypN	4.8 (3.6-5.9)	99.5 (99.0-100.0)	95.5 (90.4-100.0)	33.0 (30.9-35.1)
CRI	9.1 (7.5-10.6)	99.7 (99.2-100.0)	98.4 (96.1-100.0)	34.1 (31.9-36.2)
MISC	13.3 (11.5-15.1)	97.6 (96.4-98.8)	92.1 (88.3-96.0)	34.7 (32.4-36.9)
DN + HypN + CRI	12.6 (10.8-14.3)	98.4 (97.4-99.4)	94.3 (90.9-97.7)	34.7 (32.4-36.9)
DN + HypN + CRI + ARF	13.6 (11.8-15.5)	98.2 (97.2-99.3)	94.2 (90.9-97.5)	34.9 (32.7-37.1)
DN + HypN + CRI + MISC	20.2 (18.0-22.4)	96.1 (94.6-97.7)	91.8 (88.6-94.9)	36.2 (33.9-38.5)
DN + HypN + CRI + MISC + ARF	20.7 (18.5-22.9)	96.0 (94.4-97.5)	91.6 (88.5-94.8)	36.3 (34.0-38.7)

NOTE. Calculated for a prevalence of CKD of 67.2%.

*chronic renal insufficiency (CRI), diabetic nephropathy (DN), hypertensive nephropathy (HypN), acute renal failure (ARF), and miscellaneous other renal disease (MISC).

All of the electronic data that we receive go to a data warehouse, where they are normalized, i.e., if we receive a code that is not consistent with a codeset that code is removed, e.g., if we had an ICD-9 code of 5000.333 it would be removed since it is not consistent with an ICD-9 code; if a laboratory result is inconsistent, e.g., a potassium result that is > 20 mmol/l would not be included in the normalized data set since this is not physiologically compatible with life.

In addition, to the validation of the codes we receive, depending on the rule, we usually corroborate any administrative data we receive with supporting evidence, for example we will look for a claim and pharmacy data to support a specific condition.

2b5. What are the performance scores on the measures as specified?

ActiveHealth Management Response

Using our test data, the compliance rate of the #0626 was 86.4% (denominator: 21,693). In addition, we tested client data (for baseline purposes) and found a compliance rate of 79% (population, 162, 131; denominator, 1496). For #0627 the compliance was 32.4% (denominator: 37)

3. Usability

3.1 Indicated the measure is currently used in public reporting; in 3a1 states clients publish results publicly. Where can those performance results be accessed?

ActiveHealth Management Response

In the past we sent reports to a health plan, which then made the results available to their physicians. We are currently working with a number of clients as a part of ACO initiative and anticipate that we will support public reporting and quality initiatives.

3b1. What is rationale for usefulness for QI?

.

ActiveHealth Management Response

Renal disease occurs in 13% of the population and cardiovascular disease causes significant morbidity and mortality. There is literature to show that few patients are screened for risk factors for cardiovascular disease or treated. Based on our testing on real data we found 86 % compliance with lipid panel testing and 32 % compliance with treating for dyslipidemia. This performance gap is supported by the literature and therefore is important for QI.