**SUMMARY**
**Kidney Care Quality Alliance Conference Call**
**April 18, 2016**

A conference call of the Kidney Care Quality Alliance (KCQA) was convened on Monday, April 18, 2016.

Representatives of the following organizations participated: American Kidney Fund, American Nephrology Nurses' Association, American Society of Nephrology, American Society of Pediatric Nephrology, Centers for Dialysis Care, DaVita Healthcare Partners Inc., Dialysis Patient Citizens, Dialysis Clinic Inc., Fresenius Medical Care North America, Fresenius Medical Care Renal Therapies Group, Greenfield Health Systems, Kidney Care Council, Kidney Care Partners, National Forum of ESRD Networks, National Kidney Foundation, National Renal Administrators Association, NxStage Medical, Renal Physicians Association.

Public members present: Dr. Joseph Messana and Dr. Claudia Dahlerus, University of Michigan KECC.

**OPENING REMARKS**
Following the roll call, Dr. Allen Nissenson, KCQA Steering Committee Co-Chair, welcomed and thanked participants for joining the call and commended the Steering Committee and Medication Management Workgroup for their work to date. He noted that KCQA has developed at least one very good metric to submit to the National Quality Forum (NQF), possibly two if data issues can be resolved. Dr. Nishimi also welcomed participants and thanked DaVita, DCI, and Fresenius for participating in measure testing.

**AGENDA**
Dr. Nishimi informed call participants that, since late January, the consultants have been working with the three testing organizations, DaVita, DCI, and Fresenius, on testing the specifications for *MM-2: Medication Reconciliation for Patients Receiving Care at Dialysis Facilities* and *MM-3: Medication Reconciliation at Care Transitions for Patients Receiving Care at Dialysis Facilities.* She noted the final data pulls were recently received and results were reviewed with the Steering Committee and Medication Management Workgroup. She informed participants the purpose of today's conference call is to review the reliability testing results and the Steering Committee's recommendations for specification changes based on those results. There were no preliminary questions from participants.

**EMPIRICAL RELIABILITY TESTING RESULTS**
Dr. Nishimi advised participants empirical reliability testing at the measure score level was conducted on data pulls for MM-2 and MM-3 from DaVita, DCI, and Fresenius, as follows:

- Each testing organization pulled Q2 and Q3 data for 2015 in accordance with the specifications. Testing organizations provided their datasets for each facility (anonymized) for each month to Drs. Craig Solid and David Gilbertson, KCQA's methodology consultants.
- Dr. Solid assessed the combined dataset using the beta-binomial test for reliability, an approach commonly used for measures before NQF and hence familiar to its Standing

Committees. The approach has been characterized as a "natural model for estimating the reliability of simple pass/fail rate measures."[1]

**General Information**
Dr. Nishimi noted the number of contributing facilities differed by month, but was generally 4,700-4,800 facilities. The ranges were:

- MM-2: 4,781 (April 2015) to 4,836 (September 2015)
- MM-3: 4,699 (April 2015) to 4,759 (August 2015)

For MM-2, the general population measure, this translated to approximately 323,000 to 328,000 measurement events per month. For MM-3, the "high-risk" measure, events were about 45,000 per month.

**Results for MM-2:** ***Medication Reconciliation for Patients Receiving Care at Dialysis Facilities***
Dr. Nishimi indicated three aspects were examined during testing: performance gap/opportunity for improvement, small sample size effect, and overall reliability.

Opportunity for Improvement: MM-2
Dr. Nishimi informed call participants NQF requires demonstration that the measure identifies a performance gap. Performance on MM-2 ranged from 0-100% achievement in each of the six months. The monthly median performance ranged from 46.4% (May 2015) to 50.8% (July 2015).

In the ensuing discussion, Dr. Klemens Meyer noted he had briefly reviewed the materials and asked for clarification on what constitutes performance for the measures. Dr. Nishimi responded three criteria must be met: 1) the facility must attest medication reconciliation was performed; 2) the date of the reconciliation must be included; and 3) the name or other unique identifier of the eligible professional who performed the reconciliation must be present. Additionally, the measure defines a number of factors that must be addressed for each medication, and allergies and intolerances also must be addressed. Dr. Meyer asked if a facility will not get credit if all the defined criteria are not met. Dr. Nishimi responded yes, for that particular patient. She added the measures were crafted to require that a series of events must occur to for the facility to be able to attest yes for each patient so the measures would not be simple "check box" measures, which are generally not viewed favorably at NQF. Dr. Meyer voiced understanding, but commented that the specifications lack face validity in his opinion. Dr. Nishimi requested the face validity discussion be deferred until she reviewed the remainder of the reliability results and they are discussed.

Assessment of Small Sample Size: MM-2
Dr. Nishimi noted the beta-binomial method yields reliability statistics for each facility. Per the literature (Adams, 2009) and the consultants' experience with NQF, a reliability statistic of 0.7 is generally viewed as an acceptable threshold. For the 6-month period and using all facilities regardless of size, MM-2 reliability was:

**Table 1. MM-2 Reliability Statistics, All Facilities**

| | |
|---|---|
| Minimum | 0.4241 |
| 10th percentile | 0.9847 |
| Median | 0.9952 |

[1] Adams, JL. The reliability of provider profiling: A tutorial. RAND Health, 2009.

| | |
|---|---|
| 90th percentile | 0.9995 |
| Maximum | 1.0 |

To assess small sample and implementation issues, however, Dr. Nishimi informed participants that Dr. Solid examined the measure's reliability in the context of KCP's policy of excluding facilities with <= 25 patients (in this case patient-events) and CMS' general implementation approach of excluding facilities with <11 patients (again, patient-events for these measures). Both the percentage of facilities that would be excluded from measurement and the reliability of the measure for small facilities was analyzed:

- Using the <11 threshold, 3.3-3.6% of facilities were excluded from MM-2, depending on the month. Using the <=25 threshold, 13.4-13.8% of facilities were excluded, again depending on the month.
- When the reliability statistic for facilities with <11 and <=25 events were examined, Dr. Nishimi noted the effect of small sample size became more clear for MM-2 (Table 2).

**Table 2. MM-2 Reliability Statistics, Facilities with <11 and <=25 Events**

| | <11 (i.e., <=10) | <=25 |
|---|---|---|
| Minimum | 0.3615 | 0.3993 |
| 10th percentile | 0.6937 | 0.8964 |
| Median | 0.9174 | 0.9773 |
| 90th percentile | 1 | 1 |
| Maximum | 1 | 1 |

She noted the measure does not achieve the desired reliability threshold of 0.7 at the 10th percentile. Conversely, using KCP's policy of <=25 events does achieve sufficient reliability for all but outliers, but Dr. Nishimi noted the consultants were concerned about the impact on quality and patient safety of eliminating approximately 13.5% of facilities from measurement each month.

To achieve the desired threshold, but also include as many facilities as possible, the data were examined to determine at what sample size a 10th percentile reliability statistic of 0.70 could be achieved. Based on this additional analysis (Table 3), Dr. Nishimi advised KCQA members the target reliability threshold is reached and <=11 (i.e., <12), and so the Steering Committee recommends KCP's current policy of <=25 be modified for MM-2 and the specifications revised to reflect the <=11 threhold. She stated this excludes 3.7-3.9% of facilities each month. She noted another option would have been to collect data on a 2-month basis, but the consultants believe retaining the 1-month construct more accurately captures the Workgroup's interest in accountability in an actionable timeframe.

**Table 3. MM-2 Reliability Statistics, Facilities with <11, <=11, <=12 Events**

| | <11 (i.e., <=10) | <=11 | <=12 |
|---|---|---|---|
| Minimum | 0.3615 | 0.3622 | 0.3714 |
| 10th percentile | 0.6937 | 0.7089 | 0.7471 |
| Median | 0.9174 | 0.9177 | 0.9293 |
| 90th percentile | 1 | 1 | 1 |
| Maximum | 1 | 1 | 1 |

Overall Reliability: MM-2

When facilities with <=11 events are excluded (~3.7%), Dr. Nishimi informed participants the overall reliability of the measure is excellent (Table 4).

**Table 4. MM-2 Reliability Statistics All Facilities vs. Small Facilities Excluded (<=11)**

| | ALL | <=11 Excluded |
|---|---|---|
| Minimum | 0.4241 | 0.8166 |
| 10th percentile | 0.9847 | 0.9867 |
| Median | 0.9952 | 0.9953 |
| 90th percentile | 0.9995 | 0.9995 |
| Maximum | 1 | 1 |

**Results for MM-3:** ***Medication Reconciliation at Care Transitions for Patients Receiving Care at Dialysis Facilities***

For MM-3, Dr. Nishimi noted three aspects were examined during empirical testing: sample size issues, performance gap/opportunity for improvement, and overall reliability. However, a feasibility issue also was identified that the Steering Committee believes raises significant validity issues.

MM-3 Sample Issues

Dr. Nishimi noted it became clear as Dr. Solid initially reviewed the data that the currently specified monthly timeframe did not yield sufficient care transition events (i.e., new admissions, discharge from skilled nursing facility, and/or post-hospitalization, emergency department visits, or observation stays). For example, using the CMS threshold of <11 events (set by CMS for privacy reasons), nearly two-thirds of facilities are excluded from measurement each month (range 61.6-64.3%). She stated they considered a quarterly timeframe, but sufficient numbers and stabilized reliability statistics for MM-3 were only achieved at a 6-month time interval. Dr. Nishimi noted the information she will report used six months of aggregated data to produce the performance scores.

Dr. Nishimi indicated using the <11 threshold excludes on average 5.95% of facilities from MM-3; using <=11 excludes ~6.5%, and the <=25 threshold excludes ~20.66% of facilities.

Dr. Nishimi asked if there were any questions on the recommendation to modify the specifications to a 6-month time interval from a 1-month measurement period. None were voiced, but Dr. Don Molony remarked that six months of data will still drive change. Dr. Nishimi agreed, noting that one month is easier to react to, but six months is certainly more actionable than a year. Dr. Molony asked if the consultants had looked at a 3-month timeframe. Dr. Nishimi said they had, but the data were not stable enough at that point.

Opportunity for Improvement: MM-3

Dr. Nishimi noted performance on MM-3 ranged from 0-100% achievement (6-month period). The median performance was 23.4% (6-month period).

Overall Reliability: MM-3

Based on the preliminary analyses to determine the degree to which the MM-3 timeframe needed to be expanded, Dr. Nishimi noted they did not further analyze reliability statistics for a range of facility sizes. Rather, she indicated they focused on overall reliability statistics when facilities with <11, <=11, and <=25 events are removed. Again, per Adams and experience with NQF, a reliability statistic of 0.7 is generally viewed as an acceptable threshold. For the 6-month period, the following was found:

**Table 5. MM-3 Reliability Statistics, Small Facilities Excluded (6-month Timeframe)**

| | <11 Excluded | <=11 Excluded | <=25 Excluded |
|---|---|---|---|
| Minimum | 0.5344 | 0.5537 | 0.7262 |
| 10th percentile | 0.7483 | 0.7515 | 0.8111 |
| Median | 0.8949 | 0.8959 | 0.9055 |
| 90th percentile | 0.9667 | 0.9668 | 0.9685 |
| Maximum | 1 | 1 | 1 |

She noted MM-3's reliability is solid at the 10th percentile as a semi-annual measure when the CMS threshold (<11) is used, but the Steering Committee recommends using the <=11 level (6-month timeframe) to harmonize the specifications of MM-3 to MM-2, which NQF would view favorably. She added the consultants also believe the NQF Committee will be less concerned about the lower boundary statistic of 0.5537 for outliers when balanced against significantly raising the threshold (e.g., to <=25), which would exclude more than 20% of facilities from a safety measure.

MM-3 Feasibility and Validity

Dr. Nishimi reminded participants that MM-3 relies on accurate data capture of care transitions in the facility records. The measure specifies medication reconciliation must occur within eight days of the high-risk event, which are "transitions between care settings (e.g., discharge from hospital or other care setting)" and new admissions. For purposes of testing, specific direction was provided to the testing organizations to indicate the "e.g.," included hospital to facility transitions, as well as ED, observation stay, or skilled nursing facility transitions. Organizations were asked for an enumeration of each type, if feasible.

As anticipated by the testing preparation calls with the sites, Dr. Nishimi informed KCQA members that identifying all but new admissions was suboptimal. All three organizations noted the difficulties in the specifications, and one testing organization did not report on ED visits or observation visits at all, believing any capture simply was not representative. She stated missing data for all but new admissions presents a feasibility hurdle and is a significant threat to the validity of MM-3.

Dr. Nishimi remarked the magnitude of missing data is unknown, but especially troubling is a facility that under-reports or under-captures care transitions (e.g., hospitalizations) is likely to appear as a "good performer" when compared to a facility that reports all hospitalizations and any ensuing reconciliations, a few of which may miss the specified 8-day period. She noted this issue is not without precedent—KCP and others have historically expressed concern about similar under-reporting for the NHSN infection measure, which has led CMS to at least begin an audit study to assess the situation.

She indicated KCQA is not in a position to conduct such audits, but the consultants are attempting, on a more limited basis, to assess the degree to which facility records are missing hospitalizations and other care transition events captured in claims data. Specifically, she informed call participants that Chris Lovell, a member of the Steering Committee who also was directly involved in DCI's testing, noted facilities participating in the ESCO program have access to claims data that would permit these entities to reconcile events reported in claims (e.g., hospitalizations, ED visits, observation stays) that are missing in that facility's records. She stated ESCO sites have limitations on their use of the data but, with KCQA Co-Chair Allen Nissenson, they have been in contact with CMS for permission for ESCO sites to share aggregate percent of missing care transition data. .

Dr. Nishimi advised call participants the Steering Committee recommends that until the potential magnitude of missing care transition data (which NQF requests) can be adequately assessed, KCQA should defer submitting MM-3 to NQF at this time. She stated that once KCQA can quantify the percentage of missing events, the Steering Committee feels members can then judge the potential impact of missing data on validity and vote on submission. Dr. Nishimi added the precise timing for CMS's decision is unknown, but noted the Agency is moving forward expeditiously with the request. She further stated if permission is granted in a timeframe that means KCQA would miss the NQF measure submission deadline by merely a week or two, we would request an extension for MM-3 from NQF.

Dr. Molony commented the proposed plan to identify missing data is clever, but he suspects ESCOs will be high performers and may have less missing data than a random sample of facilities. Dr. Nishimi agreed this is a possibility that must be considered by KCQA members once the data are received, but at least members will have some concrete basis upon which to make a recommendation.

**FACE VALIDITY TESTING**

In addition to reliability testing, Dr. Nishimi informed call participants that NQF requires validity testing. To that end, the KCQA Steering Committee, KCQA Lead Representatives, and other experts identified by the Steering Committee will be asked for an assessment of face validity via surveymonkey. She noted the Workgroup, as developer of the specifications, will not be separately surveyed since members are assumed to have advanced specifications they felt achieved face validity. Individuals will be asked to respond to the following:

1. How likely is it that the measure score for MM-2/MM-3 provides an accurate reflection of medication reconciliation quality? (1=highly unlikely; 2=unlikely; 3=neither likely or unlikely; 4=likely; 5=highly likely)
2. What is the likelihood that MM-2/MM-3 can be used to distinguish good from poor quality? (1=highly unlikely; 2=unlikely; 3=neither likely or unlikely; 4=likely; 5=highly likely)

She noted the face validity survey will be launched shortly after KCQA members consider the proposed revisions to the specifications recommended as a result of reliability testing.

In the ensuing discussion, Dr. Meyer remarked he had now had the opportunity to review the materials in greater detail and is somewhat reassured about his previously stated concerns on face validity. He noted that "unknown" is an acceptable response for many of the required data elements, and remarked this detail makes the measures more reasonable. He added that without that allowance the measures are unrealistic because people often don't know what medications they're taking, who prescribed what medications, or what doses they're on. Ms. Glenda Payne indicated she was on the Workgroup, and the members struggled with this issue. She noted the reality is many elements will be marked as "unknown". Dr. Meyer agreed. He noted there might be some incongruencies, for example, in identifying who discontinued a medication—it might often be assumed to be the individual who first identified and listed it as discontinued in the medical records. Ms. Payne agreed, but remarked at least an individual has validated it has in fact been discontinued. She added the measures are intended to have a valid and updated medications list. Dr. Meyer agreed, noting the quality is not in checking all the boxes in each row, but rather is in making the effort to reconcile the medications list. He remarked these may not be perfect measures, but they will still drive improvement over current practices. Ms. Payne agreed, suggesting the measures could be viewed as "intermediate"

measures, until better records are available. Dr. Meyer concluded "unknown" must remain in the specifications; otherwise the measure could be gamed to perform well. Dr. Nishimi indicated this was why "unknown" was added to the specifications.

Dr. Leslie Spry asked what the gold standard would be for auditing the measures to confirm the specifications were met. Dr. Nishimi responded that four data points are auditable—name, date of the reconciliation, attestation that all medications were addressed, and attestation that allergies/intolerances were reviewed. Dr. Spry remarked these are all data entry points, and questioned what paper trail would exist to allow for auditing. Dr. Meyer responded there would not be a paper trail—that all information is obtained by talking with the patient and housed electronically. He added if a paper trail were required for proof, the measures would be overly burdensome and not feasible. Dr. Spry suggested the measures are checkbox indicators without a verifiable paper trail to allow for auditing. Dr. Nishimi indicated that requiring all of the multiple components of the measures be met to receive credit moves the measures beyond the realm of a simple checkbox and minimizes the propensity for gaming. Dr. Spry asked if there were sites that actually met all of the terms of the measures all of the time during testing. Dr. Nishimi responded yes, there were facilities that achieved 100% on the measure in a given month, as well as facilities that scored 0%.

Dr. Spry noted "eligible professionals" in the measures' specifications includes pharmacy technicians. He remarked medication reconciliation is only as good as the individual who performs it, and indicated he has noticed misspellings, mistakes, and unusable data when a pharmacy technician had performed a medication reconciliation in his organization. Dr. Wendy St. Peter responded that pharmacy technicians, when trained appropriately and thoroughly, are wholly capable of performing medication reconciliation competently. She indicated the Workgroup had recognized these concerns exist, but had attempted to develop measures that could be reasonably reproducible and tracked in existing medical records systems. Dr. Nishimi added the Workgroup had carefully considered the appropriate eligible professionals for each component of medication management (i.e., medication documentation, reconciliation, and review) and had developed the current list. Dr. Sharon Perlman noted it is incumbent on other providers involved in the process to bring poor or unacceptable work to the attention of the responsible party's superiors. Ms. Payne commented that nurses on the Workgroup had been vocal about the lack of pharmacists and pharmacy technicians in the dialysis facility, but opined if present, it would be assumed their training would be appropriate.

**PUBLIC COMMENT**

No public comments were made related to the discussion, although KECC representatives expressed appreciation for the opportunity to listen as public participants.

**NEXT STEPS**

Dr. Nishimi reviewed next steps, stating that immediately following the All-KCQA call, KCQA Lead Representatives will be asked whether they object to the following:

- Modify MM-2 and MM-3 specifications to indicate facilities <=11 (i.e., <12) should be excluded; and
- Modify MM-3 specifications to indicate data should be aggregated to a 6-month (not monthly) timeframe; and
- Defer NQF submission for MM-3 until the missing data issue can be quantified, after which KCQA members would assess the impact on validity and vote on further changes

and/or submission.

She also indicated that once the specifications are finalized, the surveymonkey face validity link would be forwarded to the Steering Committee, KCQA Lead Representatives, and other experts. Lastly, she stated the KCQA Steering Committee will review the results of all testing and make a final recommendation to KCQA members on NQF submission; KCQA Lead Representatives will have approximately 10 days to respond via surveymonkey.

Dr. Molony asked if an individual who is both a Lead Representative and also was on the Workgroup should defer the face validity survey to another individual within his/her organization. Dr. Nishimi responded that Lead Representatives will receive the email containing the survey link and each organization should respond as it wishes. She asked Lead Representatives who changed the designee for the face validity assessment to alert her or Dr. McGonigal about who will be responding.

Dr. Nishimi also informed call participants the KCQA medication management measures have been triaged by NQF to its Patient Safety Project, which has purview over medication management measures regardless of care setting, rather that the Renal Project.

Dr. Nissenson and Dr. Nishimi thanked participants for their time and input, and the conference call was adjourned.