TO: KCQA Members

FR: Robyn Nishimi, Lisa McGonigal

RE: Medication Management Reliability Testing

DA: 14 April 2016

Since late January we 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* (Attachment A). We recently received the final data pulls and reviewed the results with the Steering Committee and Medication Management Workgroup, which developed the specifications. This memorandum summarizes the reliability testing results, makes recommendations for specification changes based on those results, and outlines next steps.

EMPIRICAL RELIABILITY TESTING

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, our 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

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*

Three aspects were examined: performance gap/opportunity for improvement, small sample size effect, and overall reliability.

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

Opportunity for Improvement: MM-2

NQF requires demonstrating 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).

Assessment of Small Sample Size: MM-2

The beta-binomial method yields reliability statistics for each facility. Per Adams and our 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, reliability was:

Table 1. MM-2 Reliability Statistics, All Facilities

Minimum	0.4241
10th percentile	0.9847
Median	0.9952
90th percentile	0.9995
Maximum	1.0

To assess small sample and implementation issues, however, we 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). We analyzed both the percentage of facilities that would be excluded from measurement, as well as the reliability of the measure for small facilities:

- Using the <11 threshold excludes 3.3-3.6% of facilities from MM-2, depending on the month. Using the <=25 threshold excludes 13.4-13.8% of facilities from MM-2, depending on the month.
- When we examine the reliability statistic for facilities with <11 and <=25 events, the effect of small sample size becomes 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

We note that, at the 10th percentile, the measure does not achieve the desired reliability threshold of 0.7. At the same time, using KCP's policy of <=25 events does achieve sufficient reliability for all but outliers, but we are concerned about the impact on quality and patient safety of eliminating approximately 13.5% of facilities from measurement each month.

Toward this end, we examined the dataset to determine at what sample size we could achieve a 10th percentile reliability statistic of 0.70. **Based on this additional analysis (Table 3), the Steering Committee recommends KCP's current policy be modified for MM-2 to <=11 (i.e., <12) and that the specifications reflect this threshold; this excludes approximately 3.7-3.9% of facilities each month.** We note another option would have been to collect data on a 2-month basis, but we 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 or fewer events are excluded (~3.7%), 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, three aspects were examined during empirical testing: sample size issues, performance gap/opportunity for improvement, and overall reliability. We also identified a feasibility issue, however, which we believe raises significant validity issues, as discussed in a following section.

MM-3 Sample Issues

Even before examining details for MM-3, it became clear the currently specified monthly timeframe does 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%). We considered a quarterly timeframe, but only when we examined a 6-month time interval did we have both sufficient numbers and stabilized reliability statistics for MM-3. **The following sections report information for MM-3 wherein six months of data are aggregated to produce the performance score.**

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.

Opportunity for Improvement: MM-3

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, we did not further analyze reliability statistics for small-sized facilities. Rather, Table 5 presents the overall reliability statistics when 0, <11, <=11, and <=25 events are removed. Again, per Adams and our experience with NQF, a reliability statistic of 0.7 is

generally viewed as an acceptable threshold. For the 6-month period, we find the following:

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

We note that while MM-3's reliability is solid at the 10th percentile as a semi-annual measure when the CMS threshold is used, **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**. We 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

MM-3 relies on accurate data capture of care transitions in the facility records. The measure specifies medication reconciliation must occur within 8 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, we provided specific direction to indicate the "e.g.," included hospital to facility transitions, as well as ED, observation stay, or skilled nursing facility transitions. We also asked for an enumeration of each type, if feasible.

As anticipated by our testing preparation calls with the sites, identifying all but new admissions is 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 not representative. **Missing data for all but new admissions presents a feasibility hurdle and is a significant threat to the validity of MM-3.**

As we have proceeded through testing and further contemplated use of the measure, we note that while the magnitude of missing data is unknown, we are troubled 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.

We note this issue is not without precedent—KCP and others have historically expressed concern about similar under-reporting for the NHSN infection measure, which has lead CMS to at least begin an audit study to assess the situation. Although we obviously are not in a position to conduct such audits, we 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, Chris Lovell, a member of the Steering Committee who also was directly involved in DCI's testing, noted that facilities participating in the ESCO program have access to claims data that would permit these participants to reconcile events reported in claims (e.g., hospitalizations, ED visits, observation stays) that are missing in that facility's records. ESCO sites, however, have limitations on their use of the data. With KCQA Co-Chair Allen

Nissenson, we have been in contact with CMS for permission for ESCO sites to share aggregate % of missing care transition data to us, but have yet to receive permission that would let the testing organizations do so; CMS is currently assessing whether each site's Data Use Agreement would need to be amended.

The Steering Committee recommends that until we can adequately assess the potential magnitude of missing care transition data (which NQF requests), KCQA defer submitting MM-3 to NQF for endorsement consideration at this time. Once we are able to quantify the percentage of missing events, the Steering Committee notes KCQA members can then judge the potential impact of missing data on validity and vote on submission. The precise timing for CMS's decision is unknown, although we note the Agency is moving forward expeditiously with our request. If permission is granted in a timeframe that means KCQA would miss the NQF measure submission deadline by merely a week or two, we will request an extension for MM-3 from NQF.

FACE VALIDITY TESTING

In addition to reliability testing, NQF requires validity testing. We will conduct a surveymonkey-based face validity assessment of the specifications by three expert groups: KCQA Steering Committee, KCQA Lead Representatives, and other experts identified by the Steering Committee. The Workgroup, as developer of the specifications, will not be separately surveyed. 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)

We anticipate conducting the face validity survey shortly after KCQA members consider the proposed revisions to the specifications recommended as a result of reliability testing.

NEXT STEPS

Immediately following the All-KCQA call, KCQA Lead Representatives will be asked whether they object to the following (April 18-April 20):

- 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.

Once the specifications are finalized, the surveymonkey face validity link will be forwarded to the Steering Committee, KCQA Lead Representatives, and other experts (1 week period).

Lastly, the KCQA Steering Committee will review the results of the face validity assessment together with the reliability data and make a final recommendation to KCQA members on NQF submission. KCQA Lead Representatives will have ~10 days to respond via surveymonkey.

MM-2: Medication Reconciliation for Patients Receiving Care at Dialysis Facilities

Description	Percentage of patient-months for which medication reconciliation was performed and documented by an eligible professional.
Measure Type	Process.
Data Source	Facility medical record.
Level of Analysis	Dialysis facility.
Numerator	Number of patient-months for which medication reconciliation* was performed and documented by an eligible professional** during the reporting period. The medication reconciliation MUST: • Include the name or other unique identifier[1] of the eligible professional;[2] AND • Include the date of the reconciliation; AND • Address ALL known medications that are not administered intradialytically (prescriptions, over-the-counters, herbals, vitamin/mineral/dietary [nutritional] supplements, and medical marijuana); AND • Address for EACH medication: Medication name,[3] indication,[4][5] dosage,[5] frequency,[5] route of administration,[5] start and end date (if applicable),[5] discontinuation date (if applicable),[5] reason medication was stopped or discontinued (if applicable),[5] and identification of individual who authorized stoppage or discontinuation of medication (if applicable);[5] AND • List any allergies, intolerances, or adverse drug events experienced by the patient. *"Medication reconciliation" is defined as the process of creating the most accurate list of all medications that are not administered intradialytically that the patient is taking, including name, indication, dosage, frequency, and route, by comparing the most recent medication list in the dialysis medical record to one or more external list(s) of medications obtained from a patient or caregiver (including patient-/caregiver-provided "brown bag" information), pharmacotherapy information network (e.g., Surescripts®), hospital, or other provider. **For the purposes of medication reconciliation, "eligible professional" is defined as: physician, RN, ARNP, PA, pharmacist, or pharmacy technician.
Denominator	Total number of patient-months for all patients assigned to a dialysis facility during the reporting period.

[1] Testing indicated that unique provider identifiers (e.g., physician UPIN) are frequently used in lieu of the provider's name and can be reliably captured. Verbiage was modified to reflect this finding.

[2] ~~The preliminary feasibility assessment reveals that "name of the eligible professional" may present a feasibility issue for testing purposes. Accordingly, this data element may be removed from the specifications if the measure is advanced for testing.~~

[3] For patients in a clinical trial, it is acknowledged that it may be unknown as to whether the patient is receiving the therapeutic agent or a placebo.

[4] ~~The preliminary feasibility assessment reveals that "indication" is not captured or not captured in a manner amenable to testing and so also is a feasibility issue. Accordingly, this data element may be removed from the specifications if the measure is advanced for testing.~~

[5] "Unknown" is an acceptable response for this field.

Exclusions	1. Transient patients, defined as in-center patients who received < 7 hemodialysis treatments in the facility during the reporting month. 2. Facilities with < 12 (i.e., <= 11) eligible patients during the reporting month.
Reporting and Stratification	No risk adjustment or risk stratification.

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

Description	Percentage of high-risk patient-events aggregated for a semi-annual (6-month) measure timeframe for which medication reconciliation was performed and documented by an eligible professional.
Measure Type	Process.
Data Source	Facility medical record.
Level of Analysis	Dialysis facility.
Numerator	Number of high-risk patient-events* aggregated for a semi-annual (6-month) measure timeframe for which medication reconciliation** was performed and documented by an eligible professional*** within 8 days of a transition event (e.g., discharge from hospital) or admission for in-center patients and within 30 days for home patients ~~during the reporting period~~. The medication reconciliation MUST: • Include the name or other unique identifier[6] of the eligible professional;~~[7]~~ AND • Include the date of the reconciliation; AND • Address ALL known medications that are not administered intradialytically (prescriptions, over-the-counters, herbals, vitamin/mineral/dietary [nutritional] supplements, and medical marijuana); AND • Address for EACH medication: Medication name,[8] ~~indication[9]~~[10] dosage,[10] frequency,[10] route of administration,[10] start and end date (if applicable),[10] discontinuation date (if applicable),[10] reason medication was stopped or discontinued (if applicable),[10] and identification of individual who authorized stoppage or discontinuation of medication (if applicable);[10] AND • List any allergies, intolerances, or adverse drug events experienced by the patient. ~~If a facility has been unable to procure the discharge medications list from the discharging facility within the defined 8 days of the applicable event for in-center patients or 30 days for home patients, the facility must indicate the following to receive credit for the measure:[11]~~

[6] Testing indicated that unique provider identifiers (e.g., physician UPIN) are frequently used in lieu of the provider's name and can be reliably captured. Verbiage was modified to reflect this finding.

~~[7] The preliminary feasibility assessment reveals that "name of the eligible professional" may present a feasibility issue for testing purposes. Accordingly, this data element may be removed from the specifications if the measure is advanced for testing.~~

[8] For patients in a clinical trial, it is acknowledged that it may be unknown as to whether the patient is receiving the therapeutic agent or a placebo.

~~[9] The preliminary feasibility assessment reveals that "indication" is not captured or not captured in a manner amenable to testing and so also is a feasibility issue. Accordingly, this data element may be removed from the specifications if the measure is advanced for testing.~~

[10] "Unknown" is an acceptable response for this field.

	☐ ~~Attempted but unable to obtain discharge medications list from discharging facility within 8 days of discharge for in-center patient or 30 days for home patient.~~ • ~~Date of attempt to obtain discharge medications list: _____~~ • ~~Name of person who attempted to obtain discharge medications list: _____~~ • ~~Name of discharging facility: _____~~[12] * High-risk patient-events are defined as transitions between care settings (e.g., discharge from hospital or other care setting) and new admissions to the dialysis facility.[13] **"Medication reconciliation" is defined as the process of creating the most accurate list of all medications that are not administered intradialytically that the patient is taking, including name, indication, dosage, frequency, and route, by comparing the most recent medication list in the dialysis medical record to one or more external list(s) of medications obtained from a patient or caregiver (including patient-/caregiver-provided "brown-bag" information), pharmacotherapy information network (e.g., Surescripts®), hospital, or other provider. ***For the purposes of medication reconciliation, "eligible professional" is defined as: physician, RN, ARNP, PA, pharmacist, or pharmacy technician.
Denominator	Total number of high-risk patient-events for all patients assigned to a dialysis facility ~~during~~ aggregated for a semi-annual (6-month) measure timeframe~~the reporting period~~.
Exclusions	1. Transient patients, defined as in-center patients (who are NOT new admissions that month)[14] who received < 7 hemodialysis treatments in the facility during the reporting month. 2. In-center patients discharged from hospital or other care setting who are readmitted within 8 days of the discharge. (Exclusion applies to both in-center and home patients.) 3. Facilities with < 12 (i.e., <=11) eligible patient-events during the semi-annual (6-month) measure timeframe. ~~3. Home patients discharged from hospital or other care setting who are readmitted within 8 days of the discharge.~~
Reporting and Stratification	No risk adjustment or risk stratification.

[11] ~~The preliminary feasibility assessment reveals that data elements required for this "failed attempt" attestation may not be captured or not captured in a manner amenable to testing and so is a feasibility issue. Accordingly, these data elements may be removed from the specifications if the measure is advanced for testing.~~

[12] Testing indicated the data elements required for the "attempted but failed to obtain" component of the numerator are not consistently recorded in facilities' medical records and cannot be reliably captured. This finding required that the component of the measure be removed from the specifications.

[13] Implementation guidance regarding how "high-risk transition events" are defined for the measure was sought by testing organizations. The following clarifying language was included in the testing protocol: "A 'high-risk transition event' is defined as transitions from other care settings back to the dialysis facility (e.g., discharge from hospital or other inpatient care setting, emergency room discharge, observation stay discharge) and new admissions to the dialysis facility."

[14] Additional clarity on how "transient patients" are defined for the measure was sought by testing organizations, in particular to clearly differentiate newly admitted patients who received < 7 treatments during the reporting month (who are defined as "high-risk" and are thus a focus of the measure) from true transient patients. Clarifying language was added to the specifications to reflect this need.