WELCOME AND INTRODUCTIONS

Day 1 of the meeting began with a re-introduction of Committee members and National Quality Forum (NQF) staff. Measure developers from Minnesota Community Measurement were also present to sit in on the discussion. The co-chairs gave a recap of project objectives: evaluating submitted candidate standards and identifying gap areas where more performance measures are
necessary. They also gave an overview of the agenda, noting their plan to deal with candidate measures deemed as “process measures,” should this instance arise.

VOTING PROCESS OVERVIEW

Dr. Winkler, NQF clinical consultant and the Outcomes Project advisor, reviewed the importance of capturing the Steering Committee members’ votes. Mr. Corbridge, Outcomes Project manager, gave a synopsis of the meeting materials presented and instructions on how each piece of information would be used throughout the meeting. He explained the results of the pre-meeting work each workgroup conducted and how the results of their evaluation of the subcriteria (NQF Evaluation Criteria) would be used to convey information about the candidate measures to other Steering Committee members in different workgroups. The Committee then began its evaluation of the candidate measures.

Mr. Corbridge gave a brief synopsis of each candidate standard (description, numerator, and denominator) to the Committee. For each measure the Committee determined whether or not it was within the scope of the project. Measures within scope were further evaluated and discussed based on each of the four NQF evaluation criteria: importance to measure, scientific acceptability, usability, and feasibility. Those measures deemed out of scope were not considered for potential endorsement. The workgroup members assigned for preliminary review of the measures facilitated the discussion of their initial subcriteria evaluations and opened the floor for discussion with other Committee members. After the discussion around each criterion and its respective subcriteria, Committee members voted. The voting scale was completely, partially, minimally, not at all, or not applicable. Votes were cast by a raise of hand and tallied by NQF staff. Afterwards, an overall vote on whether or not the standard should be recommended for endorsement was cast.

From this exercise, these four candidate standards were deemed out of scope and not moved forward:

- **Measure OT3-005**: Services offered for psychosocial needs (paired with Measure OT3-021, Assessment of psychosocial needs)
- **Measure OT3-014**: Psychiatrist-rated assessment of psychiatric inpatients' clinical status
- **Measure OT3-017**: Percentage of eligible patients who transfer from a substance abuse treatment program to a continuing care physician for ongoing buprenorphine maintenance therapy
- **Measure OT3-021**: Assessment of psychosocial needs (paired with Measure OT3-005, Services offered for psychosocial needs)
Committee members’ input on future use, further development opportunities, and overall improvements of the measures can be viewed in the Candidate Standards Review section of this document under “Standards out of Scope.”

One candidate measure was not moved forward as it failed to meet the importance to measure and report criterion. Of the four NQF measure evaluation criteria, the importance to measure and report is a threshold criterion; those measures not meeting this criterion are tabled with no further review. The idea is to promote the development of new standards that will enhance NQF’s existing portfolio. The following measure was not moved forward for lack of importance to measure and report:

- **Measure OT3-002:** Patient attitudes toward and ratings of care for depression (PARC-D 30) questionnaire

Committee members’ input on future use, further development opportunities, and overall improvements of the measures can be viewed in the Candidate Standards Review section of this document under “Other.”

After completion of the measures assigned for Day 2, the Committee revisited three (003, 004, and 006) of the measures reviewed on Day 1, due to lack of clarity on the specifications of the measures. The measure developer joined the meeting via conference call to offer insight and answer any questions. The Committee recommended the developer explore risk-adjustment methodologies while providing further validity and reliability testing to enhance the measure’s credibility.

Details about the analysis for each individual measure, along with the results of the voting around recommendation for endorsement are listed below.

**Note:** Ms. Pitzen from MN Community Measurement along with Ms. Mayberry were present (in person) on Day 1 of the meeting to give supplemental background information on their candidate standards (OT3-011-10, OT3-012-10, and OT3-022-10).

**CANDIDATE STANDARDS REVIEW**

**STANDARDS RECOMMENDED FOR ENDORSEMENT**

Minnesota Community Measurement Depression Remission Measures
OT3-012-10: Depression remission at six months (Minnesota Community Measurement)
This measure was recommended to be paired with OT3-022-10: Depression utilization of the Patient Health Questionnaire (PHQ-9) tool.
This candidate standard was recommended for NQF endorsement and is to be paired with the Depression utilization of the Patient Health Questionnaire (PHQ-9) tool (OT3-022-10) submitted by Minnesota Community Measurement.

Vote: yes—17, no—0, abstention—1

OT3-011-10: Depression remission at 12 months (Minnesota Community Measurement)
This measure was recommended to be paired with OT3-022-10: Depression utilization of the Patient Health Questionnaire (PHQ-9) tool.
This standard was recommended for NQF endorsement and is to be paired with the Depression utilization of the Patient Health Questionnaire (PHQ-9) tool (OT3-022-10) submitted by Minnesota Community Measurement.

Vote: yes—17, no—0, abstention—1

OT3-022-10: Depression utilization of the Patient Health Questionnaire (PHQ-9) tool
(Minnesota Community Measurement)

Vote: yes—8, no—6, abstention—1

Note: This vote was for the measure as a standalone measure.

Vote: Linking this measure independently with measure numbers OT2-011-10 and OT3-012-10.

Note: This standard was also recommended to be paired with candidate standards OT2-011-10 and OT3-012-10.

Vote: yes —15, no —1

The Steering Committee was very pleased with the measures and believes the endorsement of this scale will set a new standard of care, helping to push the field forward. Two of the three measures: OT3-012-10, Depression remission at six months and OT3-011-10, Depression
remission at 12 months were identical in their constructs except for variations in their timeframes for assessing depression remission.

These measures assess a patient’s longitudinal change in the PHQ-9 score at six and twelve months. The PHQ-9 tool is a widely accepted and standardized instrument used in the diagnosis and monitoring of depression treatment. The Steering Committee acknowledged the value of the PHQ-9 to document a baseline and monitor symptoms and signs of major depression, and to catalyze standardized measurement of response and remission for depression care. The measure developer attested that the tool was currently being implemented on a voluntary basis throughout the state of Minnesota and is currently being considered for use in “pay-for-performance” models within the state.

The developer convened a technical advisory workgroup to assess risk adjustment of this standard with the severity of the PHQ-9 score used to evaluate risk adjustment. The Committee discussed in detail the time specifications outlined in the measures. The measure developer explained the rationale for selecting the six-month and twelve-month measurement points, indicating earlier tests assessing remission in timeframes less than six months were often uninformative, since insufficient time had elapsed to treat a patient adequately. The measure developer indicated that the average number of patients who continued treatment at six and twelve months were almost identical, at approximately 20 percent.

The Committee acknowledged that the Depression utilization of the PHQ-9 tool (OT3-022-10) measure is a process measure; however, the Steering Committee noted the measure forms the basis of the denominator for the two Minnesota Community Measurement depression remission measures (OT3-011-10, Depression remission at twelve months and OT3-012-10, Depression remission at six months). For this reason, the Committee recommended that it be endorsed as a paired measure to each of the two depression remission measures. The pairing of these measures is critical as it ensures that clinicians are administering the PHQ-9, building the denominator for the two depression remission measures.

The Committee deemed this standard important and easy to implement with a relatively low burden to facilities. The Committee was encouraged by the level of testing and current use of the measure and noted that the score captured from the PHQ-9 can be used for patient care as well as quality measurement. The Minnesota Community Measurement measures submitted to the NQF Mental Health Outcomes project were recommended for NQF endorsement as paired consensus standards.

**OT3-047-10: Inpatient Consumer Survey (ICS)**
The Committee acknowledged this measure addresses an area that is important to measure and publicly report. In a continued effort to expand the application of mental health measures across care settings the Committee suggested that the reliability and validity testing commence in
broader settings, and not solely at state hospitals. Despite the Committee’s suggestion regarding reliability and validity testing, the Committee was pleased with the level of testing already completed. The measure developer provided data about the current use of this survey, stating that the responses were captured at time of discharge. The developer indicated that variability in response rates ranged from 20 percent to 80 percent with an average of around 45 percent. The developer noted that facilities with large populations of patients with low health literacy would be more likely to have lower response rates, thus contributing to the variability. The Committee was very pleased that the measure was developed via consumer workgroups and that there is an existing infrastructure to support the measure. Overall, the Committee believes the measure is very strong and recommended it for NQF endorsement.

Vote: yes—15, no—0

CANDIDATE CONSENSUS STANDARDS NOT RECOMMENDED FOR ENDORSEMENT

OT3-001-10: Suicide deaths of “at risk” adult psychiatric inpatients within 30 days of discharge
The Committee believed that the measure addressed an important area, but had reservations pertaining to perceived measure limitations, specifically feasibility and usability. The Committee’s primary concern was the measure’s specifications for capturing suicide deaths at 30 days following discharge. As stipulated in the submission the measure relies on the collection of patient status through follow-up phone calls. The Committee noted this approach faced significant challenges in actually being able to follow up with patients and placed a large burden on facilities. Sections of the testing portion of submission were not very detailed, thereby making it difficult for the reviewers to give a full assessment. The developer indicated they were currently testing the measure, but had no plan to account for risk adjustment. The Steering Committee strongly suggested that risk adjustment was essential for this measure as there are many exogenous factors that can affect the outcome of an individual’s suicidal ideations or completion. Despite the challenges to the measure, the Committee acknowledged the measure addressed a critical gap area, with a population level focus. The Committee believes this measure needs additional refinement, including testing in additional settings and inclusion of risk adjustment. The Committee encourages the developer to resubmit the measure to NQF at a later date once they have addressed the Committee’s concerns. This measure was not recommended for NQF endorsement.

Vote: yes—0, no—18
Western Psychiatric Institute and Clinic of UPMC Presbyterian Shadyside submitted three measures to the NQF Mental Health Outcomes project pertaining to psychiatric readmission. The measures, 30 day readmissions (OT3-003-10), 7 day readmissions (OT3-004-10), and 48 hour readmissions (OT3-006-10), were identical in their constructs except for variations in the timeframes used for measuring readmissions. The Committee quickly came to consensus that these candidate standards were very similar in their constructs to other NQF hospital readmission measures currently in use (NQF endorsed an All-cause readmission index (risk-adjusted) [#0329] from the United Health Group). The Committee stressed the need to incorporate mental health into the broader care settings and believed these measures may ultimately have the unintended consequence of isolating mental health readmissions from other sectors of care. Furthermore, the Committee noted these standards only allow for measurement of whether or not the patient is being re-admitted to the initial psychiatric facility where the original hospitalization occurred, which the Committee regarded as a shortcoming. For these reasons, the Committee recommended that current NQF measures should consider expanding the types of readmissions to include mental health and substance use (MHSU) conditions at the time of maintenance review. Measures that delineate specific care settings inevitably create a conceptual barrier, limiting measurement and broad adoption. The Steering Committee believes the focus on strictly mental health settings runs counter to the value of integrating MHSU care into broader medical care settings, an important Committee goal.

Further deliberations on all three measures focused on the Committee’s concerns with the measure’s lack of testing and risk adjustment, in addition to the overall scientific acceptability of the measures. Throughout the two day meeting the Committee continually stressed the need for risk-adjustment modeling when dealing with outcome measures. The Committee believes these measures in particular needed to have a strong risk-adjustment model and supporting testing. The measures deal with long time intervals which might increase the likelihood of readmission rates
as a result of exogenous factors regardless of the quality of care provided during a patient’s hospital stay.

The Committee believes that the measures are potentially of great value but require additional refinement before they should be considered for public reporting. The readmission standards submitted by Western Psychiatric Institute and Clinic of UPMC Presbyterian Shadyside were not recommended for NQF endorsement.

**OT3-010-10: Milestones of Recovery Scale (MORS)**

The Committee noted the merit of this standard is its approach to examining the recovery process from the patient perspective, a point of view often overlooked in the mental health arena. The Steering Committee was pleased that the measure is currently in use in existing programs. Despite the measure’s importance and strong attributes, the Committee had substantial concerns regarding the measure’s scientific acceptability and usability. The predominate concerns of the Committee focused on the measure’s lack of testing for validity and reliability, lack of risk adjustment, and lack of attention to health disparities. The Committee was also concerned by the measure’s link between improvement and important patient-oriented outcomes and being able to assign accountability. Overall, the Committee identified positive measure attributes not seen in previously discussed measures and was enthusiastic about the potential concept of the measure. The Committee encouraged the developer to address the Committee’s suggestions and submit a revised measure to NQF at a later date. This standard was not recommended for NQF endorsement.

*Vote: yes—0, no—18*

**OT3-013-10: Time from first face-to-face treatment encounter to buprenorphine dosing**

The Committee noted this measure targeted a very substantial problem and was an important measurement area. While the Committee acknowledged this measure’s attempt to improve treatment times for patients with a substance abuse problem, they had concerns with the lack of testing of the measure and the link between this measure and patient outcomes. The Committee acknowledged the potential obvious gains from moving toward shorter time intervals between the first face-to-face encounter and the time when the first dose of buprenorphine is received, but noted the relationship to patient outcomes has not been demonstrated. The Committee asked the developer to explain the data which supports the measure’s time-interval in relation to patient outcome. The measure developer indicated the proposed time-interval for the measure comes from clinical practice guidelines. The developer informed the Committee that this standard was intended to be an intermediate outcome. The developer explained that the measure addressed an
intermediate outcome, but with no formal reliability or validity testing the Committee questioned the measure’s use in public reporting at this time. The Committee was supportive of the concept and encouraged the developer to make improvements for future submission. This standard was not recommended for NQF endorsement.

*Vote: yes—0, no—15*

**OT3-016-10: Retention in treatment**

The Committee found this measure to be in scope and acknowledged the value of assessing treatment retention, but noted the connection between patient outcomes and treatment retention had not been demonstrated. While the measure did not demonstrate the causal link, the Committee entertained the question of whether or not treatment retention alone could be used as an outcome indicator. The Committee ultimately decided that treatment retention was not a reliable assessment of patient outcomes noting that a patient can be seen multiple times (treatment retention), but if the quality of care provided is sub-optimal then patient outcomes may not improve. The measure as submitted lacked any formal testing or risk-adjustment models; however, the developer did indicate they are in the process of conducting risk-adjustment testing. Because testing, including the need to assess for risk adjustment, has not been completed, the Committee could not support moving the measure forward for endorsement at this time. The Committee is supportive of the concept and encourages the developer to make improvements for future submission. This standard was not recommended for NQF endorsement.

*Vote: yes—0, no—16*

**OT3-008-10: Fall rate per 1,000 patient days**

The Committee agreed that this candidate standard is focused in an area where performance measurement is lacking because there is no existing national database to assess fall rates among psychiatric patients. The Committee noted this standard is similar to two existing NQF measures (NQF #0141: Patient fall rates and NQF #0202: Falls with injury); however, these currently endorsed standards to not include the MHSU arena. It was the consensus of the Committee to expand the currently NQF-endorsed patient fall measures (NQF #0141 and NQF #0202) to include psychiatric settings. In expanding the currently endorsed measures, the Committee suggested the measure developer consider stratification by relevant variables such as the presence of substance abuse or medical co-morbidity. A representative from the American Nurses Association who maintains the currently endorsed measures was present at the meeting and indicated a willingness to expand the measure to include inpatient mental health settings. Because it is expected that the endorsed measure’s characteristics will be expanded, this standard was not recommended for NQF endorsement.
Vote: yes—0, no—16

Note: Given the stratification contingency the Committee proposed a motion to request that this measure be maintained by the developer with an existing related measure, and have their definition expanded (to include dementia, co-occurring diagnoses) and stratified accordingly.

Vote: yes—16, no—0

OT3-009-10: Adverse/serious event
The Committee noted that the measure was in scope and addressed an important topic area that has not been addressed from a measurement perspective in the mental health arena. While the Committee agrees that the measure targets an important area, concerns were raised about the subjectivity of the measure. At present, the measure numerator statement reads “number of adverse/serious events that patients admitted to a hospital-based inpatient psychiatric setting experienced.” This standard cites literature from the Institute of Medicine (IOM), which defines serious as injuries caused by medical management, but the standards do not specifically delineate what injuries caused by medical management are classified as serious. The Committee believes the subjectivity of the measure’s numerator may lead to potential variability of the measure between reporting entities. This in turn could make the measure difficult to understand or make quality difficult to discern. The Committee noted further concerns with the measure’s lack of adequate testing. Because of inadequate testing and a lack of standardized specifications that could cut across care settings, the Committee ultimately could not support moving the measure forward for endorsement at this time. This standard was not recommended for NQF endorsement.

Vote: yes—1, no—15

Note: The Committee made a motion that this standard be examined in another project (potentially the NQF Patient Safety Project) to attain an assessment of adverse events in an attempt to develop a national standard related to patient safety, which would span care settings.

Votes to recommend the measure for the NQF Patient Safety Project: yes—16, no—0

STANDARDS OUT OF SCOPE
The scope of the NQF Outcomes Project: Mental Health was to enlarge NQF’s portfolio of outcome measures for mental health conditions, such as depression, psychosis, and other serious mental illnesses, substance use disorders, and Alzheimer’s disease, and related illnesses. In the “Call for Measures” the Steering Committee established a broad framework for the Mental Health Outcomes Project, encompassing patient, caregiver, and population outcomes related to:
symptoms, function, health-related quality of life, social determinants of health, experience of care, etc. All measures were first evaluated to determine whether they addressed the scope of the project and were deemed either “in or out of scope.” All process measures were indicated as out of scope. Below is the list of measure identified as out of scope during the Committee’s review of the candidate standards submitted to the project:

**OT3-005: Services offered for psychosocial needs**
The Steering Committee deliberated on whether or not this standard as stipulated actually assesses a patient outcome. The Committee acknowledges housing is one of the larger deterrents of health and well-being and is an important measurement area, but it remains unclear how the measure is tied to patient outcomes. The Committee supports the measure’s focus and expressed the need to develop a measure that links the act of providing housing to better health outcomes. The Committee noted this is a major gap area and one with large political and socioeconomic factors.

Members of the Committee noted mental health systems or states control a large portion of services for housing assistance; specifically, the VA has a very high level of access to housing services. The Committee recommended the current structure and system should be re-examined, removing potential barriers to housing in order to more effectively link patients with housing while assessing patient outcomes. Despite the great need for stable housing as one of the larger deterrents of health, this standard was not moved forward because it was deemed out of the scope of the project.

*Vote: in scope—1 (intermediate outcome), out of scope—13, abstention—1*

**OT3-014: Psychiatrist-rated assessment of psychiatric inpatients' clinical status**
The Committee agreed this tool was derived via expert opinion and is not discipline-specific. The information offered in the submission form focused more on development/use of the tool than results of outcomes captured from utilizing the tool. This standard was not moved forward because it was deemed out of the scope of the project.

*Vote: in scope—0, out of scope—13*

**OT3-021: Assessment of psychosocial needs**
The Steering Committee affirmed this candidate standard addresses a very important area of care in the mental health arena; however, based on the information submitted, the Committee agreed that a more complete picture was needed to evaluate the measure. As submitted, the standard does not offer substantial enough evidence to support the measurement of a patient outcome. This standard was not moved forward because it was deemed out of the scope of the project.
Vote: in scope—0, out of scope—15

OT3-017: Percentage of eligible patients who transfer from a substance abuse treatment program to a continuing care physician for ongoing buprenorphine maintenance therapy.

The Committee agreed this candidate standard was more like a system or utilization outcome measure than a performance measure that evaluates a patient or caregiver outcome. Committee members noted there is no evidence supporting the measurement of the standard to outcomes. This standard was not moved forward because it was deemed out of the scope of the project.

Vote: in scope—11, out of scope—5

OTHER

OT3-002: Patient attitudes toward and ratings of care for depression (PARC-D 30) questionnaire

The Steering Committee acknowledges patient’s or caregiver’s attitudes toward care are key outcomes necessary to assessing value and quality within the healthcare system. As currently stipulated the Committee identified this measure as a resource tool, used to evaluate the process of assessing patient values as opposed to an actual performance measure to assess outcomes. This measurement tool attempts to report on patient attitudes but lacks the bridge necessary to link it to care received. The Committee believes the tool fails to meet the NQF importance to measure and report criterion and is more applicable to patient experience. This standard was not moved forward due to lack of importance to measure.

Vote: in scope—0, out of scope—18

Additional Recommendations

1. Develop a broad definition for mental health outcomes
   The Steering Committee supports the development of a concise definition for MHSU outcomes to be used as a standard in the field. Such a definition would enable more effective measurement of patient outcomes across care settings.

2. When appropriate, apply measures across care settings rather than developing MHSU specific measures
The Steering Committee strongly recommends measure developers consider the broadest application of measures, assuring applicability across care settings (i.e., a measure of patient fall rates should be applicable in both a mental health and other care settings). The Steering Committee recommended NQF examine their portfolio of existing outcome measures and consider stratification for the MHSU populations, thereby allowing these measures to be applied to persons with various MHSU conditions across care settings.

3. **Provide immediate support for efforts to develop Alzheimer’s and dementia outcome measures**

The Steering Committee strongly affirms the need for measure developers and the MHSU arena to develop Alzheimer’s and dementia outcome measures. With Alzheimer’s as one of the top 20 Medicare condition priorities the Steering Committee was troubled by the lack of Alzheimer’s or dementia outcome measures submitted to the project. The Steering Committee has identified potential Alzheimer’s outcome measures and encourages their submission to future NQF projects.

4. **Align measures with the National Priorities Partnership**

The National Priorities Partnership established a clear set of principles for improving the health and well-being of all Americans. The Steering Committee affirmed the need for the mental health community to align their work in the performance measurement arena with the initiatives currently underway within NQF in association with the National Priorities Partnership.

5. **Establish important measurement focus areas in the MHSU arena**

The Steering Committee identified five key measurement focus areas needed to help improve the quality and value of care in the mental health arena. Further, the Committee indicated the need to use not only individual, but population-based measures in the measurement of behavioral health outcomes.

- initiatives geared towards the inclusion of MHSU care into the broader healthcare setting;
- Alzheimer’s and dementia outcome measures;
- the relationship of environment (e.g., housing) to mental health disorders;
- evidence-based measures which address larger social determinates of health (e.g., employment or incarceration status); and
- overuse/underuse of mental health and supporting services.
The Committee suggested NQF examine their portfolio of existing outcome measures and identify those measures which could be stratified for mental health populations, thereby allowing these measures to be applied to persons with various mental health conditions. This exercise may be influential in making connections to those topic areas, populations of people, conditions, etc. identified as gap areas.

**NEXT STEPS**

Dr. Winkler told the Committee that NQF staff would write a meeting summary reflective of the voting. She also informed them of the available options for capturing any aspects of the meeting (transcripts and mp3 recordings). The Committee recapped the number of measures actually recommended for endorsement (four). She informed the Committee about the posting of a summary of their work over the past two days, and of the commenting process for the meeting summary. She also informed them of how their recommendations on the candidate standards would be considered by the Consensus Standards Approval Committee in their assessment of whether or not the standard would be endorsed.

**Future Work**

Mr. Corbridge facilitated a discussion around the current status of Alzheimer’s and dementia measures, posing the idea of convening a small workgroup outside of the standing Committee to evaluate potential Alzheimer’s measures to be submitted to NQF. Committee members volunteered to participate in this workgroup and communicate their efforts to NQF in attempts to compliment the efforts of the work being done through this project, based on the candidate standards currently in play. Steering Committee members gave input about where to solicit such measures, and how NQF can cast a broader networking net in attempts to develop relationships with developers who are oblivious to other NQF activities. It was suggested that NQF provide literature to be disseminated among these populations of measure developers to inform them of the process and upcoming opportunities to submit candidate standards.