NATIONAL QUALITY FORUM
Driving measurable health improvements together

Meeting Summary

Behavioral Health and Substance Use Standing Committee – Measure Evaluation Web Meetings

The National Quality Forum (NQF) convened the Behavioral Health and Substance Use (BHSU) Standing Committee for three web meetings on February 8, 11, and 22, 2021 to evaluate four measures. The meetings were led by NQF Senior Director Samuel Stolpe, NQF Manager Hannah Bui, NQF Manager Tamara Funk, and NQF Analyst Udobi Onyeuku.

Welcome, Introductions, and Review of Meeting Objectives

NQF welcomed the Standing Committee and participants to the web meeting. NQF staff reviewed the meeting objectives. The Standing Committee members each introduced themselves and disclosed any conflicts of interest. Standing Committee member and Co-Chair Harold Pincus disclosed a conflict with NQF #0576 *Follow-Up After Hospitalization for Mental Illness* due to his involvement on the developer's Behavioral Health Measurement Advisory Panel. He was recused from both the discussion and voting for NQF #0576.

Some Standing Committee members were unable to attend the entirety of all three meetings due to early departures and late arrivals. The vote totals reflect members present and eligible to vote. Quorum (the minimum number of Standing Committee members who must be present at the meeting for valid voting) for the BHSU Standing Committee this cycle was 16 out of 23 members (66 percent). Quorum was met and maintained for the entirety of all the meetings.

Topic Area Introduction and Overview of Evaluation Process

NQF staff provided an overview of the topic area and the current NQF portfolio of endorsed measures. There are currently 46 measures in the BHSU portfolio. Additionally, NQF reviewed the Consensus Development Process (CDP) and the measure evaluation criteria.

Measure Evaluation

During the meeting, the BHSU Standing Committee evaluated four measures, including two maintenance and two new measures for endorsement consideration. The Standing Committee's deliberations will be compiled and provided in the draft technical report. NQF will post the draft technical report on March 22, 2021, for public comment on the NQF website. The draft technical report will be posted for 30 calendar days.

A measure is recommended for endorsement by the Standing Committee when the vote margin on all must-pass criteria (Importance, Scientific Acceptability, Use), and overall, is greater than 60 percent of voting members in favor of endorsement. A measure is not recommended for endorsement when the vote margin on any must-pass criterion or overall is less than 40 percent of voting members in favor of endorsement. The Standing Committee has not reached consensus if the vote margin on any must-pass criterion or overall is between 40 and 60 percent, inclusive, in favor of endorsement. When the Standing Committee has not reached consensus, all measures for which consensus was not reached will be

released for NQF member and public comment. The Standing Committee will consider the comments and re-vote on those measures during a webinar convened after the commenting period closes.

Rating Scale: H – High; M – Medium; L – Low; I – Insufficient; NA – Not Applicable

#0576 Follow-Up After Hospitalization for Mental Illness (National Committee for Quality Assurance)

Description: The percentage of discharges for members 6 years of age and older who were hospitalized for treatment of selected mental illness or intentional self-harm diagnoses and who had a follow-up visit with a mental health provider. Two rates are reported:

1. The percentage of discharges for which the member received follow-up within 30 days after discharge.
2. The percentage of discharges for which the member received follow-up within 7 days after discharge.;

Measure Type: Process; **Level of Analysis**: Health Plan; **Setting of Care**: Inpatient/Hospital, Outpatient Services; **Data Source**: Claims

Measure Steward/Developer Representatives at the Meeting

Brittany Wade
Sarah Paliani
Jungqiu Liu
Bob Rehm

Standing Committee Votes

- Evidence: H-1; M-18; L-0; I-1 (Pass – 19/20, 95 percent)
- Performance Gap: H-6; M-13; L-1; I-0 (Pass – 19/20, 95 percent)
- Reliability: H-6; M-13; L-1; I-0 (Pass – 19/20, 95 percent)
- Validity: H-2; M-15; L-3; I-0 (Pass – 17/20, 85 percent)
- Feasibility: H-14; M-5; L-1; I-0 (Pass – 19/20, 95 percent)
- Use: Pass-19; No Pass-0 (Pass – 19/19, 100 percent)
- Usability: H-3; M-14; L-3; I-0 (Pass – 17/20, 85 percent)

Standing Committee Recommendation for Endorsement: Yes-18; No-2 (Pass – 18/20, 90 percent)

The Standing Committee recommended the measure for continued endorsement.

NQF staff provided a brief description of the measure, as presented above. Sarah Paliani from the National Committee for Quality Assurance (NCQA) represented the measure developer and provided an overview of the submission, including the specifications and highlights from the evidence and testing. The Standing Committee noted there was updated evidence from the developer, specifically a new article that supported evidence previously submitted. The Standing Committee asked the developer for justification for the seven-day and 30-day rates, to which the developer responded that the National Institute for Health and Care Excellence (NICE) guidelines cited in their submission supported this recommendation. The developer emphasized that deaths by suicide were lower for those who received

a seven-day follow up and that readmissions were reduced for patients with mental health conditions when a follow-up occurred within 30 days. During the review of the performance gap analysis, the Standing Committee requested that the measure developer perform stratification by race and ethnicity. The NCQA responded by explaining that adding this type of stratification to their measures is a current initiative of the NCQA. They hope to add this stratification to most of their measures over the next couple of years. The Standing Committee also asked the developer about improvements in performance over time. The developer suggested that the measure has indeed shown improvements since it was implemented. The Standing Committee reviewed the reliability portion of the submission: They noted that the developer used an appropriate method for conducting the testing and that their results fell within acceptable ranges. When discussing validity, the Standing Committee asked whether telehealth visits count for the numerator, which the developer confirmed. The Standing Committee also expressed concerns related to the exclusion of primary care, noting that only mental health specialist visits are counted for the measure and that rural health communities may find it more challenging to participate. The developer noted that collaborative care coding is being considered for measurement year 2021. The Standing Committee agreed that the feasibility of the measure meets an appropriate standard. In the discussion of usability and use, the Standing Committee noted that the measure exhibited some performance fluctuations over time. The developer stated that their analysis of performance over time suggests that the fluctuations were appropriate and even expected, given changes in the measure specifications. The Standing Committee stated that the measure does not appear to exhibit significant change in performance over time, and the submission still meets the usability criterion requirements.

#3205 Medication Continuation Following Inpatient Psychiatric Discharge (Centers for Medicare & Medicaid Services)

Description: This measure assesses whether patients discharged from an inpatient psychiatric facility (IPF) with major depressive disorder (MDD), schizophrenia, or bipolar disorder filled a prescription for evidence-based medication within 2 days prior to discharge and 30 days post-discharge. This measure evaluates admissions over a two-year period.; **Measure Type**: Process; **Level of Analysis**: Facility; **Setting of Care**: Inpatient/Hospital; **Data Source**: Claims

Measure Steward/Developer Representatives at the Meeting

Jason Smoot

Standing Committee Votes

- Evidence: H-1; M-16; L-1; I-0 (Pass – 17/18, 94.4 percent)
- Performance Gap: H-5; M-13; L-0; I-0 (Pass – 18/18, 100 percent)
- Reliability: H-0; M-19; L-0; I-0 (Pass – 19/19, 100 percent)
- Validity: H-1; M-11; L-4; I-3 (Pass – 12/19, 63.2 percent)
- Feasibility: H-2; M-13; L-4; I-0 (Pass – 15/19, 78.9 percent)
- Use: Pass-20; No Pass-0 (Pass - 20/20, 100 percent)
- Usability: H-1; M-15; L-2; I-0 (Pass – 16/18, 88.9 percent)

Standing Committee Recommendation for Endorsement: Yes-11; No-8 (CNR – 11/19, 57.9 percent)

The Standing Committee did not reach consensus on the vote for overall suitability for continued endorsement.

NQF staff provided a brief description of the measure, as presented above. Jason Smoot from Mathematica provided an overview of the measure and submission, outlining the measure specifications and reviewing the testing approach for scientific acceptability. The Standing Committee asked the developer for clarification on whether the medications matched the diagnoses in the measure coding, as well as if long-acting medications were accounted for. The developer responded that medications are indeed paired with diagnoses and that the long-acting medications are included appropriately as well. In consideration of the evidence submission, the Standing Committee suggested that there were no significant threats to the evidence. The Standing Committee questioned how discharges with no medications or against medical advice are handled. The developer noted that patients who are not discharged to home or home healthcare are not included in the denominator and that patients who are discharged without a prescription are still included. The Standing Committee noted that the severity of illness would typically dictate that all patients in this measure should be on medication, but the developer noted that it is difficult to determine how many patients are contraindicated from receipt of medication. Pregnant patients were noted to be excluded from the measure, which the Standing Committee noted to be a concern since postpartum depression and pregnancy with severe mental illness represent an at-risk population. The developer was encouraged to explore the possibility of including this population within measurement. In the discussion of performance gap and disparities, the Standing Committee noted a wide performance gap as well as good data submitted by the developer on disparities. The Standing Committee noted that the developer used a common approach to determining score level reliability for the measure, namely the beta-binomial signal-to-noise analysis, and that the results fell within an acceptable range. In the validity discussion, the Standing Committee noted the known-group method used for score level validity as well as abstraction methods for data element level reliability. The Standing Committee noted that the measure involves both hospitals and outpatient providers, with the accountability falling on the facility that may not have full control over this aspect of care. This was expressed to be especially challenging for mental health patients. The Standing Committee noted that the healthcare system is increasingly moving toward shared accountability for issues such as this.

The developer was encouraged to ensure that long-acting injectables administered within the facility are included in the measure since this allows for more control on the part of the facility. This was concerning from a data sourcing perspective, given that long-acting injectables would not be captured in Medicare Part D data. The Standing Committee expressed concern regarding whether individual drugs given before discharge could be easily retrieved electronically or whether they would be bundled into the diagnosis-related group (DRG), meaning a manual chart review would need to occur. The Standing Committee also expressed concern that the measure is not keeping pace with practice, both pharmacologically and therapeutically, as long-acting injectables are increasingly used within inpatient settings to allow for an extended period of therapeutic coverage post-discharge. The developer was encouraged to analyze their own measure to ensure that this element of care provision is appropriately captured within the measure to ensure that it is valid.

In the feasibility discussion, the Standing Committee noted that this measure draws on claims data but expressed further concerns related to the limitations identified in the validity discussion. The Standing Committee expressed no concerns related to use. In the discussion on usability, the Standing Committee noted that the measure has yet to be included inside of an accountability program, and thus, it has not received feedback from end users in the market or a clear line on harms that may outweigh benefits. The Standing Committee did not reach consensus on the vote for overall suitability for continued endorsement. The measure's suitability will be discussed by the Standing Committee during the post-comment web meeting with any public comments taken into account. The Standing Committee will also re-vote for the final time on whether to recommend this measure for endorsement. In order to receive

the Standing Committee's recommendation for endorsement, the measure must receive greater than 60 percent of votes to approve overall suitability for endorsement during the post-comment meeting vote.

#3589 Prescription or Administration of Pharmacotherapy to Treat Opioid Use Disorder (RTI International)

Description: This measure reports the percentage of a provider's patients who were Medicaid beneficiaries ages 18 to 64 with an OUD diagnosis who filled a prescription for, or were administered or ordered, a FDA-approved medication to treat OUD within 30 days of the first attributable OUD treatment encounter with that provider.; **Measure Type**: Process; **Level of Analysis**: Facility, Clinician : Individual; **Setting of Care**: Emergency Department and Services, Inpatient/Hospital, Outpatient Services; **Data Source**: Claims, Enrollment Data

Measure Steward/Developer Representatives at the Meeting

Tami Mark
Julie Seibert

Standing Committee Votes

- Evidence: H-15; M-3; L-1; I-0 (Pass – 18/19, 94.7 percent)
- Performance Gap: H-14; M-5; L-0; I-0 (Pass – 19/19, 100 percent)
- Reliability: H-2; M-15; L-1; I-0 (Pass – 17/18, 94.4 percent)
- Validity: H-1; M-16; L-0; I-1 (Pass – 17/18, 94.4 percent)
- Feasibility: H-1; M-17; L-0; I-0 (Pass – 18/18, 100 percent)
- Use: Pass-16; No Pass-1 (Pass – 16/17, 94.1 percent)
- Usability: H-4; M-12; L-1; I-0 (Pass – 16/17, 94.1 percent)

Standing Committee Recommendation for Endorsement: Yes-17; No-0 (Pass – 17/17, 100 percent)

The Standing Committee recommended the measure for initial endorsement.

NQF staff provided a brief description of the measure, as presented above. Tami Mark from RTI International introduced the measure, noting portions of the measure specifications, the measure type, level of analysis, and the rationale for its use. The Standing Committee noted American Society of Addiction Medicine (ASAM) guidance and other evidence to focus on provider-level practice standards and approaches to ensure appropriate opioid use disorder treatment and noted a general need for this sort of measure at this level of analysis. The developer was encouraged to include other populations beyond Medicaid beneficiaries captured in fee-for-service and encounter data. The Standing Committee noted challenges associated with patients across the country having access to providers with buprenorphine waivers, noting that accountability for a systematic problem is unlikely to be solved by a provider-level measure. The Standing Committee noted the thorough review of the evidence provided by the developers and voiced no concerns with the evidence. For performance gap, the Standing Committee noted the large number of facilities and clinicians included. In the discussion of reliability, the Standing Committee noted that the developer included appropriate analyses with appropriate performance levels. The Standing Committee discussed the nuances associated with holding clinicians accountable for both appropriate prescribing and ensuring that patients engage around their medications and pick them up. The developer informed the Standing Committee that the word *ordered* was inappropriately included in the measure description and that it will be removed from the

description. The Standing Committee reviewed the measure testing submitted by the developer for validity and noted that the convergent validity testing for the measure was appropriate. The Standing Committee discussed that the measure draws from claims data and agreed that it meets NQF feasibility requirements. In the discussion on use, the Standing Committee stated that this is currently included in the Shatterproof Atlas but only as a provider-facing metric. The Standing Committee discussed the potential impacts of having this measure serve as a patient-facing metric as well. The Standing Committee expressed no concerns related to usability.

#3590 Continuity of Care After Receiving Hospital or Residential Substance Use Disorder (SUD) Treatment (RTI International)

Description: Percentage of Medicaid discharges, ages 18 to 64, being treated for a substance use disorder (SUD) from an inpatient or residential provider that received SUD follow-up treatment within 7 or 30 days after discharge. SUD follow-up treatment includes outpatient, intensive outpatient, or partial hospitalization visits; telehealth encounters; SUD medication fills or administrations; or residential treatment (after an inpatient discharge). Two rates are reported: continuity within 7 and 30 days after discharge.; **Measure Type**: Process; **Level of Analysis**: Facility; **Setting of Care**: Inpatient/Hospital; **Data Source**: Claims, Enrollment Data

Measure Steward/Developer Representatives at the Meeting

Tami Mark
Julie Seibert

Standing Committee Votes

- Evidence: M-17; L-0; I-0 (Pass – 17/17, 100 percent)
- Performance Gap: H-13; M-4; L-0; I-0 (Pass – 17/17 100 percent)
- Reliability: H-1; M-16; L-0; I-0 (Pass – 17/17, 100 percent)
- Validity: H-0; M-17; L-0; I-0 (Pass – 17/17, 100 percent)
- Feasibility: H-5; M-10; L-1; I-0 (Pass – 15/16, 93.8 percent)
- Use: Pass-16; No Pass-0 (Pass – 16/16, 100 percent)
- Usability: H-3; M-13; L-1; I-0 (Pass – 16/17, 94.1 percent)

Standing Committee Recommendation for Endorsement: Yes-17; No-0 (Pass – 17/17, 100 percent)

The Standing Committee recommended the measure for initial endorsement.

NQF staff provided a brief description of the measure, as presented above. Tami Mark from RTI International introduced the measure, reviewing the measure specifications and the testing approach for the submission. In the discussion on evidence, the Standing Committee noted that the developer did not include a systematic review but did provide a good number of articles that suggested follow-up as an effective means to improving outcomes for psychiatric discharges that were directionally positive and of moderate quality. The Standing Committee expressed that they were not aware of any articles that suggest that follow-up is not an effective intervention. The Standing Committee discussed the time frame for the follow-up to occur, the evidence to support it, and the inconsistency across measures. The developer conceded that there are inconsistences in time frame, and this is a potential area for future work; however, the developer also suggested that earlier follow-up interventions may be better since the literature suggests that the period closest to discharge carries the most risk for poor outcomes for

patients. The Standing Committee discussed outcomes associated with follow-up, including reductions in death, criminal justice involvement, substance use, and treatment justice. In the discussion of performance gap, the Standing Committee reviewed the developer's submission, noting significant variation between providers on follow-up. It was observed that males were less likely to receive follow-up care than females, and Black patients received fewer follow-ups than White patients. The Standing Committee expressed no concerns related to performance gaps. In the discussion on reliability, the Standing Committee reviewed the measure specifications and the reliability testing approach provided in the developer's submission, noting that the developer used appropriate analytical approaches with results with moderate effect sizes. The Standing Committee emphasized the importance of inclusion of telemedicine codes within the measure, especially in light of the COVID-19 pandemic. In the discussion on validity, the Standing Committee reviewed both the developer's submission and the staff analysis and concurred that the measure has moderate validity. They further noted that it is important to have follow-up measures that match the acuity of the condition with the proposed follow-up. In the discussion on validity, the Standing Committee reviewed the analytic approach taken by the developer and found the testing to be appropriately conducted with moderate results. The Standing Committee expressed that the data elements for the measure are generated as part of the routine delivery of care. The Standing Committee specifically discussed this measure within the context of the Institutions for Mental Diseases (IMD) exclusion, which prohibits the use of federal Medicaid funds to treat enrollees ages 21–64 in psychiatric residential treatment facilities that have more than 16 beds. In 2015, the Centers for Medicare & Medicaid Services (CMS) created a streamlined application pathway for state waivers of this rule to allow Medicaid coverage for SUD treatment in residential facilities, but the Standing Committee noted that not every state has a Medicaid waiver such as the one described, and there may be some limitations associated with the IMD exclusion. The Standing Committee discussed that the measure is currently being used for external benchmarking as well as internal benchmarking for some institutions. The Standing Committee also expressed that the measure is helpful for continuous improvement and that the developer demonstrated the requirements for initial NQF endorsement related to use. The Standing Committee did not note any concerns related to usability.

Related and Competing Measures

In the discussion of related and competing measures, the Standing Committee noted that many of the measures reviewed in this cycle have overlapping features with other measures and called upon the measure developers to harmonize their measures to the extent possible.

For NQF #0576, the Standing Committee noted two related measures: NQF #2605 *Follow-Up After Emergency Department Visit for Mental Illness or Alcohol and Other Drug Abuse or Dependence* and NQF #3489 *Follow-Up After Emergency Department Visit for Mental Illness*, with all three measures developed by the NCQA. The Standing Committee acknowledged that NQF #0576 differs from the other two in that it is related to inpatient stays rather than emergency department visits but called upon the NCQA to justify the continuance of NQF #2605 and NQF #3489, given the apparent overlap between the two measures.

In the discussion of NQF #3205, the Standing Committee regarded NQF #1879 *Adherence to Antipsychotic Medications for Individuals With Schizophrenia* as related but not competing. The measure developer (Mathematica) noted that these two measures are harmonized in the medication coding.

NQF #3589 was compared to both NQF #3175 *Continuity of Pharmacotherapy for Opioid Use Disorder* and NQF #3400 *Use of Pharmacotherapy for Opioid Use Disorder*, with the same assessment that they are related but not competing. The measure developer (RTI International) noted that these measures are harmonized to the extent possible.

NQF #3590 had multiple measures that were considered in the related and competing measures discussion. These included four NCQA measures that were considered related but not competing and appropriately harmonized: NQF #0004 *Initiation and Engagement of Alcohol and Other Drug Abuse or Dependence Treatment*, NQF #0576 *Follow-Up After Hospitalization for Mental Illness*, NQF #1937 *Follow-Up After Hospitalization for Schizophrenia*, and NQF #2605 *Follow-Up After Emergency Department Visit for Mental Illness or Alcohol and Other Drug Abuse or Dependence*. Mathematica's measure titled NQF #3312 *Continuity of Care After Medically Managed Withdrawal From Alcohol and/or Drugs* was noted to have seven-day and 14-day rates, which differed from the seven-day and 30-day follow-up rates from the other measures. NQF #3453 *Continuity of Care After Inpatient or Residential Treatment for Substance Use Disorder* was noted to be at a different level of analysis and was otherwise harmonized with NQF #3590.

Measure Gaps and Priorities

The Standing Committee spent some time discussing measurement gaps within the NQF portfolio of endorsed behavioral health measures. In particular, the Standing Committee noted that the Measure Application Partnership (MAP) did not include any behavioral health measures in the Measures Under Consideration list for the 2020-21 MAP review cycle. The Standing Committee suggested that one measure gap that warrants consideration is for measures of engagement in behavioral health services, ensuring that patients are not lost from care. The Standing Committee also noted some process measures that capture engagement but that detecting relapse and risk for relapse early as a potential area to explore for certain conditions would be helpful. In addition, the Standing Committee added that there continue to be measure gaps related to transitions in care for behavioral health patients.

Public Comment

No public or NQF member comments were provided during the measure evaluation meeting. There were also no comments made online during the pre-evaluation commenting period, which opened on December 15, 2020, and closed on January 15, 2021.

Next Steps

NQF will post the draft technical report on March 22, 2021, for public comment for 30 calendar days. The continuous public comment with member support will close on April 20, 2021. NQF will reconvene the Standing Committee for the post-comment web meeting on May 24, 2021.