NATIONAL QUALITY FORUM

Meeting Summary

Behavioral Health and Substance Use Standing Committee Post-Comment Web Meeting, Fall 2018 Cycle

The National Quality Forum (NQF) convened a public web meeting for the Behavioral Health and Substance Use Standing Committee on May 3, 2019.

Welcome, Introductions, and Review of Web Meeting Objectives

Shaconna Gorham, NQF senior project manager, welcomed participants to the web meeting and reviewed the following meeting objective: to review and discuss the comments received for this review cycle.

Review and Discussion of Comments

Nicolette Mehas, NQF director, provided opening remarks and reviewed the measures that were evaluated during the fall 2018 measure cycle. During this cycle, the Behavioral Health and Substance Use Standing Committee reviewed four measures. Two maintenance measures were recommended for continued endorsement:

- 0004 Initiation and Engagement of Alcohol and Other Drug Abuse or Dependence Treatment
- 2152 Preventive Care and Screening: Unhealthy Alcohol Use: Screening & Brief Counseling

One new measure was recommended for endorsement:

- 3453 Continuity of Care After Inpatient or Residential Treatment for Substance Use Disorder (SUD)

One new measure was not recommended for endorsement:

- 3451 Non-Acute Mental Health Services Utilization for Dual Eligible Beneficiaries

The draft report for this measure cycle was posted on the project webpage for public and NQF member comment on March 11, 2019 for 30 days. The commenting period closed on April 9, 2019. Michael Abrams, NQF senior director, presented a full summary of the most salient comments received during the post-evaluation public and member comment period. NQF received 16 comments from four member organizations. The three major themes identified in the comment memo after the post-evaluation commenting period were as follows:

1. Measure specification considerations
2. Data limitations
3. Measure gaps

Theme 1 – Measure Specification Considerations

For measure 0004, at least one commenter asked for clarity as to why the post-initiation engagement period was extended from 30 to 34 days. The Committee accepted the developers' response that the extension was implemented to account for typical prescription refill cycles as well as common lags in claims processing. The Committee took no exception to a comment generally encouraging measures to use DSM-5 terminology of "substance use disorder" rather than the DSM-IV "abuse or dependence" lexicon.

For measure 3453, one comment expressed concern that peer supports or case management services alone were not counted as a numerator event. In response, the developers cited the published literature regarding peer supports as being insufficient, absent other more traditional forms of care (e.g., standard SUD outpatient care), but the developers indicated that in the future they plan to consider both case management and peer supports as qualifying events as additional supporting evidence emerges. The Committee was satisfied with that response.

Regarding a potential concern that measure 0004 psychotherapy codes are required for MAT, the Committee discussed and the developers (who were on the call) clarified that psychotherapy codes are not also required to count MAT (i.e., MAT codes in isolation are counted as a numerator event). As an example of how such measurement details are difficult to remember, just prior to this conference call, the developer wrote to NQF that psychotherapy codes were needed to include MAT codes because: "guidelines recommend the use of medication-assisted treatment (MAT), or pharmacotherapy used in conjunction with psychosocial services."

For measure 3453, a commenter expressed concern that a primary diagnosis, as opposed to any diagnosis in the claims record, was too stringent a requirement for counting persons or events for this measure. The developer responded that such stringency was only applied to the numerator (i.e., to treatments) not to the denominator (i.e., potential cases to treat). At least one Committee member did seek clarification about the meaning of "secondary" diagnostic position. The developer clarified that "secondary" position refers to any diagnosis recorded on the claim that is not labeled as "primary" diagnosis.

Theme 2 – Data Limitations

For measure 3453 and measure 2152, the Committee discussed a commenter's concern about lack of sensitivity of the numerator of these measures because some services may not be billable or otherwise visible in claims data (e.g., because of bundling or the absence of specific codes to differentiate them from coincidental services). The developer agreed that this problem represents a typical error of omission that accompanies the use of claims data as the source, and the Committee generally agreed. At least one Committee member cautioned that errors of commission (i.e., over-counting, and thus inflating performance scores) were of concern vis-à-vis the suggestion of this comment. This is because one can lower the bar too much, allowing providers or other accountable entities to credit themselves for trivial or transient services that do not really reflect good care. By the end of the discussion, the Committee generally agreed that both sensitivity and specificity issues of this measure should be tightened as much as possible moving forward. Moreover, the discourse here lead Committee members to broach the general challenges of measures of "care coordination" and related concepts.

For measure 0004, one commenter expressed concern that the lags in claims processing inhibits use of claims, per se, as a trigger for timely initiation and engagement of new SUD episodes. The developer was not asked to respond to this comment as it seemed to staff to be purely a data rather than a measure design issue. The Committee was sympathetic and even noted this problem as "universal" to many quality measure data sources.

Theme 3 – Measure Gaps

One commenter offered the general comment that serious mental illness (SMI) should now be designated as a disparities category (presumably for strategic reporting or even risk-adjustment purposes). Committee members were generally sympathetic to this suggestion especially as it could be used to tailor measures and treatments given the unique vulnerabilities faced by individuals with SMI (e.g., much higher smoker rates, higher risks of victimization, etc.). The Committee at the same time cautioned against the use of SMI as a risk-adjustment parameter, because of concern that this would give providers and other accountable entities an excuse to deliver something less than the best care.

The same commenter mentioned above noted the need for more measures that directly assess the following constructs: (1) quality of life (with specific instruments suggested), (2) the full continuum of treatment/intervention (from birth to death; and from prevention to recovery), and (3) measures to specifically address the use of long-acting injectable (LAI) antipsychotics. The first two gap suggestions were met with the Committee's agreement. The LAI suggestion was also encouraged by at least one member who noted that such methods were promising yet underutilized, and that many persons do not like taking daily oral doses of antipsychotics. Another Committee member, however, argued that patients may not prefer LAIs, which prompted the brief but important, general comment that measurement and the BHSU treatment it encourages should be tailored to patient preferences and needs as much as possible—especially given the status quo which does not measure patient preferences well.

Measure-Specific Comments

For measure 3451, commenters supported the Standing Committee's decision not to recommend this measure for endorsement but expressed the need for measures for the dual-eligible population. The Committee agreed that this is an important measure gap, but otherwise had no further discussion on this topic.

For measure 2152, one commenter questioned whether the denominator (element level) kappa statistic value of 0.31 was enough to demonstrate reliability. NQF staff clarified that there are not set thresholds for such statistics. Instead, Committees are counseled to consider such statistics in their fullest context to decide if it is sufficient for recommending endorsement. Moreover, in this case this "fair" (per Cohen) Kappa coincided with 85 percent agreement in denominator identifications (2 visits or 1 preventative visit reproduces in test-retest of claims) and resulted in the seemingly low Kappa because of the adjustment that this statistic makes for agreement by chance. More importantly, the most recent application for measure 2152 reported score-level Adam-R reliability coefficients that strongly supported the conclusion that the measure is reliable at differentiating performance between different providers. This Adams-R score level reliability supplants the element level testing and fulfills NQF's reliability testing requirements. The Committee did not take exception with this overall assessment.

Public Comment

Desmirra Quinnonez, NQF project analyst, opened the web meeting to allow for public comment. Two members of the public provided comments. The first commenter thanked the Committee for their work and inclusion of her and other public input. She conveyed three points in her comments: (1) Case management as a mode of therapy has (e.g., T1016 and H006), in her experience, been useful and should be encouraged vis-à-vis measurement; (2) she agrees with the importance of special consideration for SMI populations given the unique challenges (e.g., higher smoking rates) they face; and (3) she was pleased to hear the discourse about patient-focused treatment plans as she noted her organization's efforts in deploying motivational interviewing and stages of change approaches to its therapeutic enterprise.

Finally, a developer from Brandeis, who is working with Mathematic on measure 3453, offered thanks to the Committee and confirmed that the measure defines "secondary" diagnosis as any diagnosis recorded on the claim outside of the primary position.

Next Steps

Shaconna Gorham reviewed the remaining fall 2018 cycle timeline and highlighted the upcoming Consensus Standards Approval Committee (CSAC) in-person meeting, scheduled for June 5 and 6, 2019. The CSAC will review the Committee's measure endorsement recommendations and render a final endorsement decision. Following the CSAC decision, a 30-day appeals period will open on June 14, 2019 and will close on July 15, 2019.

During the review, the Committee asked staff if they should be concerned that this measure submission cycle "pipeline was thin" with just six measures for review, and with none being new. Staff responded that such concerns were reasonable, but that during the next cycle at least two to three measures were anticipated. Moreover, staff encouraged the Committee by stating that its gaps discourse, along with ongoing QI activities at NQF, are influential, cross-pollinating, and generative in the measurement world.