NATIONAL QUALITY FORUM
Driving measurable health improvements together

Meeting Summary

Primary Care and Chronic Illness Standing Committee—Post-Comment Web Meeting, Spring 2020 Cycle

The National Quality Forum (NQF) convened the Primary Care and Chronic Illness Standing Committee for a post-comment web meeting on September 24, 2020.

Welcome, Introductions, and Review of Meeting Objectives

Co-chairs Dale Bratzler and Adam Thompson and NQF Senior Director Sam Stolpe welcomed the Standing Committee and participants to the post-comment web meeting. NQF staff reviewed the meeting objectives and conducted roll call. 18 committee members were present for the discussion, allowing the committee to revote.

Background

During this review cycle, the Primary Care and Chronic Illness Standing Committee reviewed three measures during the June and July 2020 measure evaluation meetings. One measure was not recommended for endorsement, and the Committee did not reach consensus on two measures.

Not Recommended:

- 3570e Intervention for Prediabetes

Consensus Not Reached:

- 3569e Prediabetes: Screening for Abnormal Blood Glucose
- 3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes

NQF staff provided an overview of the process for discussing and revoting on the criterion that did not reach consensus. NQF clarified for the Committee that during the post-comment measure review, the criterion under consideration must exceed 60% Committee votes of "pass"; otherwise, it does not pass. During the meeting, the Primary Care and Chronic Illness Standing Committee voted on criteria where consensus was not reached from the spring 2020 evaluation cycle.

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

Review of Public Comments

Theme 1 - Alignment of measure exclusions

Several commenters noted that the three measures, which were evaluated, have different exclusions. Commenters also suggested adding an exclusion such as exclusion of patients who are older and/or have multiple comorbidities and limited life expectancy.

> **Committee Response:**
> Thank you for your comment. The Committee carefully reviewed the exclusions with the developer during the post comment call. The Committee agrees that for NQF #3569e there

should be more constraints around the age criteria to align with United States Preventative Services Task Force (USPSTF) recommendations.

Theme 2 - Concerns with data capture

Several commenters raised concerns over the measures' feasibility noting that currently there is no easy way to capture some of the interventions, and they are likely not well-documented in EHRs.

Committee Response:

Thank you for your comments. The Committee reviewed these comments as well as the developer's response and discussed this theme at length during the post-comment meeting. The Committee has heard the measure developer's argument that the measures' Feasibility Scorecard issues within the accuracy domain were offset by the validity testing using the parallel forms methodology. In general, the Committee has indicated that the validity testing has adequately demonstrated that the data elements necessary to calculate the measure may be represented inside of the EHRs where the measure was tested. The Committee expressed other concerns associated with the validity of NQF #3569e and #3570e. The Committee did not address validity concerns with #3571e, because the Committee did not pass the measure due to weaknesses on the evidence criterion.

Re-vote on Consensus Not Reached Measures

3569e Prediabetes: Screening for Abnormal Blood Glucose

Measure Steward/Developer Representatives at the Meeting

- Beth Tapper
- Kate Kirley
- Koryn Rubin
- Greg Wozniak
- Tannaz Moin
- Janet Williams
- Stavros Tsipas

Standing Committee Votes

- Validity: M-8; L-9; I-1

Standing Committee Recommendation for Endorsement: Not Recommended

NQF #3569e did not achieve consensus on validity during the initial measure evaluation meeting. In the discussion of comments received related to NQF #3569e, the Committee first turned to the measure developer, the American Medical Association (AMA), to ask for a summary of their responses to the comments. In response to the comment that suggested that the term "prediabetes" was inappropriate because it confers the suggestion of a disease and expressed a preference for the term "abnormal blood glucose," the developer first provided an acknowledgement that the measure title itself includes reference to abnormal glucose and that they have noted the input point. The developer also responded to comments related to the fact that this measure does not include an upper age limit exclusion, noting that this point was debated within their own Technical Expert Panel (TEP) resulting in a consensus not reached vote which lead directly in not including an upper age limit. The developer stated that not including an upper age limit is aligned with the recommendation from American Diabetes Association (ADA) guidelines. The developer also alluded to evidence that suggested that an upper age limit for the measure is not appropriate given that older patients have been shown to benefit from screening as well.

The requirement of confirmation of results was noted by AMA's TEP to not be pragmatic nor aligned with USPSTF guidelines.

The comments also reflected previous Committee discussion around validity concerns expressed during the initial July measure evaluation meeting. At that time, several Committee members had expressed reservations associated with the accuracy domain of the Feasibility Scorecard for data elements related to fasting glucose tests. The developer noted that the parallel forms validity testing that they performed resolved initial concerns related to the calculation of the measure as suggested in the Feasibility Scorecard, noting moderate to excellent crude agreement and kappa statistics between abstractors and calculations from the eCQM. The developer emphasized that the fasting glucose data element that had accuracy concerns were directly addressed through the feasibility testing. The developer was questioned on their assertion that the overwhelming majority of records were HbA1c data elements and was asked what percentage of data elements pulled were fasting blood glucose. The developer indicated that the fasting blood glucose accounted for less than ten percent of the data.

Several Committee members disagreed with the developer on not including an upper age limit, viewing the lack of the upper age limit as a threat to the validity of the measure. The Committee asked the developer to highlight the evidence that older patients benefit from such interventions. The developer reviewed their references included in responses to comments with the Committee, including ADA guideline screening recommendations, a smaller study by Kramer, et al., and evidence from the National Diabetes Prevention Program.

AMA also noted that their TEP felt that having measure exclusions for patients with limited functional status or limited life expectancy were sufficient to identify those who should not be screened. One member agreed with the developer that not having an upper age limit was appropriate based on experiences managing lifestyle change programs. Another member noted this but added that they were concerned that during an appointment with especially older patients, a clinician may be required to perform a screening that they did not consider appropriate or face being penalized on their measure performance. Other members pointed out that the developer could simply adopt the 40-to-70-year age group suggested by the comments from American Academy of Family Physicians (AAFP) and American Geriatrics Society (AGS). This solution would not prohibit clinicians from still screening patients who were older but rather provides a known and supported age span for the purposes of accountability, allowing for a consistent measure denominator. The developer responded that the expectation for performance on the measure is not to achieve perfection and that performance cut points can be used to account for instances where clinicians may determine that it may not be appropriate for certain patients to be screened.

The Committee asked if age-range concerns were appropriate to consider within a validity discussion, noting reservations around supporting the measure with the current age limits. NQF staff reaffirmed that if the Committee felt that the definitions that were used to capture the patient population within the measure do not align with clinical recommendations, it has direct bearing on the question of whether the metric does in fact measure what it purports to measure, which is the central question of validity.

The Committee did not support the measure on validity, a must-pass criterion. The Committee did not recommend the measure for endorsement.

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes

Measure Steward/Developer Representatives at the Meeting

- Beth Tapper
- Kate Kirley
- Koryn Rubin
- Greg Wozniak
- Tannaz Moin
- Janet Williams
- Stavros Tsipas

Standing Committee Votes

- Evidence: M-7; L- 7; I-4
- Exception to Evidence: Yes-9; No-8
- Validity: N/A

Standing Committee Recommendation for Endorsement: Not Recommended

NQF staff noted that consensus was not reached during the measure evaluation meeting for NQF #3571e on evidence and validity, both must-pass criteria. The comments received reflected concerns associated with evidence on the screening interval of one year, saying that exclusions for this measure differed from the other eCQMs submitted by AMA (e.g. comfort care not included in this measure), and that testing should include a variety of tests, a specific time frame, and include considerations associated with access. The developer responded to those concerns by noting that public comments were generally supportive of an exception to evidence and emphasized that their validity testing performed addressed many of the issues raised related to the accuracy of data elements.

The Committee began the discussion by reviewing a comment and response from AAFP focused on the screening interval. AAFP asserted that a three-year interval would be appropriate. In their response, the developer noted that a three-year interval is appropriate for a normal glucose readings (USPSTF), but that annual testing is appropriate if an abnormal glucose result is obtained (ADA). A Committee member suggested that there is not sufficient evidence that supports the one-year rescreening interval, because it has not been directly correlated with better outcomes but is based on expert opinion. Another Committee member countered that the test itself is not overly burdensome and seems appropriate.

NQF staff then reviewed the evidence discussion in the July measure evaluation meeting, noting that the developer cited the USPSTF and ADA guidelines as evidence for the measure. The developer noted that the annual testing recommendation came directly from the ADA guideline where it was given an “E” grade, meaning that it is based on expert opinion. Staff then reviewed the NQF criteria for evidence submissions, including a detailed walkthrough of the evidence algorithm found in NQF’s 2019 measure evaluation criteria, highlighting the pathway of exception to evidence for measures rated as “insufficient” because they are based on expert opinion. The Committee asked the developer if there was a systematic review associated with benefits and risks of the intervention as part of the expert opinion recommendations. The developer referred to the ADA guidelines and their own TEP review of the measure to indicate that a careful review of existing evidence was conducted prior to providing that expert opinion.

The developer was asked if patients who were prediabetic and found to be stable for a lengthy period of time would be excluded, but it noted that there is not a ready approach to guide such an exclusion.

The Committee did not support the measure on evidence, a must-pass criterion. The Committee did not discuss validity, which also did not reach consensus during the measure evaluation meeting, because it did not pass on evidence. The Committee did not recommend the measure for endorsement.

Request for Reconsideration

3570e Intervention for Prediabetes

Measure Steward/Developer Representatives at the Meeting

- Beth Tapper
- Kate Kirley
- Koryn Rubin
- Greg Wozniak
- Tannaz Moin
- Janet Williams
- Stavros Tsipas

Standing Committee Votes

- Request for Reconsideration: Yes-2; No-14

Standing Committee Recommendation for Endorsement: Not Recommended

NQF staff noted that the measure did not pass during the initial measure evaluation meeting and that the developer had since provided a reconsideration request. The developer suggested within that request that the Committee had been inconsistent in the application of NQF criteria and that the Committee had conflated validity and feasibility. Moreover, the developer suggested that it was not clear why it was that the measure passed on validity where the other two did not achieve consensus on validity and did not pass on feasibility. The developer also noted that the measure passed all must-pass criteria but did not receive overall endorsement.

NQF staff summarized the comments received, noting that some commenters called into question the interventions contained within the measure specifications, namely either referral to CDC-recognized diabetes prevention program (DPP), referral to medical nutrition therapy with a registered dietician, or prescription of metformin. One commenter noted that intensive behavioral counseling or other interventions are not adequately represented in this measure, making the measure poorly aligned with current guidelines and best practices. Commenters noted that the options of prescribing metformin or referring patients out will either be burdensome and drive up cost, or result in a narrow, specific pharmacotherapeutic option. It was also noted that DPPs are not widely available through the entire country.

One Committee member noted that programs based on DPP protocols are fairly well available throughout the country. Another member added that poor bandwidth is now the primary barrier, but telehealth and virtual dashboards are beginning to address access challenges for rural areas, also noting that may health plans are covering the service. Another Committee member noted that they had a challenge in accessing this service himself under his insurance carrier unless he was coded as diabetic. Another noted that there are provisions for Medicare beneficiaries that makes DPP widely accessible. Other Committee members expressed concern that the measure equates the three interventions when evidence suggests that behavioral interventions are stronger than metformin. Another Committee member expressed support for this remark and added that from a feasibility perspective this fact alone creates a lot of challenges. Committee members expressed concern associated with the unintended consequences of driving a greater utilization of metformin.

The Committee then invited the measure developer to outline their rationale for their reconsideration request. The AMA noted that the measures were submitted according to NQF measure evaluation criteria and with significant effort put into the development and testing of the measures. The AMA stated that they are concerned that the criteria for feasibility, scientific acceptability (particularly validity), and usability were not applied appropriately. Related to the feasibility and validity concerns, the developer stated that the results of the validity testing demonstrate that the results of the parallel forms tested resulted in kappa statistics indicating moderate to near-perfect agreement. AMA further acknowledged that not every data element was captured in the two EHR systems tested but noted that the validity testing showed that the results produced were acceptable. Moreover, AMA noted that EHR systems will improve to better capture the data elements needed as organizations begin working to implement and track these measures. AMA referred to previous dialogue related to fasting blood glucose, which was relevant for all three measures. The concern that this element was not captured in structured data fields was not found to be problematic within AMA's data element validity testing with what AMA characterized as a nearly zero percent occurrence. A representative endocrinologist from AMA's TEP added additional commentary that the measure does not say that the three interventions are equivalent but that there are different options. Further, she noted that the Committee emphasized that within the DPP study, intensive lifestyle interventions were the most efficacious, but there are other studies concluding that metformin is equivalent for certain populations and certain conditions. It was also emphasized that comparative effectiveness studies of virtually delivered DPP interventions have shown similar weight loss outcomes.

One Committee member noted that the measure is not doing enough to improve patient outcomes, adding that while there are options, the measure does treat the three interventions as equivalent and that there may be unintended consequences associated with that. The member further suggested that if it were framed as all of these options being offered to the patient, then that would be different, but as the measure is constructed there is only one box that can be checked, and this will not necessarily lead to the same results.

The Committee then voted on AMA's reconsideration request, which was not approved by the Committee.

Public Comment

No public or NQF member comments were provided during the measure evaluation meeting.

Next Steps

NQF staff informed the Committee that the Consensus Standards Approval Committee (CSAC) will review measures #3569e, #3570es and #3571e during its November 17-18, 2020 meeting. Afterwards, the 30-day appeals period will be from November 23, 2020 to December 22, 2020.