



Cancer Standing Committee – Spring 2020 Measure Evaluation Web Meeting

The National Quality Forum (NQF) convened the Cancer Standing Committee for a web meeting on July 10, 2020 to evaluate one measure.

Welcome, Introductions, and Review of Meeting Objectives

NQF welcomed the Standing Committee and participants to the web meeting. NQF thanked the Committee members for their continued dedication and time during this pandemic, given the COVID-19 impact on the healthcare enterprise. NQF staff reviewed the meeting objectives. Committee members each introduced themselves and disclosed any conflicts of interests. Committee member Dr. Robert Rosenberg was recused from measure #0508 due to his work as an employee of the measure developer.

Some Committee members were unable to attend the entire meeting. There were early departures and late arrivals. The vote totals reflect members present and eligible to vote. Quorum was met and maintained for the entirety of the meeting.

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 16 NQF-endorsed measures in the Cancer Standing Committee's portfolio. Additionally, NQF reviewed the Consensus Development Process and the measure evaluation criteria.

Measure Evaluation

During the meeting, the Cancer Standing Committee evaluated one maintenance measure for endorsement consideration. A summary of the Committee deliberations will be compiled and provided in the draft technical report. NQF will post the draft technical report on August 14, 2020, for member and public comment on the NQF website. The draft technical report will be posted for 30 calendar days.

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

0508 Diagnostic Imaging: Inappropriate Use of “Probably Benign” Assessment Category in Screening Mammograms (American College of Radiology)

Measure Steward/Developer Representatives at the Meeting

Karen Orozco, Judy Burleson

Standing Committee Votes

- Evidence: H-1; M-16; L-0; I-0
- Performance Gap: H-0; M-7; L-10; I-0 (*consensus not reached*)
- Reliability: H-15; M-1; L-0 I-0
- Validity: H-0; M-4; L-5; I-8

- Feasibility: N/A
- Use: N/A
- Usability: N/A

Standing Committee Recommendation for Endorsement: Measure did not pass on validity

The Standing Committee did not vote on the recommendation for endorsement because the Committee did not pass the measure on validity—a must-pass criterion. This measure reports the percentage of final reports for screening mammograms that are classified as “probably benign.”

The Committee began its discussion with evidence, which was updated by the developer to include the ACR B-RADS Atlas, which provides guidance on using a “probably benign” category versus other categories. The Committee also discussed the logic model presented within the evidence document to describe the steps/decision process when implementing this measure. Specifically, the logic model states that if an abnormality is not malignant and the radiologist is also not 100% sure that it is benign, an evaluation of a patient's prior mammography exams is required, rather than an additional diagnostic scan. Clarification was provided by the developer, which mentioned that the recommendation is to use prior mammography exams to resolve issues. Based on that information, a Committee member noted that it is important to capture that this measure is applicable to follow up mammograms rather than first-time mammograms. Overall, the Committee agreed with the evidence provided.

At the outset of the discussion on performance gap, NQF staff shared the preliminary analysis rating of low for this criterion, which indicates the measure is topped out (mean performance reported was 2.93%, lower score is better). NQF staff noted that such a high-performance rate allowed the Committee to consider this measure for Reserve Status. The purpose of Reserve Status is to retain endorsement of reliable and valid measures that have overall high levels of performance so that performance can be monitored, as necessary, to ensure that performance does not decline. NQF staff noted that Reserve Status should be applied only to highly credible, reliable, and valid measures that have high levels of performance due to quality improvement actions (e.g., not due to documentation practices only).

During the discussion on performance gap, the Committee reviewed the data presented from the developer from more than 100,000 providers with at least 10 patients who received a mammogram each year between 2015-2018. The average performance across all physicians was 0.52%. The measure developer clarified the interpretation of the performance rate, which uses inverse terminology and therefore *not* meeting the measure is the correct quality action. Providers’ case volume as it relates to implementation of this measure also was noted by the Committee; rates of compliance for providers with a low case volume could be very different. During the vote on this must-pass criterion, the Committee did not reach consensus on performance gap.

The Committee reviewed and discussed the measure’s reliability testing; a beta-binomial model measuring the ratio of signal-to-noise was provided showing a reliability statistic of 0.99 for physicians having a minimum of 10 events for the period 2015-2018, suggesting the measure has high reliability. This Committee agreed with this assessment, citing it to be very reliable.

During the discussion on validity, NQF staff noted the preliminary analysis rating was insufficient. They stated that the developer conducted construct validity, calculating Pearson’s coefficients. Staff noted, however, that the developer was unable to find a correlation of this measure with two other process measures (including an NQF-endorsed measure), having hypothesized that good performance on this measure likely indicates physicians who follow guidelines are working within practices that have good systems for tracking patients or do not unnecessarily recall patients. The Committee agreed this

measure has high face validity, but acknowledged that is not the preferred validity for maintenance measures, since NQF requires empiric validity testing. The Committee did not pass this measure on validity, and therefore this measure is not recommended for endorsement.

Public Comment

No public or NQF member comments were provided during the pre-comment period, held from May 11–June 19, 2020, or during the measure evaluation meeting on July 10, 2020.

Next Steps

NQF will post the draft technical report on August 14, 2020 for public comment for 30 calendar days. The continuous public comment with member support will close on September 14, 2020. NQF will reconvene the Standing Committee for the post-comment web meeting on September 23, 2020.