2024 Pre-Rulemaking Measure Review
Preliminary Assessment

MUC ID	Title
MUC2024-082	Cancer Screening and Counseling Patient-Reported Outcome-Based Measure (PRO-PM)
Measure Steward & Developer	**Proposed CMS Programs**
Centers for Medicare & Medicaid Services (CMS); Yale Center for Outcomes Research and Evaluation (CORE)	Merit-based Incentive Payment System–Quality

Measure Overview
Developer-provided rationale: The goal of this measure is to evaluate whether clinicians have provided high-quality counseling for breast, cervical, colon, and lung cancer screenings for all patients. Enhanced clinician counseling can reduce disparities and promote equity in cancer screenings.
CMS-provided program rationale: CMS may add the measure Cancer Screening and Counseling Patient-Reported Outcome-Based Measure (PRO-PM) to the MIPS quality measure set as a new clinical quality measure. The PRO-PM focuses on incentivizing high-quality counseling services to reduce disparities in screenings for four cancer types: 1) breast, 2) cervical, 3) colorectal, and 4) lung. The PRO survey instrument includes questions focused on the quality of clinician counseling for cancer screening and its impact on decisions for cancer screening. This measure has the potential for future inclusion in the Value of Primary Care and Advancing Cancer Care MIPS Value Pathways (MVPs) and aligns with CMS's Meaningful Measures 2.0 framework by incorporating the patient perspective into the measurement of their clinical care.
Description: A PRO-PM to assess the quality of clinician counseling for patients eligible for select cancer screenings. The PRO-PM focuses on incentivizing high-quality counseling services to reduce disparities in screenings for four cancer types: 1) breast, 2) cervical, 3) colorectal, and 4) lung cancer. The PRO-PM requires use of a novel PRO survey instrument to collect the outcome data from patients while minimizing the burden of data collection on providers and patients and optimizing response rates. The PRO survey instrument includes questions focused on the quality of clinician counseling for cancer screening and its impact on decisions to screen for cancer.

The analyses upon which this publication is based were performed under Contract Number 75FCMC23C0010, entitled, "National Consensus Development and Strategic Planning for Health Care Quality Measurement," sponsored by the Department of Health and Human Services, Centers for Medicare & Medicaid Services.

Measure Overview

Measure background: New measure, never reviewed by Measure Applications Partnership (MAP) Workgroup or Pre-Rulemaking Measure Review (PRMR) or used in a Medicare program.

Numerator: The sum of all individual scores for eligible respondents.

Exclusions: N/A

Denominator: The total number of eligible respondents.

The target population for the survey is patients aged 21-84 who had an encounter with a qualifying clinician (clinicians who may provide primary care or counseling) meeting one of the following criteria:

- Comprehensive medicine evaluation & management (E/M)
- Medicine counseling and/or risk reduction intervention
- Pelvic & clinical breast screening exam
- Other outpatient or telephone E/M visit with one of the following qualifying diagnosis codes:
    - General exam or general gynecologic exam
    - Diagnosis related to family history or personal risk for cancer

Exclusions:

- People already diagnosed or with personal history of cancer and people presenting with symptoms of cancer or referred for diagnostic screening. This exclusion is defined using claims, and these patients are excluded from the target population (i.e., they are not sent the survey).
- Patients completing the survey who indicate they did not discuss one of the screenings of focus for this measure and selected one or more of the following reasons (listed below) are routed directly to the demographic question using skip-logic and are removed from the denominator:
    - History of cancer
    - All screenings already up to date
    - They do not need screening
    - They did not want to discuss screening

Exceptions: N/A

Measure type: Patient-Reported Outcome Performance Measure (PRO-PM)	**Measure has multiple scores:** No **Measure is a composite:** No **Measure is digital and/or an eCQM:** No **Measure is a paired or group measure:** No
Level of analysis: Clinician: Group	**Data source(s):** Digital-Applications: Patient-Reported Health Data or Survey Data (electronic)

Measure Overview	
Care setting(s): Ambulatory/office-based care	**Risk adjustment or stratification:** None
CBE endorsement status: Never submitted	**CBE endorsement history:** N/A
Is measure currently used in CMS programs? No	**Measure addresses statutorily required area?** No

Meaningfulness

Importance	
Type of evidence:	Peer-Reviewed Systematic Review; Empirical Data [Source: Measures Under Consideration (MUC) Entry/Review Information Tool (MERIT) Submission Form]
Importance: The developer reported performance scores for 12 clinician groups: min., 2.0; 10th percentile, 2.3; median, 2.9; mean, 3.0; 90th percentile, 3.3; max., 5.2; standard deviation, 0.8. The peer-reviewed literature presented focuses on disparities in screening rates, patients' access and barriers to screening (including social risk factors), and justification for selection of the four cancer screening types included. The submission briefly mentions the role of counseling and clinical interventions in increasing cancer screening rates. The submission mentions cancer screening guidelines in general terms but specific, relevant guidelines are not provided. Nine out of 11 patient work group members who voted agreed (3, 27%) or strongly agreed (6, 55%) that the measure is meaningful and would produce information that is valuable in making care decisions.	
Rating: Met	

Measure Performance

Table 1 shows deciles (i.e., the data sorted and broken into 10 equal parts) by performance score based on the data provided in the testing submission for the 12 clinician groups.

Interpretation: The mean score for the 12 entities described in the testing submission for this measure was 3.0. For this continuous variable measure, a higher score indicates better quality of care.

Table 1. MUC2024-082 Performance Score Deciles

	Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max
Mean Score (SD)	3.0 (0.8)	2.0	2.3	2.6	2.7	2.8	2.9	3.1	3.1	3.2	3.3	5.2	5.2
Number of Entities	12	1	2	1	1	1	1	1	2	0	2	1	1

Conformance
Measure alignment with conceptual intent: It is unclear how well the measure specifications align with the intent of the measure. The developer submitted the survey instrument; however, the submission does not explain the algorithm for calculating the numerator (i.e., the developer does not state which survey item or items are included, how they are scored and summed to yield the numerator, or the range in possible numerator values). Through additional discussion, the developer said that a scoring guide is available but not able to be included in the measure review package due to its exclusion in the original submission to MERIT. The committee should seek clarification on the specification of the measure, including how the algorithm handles complex situations, such as when one type of screening may be due but another is not or for subgroups with different screening needs, and whether the algorithm reflects current screening guidelines for each type of cancer. The committee should also consider whether the performance characteristics of the instrument are sufficient for the purposes of the proposed PRO-PM.
Rating: Not Met but Addressable

Feasibility	
eCQM feasibility testing conducted:	No [Source: MERIT Submission Form]
Feasibility: This measure relies on a web interface format. As such, no data elements are available in defined fields from electronic sources such as electronic health records (EHRs). Communication with the developer clarified that workflow changes are not necessary for measure implementation. (Note: The original MERIT submission indicated “not applicable” due to a submission error.) The committee should consider resources and workflows required for data collection of this PRO-PM.	
Rating: Met	

Validity	
Validity testing:	Face Validity [Source: MERIT Submission Form]
Testing level(s):	N/A
Validity: Face validity was evaluated by a vote of technical expert panel (TEP) members spanning backgrounds in clinical practice, patient counseling, psychometrics and performance measurement, quality improvement, health care disparities, and patient perspectives; nine of the 15 TEP members voted on the question of whether the measure “could differentiate good from poor quality care among providers.” Of the nine voting members, three (33%) strongly agreed and five (56%) agreed with this statement. The remaining member voted neutral.	
Threats to validity: The lack of precision and detail noted in the description of the measure specifications is a potential threat to validity. In addition, this PRO-PM is not risk adjusted or recommended for stratification. The committee should consider whether risk adjustment or stratification is warranted for this measure.	
Rating: Met	

Reliability	
Reliability testing method(s):	Signal-to-Noise [Source: MERIT Submission Form]
Testing level:	Clinician — Group
Reliability discussion: The dataset consists of 1,921 patients across eight clinician groups. The median reliability is 0.7363, and the minimum reliability is 0.5594. At least 75% of the entities have a reliability >0.6, indicating acceptable reliability and that only 25% of entities may not be able to distinguish good from poor quality care.	
Additional Reliability analyses: For Table 2, Battelle used the provided performance and reliability data and approximated decile averages by interpolation.	
Rating: Met	

Reliability Table

Table 2 shows deciles by reliability (calculated using signal-to-noise method) based on the data provided in the testing submission for the eight clinician groups. Battelle created this table to provide reviewers with a standardized format to assess reliability.

Interpretation: At least 75% of the entities have a reliability >0.6, indicating acceptable reliability and that only 25% of entities may not be able to distinguish good from poor quality care.

Table 2. MUC2024-082 Mean Reliability (by Reliability Decile)

Mean	SD	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max	IQR
0.7366	0.1400	0.5594	0.5594	0.6104	0.6557	0.6918	0.7243	0.7557	0.7882	0.8243	0.8696	0.9348	0.9348	0.1686

Usability	
Usability considered in application:	Yes
Usability discussion: In discussions during the development of the PA, the developer said that this measure was developed with input from a person and family engagement workgroup. This measure is not currently in use, but developers cite potential unintended consequences such as overuse of screening tests, overdiagnosis, false positives, or inadvertently exacerbating existing disparities.	
Rating: Met	

External Validity	
Was this measure tested in the same target population as the CMS program?	Yes
External validity discussion: The developer tested this measure in an ambulatory care setting in an all-payer target population. Cancer screening applies to a broad population base, appropriate for MIPS.	
Rating: Met	

Appropriateness of Scale

Similar or related measures in program(s):	None
Measure appropriateness, equity, and value across target populations/measured entities: The committee should consider whether the measure, as specified, (i.e., without risk adjustment or stratification) can address equity. Regarding equity of this measure's performance and benefit across populations, the developer's literature review and analysis do not provide sufficient information to assess the potential for differential benefit or harm to specific subgroups of participating entities or their patient populations. The committee should consider the distribution of benefit and risks/burdens of the measure within the proposed program population.	

Time to Value Realization

Plan for near- and long-term impacts after implementation:	The submission did not explore measure impacts in any detail.
Measure implementation impacts over time: There is a need for a full examination of near- and long-term impacts of this measure after implementation across patient and measured entities. Questions for the committee to consider: • What are the potential near- and long-term impacts of this measure on measured entities, MIPS, and patient populations? • Will benefits and burdens associated with this measure be realized within an appropriate implementation time frame? • How will this measure mature through revisions in the future if added to the MIPS measure set?	