2024 Pre-Rulemaking Measure Review
Preliminary Assessment

MUC ID	Title
MUC2024-028	Screening for Abnormal Glucose Metabolism in Patients at Risk of Developing Diabetes
Measure Steward & Developer	**Proposed CMS Programs**
American Medical Association	Merit-based Incentive Payment System (MIPS)–Quality

Measure Overview
Rationale (excerpt from submission): This measure is critical to identifying patients with prediabetes who may benefit from interventions to prevent type 2 diabetes and to identify patients with undiagnosed type 2 diabetes. The Centers for Disease Control and Prevention (CDC) estimates that approximately 98 million American adults have prediabetes [CDC, 2024]. They note that more than 80% of adults with prediabetes are not aware that they have the condition. Regular glycemic screening is a critical first step to identifying patients with prediabetes and helping patients avoid the disability and costs associated with progression to type 2 diabetes.
CMS-provided program rationale: CMS may add the Screening for Abnormal Glucose Metabolism in Patients at Risk of Developing Diabetes measure to the MIPS quality measure inventory as a new electronic clinical quality measure. This measure supports the care of identified prediabetic patients who may benefit from intervention to prevent type 2 diabetes and identifies patients with undiagnosed type 2 diabetes. Regular glycemic screening is an important first step to identifying patients with prediabetes and helping patients avoid the disability and costs associated with progression to type 2 diabetes. This measure is fully tested and developed. MIPS does not have any related measures that examine the rate of screening for abnormal glucose metabolism in patients at risk for diabetes. This measure fulfills a gap in MIPS for treatment of patients with prediabetes and may be a potential future addition to the Primary Care MIPS Value Pathway (MVP).
Description: Percentage of adult patients with risk factors for type 2 diabetes who are due for glycemic screening for whom the screening process was initiated during the measurement period.
Measure background: Submitted previously, but not included in Measures Under Consideration (MUC) List.

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	
Numerator: Patients who had a glycemic screening test performed and result documented during the measurement period. Glycemic screening tests include: fasting plasma glucose, glucose tolerance test, hemoglobin A1c. **Exclusions:** N/A	
Denominator: All patients with at least two office visits or one preventive visit during the measurement period who have the following risk factors for type 2 diabetes: Most recent BMI greater than or equal to 25 kg/m2 (BMI greater than or equal to 23 kg/m2 for Asian patients) during measurement period; AND Age 35-70 at start of measurement period **Exclusions:** Patient is pregnant during measurement period Patient with diagnosis of advanced illness or limited life expectancy during measurement period Patient with diagnosis of diabetes during 2-year look-back period Patient with diagnosis of prediabetes during 2-year look-back period Patient with glycemic screening performed during 2-year look-back period **Exceptions:** N/A	
Measure type: Process	**Measure has multiple scores:** No **Measure is a composite:** No **Measure is digital and/or an eCQM:** Yes **Measure is a paired or group measure:** No
Level of analysis: Clinician: Individual and Group	**Data source(s):** Digital-Electronic Health Record (EHR) Data
Care setting(s): Ambulatory/office-based care	**Risk adjustment or stratification:** No
CBE endorsement status: Never submitted	**CBE endorsement history:** Never submitted
Is measure currently used in CMS programs? No	**Measure addresses statutorily required area?** No

Meaningfulness

Importance	
Type of evidence:	Clinical Guidelines or U.S. Preventive Services Task Force (USPSTF) Guidelines [source: MUC Entry/Review Information Tool (MERIT) Submission Form, MIPS Peer-Reviewed Journal Article Form]
Importance: The purpose of this measure is to ensure that patients who are at risk of developing diabetes have a screening process initiated for abnormal glucose metabolism in accordance with the USPSTF guideline recommendations as well as to address a recommendation from the National Clinical Care Commission (NCCC) to Congress and the Secretary of Health and Human Services (HHS) that called for adopting the screening measure developed by the American Medical Association as part of a strategy to prevent diabetes among individuals at higher risk. This aligns with the Meaningful Measures domain of Wellness and Prevention. Improved performance on this measure will have significant impact on the clinical practice of providers enrolled and the patient population served by providers participating in MIPS.	
Rating: Met	

Measure Performance

Table 1 shows performance score deciles (i.e., the data sorted and broke into 10 equal parts) based on the data provided in the testing submission for the 48 entities with 20 or more eligible cases.

Interpretation: The mean score for the 48 entities described in the testing submission for this measure was 45.9. For this proportion measure, a higher score indicates better quality of care.

Table 1. MUC2024-028 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)	45.9 (18.1)	19.0	20.0	25.9	29.7	37.4	42.9	47.0	55.3	60.3	68.3	78.1	87.7
Number of Entities	48	1	5	5	5	5	4	5	5	5	5	4	1

Conformance
Measure alignment with conceptual intent: As outlined in the MIPS Peer-Reviewed Journal Article Form submitted, this measure's specification is appropriate and aligned with the measure target (initiation of glycemic screening process) among adult patients with risk factors for type 2 diabetes seen in primary care.
Rating: Met

Feasibility	
eCQM Feasibility testing conducted:	Yes, provided in the eCQM Scorecard.
Feasibility: As this measure is an eCQM, the measure developer conducted eCQM testing and submitted a feasibility scorecard. Results on this scorecard address the following domains/question areas: • Data availability: Is the data readily available in a structured format, i.e., resides in fixed fields in EHR? • Data accuracy: What is the accuracy of the data element in EHRs under normal operating conditions? Are the data source and recorder specified? • Data Standards: Is the data element coded using a nationally accepted terminology standard? • Workflow: Is the data captured during the course of care? And how does it impact workflow for the user? Across the three EHR systems (Epic, Next Gen, and Athena), the measure demonstrated high feasibility of data elements. The following data elements were feasible at two-thirds of the sites: Encounter, Performed: Preventative Care; Encounter, Performed: Office Visit; Diagnoses: Advanced Illness; and Diagnosis: Limited Life Expectancy. The feasibility plan provided in the feasibility scorecard further describes specific code considerations for these elements and, ultimately, these elements are feasible and do not require additional changes.	
Rating: Met	

Validity	
Validity testing:	Face Validity [Source(s): MERIT Submission Form, MIPS Peer-Reviewed Journal Article Form]
Testing level(s):	Individual Clinician
Validity: Face validity testing for the measure was assessed across 11 individuals on the measure developer's technical expert panel (TEP), comprised of clinicians and patients/caregivers. Ten out of the 11 members voted "Yes" when asked, "Do you agree that the performance scores resulting from the Screening for Abnormal Glucose Metabolism measure can be used to distinguish good from poor clinician-level performance?" Both patient/caregiver members unanimously voted in favor of the measure. The developer did not conduct empiric validity testing of this measure.	
Threats to validity: The measure is not risk adjusted or recommended for stratification. Exclusions for the measure reduce likelihood of introducing confounders and are appropriate for the measure specification and intent.	
Rating: Met	

Reliability	
Reliability testing method(s):	Signal-to-Noise [Source(s): MERIT Submission Form, MIPS Peer-Reviewed Journal Article Form, Methodology Supplemental]
Testing level:	Individual Clinician

Reliability discussion: The numerator and denominator for this measure are well defined. Entity-level reliability is calculated from a full year of data (2021). The dataset consists of 4,530 patients across 48 clinicians (each with 20 or more eligible cases). The median reliability is 0.897, and the minimum reliability is 0.713. Of the entities, 100% have a reliability >0.6, suggesting that this measure is capable of differentiating entities by quality of performance.
 	
Inter-rater agreement was used to assess encounter-level reliability of seven data elements, sometimes at two different sites.
• A kappa value of 1.0 was observed at one or both of the sites (when two were tested) for five of the seven data elements:
  ○ "Is the patient 18 years of age or older before the start or during the measurement period?"
  ○ "Does the medical record indicate an office visit during the measurement period?"
  ○ "Does the medical record indicate at least two office visits or at least one preventive visit during the measurement period?"
  ○ "Is the patient ≥43 years of age before the start or during the measurement period?"
  ○ "Does the medical record indicate an HbA1c laboratory test during or 3 years before the measurement period?"
• The data element "Does the medical record indicate a BMI ≥ 25 at encounter during measurement period?" resulted in a kappa of 0.91 at Site 1, but a kappa of 0 at Site 2. (A kappa of 0 occurs even if the percent agreement is high if one rater rates every record the same but the other rater does not.)	

Reliability
• The data element "Does the medical record indicate an active diabetes diagnosis with an encounter such that the diagnosis overlaps after the encounter?" was evaluated on 36 records at one site and yielded a kappa of 0. Except for the data elements that resulted in a kappa of 0, the kappa values are above 0.6 for all data elements. The kappa of 0.91 for "Does the medical record indicate a BMI ≥ 25 at encounter during measurement period?" suggests that the data element is reliable, and the kappa of 0 at Site 2 could be improved with a larger sample of records. Little can be said about the reliability of "Does the medical record indicate an active diabetes diagnosis with an encounter such that the diagnosis overlaps after the encounter?" without reassessing with a larger sample of records. Overall, the inter-rater agreement in this testing sample indicates strong encounter-level reliability for these data elements.
Additional reliability analyses: The data provided in MUC2024-028-Methodology-PerformanceScores-20240510.pdf (with a denominator greater than or equal to 20) was used to fill Table 2 below (results match those provided in the submission).
Rating: Met

Reliability Tables

Table 2. MUC2024-028 Mean Reliability (by Reliability Decile)

Interpretation: All of the entities have a reliability >0.6, suggesting that this measure is capable of differentiating entities by quality of performance.

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
0.897	0.067	0.713	0.756	0.813	0.852	0.876	0.904	0.916	0.934	0.947	0.958	0.968	0.977	0.096

Usability	
Usability considered in application:	Yes [Sources: MERIT submission form, Methodology Supplement]
Usability discussion: Based on submission documents, there is an opportunity for improvement on the measure target among clinician and clinician groups participating in MIPS. The developer did not identify any external program-level factors that may present barriers to measure. The measure submission provides a thoughtful discussion of potential unintended consequences of the measure within MIPS, including the possibility of overuse of glycemic screening (if the measure were to be specified too narrowly) or undue credit for screening with inappropriate methods (if the measure were specified too broadly). The risk-benefit of these potential consequences was discussed with the measure developer's TEP and deemed acceptable for the patient-level benefits. Overall, this measure seems to have high usability within MIPS.	
Rating: Met	

External Validity	
Was this measure tested in the same target population as the CMS program?	Yes
External validity discussion: The measure testing for this measure was conducted in clinician populations and care sites representative of the MIPS population and indicates that this measure has suitable external validity.	
Rating: Met	

Appropriateness of Scale

Similar or related measures in program(s):	None
Measure appropriateness, equity, and value across target populations/measured entities: While the developer did not indicate any similar or related measures in MIPS, they did suggest that this measure aligns with multiple MIPS improvement activities including (but not limited to) Chronic Care and Preventative Care Management for Empaneled Patients and Glycemic Screening Services. Regarding equity of this measure's performance and benefit across populations, the measure developer's literature review and sub-analysis do not suggest differential benefit or harm to specific subgroups of MIPS-participating clinicians or their patients. The committee should consider if, based on their professional and patient experience, there is a chance for variation in distribution of benefit or burden across provider and patient populations.	

Time to Value Realization

Plan for near- and long-term impacts after implementation:	No
Measure implementation impacts over time: While the measure developer makes brief mention of potential outcomes for their measure on clinician and patient population, there is a need for further examination of near- and long-term impacts of this measure after implementation across provider and patient populations. Questions for the committee to consider include: • 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 timeframe? • How will this measure mature through revisions in the future if added to the MIPS quality measure inventory?	