2024 Pre-Rulemaking Measure Review
Preliminary Assessment

MUC ID	Title
MUC2024-088	Depression Screening and Follow-Up for Adolescents and Adults (DSF)
Measure Steward & Developer	**Proposed CMS Programs**
National Committee for Quality Assurance (NCQA)	Part C Star Ratings

Measure Overview
Developer-provided rationale: The DSF measure assesses the percentage of eligible Medicare Advantage plan members who were screened for clinical depression using a standardized instrument and, if screened positive, received follow-up care within 30 days. The measure provides health plan-level data related to screening for depression and follow-up treatment such as psychotherapy or pharmacotherapy. Plans may use these data to target education and outreach efforts and strengthen patient access to treatment for depression.
CMS-provided program rationale: The intent of this measure is to encourage Medicare Advantage health plans to screen for clinical depression and provide follow-up care within 30 days.
Description: The percentage of Medicare Advantage plan members 12 years of age and older who were screened for clinical depression using a standardized instrument and, if screened positive, received follow-up care within 30 days.
Measure background: Measure currently used in a Medicare program and is being submitted without substantive changes for a new or different program.
Numerator: Depression Screening: Medicare Advantage plan members aged 12 and older with a documented result for depression screening, using an age-appropriate standardized instrument, performed between January 1 and December 1 of the MP. Follow-up on Positive Screen: Medicare Advantage plan members aged 12 and older who received follow-up care on or up to 30 days after the date of the first positive screen. **Exclusions:** Medicare Advantage plan members with a history of bipolar disorder any time during the member's history through the end of the year prior to the MP; Medicare Advantage plan members with depression that starts during the year prior to the MP; Medicare Advantage plan members in hospice or using hospice services any time during the MP.
Denominator: Depression Screening: Medicare Advantage plan members 12 years of age and older at the start of the MP who also meet criteria for inclusion in the denominator (i.e., they are not excluded due to a history of bipolar disorder, depression that started in the year prior to the MP, receipt of hospice services during the MP, or not being alive at the end of the MP). Medicare Advantage plan members 12 years of age and older at the start of the MP who were enrolled in a Medicare Advantage plan throughout the period with no more than one gap in enrollment of up to 45 days. Follow-up on Positive Screen: Medicare Advantage plan

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	
members 12 years of age and older at the start of the MP who also meet criteria for inclusion in the denominator (i.e., they are not excluded due to a history of bipolar disorder, depression that started in the year prior to the MP, receipt of hospice services during the MP, or not being alive at the end of the MP) with a positive depression screen finding between January 1 and December 1 of the MP. Medicare Advantage plan members 12 years of age and older at the start of the MP who were enrolled in a Medicare Advantage plan throughout the period with no more than one gap in enrollment of up to 45 days with a positive depression screen finding between January 1 and December 1 of the MP. **Exclusions:** Medicare Advantage plan members with a history of bipolar disorder any time during the member's history through the end of the year prior to the MP; Medicare Advantage plan members with depression that starts during the year prior to the MP; Medicare Advantage plan members in hospice or using hospice services any time during the MP. **Exceptions:** None	
Measure type: Process	**Measure is a composite:** No **Measure is digital and/or an eCQM:** No **Measure is a paired or group measure:** No
Level of analysis: Health Plan	**Data source(s):** Digital-Administrative systems, Administrative Data (non-claims); Digital-Administrative systems: Claims Data; Digital-Case Management Systems; Digital-Electronic Health Record (EHR) Data; Digital-Health Information Exchanges (HIE) Data.
Care setting(s): Ambulatory/office-based care	**Risk adjustment or stratification:** No
CBE endorsement status: Not CBE endorsed.	**CBE endorsement history:** Never submitted for CBE endorsement.
Is measure currently used in CMS programs? The measure is currently being used in the End-Stage Renal Disease (ESRD) Quality Incentive Program; Medicare Shared Savings Program (MSSP); and the Merit-based Incentive Payment System (MIPS). The measure is also a part of the Medicare Adult Core Set.	**Measure addresses statutorily required area?** No

Meaningfulness

Importance	
Type of evidence:	Clinical Guidelines or U.S. Preventive Services Task Force (USPSTF) Guidelines [Source: Measures Under Consideration (MUC) Entry/Review Information Tool (MERIT) Submission Form]
Importance: The USPSTF recommends screening for depression in the adult population. This recommendation is made to reduce disparities in depression-associated morbidity and to achieve the benefit of screening, which includes further evaluation and/or diagnosis as well as a potential referral for evidence-based care. USPSTF rated the evidence for this guideline as a grade B and included 105 published studies. The measure developers did not submit an independent literature review to support the evidence base for this measure. The committee should consider importance of this measure to Part C Star Ratings-participating entities and patient populations.	
Rating: Met	

Measure Performance

Tables 1a and 1b show deciles by performance score based on the data provided in the testing submission for the 399 and 117 entities, respectively.

Interpretation: For the Depression Screening indicator, the mean score for the 399 entities described in the testing submission for this measure was 13.4. For the Follow-Up indicator, the mean score for the 117 entities described in the testing submission for this measure was 74.2. For this proportion measure, a higher score indicates better quality of care.

Table 1a. MUC2024-088 Performance Score Deciles – Depression Screening

	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	13.4	0.00	0.0	0.0	0.2	0.8	2.3	5.1	10.1	18.6	32.9	60.3	99.0
Entities	399	100	40	40	40	40	40	40	40	40	40	39	1

Table 1b. MUC2024-088 Performance Score Deciles – Follow-Up on Positive Screen

	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	74.2	11.5	43.8	56.4	63.9	69.6	74.4	78.7	82.7	86.6	90.7	95.4	99.9
Entities	117	1	12	12	12	12	12	12	12	11	11	11	1

Conformance
Measure alignment with conceptual intent: Measure specification is aligned with the measure focus (documented result for depression screening with an age-appropriate screening tool and appropriate follow-up) among Medicare Advantage plan members 12 years of age and older who do not meet exclusion criteria. Numerator and denominator populations are appropriate and exclusions align with clinical evidence.
Rating: Met

Feasibility	
eCQM feasibility testing/analysis conducted:	No [Source: MERIT Submission Form]
Feasibility: The developer reports that all data elements are in defined fields in electronic sources and some data elements align with United States Core Data for Interoperability (USCDI)/USCDI+ quality standard definitions. No reported workflow changes are needed for measure implementation. The submission mentions field testing as a potential part of face validity testing in the MERIT submission form but did not provide field testing feasibility results.	
Rating: Met	

Validity	
Validity testing method(s):	Face Validity & Empiric Validity [Source: MERIT Submission Form]
Testing level(s):	Health Plan
Validity: The developer established face validity "using NCQA's standardized process called the HEDIS measure life cycle," which involved review of literature related to measure importance, scientific acceptability, feasibility and gaps in care. The developer presented this information along with the measure to relevant clinical measure advisory panels. The Committee for Performance Measurement unanimously voted to approve the measure for the Healthcare Effectiveness Data and Information Set (HEDIS) health plan reporting. The National Committee for Quality Assurance (NCQA) Behavioral Health Measurement Advisory Panel and Geriatric Measurement Advisory Panel also reviewed the measure, but the submission did not provide results of any subsequent voting. The developer assessed empiric validity of the results using Pearson correlation to demonstrate construct validity. For Medicare plans, the correlation between the Depression Screening and Follow-up measure rates with the Utilization of the PHQ-9 measures was found to be very strong at 0.81. This indicates that this measure has strong correlation with currently in use utilization measures and has the potential for validity within the program population.	
Threats to validity: The developers considered potential threats to measure validity and suggest stratification by patient age (12-17 years, 18-64 years, 65 years and older). The committee should consider if any additional patient-level factors may be appropriate for stratification.	
Rating: Met	

Reliability	
Reliability testing method(s):	Signal-to-Noise [Source: MERIT Submission Form, Measure Information Form]
Testing level:	Health Plan
Reliability discussion: The numerator and denominator for this measure are well defined. The dataset consists of 11,265,473 patients across 399 health plans for depression screening and 36,755 patients across 117 health plans for follow-up on positive screen. For depression screening, the median reliability is 0.99. For follow-up on positive screen, the median reliability is 0.96, and the minimum reliability is 0.76. The developer reported that 100% of entities have a reliability >0.6 for depression screening and follow-up on positive screen, suggesting that these measures are capable of differentiating entities by quality of performance.	
Additional reliability analyses: For Table 2, Battelle used the performance and reliability data provided and approximated decile averages by interpolation for the Follow-Up on a Positive Screen indicator. For the Depression Screening indicator, only a single estimate for reliability is required; therefore, interpolated decile averages of the reliability data were not generated.	
Rating: Met	

Reliability Table

Table 2 shows deciles by reliability (calculated using a signal-to-noise method) based on the data provided in the testing submission for the 117 health plans measuring follow-up on positive screen. Battelle created this table to provide reviewers with a standardized format to assess reliability. Because a single estimate was reported for the reliability of depression screening, there is insufficient information to interpolate the decile averages.

Interpretation: For this measure, 100% of entities had reliability >0.6, suggesting that the measure is capable of differentiating entities by quality of performance.

Table 2. MUC2024-088 Mean Reliability (by Reliability Decile) – Follow-up on Positive Screen

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
0.96	0.03	0.76	0.90	0.92	0.94	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	0.04

Usability	
Usability considered in application:	Yes
Usability discussion: This measure is currently in use in the End-Stage Renal Disease (ESRD) Quality Incentive Program, Medicare Shared Savings Program (MSSP), and the Merit-based Incentive Payment System (MIPS). The developer reported that there have been no unintended consequences for this measure during testing or since implementation. The committee should consider usability of this measure at the health plan level within Part C Star Ratings.	
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 within health plans that align with the Part C Star Ratings program.	
Rating: Met	

Appropriateness of Scale

Similar or related measures in program(s):	None
Measure appropriateness, equity, and value across target populations/measured entities: The measure developer did not identify any related measures within Part C Star Ratings. 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:	No
Measure implementation impacts over time: While the measure developer briefly mentions potential outcomes for their measure on patient populations, there may be a need for further examination of near- and long-term impacts of this measure for measured entities and patients after implementation. Questions for the committee to consider: • What are the potential near- and long-term impacts of this measure on measured entities, proposed CMS program, 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 proposed CMS program?	