# 2024 Pre-Rulemaking Measure Review Preliminary Assessment

| MUC ID | Title |
|---|---|
| MUC2024-026 | Person-Centered Outcome Measures: Goal-Identification, Follow-Up, and Goal Achievement |
| **Measure Steward & Developer** | **Proposed CMS Programs** |
| National Committee for Quality Assurance (NCQA) | Merit-based Incentive Payment System (MIPS)–Quality |

| Measure Overview |
|---|
| **Developer-provided rationale:** There is broad agreement that patient goals and priorities should guide care and quality measures used to evaluate care. For older adults with multiple chronic conditions and functional limitations, clinical guidelines from the American Geriatrics Society have pointed to the importance of providing goal-based care. For this complex population, goal setting has been shown to reduce patient-reported treatment burden and receipt of unwanted care and correlates with greater physical and social wellbeing and care satisfaction.<br><br>CMS supports aligning care with patients' goals as demonstrated by the "Meaningful Measures" initiative, which calls for quality measures where "care is personalized and aligned with patient's goals." |
| **CMS-provided program rationale:** CMS may add the Person-Centered Outcome Measures: Goal-Identification, Follow-Up, and Goal Achievement to the MIPS quality measure inventory as a new clinical quality measure. This measure promotes goal-based care for adult patients with complex care needs, aligning with evidence supporting the efficacy of person-centered care with personalized goal setting for achieving positive health and functioning outcomes and improvements. This measure aligns with CMS's Meaningful Measures 2.0 framework and fits into the priority areas of person-centered care, chronic conditions, and behavioral health. For this complex population, goal setting has been shown to reduce patient-reported treatment burden and receipt of unwanted care and correlates with greater physical and social well-being and care satisfaction. This measure has the potential to be in several MIPS Value Pathways (MVPs) due to the clinical focus of this measure being generalized as it is based on patient-directed goal identification and achievement. |
| **Description:** The percentage of individuals 18 years of age and older with a complex care need who identified and documented person-centered goal and action plan, followed up with the identified goal, and achieved the identified goal.<br><br>Three rates are reported: |

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

- Goal Identification: percentage of individuals aged 18 or above with complex care need who had a person-centered outcome goal identified resulting in completion of goal attainment scaling (GAS) or patient-reported outcome measure (PROM) and development of an action plan.
- Follow-up: percentage of individuals aged 18 or above with complex care need who received follow-up on their person-centered outcome goal within two weeks to six months of when the person-centered outcome goal and GAS or PROM were identified.
- Achievement: percentage of individuals aged 18 or above with complex care need who achieved their person-centered outcome goal within two weeks to six months of when the person-centered outcome goal and GAS or PROM were identified.

**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:** Numerator one: Individuals in the denominator who had a person-centered outcome goal identified resulting in completion of GAS or PROM and development of an action plan during the intake period.

Numerator two: Individuals in the denominator who received follow-up on their person-centered outcome goal within two weeks to six months of the encounter date during the intake period where a goal and a GAS or PROM score were identified.

Numerator three: Individuals in the denominator who achieved their person-centered outcome goal within two weeks to six months of the encounter date during the intake period where a goal and a GAS or PROM score were identified.

**Exclusions:** N/A

**Denominator:** Individuals 18 years of age and older at the start of the measurement period with complex care needs. A complex care need is defined as a need representing two or more concurrent chronic conditions, behavioral health diagnoses, and/or social challenges. Individuals may have multiple complex care needs.

**Exclusions:**

1. Episodes for persons with a date of death in the measurement period.
2. Episodes for persons living in institutionalized long-term care (LTI).
3. Episodes for persons in hospice or using hospice services.

| | |
|---|---|
| **Measure type:** Patient-Reported Outcome Performance Measure (PRO-PM) or Patient Experience of Care | **Measure has multiple scores:** Yes<br>**Measure is a composite:** No<br>**Measure is digital and/or an eCQM:** No<br>**Measure is a paired or group measure:** No |

| Measure Overview | |
|---|---|
| **Level of analysis:** Clinician: Individual and Group | **Data Source(s):** Digital-Applications: Patient-Reported Health Data or Survey Data (electronic); Digital-Electronic Clinical Data (non-EHR) or Social Needs Assessments: Social Needs Assessment; Digital-Standardized Patient Assessment Data (electronic); Non-Digital-Standardized Patient Assessments (paper-based); Non-Digital-Patient-Reported Health Data or Survey Data (telephonic or paper-based) |
| **Care setting(s):** Behavioral health clinic; Federally qualified health center (FQHC); Other (area agency on aging; homebased primary care; community-based organization; behavioral health home) | **Risk adjustment or stratification:** Yes |
| **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; Peer-Reviewed Systematic Review; Peer-Reviewed Original Research [source: MERIT Submission Form] |
| **Importance:** This measure promotes goal-based care for adult patients with complex health care needs, aligning with literature supporting the efficacy of person-centered care with personalized goal setting for achieving positive outcomes and improvements in health and functioning. This measure aligns with CMS Meaningful Measures 2.0 and fits into priorities: person-centered care, chronic conditions, and behavioral health. The measure's requirements (completion of setting individualized goal, development of an action plan, and timely follow-up to assess the goal) align with the American Geriatrics Society guideline "Person-Centered Care: A Definition and Essential Elements," and the submission's literature review affirms this is an area of concern to patients with complex medical needs. | |
| **Rating:** Met | |

## Measure Performance

Tables 1a through 1b show performance score deciles (i.e., the data sorted and broken into 10 equal parts) for the three indicators: Goal Identification, Goal Follow-up, and Goal Achievement. These tables are based on the data provided in the testing submission for the 10 clinician groups with between 48 and 2,495 eligible cases and the 101 individual clinicians with between 30 and 371 eligible cases.

Interpretation: For the three domains, the mean scores are 75.1 (Goal Identification), 25.5 (Goal Follow-Up), and 10.2 (Goal Achievement). As indicated in the measure submission, better quality of care is associated with higher scores on these domains.

**Table 1a. MUC2024-026 Performance Score Deciles – Clinician Group Level**

| | 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 – Goal Identification | 75.1 | 18.1 | 18.1 | 38.3 | 54.9 | 67.9 | 86.5 | 92.5 | 95.3 | 97.8 | 100.0 | 100.0 | 100.0 |
| Mean Score – Goal Follow-up | 25.5 | 2.6 | 2.6 | 10.9 | 12.0 | 15.0 | 18.0 | 19.6 | 30.1 | 36.9 | 49.2 | 60.6 | 60.6 |

| | 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 – Goal Achievement | 10.2 | 0.0 | 0.0 | 2.9 | 4.9 | 5.4 | 7.9 | 10.0 | 11.5 | 11.8 | 12.1 | 35.7 | 35.7 |
| Entities | 10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

**Table 1b. MUC2024-026 Performance Score Deciles – Individual Clinician Level**

| | 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 – Goal Identification | 83.4 | 0.0 | 4.6 | 47.3 | 85.0 | 97.2 | 99.6 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| Mean Score – Goal Follow-up | 23.5 | 0.0 | 0.0 | 1.8 | 4.4 | 8.2 | 13.1 | 19.2 | 27.0 | 37.0 | 50.3 | 71.2 | 95.8 |
| Mean Score – Goal Achievement | 10.0 | 0.0 | 0.0 | 0.3 | 0.9 | 1.9 | 3.5 | 5.9 | 9.3 | 14.2 | 22.1 | 38.2 | 76.6 |
| Entities | 101 | -- | 11 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | -- |

## Conformance

**Measure alignment with conceptual intent:** As outlined in the Measure Information Form (MIF) and MIPS Peer-Reviewed Journal Article Form submitted, this measure's specification is appropriate and aligned with the measure target (identification, documentation, ongoing review and assessment of person-centered goal and action plan) among patients aged 18 and older with complex care needs in a variety of care settings.

**Rating:** Met

## Feasibility

| | |
|---|---|
| **eCQM Feasibility testing conducted:** | No [Source: MERIT Submission Form] |

| Feasibility |
|---|
| **Feasibility:** In terms of data collection, the MUC Entry/Review Information Tool (MERIT) submission indicates that implementation of this measure will require changes to provider workflows. The submission indicates that some of the required 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. Other sources of data collection include, "Manual abstraction, other digital methods, or combination." The measure uses two types of tools: GAS and PROMs, which will require additional inputs from both patients and providers during measure implementation.<br><br>The committee should consider the potential feasibility of this measure at MIPS-participating clinical sites. Does this measure have the same feasibility in resource-constrained settings as in non-resource-constrained settings? |
| **Rating:** Met |

| Validity | |
|---|---|
| **Validity testing:** | Face Validity [Sources: MERIT Submission Form, MIPS Peer-Reviewed Journal Article Form] |
| **Testing level(s):** | Clinician/Clinician Group Level |
| **Validity:** This measure underwent face validity testing through voting among experts from the PCO advisory panel at primary care, long-term service and supports, and certified community behavioral health clinic settings. From the submission, "for Primary care/long-term services and supports (LTSS) settings, out of 12 voters, 10 agreed, 2 neither agreed nor disagreed. For certified community behavioral health clinic (CCBHC) settings, out of 10 voters, 5 agreed, 5 neither agreed nor disagreed." With the current information, 69% of respondents agreed that this measure has face validity. None of the respondents disagreed that the measure had face validity. | |
| **Threats to validity:** The developer indicates that they recommend stratification by clinician group type, as their results indicate "that the demographic of the participants in CCBHCs is different from that in the primary care/LTSS site: the participants in CCBHCs are younger, and more uninsured compared to participants in primary care/LTSS sites." This suggests that there are meaningful differences in the populations served by the care settings indicated for this measure. Risk stratification or consideration of a risk-adjustment model are both ways to begin to address these population differences. However, further validity testing, including a larger sample size representative of the general provider population for face validity and at least small-scale empirical validity with description of methods at one or more care settings is encouraged to prove scientific acceptability of this pioneering measure. The developer may want to consider initial implementation of this measure within a narrower group of clinical settings to further examine performance. | |
| **Rating:** Met | |

| Reliability | |
|---|---|
| **Reliability testing method(s):** | Signal-to-Noise [Sources: MERIT Submission Form, MIPS Peer-Reviewed Journal Article Form, Attachment One: Group-Level Testing Results] |
| **Testing level:** | Individual Clinician and Group-Level Clinician |
| **Reliability discussion:**<br>*Clinician Group Level:*<br>The numerator and denominator for these measures are well defined. The dataset consists of 7,867 patients across 10 clinician groups (each with between 48 and 2,495 eligible cases). The rates of goal identification, goal follow-up, and goal achievement are measured for each clinician group. For goal identification, the median reliability is 1.00, and the minimum reliability is 0.96. For goal follow-up, the median reliability is 0.97, and the minimum reliability is 0.84. For goal achievement, the median reliability is 0.98, and the minimum reliability is 0.84. Of the entities, 100% have a reliability >0.6 for goal identification, goal follow-up, and goal achievement. The reliability testing was done on only 10 clinician groups. The reliability could be much lower for clinician groups with small denominators or if these measures are calculated with data for a shorter time period.<br><br>*Individual Clinician Level:*<br>The numerator and denominator for these measures are well defined. The dataset consists of 7,170 patients across 101 clinicians (each with between 30 and 371 eligible cases). The rates of goal identification, goal follow-up, and achievement are measured for each clinician. For goal identification, the median reliability is 1.00, the minimum reliability is 0.61. For goal follow-up, the median reliability is 0.96, the minimum reliability is 0.61. For goal achievement, the median reliability is 0.95, the minimum reliability is 0.34. For goal identification and follow-up, 100% of entities have a reliability >0.6 while for goal achievement, at least 90% of the entities have a reliability >0.6, suggesting that these measures are capable of differentiating entities by quality of performance. | |
| **Additional reliability analyses:** For Tables 2a through 2b, Battelle used the performance and reliability data provided and approximated decile averages by interpolation. | |
| **Rating:** Met | |

## Reliability Tables

Tables 2a and 2b show deciles by reliability for the three domains (Goal Identification, Goal Follow-up, and Goal Achievement) based on the data provided in the testing submission for the 10 clinician groups with between 48 and 2,495 eligible cases and the 101 individual clinicians with between 30 and 371 eligible cases. Battelle created these tables to provide reviewers with a standardized format to assess reliability.

Interpretation: Mean reliability for the three domains is well above 0.6, indicating high reliability and suggesting that this measure is capable of differentiating entities by quality of performance.

**Table 2a. MUC2024-026 Mean Reliability (by Reliability Decile) – Clinician Group Level**

| Outcome | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Goal Identification | 0.997 | 0.003 | 0.960 | 0.960 | 0.991 | 0.998 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.002 |
| Goal Follow-up | 0.982 | 0.080 | 0.840 | 0.840 | 0.855 | 0.960 | 0.970 | 0.970 | 0.975 | 0.980 | 0.985 | 0.992 | 1.000 | 1.000 | 0.025 |
| Goal Achievement | 0.985 | 0.080 | 0.840 | 0.840 | 0.875 | 0.960 | 0.975 | 0.980 | 0.987 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.040 |

**Table 2b. MUC2024-026 Mean Reliability (by Reliability Decile) – Individual Clinician Level**

| Outcome | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Goal Identification | 0.980 | 0.005 | 0.610 | 0.925 | 0.970 | 0.986 | 0.993 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.014 |
| Goal Follow-up | 0.940 | 0.010 | 0.610 | 0.855 | 0.919 | 0.938 | 0.950 | 0.960 | 0.969 | 0.976 | 0.982 | 0.989 | 1.000 | 1.000 | 0.045 |
| Goal Achievement | 0.920 | 0.020 | 0.340 | 0.708 | 0.829 | 0.887 | 0.923 | 0.949 | 0.968 | 0.981 | 1.000 | 1.000 | 1.000 | 1.000 | 0.113 |

| **Usability** | |
|---|---|
| **Usability considered in application:** | Yes. The submission discussed a potential unintended consequence: “A potential unintended consequence if the measure is implemented is resource allocation challenges. It may require additional resources, including staff time. It may be challenging in resource-constrained healthcare settings.” [Source: MERIT Submission Form] |

| Usability |
|---|
| **Usability discussion**: While the submission materials do not explicitly discuss the usability of the measure within the MIPS program, the developer does provide some elements that can be used to start assessment of usability. The potential unintended consequence of challenges to resource allocation and potential barriers to measure implementation in resource-constrained settings is identified with the consideration that staffing time and resource allocation should be considered as part of the implementation plan. Based on materials submitted, there is an opportunity for improvement on this measure target in the clinical settings indicated, but more information on workflow changes and manual data collection or abstraction required for measure implementation would be helpful for further consideration of usability. Additionally, having more information on how the measure target population and implementation steps align with the MIPS program population, alignment with other MIPS measures, and "related cost measures and improvement activities" as requested by the MERIT submission form would support the measure's use within MIPS. |
| **Rating:** Met |

| External validity | |
|---|---|
| **Was this measure tested in the same target population as the CMS program?** | Yes |
| **External validity discussion:** While this measure submission would be strengthened by providing more information on the testing sites and patient populations serviced, there is at least marginal external validity of this measure given that the types of providers included in testing are also represented within MIPS. | |
| **Rating:** Met | |

## Appropriateness of Scale

| | |
|---|---|
| **Similar or related measures in program(s):** | The developer did not identify any similar or related measures. |
| **Measure appropriateness, equity, and value across target populations/measured entities:** A review of active MIPS measures provided by the developer did not identify any similar or competing measures, suggesting that this measure would fill a gap within the current MIPS quality measure inventory. Regarding equity of this measure's performance and benefit across populations, the literature review and analysis provided by the developer in submission materials do not provide sufficient information to assess the potential for differential benefit or harm to specific subgroups of MIPS-participating clinicians or their patients beyond the consideration that the measure may perform at a lower level within behavioral health clinics. | |

## Time to Value Realization

| | |
|---|---|
| **Plan for near- and long-term impacts after implementation:** | No |
| **Measure implementation impacts over time:**<br>There is a need for further examination of near- and long-term impacts of this measure after implementation across multiple levels. Questions for the committee to consider include:<br>• What are the potential near- and long-term impacts of this measure on measured entities, MIPS, and patient populations?<br>• Will benefits and burdens associated with this measure be realized within an appropriate implementation time frame?<br>• How will this measure mature through revisions in the future if added to the MIPS quality measure inventory?<br>• Given the concern expressed about resource allocation, how might measured entities plan for and address these concerns in the near and long term? | |