

# 2024 Pre-Rulemaking Measure Review

## Preliminary Assessment Template

**Note:** This template for the 2024 preliminary assessments (PA) should guide committee members in their review of the PAs by providing an overview of each section as well as links to the Measures Management System (MMS) [Blueprint Measure Lifecycle](#) to support further study and explanation of criteria.

MUC ID	Title
MUC2024-xyz	Measure Title
Measure Steward & Developer	Proposed CMS Programs
Steward & Developer	Centers for Medicare & Medicaid Services (CMS) Program(s), such as the Hospital Readmission Reduction Program, to which the measure has been submitted

Measure Overview
<p><b>Developer-provided rationale:</b> The measure developer provides the text in this box in the CMS Measures Under Consideration (MUC) Entry/Review Information Tool (MERIT) submission. CMS MERIT is the web-based tool that developers use to provide information to CMS about their measures.</p>
<p><b>CMS-provided program rationale:</b> CMS leads provide text in this box to discuss potential program fit for the measure.</p>
<p><b>Description:</b> Developer-provided description of the measure.</p>
<p><b>Measure background:</b> Brief, standardized measure history from the MUC List.</p>
<p><b>Numerator:</b> “The <a href="#">numerator statement</a> describes the process, condition, event, or outcome that satisfies the measure’s focus or intent. The numerator statement includes parameters such as the event or events that will satisfy the numerator inclusion criteria and performance period or time interval in which the numerator event must occur, if it is different from that used for identifying the denominator.” (<a href="#">MMS Blueprint Measure Lifecycle</a>)</p>
<p><b>Exclusions:</b> “The measure developer should use numerator exclusions only in ratio and proportion measures to define elements excluded from the numerator data.”</p>
<p><b>Denominator:</b> “The <a href="#">denominator statement</a> describes the population evaluated by the individual measure. The population defined by the denominator can be the same as the target/initial population, or it can be a subset of the target/initial population to further constrain the population for the measure.”</p>

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.

<b>Measure Overview</b>	
<p><b>Exclusions:</b> “Denominator exclusions refer to criteria that result in removal of patients or cases from the denominator before calculating the numerator.”</p> <p><b>Exceptions:</b> “A denominator exception allows the measured entity to get credit when the measured entity performs the quality action, but not penalized if not done for an appropriate reason.”</p>	
<p><b>Measure type:</b> Type that applies to this specific measure. Example <a href="#">measure types</a>.</p>	<p><b>Measure has multiple scores:</b> Yes/No</p> <p><b>Measure is a composite:</b> Yes/No</p> <p><b>Measure is digital and/or an eCQM:</b> Yes/No</p> <p><b>Measure is a paired or group measure:</b> Yes/No</p>
<p><b>Level of analysis:</b> This is the level of analysis at which the measure is specified and intended for use. Examples include: facility, clinician-individual, clinician-group, health plan.</p>	<p><b>Data source(s):</b> Example <a href="#">data sources</a>: Digital-Administrative systems: Administrative Data (non-claims); Digital-Administrative systems: Claims Data</p>
<p><b>Care setting(s):</b> This is the care setting in which the measure was tested and where it is intended to be used. Examples include: hospital inpatient facility, primary care outpatient, skilled nursing facility.</p>	<p><b>Risk adjustment or stratification:</b> Yes/No, discussed more in validity section.</p>
<p><b>CBE endorsement status:</b> Endorsed, Endorsement Removed, Never Submitted</p>	<p><b>CBE endorsement history:</b> Example: Endorsed Fall Cycle 2020</p>
<p><b>Is measure currently used in CMS programs?</b> If yes, program name, with CMS Measures Inventory Tool (CMIT) link.</p>	<p><b>Measure addresses statutorily required area?</b> Yes/No. If yes, the relevant statute will be provided.</p>

## Meaningfulness

Importance	
<b>Type of evidence:</b>	<i>Peer-Reviewed Original Research, Empirical Data, Grey Literature [Source: Measures Under Consideration (MUC) Entry/Review Information Tool (MERIT) Submission Form, Evidence Attachment]</i>
<b>Importance:</b> This section includes an overview of the <a href="#">evidence</a> provided by the measure developer/steward in the MERIT submission such as clinical guidelines and published literature as well as performance data and any <a href="#">empirical evidence</a> as applicable.	
<b>Rating:</b> <span style="background-color: #d9ead3;">Met</span> , <span style="background-color: #d9ead3;">Not Met but Addressable</span> , <span style="background-color: #d9ead3;">Not Met</span> (Rating will be assigned by Battelle as appropriate)	

## Measure Performance

Example: The importance data provided (399 hospitals) corresponds to Table 4.2.5 in the Supplemental Methodology Report, which includes hospitals with at least 25 admissions. Because reliability has been calculated on hospitals with at least 25 admissions, Battelle used the mean and standard deviation from Table 4.x.y (“with ICD-10-based risk variables” columns) to estimate the importance deciles in Table 1 below including hospitals with at least 25 admissions (for simplicity, a normal distribution was assumed).

Interpretation: The mean performance on this measure for 399 entities was 74.2, with a range of performance from 43.8 to 99.9. A higher score for this measure indicates better quality care.

**Table 1. MUC2024-xyz 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)	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
Number of Entities	399	1	40	40	40	40	40	40	40	40	40	39	1

Conformance	
<b>Measure alignment with conceptual intent:</b> This section includes an overview of measure specification and discussion of alignment of the specification with measure conceptual intent.	
<b>Rating:</b> Met, Not Met but Addressable, Not Met	

Feasibility	
<b>eCQM Feasibility testing conducted:</b>	No [Source: MERIT Submission Form]
<b>Feasibility:</b> This section includes an overview of the <a href="#">feasibility of a measure</a> based on information submitted by the developer/steward in the MERIT submission. If the measure is an electronic clinical quality measure (eCQM), this section also includes analysis of the feasibility scorecard submitted.	
<b>Rating:</b> Met, Not Met but Addressable, Not Met	

Validity	
<b>Validity testing method(s):</b>	Examples: Face Validity & Empiric Validity [Sources: MERIT Submission Form, Methodology, Methodology Supplemental, Methodology Technical Expert Panel (TEP), Social Risk Factor Testing]
<b>Testing level(s):</b>	Examples: Facility, Clinician-Group, Clinician-Individual
<b>Validity:</b> This section includes overview of any <a href="#">validity testing</a> conducted by the measure developer/steward and discussion of identified or potential threats to validity.	
<b>Threats to validity:</b> If applicable, this section will also discuss appropriateness of the <a href="#">risk adjustment model or stratification</a> .	
<b>Rating:</b> Met, Not Met but Addressable, Not Met	

Reliability	
<b>Reliability testing method(s):</b>	Examples: Signal-to-noise [Sources: MERIT Submission Form, Methodology, Methodology Supplemental, Methodology TEP, Social Risk Factor Testing]
<b>Testing level:</b>	Examples: Facility, Clinician-Group, Clinician-Individual
<b>Reliability discussion:</b> This section includes overview of <a href="#">measure reliability</a> at the accountable entity (or measure score) level. If accountable entity level reliability is not available, data element level reliability will be discussed, if provided in the MERIT submission.	
<b>Additional reliability analyses:</b> When measure score-level reliability testing and performance score results are available in submission materials, Battelle analysts simulate a dataset that closely mirrors the provided median, standard deviation, and percentile information. This simulated dataset is used to generate tables showing performance scores by decile and the distribution	

Reliability
of reliability across measured entities (by reliability decile). The distribution of reliability across entities is important because denominator size (generally, a population of patients) has a substantial impact on reliability estimates for a single entity.
<b>Rating:</b> Met, Not Met but Addressable, Not Met

### Reliability Tables

Example: Table 2 shows deciles by reliability based on the information provided for the performance score (Table 4.3.6 in the Supplemental Methodology Report) and calculated reliability for the subset of 399 entities described in the testing submission. Battelle created this table to provide reviewers with a standardized format to assess reliability.

Interpretation: The mean reliability for the 100 entities measured is 0.53, below the 0.6 threshold for acceptable reliability. With approximately 30% of facilities with reliability estimates above the 0.6 threshold, 70% of these entities have a higher risk for misclassification.

**Table 2. MUC2024-xyz 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
0.53	0.33	0.102	0.14	0.20	0.27	0.32	0.42	0.56	0.64	0.71	0.76	0.86	0.959	0.44

Usability	
<b>Usability considered in application:</b>	Yes [Sources: MERIT Submission Form]
<b>Usability discussion:</b> This section includes overview of factors related to <a href="#">usability</a> of the measure within the selected CMS programs based on MERIT submission information.	
<b>Rating:</b> Met, Not Met but Addressable, Not Met	

External Validity	
<b>Was this measure tested in the same target population as the CMS program?</b>	Yes
<b>External validity discussion:</b> This section will briefly discuss whether the testing population is the same or generalizable to the program population for the selected CMS program.	
<b>Rating:</b> Met, Not Met but Addressable, Not Met	

### Appropriateness of Scale

<b>Similar or related measures in program(s):</b>	<ul style="list-style-type: none"> <li>Any related, similar, or competing measures currently in the CMS program? CMIT links will be provided for related, similar, or competing measures.</li> </ul>
<p><b>Measure appropriateness, equity, and value across target populations/measured entities:</b> This section includes a discussion of considerations related to the criterion such as: How might different populations see different benefits or burdens from use? Do focus and target population align with the selected program to make it an appropriate addition to the CMS program? Is there misalignment? Will measure have same impact in rural or low-resource settings as it will in urban or high-resource settings? Does the developer address barriers to use of measure for specific populations? Is measure more appropriate for some care settings vs others within the program?</p>	

### Time to Value Realization

<b>Plan for near- and long-term impacts after implementation:</b>	Excerpt of any submission materials addressing near-and long-term impacts of measure use.
<p><b>Measure implementation impacts over time:</b> 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 after implementation for measured entities and patients.</p> <p>Questions for the committee to consider:</p> <ul style="list-style-type: none"> <li>What are the potential near- and long-term impacts of this measure on measured entities, proposed CMS program, and patient populations?</li> <li>Will benefits and burdens associated with this measure be realized within an appropriate implementation time frame?</li> <li>Given that the measure has been in use, what rationale exists for expectations of continued measure improvement?</li> </ul>	