2024 Pre-Rulemaking Measure Review Preliminary Assessment

MUC ID	Title
MUC2024-051	Prevalent Standardized Waitlist Ratio (PSWR)
Measure Steward & Developer	**Proposed CMS Programs**
Centers for Medicare & Medicaid Services (CMS); University of Michigan	Merit-based Incentive Payment System (MIPS)–Quality

Measure Overview
Developer-provided rationale (excerpt from submission): This measure will evaluate and encourage rapid attention from dialysis practitioner groups to the optimization of health of patients to ensure early access to the waitlist, which has been demonstrated to be particularly beneficial. Given that many patients may not be ready for transplant candidacy immediately following initiation of dialysis, this measure encourages ongoing attention to transplant candidacy throughout the period following dialysis initiation.
CMS-provided program rationale: CMS may add the Prevalent Standardized Waitlist Ratio (PSWR) measure to the MIPS quality measure set as a new clinical quality measure. MIPS does not have any other related measures that address attention to transplant candidacy through the period following dialysis initiation. This measure tracks the outcomes of new placement on the kidney or kidney-pancreas transplantation waitlist or receipt of a living donor transplant following dialysis initiation, with the intended objective of improving the overall health of dialysis patients. This measure captures additional patients from the First Year Standardized Waitlist Ratio (FYSWR) measure finalized for MIPS for PY2025, which looks at patients who continue to be on the waitlist beyond the first year of dialysis initiation and supports ongoing attention/communication between the provider and waitlist candidate. This outcome measure has the potential for future inclusion in the Optimal Care for Kidney Health MIPS Value Pathway (MVP).
Description: The PSWR measure tracks the number of prevalent dialysis patients in a practitioner (inclusive of physicians and advanced practice providers) group who are under the age of 75 and were listed on the kidney or kidney-pancreas transplant waitlist or received a living donor transplant. For each practitioner group, the PSWR is calculated to compare the observed number of waitlist events in a practitioner group to its expected number of waitlist events. The PSWR uses the expected waitlist events calculated from a Cox model, adjusted for patient age, incident and prevalent comorbidities, previous waitlisting and transplant, dual eligibility, ADI, and transplant center characteristics.

Measure Overview

Measure background: Submitted previously but not included in MUC List.

Numerator: Number of prevalent dialysis patients in the practitioner group listed on the kidney or kidney-pancreas transplant waitlist or who received living donor transplants within each calendar year.

Exclusions: N/A

Denominator: The denominator for the PSWR is the expected number of waitlist or living donor transplant events in the practitioner group according to each patient's treatment history, adjusted for patient age, incident and prevalent comorbidities, previous waitlisting and transplant, dual eligibility, Area Deprivation Index (ADI), and transplant center characteristics, among patients under 75 years of age.

Exclusions: Patients with the below conditions are excluded from the measure:

- Patients were excluded when turning 75.
- Patients who were admitted to a skilled nursing facility (SNF) were excluded from that period.
- Patients were excluded if determined to be in hospice in the prior 365 days.
- Patients with dementia.

The noted exclusions represent conditions for which transplant waitlist candidacy is highly unlikely, and which can be identified readily with available data. Patients who were attributed to dialysis practitioner groups with fewer than 11 patients or 2 expected events are not excluded from the measure. If a provider cannot be matched to a TIN, patients will be grouped into a separate 'null' TIN and still included in the models but are not summarized to any valid individual TINs. All patients who meet the denominator inclusion criteria are included and used to model a given dialysis practitioner group's expected waitlist rate. If a dialysis practitioner group has fewer than 11 patients or 2 expected events, then the dialysis practitioner group is excluded from reporting outcomes.

Exceptions: N/A

Measure type: Outcome	**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-Administrative systems: Administrative Data (non-claims); Digital-Administrative systems: Claims Data
Care setting(s): Dialysis facility	**Risk adjustment or stratification:** Risk adjustment

Measure Overview	
CBE endorsement status: Not Endorsed, CBE ID 3719	**CBE endorsement history:** Submitted for endorsement Fall 2022
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: MUC Entry/Review Information Tool (MERIT) Submission Form]
Importance: Performance scores across 2,022 clinician groups (2017-2019): 0.25, 10th percentile; 0.89, 50th; 1.55, 90th; standard deviation, 0.55. The developer reports that differences in performance scores by subgroups were statistically significant and histogram plots are provided in the Measure Submission Form; committee members could request clarification of the interpretation of significant differences between subgroups. The developer presents one systematic review, which found that kidney transplantation is associated with reduced mortality, lower risk of cardiovascular events, and increased quality of life. Findings from a review of eight articles report that clinicians have substantial influence over the waitlisting process and their activities are required for Medicare certification, that regional and facility variations in waitlisting rates exist, and that optimization of patients' health for rapid waitlisting is associated with patient survival. While the developer indicates that meaningfulness to patients was not evaluated, they cite a convenience-sample study (n=409) that reported patients ranked waitlisting characteristics highest among their priorities when selecting a transplant center.	
Rating: Met	

Measure Performance

Table 1 shows performance scores by deciles (i.e., the data sorted and broken into 10 equal parts) based on the data provided in the testing submission for the 2,022 entities.

Interpretation: The mean score for the 2,022 practitioner groups described in the testing submission was 1.03. For this ratio measure, a higher score represents better quality of care.

Table 1. MUC2024-051 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)	1.03 (0.55)	0.00	0.25	0.52	0.67	0.78	0.89	1.01	1.14	1.30	1.55	2.16	6.84

	Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max
Number of Entities	2,022	1	203	203	202	202	202	202	202	202	202	202	1

Conformance
Measure alignment with conceptual intent: This measure's specification is appropriate and aligned with the measure focus (number of prevalent dialysis patients in the practitioner group listed on the kidney or kidney-pancreas transplant waitlist or who received living donor transplants within each calendar year) among eligible Medicare Fee for Service patients aged 74 and younger. Numerator and denominator populations are appropriate and exclusions align with clinical evidence.
Rating: Met

Feasibility	
eCQM Feasibility testing conducted:	No [Source: MERIT Submission Form]
Feasibility: Developers report that all data elements required are currently available from electronic sources. None of the data elements align with USCDI/USCDI+ Quality standard definitions. Aligning with USCDI standards for data elements can promote interoperability and improve feasibility. The developers reported that no changes to provider workflows are required for measure implementation.	
Rating: Met	

Validity	
Validity testing:	Empiric Validity [Source: MERIT Submission Form, MIPS Peer-Reviewed Journal Article Form]
Testing level(s):	Clinician – Group
Validity: The developer tested empiric validity at the clinician group level by comparing the measure score with transplantation rates, hypothesizing that transplant rates would be higher for entities with higher measure scores, and with mortality rates, hypothesizing that mortality would be lower for entities with higher measure scores. The measure score was divided into tertiles (worst, middle, best) and the Cochran-Armitage trend test was applied to evaluate the relationship between the tertile grouping and these practitioner group-level outcomes. Spearman correlations between the measure score and transplant and mortality rates	

Validity

were also calculated. The developers reported statistically significant moderate positive correlation (r=0.41; p<.001) and trend test (p<0.001) for the relationship of measure score with transplant.

Better PSWR performance correlated with higher transplant rate, with clear separation of transplant rates across practitioner group tertiles of performance. The direction of the relationship with mortality was as expected (modest negative correlation with better PSWR performance), with the numerically highest mortality in the lowest performance tertile of the PSWR measure, though results did not achieve statistical significance. The committee should consider if the lack of statistical significance for the lowest performance tertile of the measure is meaningful for validity of the measure within MIPS.

Threats to validity: The measure score is risk adjusted for social risk factors such as dual-eligibility and ADI, functional and clinical risk factors shown to be associated with poor outcomes following transplantation, and transplant center characteristics (transplant center waitlist mortality rate, and transplant center transplant rate), which were assigned to patients based on ZIP code. C-statistic shows acceptable model calibration, and the variables for risk adjustment are supported by the TEP and sufficient statistical testing.

The developer explained that they included social risk factors because financial resources to cover needs such as medication and transportation costs are considerations when evaluating patients for waitlisting. They developer explained that "the Kidney Disease Improving Global Outcomes [KDIGO] Clinical Practice Guideline on the Evaluation and Management of Candidates for Kidney Transplantation recommends psychosocial assessment and evaluation of adherence" [Source: MIPS Peer-Reviewed Journal Article Form].

Rating: Met

Reliability

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

Reliability discussion: The numerator and denominator for this measure are well defined. The developer calculated the reliability results from 3 years of data (January 1, 2017-December 31, 2019), which consisted of 2,022 practitioner groups with at least 11 patients and two expected events during the entire year. The developer estimated the inter-unit reliability (IUR) using a bootstrap approach to estimate the within practitioner group variation that cannot be directly estimated by analysis of variance (ANOVA). The IUR was estimated as 0.56. While this reliability is below the 0.6 threshold set for PRMR PA review to indicate that the measure is capable of differentiating entities by quality of performance, it does exceed the 0.4 required by MIPS.

Additional reliability analyses: Only a single estimate for reliability is required; therefore, interpolated decile averages of the reliability data were not generated.

Rating: Met

Usability

Usability considered in application:	No
Usability discussion: This measure has not yet been implemented. The developer did not collect input from measured entities and state that no unintended consequences are expected. This measure is not currently in use in a Medicare program.	
Rating: Not Met but Addressable	

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 a dialysis care setting with a target population of Medicare fee-for-service patients.	
Rating: Met	

Appropriateness of Scale

Similar or related measures in program(s):	The developer did not identify any related or competing measures.
Measure appropriateness, equity, and value across target populations/measured entities: The developer's review of active MIPS measures did not identify any similar or competing measures, suggesting that this measure would fill a gap within the current MIPS quality measure inventory. The focus and target population of this measure overall align with the intent and population of the program. However, the committee could discuss whether risk adjustment for factors related to financial resources supports equity. The committee should consider if this measure may have variation in benefit or burden to different populations.	

Time to Value Realization

Plan for near- and long-term impacts after implementation:	In their evidence review and validity analyses, the developer identifies impacts of this measure, including increased transplant rates, decreased morbidity and mortality, and increased quality of life for patients with transplants.
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 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?