2024 Pre-Rulemaking Measure Review
Preliminary Assessment

MUC ID	Title
MUC2024-075	Emergency Care Capacity and Quality (ECCQ)
Measure Steward & Developer	**Proposed CMS Programs**
Centers for Medicare & Medicaid Services (CMS)	Hospital Outpatient Quality Reporting Program

Measure Overview
Developer-provided rationale: This measure aims to reduce patient harm and improve outcomes for patients requiring emergency care in an ED by addressing the variation of emergency care and measuring the capacity and quality of emergency care. There are long-standing concerns about parameters that impact the quality and timeliness of care in the ED. Currently, there are no national metrics to assess the proportion of patients impacted by the quality of timely ED care.
CMS-provided program rationale: CMS is considering including this quality measure into the Hospital Outpatient Quality Reporting Program as the measure supports CMS's efforts to prevent patient harm and improved outcomes for emergency department (ED) patients by addressing the variation of emergency care and measuring the capacity and quality of emergency care in hospital outpatient departments and rural emergency hospitals. The measure captures variation in the capacity and quality of emergency care to support hospital quality improvement and improve patient outcomes. The measure also aligns with the Meaningful Measures Framework 2.0's prioritization of digital quality measurement, as well as the measurement priority areas of safety and patient-centered care. Limitations in capacity and quality of emergency care (including long wait times and ED boarding and crowding) have been shown to be associated with increases in mortality, delays in care, preventable errors, poor patient experience, and staff burnout. There are also disparities in boarding, with high-acuity black patients and patients with mental health diagnoses experiencing longer boarding times compared to white patients.
Description: This measure captures the proportion of Emergency Department (ED) visits where patients (all ages, all payers) experienced any one of four quality gaps in access: 1. The patient waited longer than 1 hour to be placed in a treatment room or dedicated treatment area that allows for audiovisual privacy during history-taking and physical examination, or 2. The patient left the ED without being evaluated by a physician/advanced practice nurse/physician's assistant, or 3. The patient boarded (time from Decision to Admit [order] to ED departure for admitted

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	
patients) in the ED for longer than 4 hours, or 4. The patient had an ED length of stay (LOS) (time from ED arrival to ED physical departure as defined by the ED depart timestamp) of longer than 8 hours.	
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: The numerator is comprised of any ED visit in the denominator with a quality gap in access; if the patient experiences any of the following during a visit, the visit is included in the numerator: 1. The patient waited longer than 1 hour to be placed in a treatment room or dedicated treatment area that allows for audiovisual privacy during history-taking and physical examination, or 2. The patient left the ED without being evaluated by a physician/advanced practice nurse/physician's assistant, or 3. The patient boarded (time from Decision to Admit (order) to ED departure for admitted patients) in the ED for longer than 4 hours, or 4. The patient had an ED length of stay (LOS) (time from ED arrival to ED physical departure as defined by the ED depart timestamp) of longer than 8 hours. Patients can have multiple visits during a performance period; each visit is eligible to contribute to the numerator and denominator. **Exclusions:** Patients who are placed in ED observation status will be included in the measure's denominator, however they will be removed from the numerator for the boarding and ED length of stay components.	
Denominator: All ED visits associated with patients of all ages, for all-payers, during the performance period. Patients can have multiple visits during a performance period; each visit is eligible to contribute to the numerator and denominator. **Exclusions:** None **Exceptions:** None	
Measure type: Intermediate Outcome	**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: Facility	**Data source(s):** Digital-Electronic Health Record (EHR) Data

Measure Overview	
Care setting(s): Emergency Department; Hospital Outpatient Department (HOPD), Rural Emergency Hospital (REH)	**Risk adjustment or stratification:** Yes
CBE endorsement status: Submitted for Fall 2024 cycle	**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 USPSTF (U.S. Preventive Services Task Force) Guidelines; Peer-Reviewed Systematic Review; Peer-Reviewed Original Research; Empirical data [Source: Measures Under Consideration (MUC) Entry/Review Information Tool (MERIT) Submission Form; eCQM-Evidence-20240510 Attachment]

Importance: There are long-standing concerns about parameters that impact the quality and timeliness of care in the ED. This measure addresses the variation of emergency care and assesses the capacity and quality of emergency care to reduce patient harm and improve outcomes for patients requiring emergency care in an ED. The developer supports the importance of this measure with a mix of systematic reviews, benchmarking data, EHR analysis, registry-based studies, and clinical guidelines.

An extensive literature review on the four components of the measure provided by the measure developer supports the evidence base for this measure and provides additional considerations for the measure's use among special population such as older patients and those seen in the ED for mental health concerns.

- Component 1: The patient waited for longer than 1 hour to be placed in a treatment space.
  - The developer highlights the increasing trend in wait times from arrival to being placed in a treatment space, with data showing a significant percentage of patients experiencing wait times over 1 hour. This delay is associated with patient harm, including increased risks of adverse events and re-visits.
- Component 2: The patient left the ED without being evaluated by a licensed clinical professional
  - The developer notes an upward trend in the percentage of patients leaving the ED without complete evaluation or treatment, which poses significant risks as many of these patients require subsequent urgent care.
- Component 3: The patient boarded (time from decision to admit order to patient departure from the ED for admitted patients) in the ED for longer than 4 hours.
  - The developer notes a lack of improvement in boarding times despite previous measures, with recent data showing an increase in median boarding times, significantly exceeding the 4-hour threshold in many cases.
- Component 4: The patient had an ED LOS (time from ED arrival to ED departure) of longer than 8 hours.
  - The developer notes a steady increase in the median ED LOS, with a significant proportion of visits exceeding 8 hours. Various studies suggest that longer ED LOS is associated with increased mortality and other adverse outcomes.

Based on the submission materials, this measure aligns with The Joint Commission's accreditation requirements (EP 6 within Standard LD.04.03.11): "The hospital should set its goals with attention to patient acuity and best practice; it is recommended that boarding time frames not exceed 4 hours in the interest of patient safety and quality of care." The developer provided evidence of a performance gap for each component of the ECCQ measure among EDs, as well as associated harms.

Importance
In an assessment of measure importance to patients, 100% of patients/caregivers consulted either strongly agreed or agreed that the measure is meaningful and produces information that is valuable in making care decisions. One patient/caregiver responded, "strongly agree" and one patient/caregiver responded "agree." Overall, this measure seems of high importance to rural patient populations and measured entities. The developer also provided evidence on how this measure may address disparities in ED experiences and outcomes among special populations, including those with behavioral health conditions, different races and ethnicities, and older patients.
Rating: Met

Measure Performance

Table 1 shows that there is a wide range of unadjusted measure scores across strata in the two datasets. For example, for Dataset A, 2 years (N=40 EDs), measure scores for the overall measure ranged from 2.91% to 55.91%, with a mean of 26.60% and a median of 30.36%; the 25th percentile was 10.36% and the 75th percentile was 39.96%. Measure score ranges are similar for the other strata but are slightly wider for the adult mental health strata and somewhat smaller for the pediatric non-mental health strata. For this proportion measure, a lower score indicates better quality of care.

Table 1. Distribution of unadjusted measure scores in Dataset A and Dataset B 2023

Measure Score	Mean (SD) (%)	Median (IQR) (%)	Range (min-max) (%)
Dataset A (2 years)			
EDs Overall (N=40)	26.60 (16.07)	30.36 (10.36-39.96)	(2.91-55.91)
EDs Entire Cohort, 2022 (N=20)	28.28 (16.63)	34.28 (10.83-39.83)	(3.52-55.91)
EDs Entire Cohort, 2023 (N=20)	24.92 (15.75)	26.30 (10.36-40.19)	(2.91-52.13)
Adult Non-Mental Health Strata (N=20)	28.02 (17.01)	32.47 (10.84-40.59)	(3.68-59.53)
Adult Mental Health Strata (N=20)	32.67 (19.85)	29.60 (14.78-45.91)	(8.52-70.80)
Pediatric Non-Mental Health Strata (N=20)	18.22 (12.50)	15.28 (8.94-27.36)	(1.61-40.73)
Pediatric Mental Health Strata (N=20)	22.90 (12.08)	20.54 (13.74-32.06)	(2.75-50.00)

Measure Score	Mean (SD) (%)	Median (IQR) (%)	Range (min-max) (%)
Dataset B 2023			
EDs Entire Cohort, 2023 (N=12)	23.87 (5.36)	24.07 (20.28-27.97)	(15.91-32.21)
Adult Non-Mental Health Strata (N=12)	23.59 (4.82)	23.54 (20.23-27.30)	(15.90-30.90)
Adult Mental Health Strata (N=12)	49.93 (10.55)	52.27 (41.35-57.57)	(34.57-66.48)
Pediatric Non-Mental Health Strata (N=12)	16.67 (10.15)	14.94 (10.04-24.37)	(2.98-34.07)
Pediatric Mental Health Strata (N=12)	52.62 (10.89)	52.19 (46.59-58.54)	(33.82-71.62)

Conformance
Measure alignment with conceptual intent: The measure specification aligns with the focus (patient experience of quality gap during ED visit across four domains) among all payers and all ages who visit the ED during the measurement period. Numerator and denominator populations are appropriate and exclusions align with available evidence and feasibility concerns.
Rating: Met

Feasibility	
eCQM feasibility testing conducted:	Yes [Sources: Bonnie Testing; Feasibility Scorecard]
Feasibility: As this measure is an electronic clinical quality measure (eCQM), the measure developers conducted Bonnie testing and submitted a feasibility scorecard. Results on this scorecard address the following domains: • 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? The feasibility assessment shows that while data elements present minimal challenges to data availability, accuracy, and workflow, there are concerns with facilities not reporting standardized terminology. The data feasibility plan outlines steps facilities can take to improve data standardization, with 27% of data elements requiring review. The committee should consider if the feasibility plan outlined in the eCQM scorecard presents a viable path forward for use of this measure within the program.	
Rating: Met	

Validity	
Validity testing:	Empiric Validity, Face Validity, and Data Element Validity [Sources: ECCQ eCQM HOQR Testing Analyses]
Testing level(s):	Facility
Validity: Face validity: The developer assessed face validity to determine if the measure effectively differentiates between good and poor quality of care among facilities. This assessment involved soliciting experts' and patients/caregivers' agreement with the following statement: "The Emergency Care Capacity and Quality eCQM for the HOQR Program could differentiate good from poor quality of care among facilities." Out of a total of 16 technical experts, 12 agreed that the measure could effectively differentiate between good and poor quality of care. The remaining four disagreed, citing concerns that factors influencing boarding times and ED length of stay might be beyond the control of the facilities, thus questioning the measure's ability to accurately reflect quality of care. Empiric validity testing: The developer used construct validity, which is the extent to which the measure accurately assesses what it is intended to assess. This analysis involved 32 hospital-based ED facilities from two datasets. Using the Pearson's correlation coefficient, the developer examined the association between measure score performance and broadly available and validated hospital quality measures (see similar and related measures in Appropriateness of Scale section). Correlations of the measure with Overall Hospital Quality Star Rating were -0.56 (dataset A) and -0.55 (dataset B), indicating a moderate correlation. The results supported this hypothesis, indicating that hospitals with higher Star Ratings also tended to score well on the new ECCQ eCQM between the measure scores and the components of the Star Ratings components.	
Threats to validity: The developer considered threats to validity and developed the recommendation to stratify this measure by age and principal diagnosis of a mental health condition. Mental health diagnoses are identified using an established code set of International Classification of Diseases (ICD)-10 and Systematized Nomenclature of Medicine (SNOMED) codes that identify "psychiatric and mental health diagnoses" but do not include diagnosis for substance abuse disorder. The measure's outcome may also be stratified (pending additional testing) by race and ethnicity, primary language, and insurance status to best address equity of emergency care.	
Rating: Met	

Reliability	
Reliability testing method(s):	Signal-to-Noise [Sources: ECCQ eCQM HOQR Testing Analyses]
Testing level:	Facility
Reliability discussion: The numerator and denominator for this measure are well defined. The developer calculated the reliability results from a combined dataset with Dataset A consisting of 20 hospital-based ED facilities and Dataset B consisting of 12, for a total of 32 hospital-based ED facilities from 2023 (January 1-December 31). The median reliability is 0.9999, and the minimum	

Reliability
reliability is 0.9997. Of the entities in the testing set, 100% have a reliability >0.6, suggesting that this measure is 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.
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 32 hospital-based ED facilities. Battelle created this table to provide reviewers with a standardized format to assess reliability.

Interpretation: Of the entities in the testing set, 100% have a reliability >0.6, suggesting that this measure is capable of differentiating entities by quality of performance.

Table 2. MUC2024-075 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
0.9999	0.0001	0.9997	0.9997	0.9998	0.9999	0.9999	0.9999	0.9999	1.000	1.000	1.000	1.000	1.0000	0.0001

Usability	
Usability considered in application:	Yes, the submission materials documented and explored measure usability in the selected program and settings.
Usability discussion: The developer considered usability of the measure in the submission, stating that “public comment was sought on the measure concept and specifications, receiving 677 comments in total. 630 were from individuals including patients, providers, and caregivers; 15 hospitals or healthcare organizations; 13 professional associations; 3 private companies; 2 non-profit organizations; 1 committee; and 13 internal stakeholder members. Overwhelmingly, the public supports this measure's importance, reaffirming its meaning for providers, patients, and all consumers of healthcare. The information this intends to measure is valuable in making care decisions and improving the quality of healthcare for the public. More than 95% of commenters supported the measure’s description and goals.” Based on the discussion of the measure in the MUC List submission documents, patients, caregivers, providers, and health care organizations provided strong supportive input. They emphasized the measure’s relevance and value in improving care decision-making. The widespread support reinforces the measures utility across various settings and patient populations. However,	

Usability
potential unintended consequences, including premature discharge from the ED, gaming, inappropriate reduction in inpatient admission, increase in staff burnout, and worse disparities of care, need to be monitored. Despite these unintended consequences, the developer suggests that the benefit of implementing the measure outweighs the potential scope and magnitude of these concerns.
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 on EDs in a range of health systems and geographic locations, which cover the same target population as the Hospital Outpatient Quality Reporting Program (patients receiving numerous types of health services such as ED services).	
Rating: Met	

Appropriateness of Scale

Similar or related measures in program(s):	• 00427-01-C-HOQR Median Time from ED Arrival to ED Departure for Discharged ED Patients (OP-18) • 00410-01-C-HOQR Left Without Being Seen (OP-22)
Measure appropriateness, equity, and value across target populations/measured entities: The developer notes that the four numerator components and the use of mean performance score (vs. median) will provide improved visibility into overall performance variation and performance gap. In reviewing related measures currently in use, the developer reported that this measure adds value because (1) it captures transfer boarding, a metric currently not captured by any CMS measure, (2) it is an eCQM, which reduces reporting burden compared to the current measures, and (3) it combines four ED capacity components into one measure to help minimize "gaming." Regarding equity of this measure's performance and benefit across populations, the developer's extensive literature review and analysis do provide relevant information to assess the potential for differential benefit or harm to specific subgroups identified in the literature review, including those seeking mental health care at the ED, older patients, and racial or ethnic groups that routinely experience disparities in ED care. The committee should consider if, based on their professional and patient experience, there is a chance for variation on 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 briefly mentions potential outcomes for their measure on patient populations, there is a need for further examination of near- and long-term impacts of this measure across 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, the Hospital Outpatient Quality Reporting 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 the Hospital Outpatient Quality Reporting Program measure set?	