Partnership for Quality Measurement
Powered by Battelle

Fall 2023 Endorsement and Maintenance (E&M) Committee Independent Review Summary

MANAGEMENT OF ACUTE EVENTS AND CHRONIC CONDITIONS COMMITTEE

Prepared by:

Battelle
505 King Avenue
Columbus, Ohio 43201
January 2024

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.

image by rawpixel.com

Table of Contents

Summary of Committee Independent Reviews

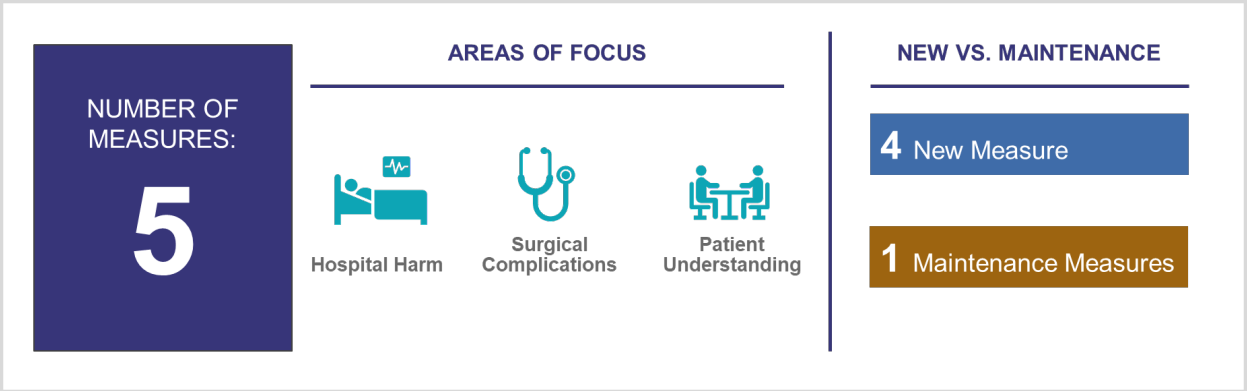
Independent E&M Committee Member Reviews Overview

At least three (3) weeks prior to an E&M committee endorsement meeting, the Recommendations Group and the Advisory Group of each E&M committee receive the full measure submission details for each measure up for review, including all attachments, the Partnership for Quality Measurement (PQM) Measure Evaluation Rubric, the public comments received for the measure(s) under review, and the E&M team preliminary assessments.

Members of both groups were asked to review each measure, independently, against the PQM Measure Evaluation Rubric. Committee members assigned a rating of "Met," "Not Met but Addressable," or "Not Met" for each domain of the PQM Measure Evaluation Rubric. In addition, committee members provided associated rationale for each domain rating, which is based on the rating criteria listed for each domain. Battelle staff aggregated and summarized the results and distributed them back to the committee, and to the respective measure developers and/or stewards, for review within one (1) week of the endorsement meeting.

These independent committee member ratings are compiled and used by Battelle facilitators and committee co-chairs to guide committee discussions.

Figure 1. Fall 2023 Measures for Committee Review



For the Fall 2023 cycle, the Management of Acute Events and Chronic Conditions committee received five (5) measures, four (4) new measures and one (1) measure undergoing maintenance endorsement review (Figure 1). The measures focused on hospital harm, surgical complications, and patient understanding.

Measure-Specific Summaries

The following brief summaries include themes and considerations gathered from the committee's independent reviews for each of the five domains of the PQM Measure Evaluation Rubric. Themes were assessed and categorized with respect to the strengths and limitations of the measure(s) under endorsement review. Corresponding to the themes are the number of committee reviews received and stratified by the ratings of "Met," "Not Met," and "Not Met, but Adressable."

CBE #4210 – Patient Understanding of Key Information Related to Recovery After a Facility-Based Outpatient Procedure or Surgery, Patient Reported Outcome-Based Performance Measure

Number of Committee Reviews: 22

Importance (n=22)	Strengths	Limitations
Consensus 86% Met 9% Not Met, but Addressable 5% Not Met	• Logic model clearly connects measure performance (patient understanding of instructions) with health outcomes such as ED visits, readmissions, etc. • Committee members' own clinical experience reflects the importance of providing effective patient education to reduce poor outcomes • Measure adds value to existing portfolio and aligns with CMS's National Quality Strategy	• Literature review could have been more robust; it does not clearly connect improved patient understanding with improved outcomes following outpatient surgery.

Feasibility (n=22)	Strengths	Limitations
Consensus 82% Met 5% Not Met, but Addressable	• Multi-modal survey administration is an advantage • Fewer survey items reduce burden on patients	• Measure would be more effective if administered before discharge, enabling facilities to address concerns (mailed surveys would not be ideal for timing)

Feasibility (n=22)	Strengths	Limitations
14% Not Met		• A text survey mode might get better response rate than email • Burden on facilities may be significant; e.g., Medicaid agencies may struggle to collect data and calculate the measure • Testing may not have been adequate to reveal true burden on facilities • Lack of clarity in how the measure is scored from survey items

Scientific Acceptability (n=22)	Strengths	Limitations
Consensus 9% Met 82% Not Met, but Addressable 9% Not Met	• None	• Testing sample was too small to report reliability by decile • More clarity is needed in how the denominator is defined/calculated, specifically regarding removal of incomplete surveys and the use of 'two midnights' to define episodes • Face validity was assessed by TEP; additional information regarding the panel composition and voting outcomes is desired

Scientific Acceptability (n=22)	Strengths	Limitations
		• Validity test comparing measure to similar OAS CAHPS measure had limited sample size and non-significant results • Not risk-adjusted

Equity (n=22)	Strengths	Limitations
Consensus 14% Met 9% Not Met, but Addressable 77% Not Met	• None	• This optional domain not addressed by the developer • Unclear whether survey was validated in diverse populations, e.g., among Spanish speakers • Translation quality can impact patient's understanding of discharge instructions, if not English speaking • Equity should be a focus of this measure given differences in how patient education is received as well as differences in ability to respond to the survey (e.g., literacy, language)

Use and Usability (n=22)	Strengths	Limitations
Consensus 91% Met	• Planned for use in HOQR	• Measure may not be sufficiently generalizable to patient populations to be ready for public reporting

Use and Usability (n=22)	Strengths	Limitations
0% Not Met, but Addressable 9% Not Met	• Support for initial rollout in HOQR, though not all support the measure for public reporting	• Unclear how data would be furnished to providers to help improve care

CBE #4130e – Hospital Harm- Postoperative Respiratory Failure

Number of Committee Reviews: 18

Importance (n=18)	Strengths	Limitations
No Consensus 28% Met 67% Not Met, but Addressable 6% Not Met	• The definition of post-operative respiratory failure (PRF) used in the measure has been widely adopted • Literature review and logic model support the business case for the measure, even if lacking consensus on PRF definition • PRF is a common, serious post-operative complication and a performance gap exists	• Lack of a consensus on the definition of PRF • Available interventions are limited and have mixed evidence

Feasibility (n=18)	Strengths	Limitations
Consensus 89% Met 6% Not Met, but Addressable 6% Not Met	• Nearly all data elements are captured in structured fields • No significant barriers were identified	• Documentation of mechanical ventilation is not standardized • Some EHRs may require adjustment, which can be expensive and burdensome

Scientific Acceptability (n=18)	Strengths	Limitations
Consensus 100% Met 0% Not Met, but Addressable 0% Not Met	• Approaches for testing reliability, validity and risk adjustment models are clearly described and appropriate	• None

Equity (n=18)	Strengths	Limitations
Consensus 100% Met 0% Not Met, but Addressable 0% Not Met	• Developers used extensive evidence base to evaluate disparities and design the risk adjustment model	• None

Use and Usability (n=18)	Strengths	Limitations
Consensus 89% Met	• Planned for use in IQR • Agreement among patients and entities that measure is useful for decision-making	• Concern about the measure's use in payment models and evidence cited showing that reduction of PRF is challenged by unclear evidence regarding risk factors and application of interventions

Use and Usability (n=18)	Strengths	Limitations
11% Not Met, but Addressable 0% Not Met		

CBE #4120e – Hospital Harm- Falls with Injury

Number of Committee Reviews: 19

Importance (n=19)	Strengths	Limitations
Consensus 95% Met 5% Not Met, but Addressable 0% Not Met	• Ample, clear evidence is provided showing that falls in inpatient settings are common, preventable, and costly, and harm increases with age • Interventions are known to improve patient safety and reduce costs • A substantial performance gap has been demonstrated • The measure is supported by patients	• Measure Information Form's importance section should be updated to document screening and mitigation strategies described elsewhere in the submission

Feasibility (n=19)	Strengths	Limitations
Consensus 95% Met 5% Not Met, but Addressable 0% Not Met	• All data elements exist in structured fields in the EHR systems tested (Epic, Allscripts) • Only one hospital required a workflow change to collect all data elements	• Feasibility testing did not include hospitals using other significant EHR systems (Cerner, Meditech) • Information about falls may reside in text fields

Scientific Acceptability (n=19)	Strengths	Limitations
Consensus 89% Met 11% Not Met, but Addressable 0% Not Met	• Reliability analysis approach is appropriate, and findings meet the threshold for reliability (>0.6) except for the two hospitals with the smallest samples • The measure has strong face validity and numerator and denominator elements have excellent positive predictive value (>98%)	• Sample used for reliability testing is small (12 hospitals) • Sample is not representative; it includes only large teaching hospitals; smaller, rural, and non-teaching hospitals are not included

Equity (n=19)	Strengths	Limitations
Consensus 95% Met 5% Not Met, but Addressable 0% Not Met	• A social disparities analysis evaluated differences by race/ethnicity and insurance status to determine appropriate approach to stratification	• Sample included only urban hospitals, so evaluation of disparities may be limited

Use and Usability (n=19)	Strengths	Limitations
Consensus 95% Met	• Measure is planned for public reporting in IQR • There exist evidence-based protocols for falls prevention hospitals can use to improve	• Public comments called attention to a potential unintended consequence of decreased patient mobilization

Use and Usability (n=19)	Strengths	Limitations
5% Not Met, but Addressable 0% Not Met		

CBE #4125 – Thirty-day Risk-Standardized Death Rate among Surgical Inpatients with Complications (Failure-to-Rescue)

Number of Committee Reviews: 18

Importance (n=18)	Strengths	Limitations
Consensus 94% Met 6% Not Met, but Addressable 0% Not Met	• Strong evidence base showing multiple, facility-level interventions (e.g., nursing staffing ratios, nursing education, residency training programs) that can influence patient outcomes • Failure to rescue is a key gap area for clinical quality measurement	• Some of the evidence related to staffing is more than 30 years old

Feasibility (n=18)	Strengths	Limitations
Consensus 100% Met 0% Not Met, but Addressable 0%# Not Met	• Claims-based measure with no burden on facilities or providers • Measure is modeled after Patient Safety Indicator 04 (PSI 04), a highly feasible measure	• None

Scientific Acceptability (n=18)	Strengths	Limitations
Consensus 11% Met 83% Not Met, but Addressable 6% Not Met	• The measure has strong face validity (90% agreement among the TEP) • Convergent validity for the measure is stronger than for PSI 04	• About half of facilities have reliability below the threshold (0.6) • Higher reliability may be difficult to achieve given limitations of claims data (most underlying clinical factors are not available), and raising the minimum case threshold or extending the reporting time frame may not achieve the desired outcome

Equity (n=18)	Strengths	Limitations
Consensus 100% Met 0% Not Met, but Addressable 0%# Not Met	• Potential disparities associated with race/ethnicity, age and sex were evaluated and reported; as none were found when adjusting for clinical risk factors, these social risk factors were not included in the risk adjustment model	• None

Use and Usability (n=18)	Strengths	Limitations
Consensus 94% Met	• Measure planned for use in public reporting • The literature review cites a range of facility-level interventions hospitals can use to improve performance	• Concerns were expressed about public reporting before the reliability limitations are addressed

Use and Usability (n=18)	Strengths	Limitations
0% Not Met, but Addressable 6% Not Met		• The measure might not be useful for facilities without a sizable Medicare population, and should be reworked to include all payers

CBE #0694 – Hospital Risk-Standardized Complication Rate Following Implantation of Implantable Cardioverter-Defibrillator (ICD)

Number of Committee Reviews: 19

Importance (n=19)	Strengths	Limitations
No Consensus 11% Met 16% Not Met, but Addressable 74% Not Met	• This remains a meaningful quality of care measure, although there are barriers to demonstrating importance	• Many of the literature sources cited are older and may be outdated • Performance data have not been updated in more than 15 years and continuing existence of a gap cannot be evaluated

Feasibility (n=19)	Strengths	Limitations
Consensus 79% Met 11% Not Met, but Addressable 11% Not Met	• All data elements are available in defined fields in claims and the registry	• Not all hospitals report registry data and the extent of data missingness is not provided

Scientific Acceptability (n=19)	Strengths	Limitations
Consensus 0% Met 5% Not Met, but Addressable 95% Not Met	• Good data element validity – agreement between chart and other sources	• Data used to test reliability is more than 10 years old • Available reliability testing results show low reliability (split-half reliability ICC = 0.1494)

Equity (n=19)	Strengths	Limitations
Consensus 0% Met 0% Not Met, but Addressable 100% Not Met	• This optional criterion was not addressed	• This optional criterion was not addressed

Use and Usability (n=19)	Strengths	Limitations
Consensus 0% Met	• None	• Measure is not currently in use • No recent performance data so improvement on the measure cannot be evaluated

Use and Usability (n=19)	Strengths	Limitations
0% Not Met, but Addressable 100% Not Met		


Partnership for
Quality Measurement
Powered by Battelle