Partnership for Quality Measurement
Powered by Battelle

Fall 2023 Initial Recognition and Management Committee Endorsement Meeting

February 9, 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 (CMS).

Partnership for Quality Measurement
Powered by Battelle

Welcome

Meeting Objectives

The purpose of today's meeting is to:

- Review and discuss candidate measures submitted to the Initial Recognition and Management committee for the Fall 2023 cycle;
- Review public comments received for the submitted candidate measures; and
- Render endorsement decisions for the submitted candidate measures.

Housekeeping Reminders for Recommendations Group*

- The system will allow you to mute/unmute yourself and turn your video on/off throughout the event
- Please raise your hand and unmute yourself when called on
- Please lower your hand and mute yourself following your question/comment
- Please state your first and last name if you are a Call-In User
- We encourage you to keep your video on throughout the event
- Feel free to use the chat feature to communicate with Battelle staff
- If you are experiencing technical issues, please contact the project team via chat on the virtual platform or at PQMsupport@battelle.org.

*Advisory Group members are asked to refrain from using the chat and the raise hand feature, as Advisory Group members will be listening to the Recommendations Group discussions and will cast their vote once discussions cease.

Meeting Ground Rules

- Be prepared, having reviewed the meeting materials beforehand
- Respect all voices
- Remain engaged and actively participate
- Base your evaluation and recommendations on the measure evaluation rubric
- Keep your comments concise and focused
- Be respectful and allow others to contribute
- Share your experiences
- Learn from others

Project Team

- Nicole Brennan, MPH, DrPH, Executive Director
- Brenna Rabel, MPH, Deputy Director
- Jeff Geppert, Measure Science Team Lead
- Quintella Bester, PMP, Senior Program Manager
- Matthew Pickering, PharmD, Principal Quality Measure Scientist
- Beth Jackson, Social Scientist IV
- Amanda Overholt, MPH, Social Scientist III
- Stephanie Peak, Social Scientist III
- Isaac Sakyi, MSGH, Social Scientist III
- Lydia Stewart-Artz, PhD, Social Scientist III
- Jessica Ortiz, MA, Social Scientist II
- Kelsey Conner, Social Scientist I
- Olivia Giles, MPH, Social Scientist I
- Elena Hughes, MS, Social Scientist I
- Sarah Rahman, Social Scientist I
- Alex Valdez-Alvarez, Social Scientist I

Agenda

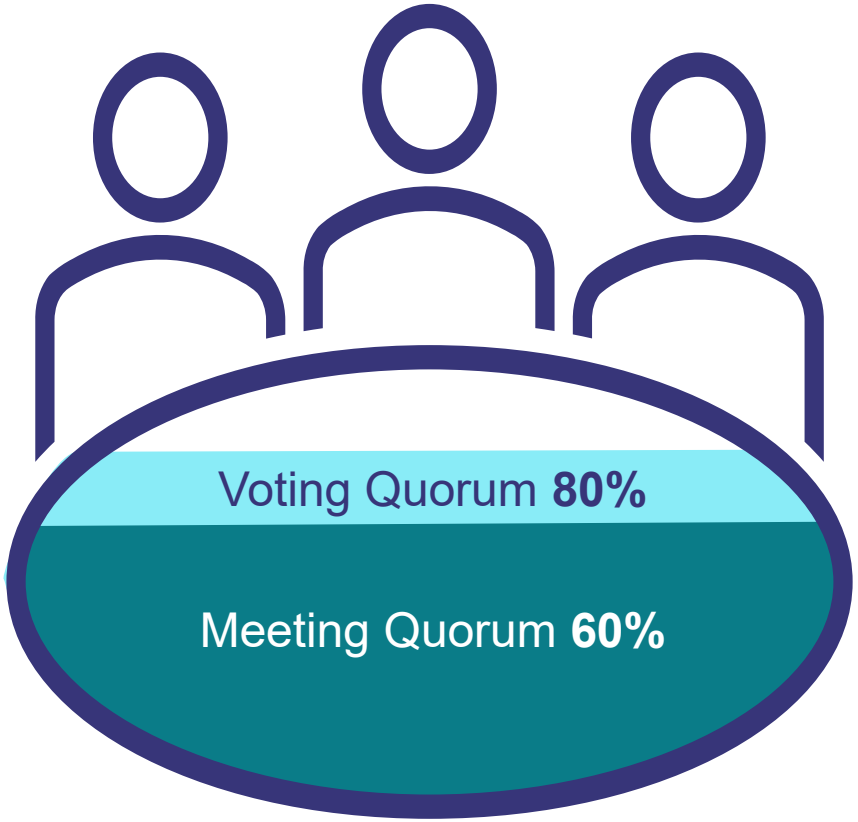
- Welcome and Review of Meeting Objectives
- Roll Call with Disclosures of Interest
- Overview of Evaluation Procedures and Measures for Endorsement Consideration
- Test Vote
- Evaluation of Candidate Measures
- Additional Measure Recommendations Discussion (if time permits)
- Opportunity for Public Comment
- Next Steps
- Adjourn

Roll Call with Disclosures of Interest

Quorum

- Meeting quorum requires that 60% of the Recommendations Group members are present during roll call at the beginning of the meeting.

- Endorsement decisions are rendered via a vote after Recommendations Group discussions. Voting quorum is at least 80% of active committee members (Recommendations Group + Advisory Group), who are not recused.



Initial Recognition and Management Fall 2023 Cycle Committee – *Recommendations Group*

- Matt Austin, PhD (***Non-Patient Co-Chair***)
- Patricia Merryweather-Arges, MA (***Patient Co-Chair***)
- Carol Sakala, PhD, MSPH
- Cecilia Purcell
- Danny Barker, MBA, RRT
- Edward Bailly, MSHCDL, MSN, FNP-BC
- Geeta Sood, MD, ScM
- Helen Haskell, MA
- Jennifer Bailit, MD, MPH
- Karen Fernandes, RN, CPHQ
- Karen Johnson, PhD
- Karen Wilding, BS
- Kyle Campbell, PharmD
- Pranali Trivedi, CPHQ
- Sherly Binu, MBA, MS, RN
- Tamaire Ojeda, MHSA, RDN, LD

**Denotes committee member is under Inactive status for the current cycle.*

Initial Recognition and Management Fall 2023 Cycle Committee – *Advisory Group*

- Abraham Jacob, MD, MHA
- Anne Llewellyn, MS, BHSA, RN, CMGT-BC, CRRN, BCPA, CMF
- Arjun Venkatesh, MD, MBA, MHS
- Ashley Comiskey, MSN, RN, CCDS
- Barbara Kivowitz, MA, MSW
- Billy Caceres, PhD, RN
- Carole Hemmelgarn, MS, MS
- Gregary Bosci, DO
- Hannah Ingber, MPH
- Janet Hurley, MD, FAAFP
- Janice Young, DNP, RN, HRM, CPHQ, CPPS
- Jean-Luc Tilly, MPA, PMP
- Jill Blaxier. MSN-ED, RN, CPHQ
- Juliet Bartsch, RN
- Kent Bream , MD
- Kobi Ajayi, PhD, MPH, MBA
- Kory Anderson, MD, FACP, CHCQM
- Lisa Leckrone, MHA, CPHQ, ASCP
- Marianne Kraemer, RN, MPA, M. Ed., CENP, CCRN-K-emeritus
- Mark Ellison, BA
- Oren Guttman, MD, MBA
- Raymond Dantes, MD, MPH
- Selena McCord, MPH
- Sheila Owens-Collins, MD, MPH, MBA
- Talia Sasson, MD, FSIR
- Tammy Love, MSN, RN-BC, CPPS, LSSGB
- Thomas Spiegel, MD, MBA, MS, FACEP
- Tracy Brasel, RN, MBA-HM
- Usha Venugopal, MD FACP, CPHQ
- Zainab Jah. MPH

**Denotes committee member is under Inactive status for the current cycle.*

Overview of Evaluation Procedures

Roles of the Committee During the Endorsement Meeting

- **Evaluate** each measure against each domain of the Partnership for Quality Measurement Measure Evaluation Rubric
- **Indicate** the extent to which each criterion is met and the rationale for the rating
- **Review** comments submitted during the public comment period
- **Render endorsement decisions** for candidate measures

Roles of the Committee Co-Chairs During the Endorsement Meeting

	Collaborate with Battelle
• **Co-facilitate** virtual endorsement meetings, along with Battelle staff	●
• **Participate on the committee** as a full voting member for the entirety of your term	
• **Serve on the Appeals committee** ▪ Includes attending the half- to full-day virtual Appeals committee meeting at the end of every E&M cycle (contingent upon whether an appeal is received)	
• Work with Battelle staff to **achieve the goals** of the project	●
• Assist Battelle staff in **anticipating questions and identifying additional information** that may be useful to the committee	●

Roles of the Committee Co-Chairs During the Endorsement Meeting, *continued 1*

Patient Representative Co-Chair

Ensure the patient community voice is considered

Non-Patient Representative Co-Chair

Ensure the Advisory group voice is considered

Evaluation and Voting Process

Non-consensus Measures

Step	Description	Interested Party
1	**Introduction of the measure in which consensus was lacking** • Presentation of the PQM Rubric domain rating results from the committee independent assessments and a summary of the committee's independent review, noting both strengths and limitations, and any potential conditions, as appropriate. • Summation of any public comments received prior to the endorsement meeting.	Battelle Staff
2	**Floor is open for any additional public comments with respect to the measure under review** • Commenters are kindly asked to keep their comments to two (2) minutes or less. • The committee does not respond directly to commenters, rather comments are shared for the committee's endorsement discussion.	Battelle Staff and Co-chairs
3	**Three-to-five (3-5) minute, high-level overview of the measure** • Presenters will kindly be asked to stop presenting if the time is over five (5) minutes. • Please refrain from using slides or screensharing of materials. • Overview may include initial Reponses to committee independent reviews and/or public comments	Developer and/or Steward

Evaluation and Voting Process

Non-consensus Measures, Continued 1

Step	Description	Interested Party
4	**Round-robin for clarifying questions** • Non-patient representative co-chair to confirm whether questions from A-group members (via independent assessments) have been considered. • Patient representative co-chair to confirm whether the patient partner questions have been considered. • After all questions have been collected, the developer/steward addresses measure-specific questions.	R-group discusses A-group listens Battelle Staff to facilitate with Co-chairs Developer and/or Steward
5	**Committee discussion of the measure elements in which consensus was lacking** • Facilitated discussion measure strengths and limitations based on PQM Measure Evaluation Rubric domain. • Determine potential resolutions that lead to committee consensus and any recommendations placed on the measure for the developer/steward to consider in the future. • The developer/steward may respond to questions posed by the committee. • Subject matter experts (SMEs) are called upon, accordingly, to address committee questions and to provide context and relevance about the measure for to the committee's consideration.	R-group discusses A-group listens Battelle Staff to facilitate with Co-chairs Developer and/or Steward SMEs

R-group: Recommendations group; A-group: Advisory group

Evaluation and Voting Process

Non-consensus Measures, Continued 2

Step	Description	Interested Party
6	**Responses to committee discussion** • After the committee discussion has concluded, prior to voting, the developer/steward is given a final opportunity to respond to the committee's discussion before the committee moves to a vote on endorsement. • Please try to keep responses brief, referring to information in the measure submission, as appropriate. • Please refrain from using slides or screensharing of materials.	Developer and/or Steward
7	**Committee vote** • Any conditions or recommendations are summarized prior to voting. • If consensus is not reached, based on the 75% threshold, the measure is not endorsed.	R-group and A-group Battelle Staff and Co-chairs summarize voting conditions

R-group: Recommendations group; A-group: Advisory group

Evaluation and Voting Process

Conditions for Voting Example

Step	Description	Interested Party
7	**Committee vote** • Any conditions or recommendations are summarized prior to voting. • If consensus is not reached, based on the 75% threshold, the measure is not endorsed.	R-group and A-group Battelle Staff and Co-chairs summarize voting conditions

Example: Some committee members raised concern with the measure testing occurring in only two or three U.S. states and recommended to see additional testing across are larger, more generalizable population, then:

- A vote to **Endorse** the measure means the committee agrees that the evidence provided to support the measure fully substantiates the measure claims.
- A vote to **Endorse with Conditions**, means the committee agrees that the evidence provided to support the measure doesn't fully substantiate the measure claims due to limited testing within 2-3 states. Therefore, the committee votes to endorse the measure with the condition that additional testing across a larger, more generalizable population be conducted by the next maintenance review.
- A vote to **Not Endorse/have Endorsement Removed**, means the committee agrees that the evidence provided to support the measure does not substantiate the claims for scientific acceptability due to the limited testing in only 2-3 U.S. states. Therefore, the committee raised concern with respect to the generalizability of the testing results. In addition, there are no reasonable changes to the measure (e.g., specifications, testing, evidence) that would allow the measure to receive conditional endorsement.

Evaluation and Voting Process

Consensus Measures

Step	Description	Interested Party
1	**Introduction of the measure in which consensus was reached** • Presentation of the PQM Rubric domain rating results from the committee independent assessments and a summary of the committee's independent review, noting both strengths and limitations, and any potential conditions, as appropriate. • Summation of any public comments received prior to the endorsement meeting.	Battelle Staff
2	**Floor is open for any additional public comments with respect to the measure under review** • Commenters are kindly asked to keep their comments to two (2) minutes or less. • The committee does not respond directly to commenters, rather comments are shared for the committee's endorsement discussion.	Battelle Staff and Co-chairs
3a	**Committee discussion of measures with consensus to endorse** • Confirm the measure strengths outweigh any limitations identified • Confirm if any conditions for endorsement • Co-chairs confirm the Advisory Group and the patient community voice have been considered (via independent assessments)	R-group discusses A-group listens Battelle Staff to facilitate with Co-chairs

Evaluation and Voting Process

Consensus Measures, Continued 1

Step	Description	Interested Party
3b	**Committee discussion of measures with <u>consensus to not endorse/remove endorsement</u>** • Confirm the measure limitations outweigh the strengths • Identify potential recommendations for the developer to improve the limitations • Co-chairs confirm the Advisory Group and the patient community voice have been considered (via independent assessments) • After the committee discussion, the developer/steward is given the opportunity to respond to the committee's review and discussion.	R-group discusses A-group listens Battelle Staff to facilitate with Co-chairs Developer and/or Steward
4	**Committee vote** • Any conditions or recommendations are summarized prior to voting. • If consensus is not reached, based on the 75% threshold, the measure is not endorsed.	R-group and A-group Battelle Staff and Co-chairs summarize voting conditions

Endorsement Decision Outcomes

Decision Outcome	Description	Maintenance Expectations
Endorsed	**Applies to new and maintenance measures.** There is 75% or greater agreement for endorsement by the E&M committee	Measures undergo maintenance of endorsement reviews every 5 years with an annual update review at 3 years.
Endorsed with Conditions	**Applies to new and maintenance measures**. There is 75% or greater agreement that the measure can be endorsed as it meets the criteria, but there are recommendations/areas committee reviewers would like to see when the measure comes back for maintenance. If these recommendations are not addressed, then a rationale from the developer/steward should be provided for consideration by the E&M committee review.	Measures undergo maintenance of endorsement reviews every 5 years with an annual update at 3 years, unless the condition requires the measure to be reviewed earlier. The E&M committee evaluates whether conditions have been met, in addition to all other maintenance endorsement minimum requirements.
Not Endorsed	**Applies to new measures only.** There is 75% or greater agreement to not endorse the measure by the E&M committee.	None
Endorsement Removed	**Applies to maintenance measures only.** Either: • There is 75% or greater agreement for endorsement removal by the E&M committee; or • A measure steward retires a measure (i.e., no longer pursues endorsement); or • A measure steward never submits a measure for maintenance and there is no response from the steward after targeted outreach; or • There is no longer a meaningful gap in care, or the measure has plateaued (i.e., no significant change in measure results for accountable entities over time)	None

Decision Outcomes:
Endorsed with Conditions

The types of conditions that may be placed on a measure include:

- Conducting/providing additional testing across a larger population, accountable entity-level, and/or different level of analysis
- Expanding the measure use beyond quality improvement and into an accountability application
- Providing implementation guidance or a near-term path forward for implementing the measure; providing clear system requirements for implementation of the measure

Battelle has identified several non-negotiable areas, meaning if a measure meets one or more of the following criteria, the measure cannot be endorsed, even with conditions:

- Lack of or unclear business case
- Lack of evidence supporting the business case
- Significantly poor feasibility for the measure to be implemented due to challenges, e.g., data availability or missingness
- Inappropriate methodology, calculations, formulas, or testing approach used to demonstrate reliability or validity
- Specifications, testing approach, results, or data descriptions are insufficient
- If a measure with an “Endorsed with Conditions” designation is evaluated for maintenance, but it has not met the prior conditions

What is the PQM Measure Evaluation Rubric?

The PQM Measure Evaluation Rubric (Rubric) consists of five (5) major domains:

1. **Importance** - Extent to which the measure is evidence-based AND is important for making significant gains in health care quality or cost where there is variation in or overall, less-than-optimal performance.

2. **Feasibility** - Extent to which the measure specifications (i.e., numerator, denominator, exclusions) require data that are readily available OR could be captured without undue burden AND can be implemented for performance measurement.

3. **Scientific Acceptability [i.e., Reliability and Validity]** - Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented.

4. **Equity (optional)** - Extent to which the measure can identify differences in care for certain patient populations, which can be used to advance health equity and reduce disparities in care.

5. **Use and Usability** - Extent to which potential audiences (e.g., consumers, purchasers, providers, and policymakers) are using or could use measure results for both accountability and performance improvement to achieve the goal of high quality, efficient health care for individuals or populations.

Consensus Voting for Final Determinations

Endorse (A)	Endorse with Conditions (B)	Do Not Endorse (C)	Consensus Voting Status
75% or More	0%	Less than 25%	A
75% or More		Less than 25%	B
Less than 25%		**75% or More**	C
26% to 74%		26% to 74%	No consensus

If no consensus is reached, based on the 75% threshold, the measure is not endorsed.

Overview of Fall 2023 Measures for Endorsement Consideration

Fall 2023 Measures for Committee Review

Three measures were submitted to the Initial Recognition and Management committee for endorsement consideration.

NUMBER OF MEASURES:

3

AREAS OF FOCUS

- **Breast Cancer Screening**
- **Radiation Screening in Stroke Patients**
- **Waveform Capnography**

NEW VS. MAINTENANCE

- **2** New Measures
- **1** Maintenance Measure

Fall 2023 Measures for Committee Review

CBE ID	Title	Importance (n)	Feasibility (n)	Scientific Acceptability (n)	Equity (n)	Use & Usability (n)
CBE #0661	Head CT or MRI Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Patients who Received Head CT or MRI Scan Interpretation within 45 minutes if ED Arrival	**Consensus (21)** 86% Met 14% Not Met, but Addressable 0% Not Met	**Consensus (21)** 86% Met 10% Not Met, but Addressable 5% Not Met	**No Consensus (21)** 29% Met 71% Not Met, but Addressable; 0% Not Met	**Consensus (21)** 10% Met 10% Not Met, but Addressable 81% Not Met	**No Consensus (21)** 24% Met 71% Not Met, but Addressable 5% Not Met
CBE #4220	Breast Cancer Screening Recall Rates	**Consensus (24)** 17% Met 83% Not Met, but Addressable 0% Not Met	**Consensus (24)** 92% Met 8% Not Met, but Addressable 0% Not Met	**No Consensus (24)** 71% Met 29% Not Met, but Addressable 0% Not Met	**Consensus (24)** 92% Met 4% Not Met, but Addressable 4% Not Met	**No Consensus (24)** 21% Met 67% Not Met, but Addressable 13% Not Met
CBE #4045	Waveform Capnography in Ventilated Patients: percent of patient transport contacts with advanced airways in whom continuous waveform capnography was used	**No Consensus (21)** 29% Met 71% Not Met, but Addressable 0% Not Met	**No Consensus (21)** 29% Met 67% Not Met, but Addressable 5% Not Met	**Consensus (21)** 5% Met 95% Not Met, but Addressable 0% Not Met	**No Consensus (21)** 14% Met 14% Not Met, but Addressable 71% Not Met	**No Consensus (21)** 62% Met 33% Not Met, but Addressable 5% Not Met

Legend:
C – Consensus; NC – No consensus; n – number of committee independent reviews

Test Vote

Consideration of Candidate Measures

CBE #4220 – Breast Cancer Screening Recall Rates

Item	Description
Measure Description	• The Breast Cancer Screening Recall Rates measure calculates the percentage of beneficiaries with mammography or digital breast tomosynthesis (DBT) screening studies that are followed by a diagnostic mammography, DBT, ultrasound, or magnetic resonance imaging (MRI) of the breast in an outpatient or office setting within 45 days.
Developer/Steward	• The Lewin Group/Centers for Medicare & Medicaid Services
New or Maintenance	• New
Current or Planned Use	• Public Reporting • Quality Improvement

Measure Type	Target Population(s)	Care Setting	Level of Analysis
Process	Medicare beneficiaries	Hospital: Outpatient	Facility

CBE #4220
Public Comments

No comments received

CBE #4220 – Breast Cancer Screening Recall Rates, continued 1

Importance - Extent to which the measure is evidence-based AND is important for making significant gains in health care quality or cost where there is variation in or overall, less-than-optimal performance.

Importance (n=24)	Strengths	Limitations
Consensus 17% Met 83% Not Met, but Addressable 0% Not Met	• **Importance of the Measure:** The measure is deemed important for preventative and diagnostic mammograms. It's seen as a balancing measure to tracking rates of screening alone. • The information provided supports the importance of the measure. It's particularly important as it focuses on monitoring "rates of recall following screening imaging instead of rates of breast cancer imaging." • This is a new measure under PQM, but it is a re-specification measure for the "Mammography Follow Up Rates" under OQR. • **Benefits of Screening and Follow-up:** There is strong evidence presented about the benefits of screening for breast cancer and follow-up evaluations when needed. It is further presented that there are negative consequences when mammography and digital breast tomosynthesis (DBT) recall rate is either too high or too low. • The developer cites evidence and guidelines from the American College of Radiology.	• **Recall Rate Range:** Questions are raised about the rationale for the 5%-12% range. There's a suggestion to consider if current conditions warrant a shift in the range or having different ranges for special circumstances. • **Patient Engagement:** There's a concern about limited patient engagement on the importance of the measure. • **Improved Outcomes:** Unclear how this measure will support improved outcomes. Connection between this measure and reduction of radiation is unclear. • The current measure may not adequately address the need for appropriate clinical follow-up without overuse. The impact of adding DBT on the ACR screening target recall rate wasn't addressed. • **Quality Improvement Strategies:** The measure does not address how the screening recall rate can be used to improve performance. • **Performance Gaps:** While gaps are reported in deciles with 4.9% and 13.0% as the closest deciles, this does not allow precise assessment of the gap from the 5-12% suggested. In addition, the high recall rate category (over 13%) has a performance level mean of 13.7% which demonstrates a measurable but not clearly meaningful performance gap in over utilization based on consensus guidelines.

CBE #4220 – Breast Cancer Screening Recall Rates, continued 2

Feasibility - Extent to which the measure specifications (i.e., numerator, denominator, exclusions) require data that are readily available OR could be captured without undue burden AND can be implemented for performance measurement.

Feasibility (n=24)	Strengths	Limitations
Consensus 92% Met 8% Not Met, but Addressable 0% Not Met	• **Feasibility and Data Collection:** The measure is feasible and does not present an undue burden on hospitals for data collection. • The data required is routinely generated from patient encounters, claims, and the EHR utilizing value sets. The measure is already being used in the CMS Outpatient Imaging Efficiency program. • The burden of reporting this measure was directly addressed by the developer, and no proprietary data is needed.	• **Feasibility and Data Collection:** There were some concerns about the representativeness of the sample used to evaluate feasibility. The geographic characteristics of the individuals were not reported, which could be important as professionals in rural areas may report an undue burden with collecting this data. • **Patient Representation:** Developer evaluated feasibility amongst 32 individuals. However, the patient representation relative to health care staff/professionals could undermine the feasibility assessment result. It would be appropriate for the developers to consider a patient-only group to ensure accurate representation.

CBE #4220 – Breast Cancer Screening Recall Rates, continued 3

Scientific Acceptability [i.e., Reliability and Validity] - Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented.

Scientific Acceptability (n=24)	Strengths	Limitations
No Consensus 71% Met 29% Not Met, but Addressable 0% Not Met	• **Reliability:** Most facilities exceeded the accepted threshold of 0.6. The median reliability score was 0.95. • **Data Adequacy:** The data used for testing were adequate, with good representation between facilities/patients. The measure uses the same data process as Outpatient Imaging Efficiency measures using claims data. • **Measure Design:** The measure specifications are well-defined and precise. The measure as outlined was well designed. • **Validity:** The developer facilitated qualitative assessments of the measure's validity. Face validity results are acceptable. Validity testing was completed and consensus was reached.	• **Outliers:** The developers did not describe the reliability of 0.41, which is a steep difference compared to the mean of 0.9 and the next lowest value of 0.81. • **Risk Adjustment:** The measure is not risk-adjusted as it is a process measure. This could impact recall rates, potentially leading to unnecessary exposure to radiation or follow-up testing in low-risk populations. • If there is a high risk of cancer in the population then the appropriate recall rate may be higher and even outside the suggested 5-12% range. Similarly, in a low-risk population, not adjusting the pass rate for the measure would increase the exposure to radiation or follow up testing (biopsy) unnecessarily. • **Validity Methods:** The validity methods rely on a consensus of 32 individuals, but it's unclear if they represent broad stakeholders in breast cancer care. • **Inclusions/Exclusions:** There are concerns about including men in the denominator, since they are typically screened only if they have significant risk factors or symptoms. There are concerns about the rationale for removing MRI as a follow-up imaging modality.

CBE #4220 – Breast Cancer Screening Recall Rates, continued 4

Equity (optional) - Extent to which the measure can identify differences in care for certain patient populations, which can be used to advance health equity and reduce disparities in care.

Equity (n=24)	Strengths	Limitations
Consensus 92% Met 4% Not Met, but Addressable 4% Not Met	• **Equity Assessment:** The developers conducted a thorough assessment of equity for sex, race/ethnicity, age, and dual eligibility status. They used performance data to calculate the rate of recall by these factors and found overall significance.	• **Limited Scope:** There are concerns that the Medicare FFS measure will not capture many of the patients in populations at risk for receiving inequitable recall rates. It would have been beneficial to see the ethnic/race data based on rural and urban settings and size of facilities. • **Clinical Significance:** While statistical testing indicates a statistically significant difference in recall rates, there is no indication if this is clinically significant, appropriate based on risk, or equitable/inequitable. • **Addressing Differences:** The measure developer can establish a difference in the measure across different patient groups. However, it does not clearly establish how the measure supports addressing these differences.

CBE #4220 – Breast Cancer Screening Recall Rates, continued 5

Use and Usability - Extent to which potential audiences (e.g., consumers, purchasers, providers, and policymakers) are using or could use measure results for both accountability and performance improvement to achieve the goal of high quality, efficient health care for individuals or populations.

Use and Usability (n=24)	Strengths	Limitations
No Consensus 21% Met 67% Not Met, but Addressable 13% Not Met	• **Measure Performance and Improvement:** The measure is seen as a useful tool for internal quality measurement and can help identify opportunities for improvement. • Most of the multi-stakeholder group agreed the measure could be used by entities for QI and decision-making (77.4%) and that it would provide consumers and providers with actionable information (80.6%). • **Patient Care and Outcomes**: The measure could potentially increase use and usability, which is important for patients who have abnormal mammograms and often wait long for further testing. • **Pay for Performance Programs:** Developer plans for the measure to be used in CMS's Hospital Outpatient Quality Reporting (HOQR) Program, a pay-for-quality program.	• **Measure Performance and Improvement**: The developer did not address how an organization could take actions to improve performance if rates fell outside of the 5% and 12% parameters. • There's uncertainty about what follow-up action medical facilities should take when their scores are higher than expected. • **Patient Care and Outcomes**: There's concern about the interpretation of results for facilities with too low or too high rates. It's unclear whether patients will use this data. • **Pay for Performance Programs**: The measure is seen as inappropriate for pay for performance programs due to the potential for unintended consequences. The use of a "range" in pay for performance could incentivize inappropriate actions. • **Variability and Differences**: There's concern that detection of variability or clusters due to geographic, cultural, or socioeconomic factors could be discouraged or misinterpreted. Concerned about facilities in underserved areas.

CBE #661 – Head CT or MRI Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Patients who Received Head CT or MRI Scan Interpretation within 45 minutes of ED Arrival

Item	Description
Measure Description	• This measure calculates the percentage of acute ischemic stroke or hemorrhagic stroke patients who arrive at the emergency department (ED) within two hours of the onset of symptoms and have a head computed tomography (CT) or magnetic resonance imaging (MRI) scan interpreted within 45 minutes of ED arrival. The measure is calculated using chart abstracted data, on a rolling, quarterly basis and is publicly reported, in aggregate, for one calendar year. The measure has been publicly reported, annually, by CMS as a component of its Hospital Outpatient Quality Reporting (OQR) Program since 2012.
Developer/Steward	• Centers for Medicare & Medicaid Services/ Lantana Group
New or Maintenance	• Maintenance
Current or Planned Use	• Payment Program • Public Reporting • Regulatory and Accreditation Programs • Quality Improvement with Benchmarking

Measure Type	Target Population(s)	Care Setting	Level of Analysis
Process	Acute ischemic stroke or hemorrhagic stroke patients	Emergency Department Hospital: Outpatient	Facility

CBE #661
Public Comments

No comments received

CBE #661 – Head CT or MRI Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Patients who Received Head CT or MRI Scan Interpretation within 45 minutes of ED Arrival, continued 1

Importance - Extent to which the measure is evidence-based AND is important for making significant gains in health care quality or cost where there is variation in or overall, less-than-optimal performance.

Importance (n=21)	Strengths	Limitations
Consensus 86% Met 14% Not Met, but Addressable 0% Not Met	• Strong evidence for this measure for the care of suspected stroke patients in the ED with multiple studies and clinical guidelines showing the importance of timely imaging and intervention in acute ischemic, supporting the measure. • Comments emphasize the importance of early stroke identification and treatment, with questions about potential barriers to rapid access to scanning. • Patients surveyed find the measure valuable.	• Measure performance has stagnated, with no improvement seen in the last five years. • The developer presents the existing quality measure including measure characteristics and specifications. While this has a high level of detail, it does not outline the importance of this measure. • The description of patient input does not support the conclusion that time-to-interpretation is meaningful for patients. • What is the rationale for the 45-minute time limit for the CT scan or MRI, as it is unable to find it in the literature cited

CBE #661 – Head CT or MRI Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Patients who Received Head CT or MRI Scan Interpretation within 45 minutes of ED Arrival, continued 2

Feasibility - Extent to which the measure specifications (i.e., numerator, denominator, exclusions) require data that are readily available OR could be captured without undue burden AND can be implemented for performance measurement.

Feasibility (n=21)	Strengths	Limitations
Consensus 86% Met 10% Not Met, but Addressable 5% Not Met	• Data for the measure are generated during care and uses data from EHRs or other electronic sources. • The measure is currently being implemented and has been in use for a long time, demonstrating feasibility. • The measure appears to be feasible as data elements are collected in the normal course of care.	• The measure requires chart abstraction to report as specified, which may be a significant fee or cost. • If a provider/facility does not have software to do the abstraction (which could be expensive), they will need manual extraction, which can also be expensive and time consuming.

CBE #661 – Head CT or MRI Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Patients who Received Head CT or MRI Scan Interpretation within 45 minutes of ED Arrival, continued 3

Scientific Acceptability [i.e., Reliability and Validity] - Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented.

Scientific Acceptability (n=21)	Strengths	Limitations
No Consensus 43% Met 57% Not Met, but Addressable 0% Not Met	• The measure is well-defined, and the data are stable without improvement. • The specifications are clear, and the reliability results are fair overall. • The measure is suitable and accurately specified. • Data element validity is clearly established. The validity testing results were reassuring that validity is adequate and specific threats to validity weren't identified.	• There are concerns about the reliability testing, with approximately 30-35% of entities having reliability less than 0.6. • Several suggestions for improvement are made, including increasing the minimum case volume, extending the timeframe for the measure, and considering a mitigation strategy for facilities with a low denominator. • There are concerns about the hypothesis testing confirming a difference between before performance for male vs female patients. • Some would have preferred a different approach to empirical validity of the measure score. • There's a suggestion to know more about any issues with data element "Head CT/MRI Scan Interpretation Time".

CBE #661 – Head CT or MRI Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Patients who Received Head CT or MRI Scan Interpretation within 45 minutes of ED Arrival, continued 4

Equity (optional) - Extent to which the measure can identify differences in care for certain patient populations, which can be used to advance health equity and reduce disparities in care.

Equity (n=21)	Strengths	Limitations
Consensus 10% Met 10% Not Met, but Addressable 81% Not Met	• None	• Developer did not address this optional criterion. • Comments note the importance of studying and addressing differences in race, ethnicity, and gender.

CBE #661 – Head CT or MRI Scan Results for Acute Ischemic Stroke or Hemorrhagic Stroke Patients who Received Head CT or MRI Scan Interpretation within 45 minutes of ED Arrival, continued 5

Use and Usability - Extent to which potential audiences (e.g., consumers, purchasers, providers, and policymakers) are using or could use measure results for both accountability and performance improvement to achieve the goal of high quality, efficient health care for individuals or populations.

Use and Usability (n=21)	Strengths	Limitations
No Consensus 24% Met 71% Not Met, but Addressable 5% Not Met	• The measure is in use in the Outpatient Quality Reporting Program. • The usability of this measure is well established by the availability for reporting after several years. • Feedback has been solicited on the measure. The measure has a feedback mechanism with no unintended consequences identified.	• Performance scores continue to show room for improvement but have remained largely stable from 2015-2021. It would be interesting to know how many sites have increased over the 6 years or remain at a stable level or have declined. • The reasons for the lack of improvement are not clearly articulated. There's a need for the measure steward to address barriers to improvement. • Question of the measure's utility for low-volume providers and whether the resources invested in it are positively impacting quality. • There has been no substantial feedback or indications of unexpected findings. The only recommended intervention is training providers, but there's a need for potential Quality Improvement (QI) mechanisms such as providing performance reports to providers.

Lunch

CBE #4045– Waveform Capnography in Ventilated Patients: Percent of patient transport contacts with advanced airways in whom continuous waveform capnography was used

Item	Description
Measure Description	• This metric is designed to measure the critical care transport team's utilization of waveform capnography during critical care medical transport. Waveform capnography has evolved as the standard for the safe placement and maintenance of advanced airways (e.g., endotracheal tubes) in adult and pediatric patients. The metric specifically focuses on transported patients with advanced airways in whom continuous waveform capnography is appropriately used. This metric is stratified by age into the following three categories: neonatal (defined as infants <29 days), pediatric (defined as patients aged 29 days to <18 years), and adults (defined as age 18 or older). This metric is reported as “Percent of patients with advanced airways in whom waveform capnography was utilized." Transport programs track this metric for each applicable transport and report their average utilization percentage monthly.
Developer/Steward	• GAMUT Quality Improvement Collaborative
New or Maintenance	• New
Current or Planned Use	• Public Reporting; Quality Improvement; Quality Improvement with Benchmarking

Measure Type	Target Population(s)	Care Setting	Level of Analysis
Process	Any person, regardless of age with an advanced airway	Emergency Medical Service/ Ambulance	Facility

CBE #4045
Public Comments

No comments received

CBE #4045– Waveform Capnography in Ventilated Patients: Percent of patient transport contacts with advanced airways in whom continuous waveform capnography was used, continued 1

Importance - Extent to which the measure is evidence-based AND is important for making significant gains in health care quality or cost where there is variation in or overall, less-than-optimal performance.

Importance (n=21)	Strengths	Limitations
No Consensus 29% Met 71% Not Met, but Addressable 0% Not Met	• **Waveform Capnography:** Standard of care for safe airway positioning and beneficial for patients. • Support for using it in transport includes empirical studies and three consensus statements. • The developer demonstrated the importance of using waveform capnography to recognize dislodgment of airways and to prevent serious complications to patients.	• **Unclear Specifications:** Unclear numerator qualifications • **Lack of/Limited Evidence:** There is no systematic review of waveform capnography in transport. • There is not a clear relationship in the logic model or the evidence for importance for how “higher score” translates into better quality. • **Performance Gap:** Measure performance since 2014 ranges from 89.2% to 95.6% (87.5-94.1% among pediatric patients and 94-97.8% among adults) and may have limited room for improvement for some groups. • **Meaningfulness to Patients:** Meaningfulness to patients has not been established.

CBE #4045– Waveform Capnography in Ventilated Patients: Percent of patient transport contacts with advanced airways in whom continuous waveform capnography was used, continued 2

Feasibility - Extent to which the measure specifications (i.e., numerator, denominator, exclusions) require data that are readily available OR could be captured without undue burden AND can be implemented for performance measurement.

Feasibility (n=21)	Strengths	Limitations
No Consensus 29% Met 67% Not Met, but Addressable 5% Not Met	• **Data Availability and Collection:** The measure is already in routine use by a substantial number of GAMUT users, which is compelling evidence that data is available and able to be captured. • Required data are routinely generated and used during care and are available in EHRs or other electronic sources. • **Feasibility:** Feasibility is demonstrated through normative evaluations from the consensus developers as well as the quality evaluators.	• **Data Availability and Collection:** The most significant challenge is data capture. Data for this measure are not routinely generated from electronic sources but must be manually abstracted by hand from the EHR. • There's also a concern about manual abstraction and not using the available EHR or mobile capabilities to interface/upload to be able to pull the data necessary for this reporting measure. • **Feasibility:** There's no assessment if the tools for capnography are easily available to transport service providers and if they are sufficiently easy to implement. The developers did not describe the process and how feasibility was done. It's unclear why the sites that have implemented this measure have not been able to develop documentation in the EHR or use mobile capabilities to have the data available to limit manual abstraction.

CBE #4045– Waveform Capnography in Ventilated Patients: Percent of patient transport contacts with advanced airways in whom continuous waveform capnography was used, continued 3

Scientific Acceptability [i.e., Reliability and Validity] - Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented.

Scientific Acceptability (n=21)	Strengths	Limitations
Consensus 10% Met 86% Not Met, but Addressable 5% Not Met	• **Measure Definition:** The measure is well-defined and precisely specified. Details for measure calculation were clear and thorough. • **Reliability:** The reliability of the measure appears satisfactory. The results suggest the measure is reliable with 100% agreement for one of the data elements and a Kappa value of 0.79 indicating substantial agreement for the other data element. • **Helpful for Patients and Providers:** The measure appears to be valid and helpful for patients and healthcare providers.	• **Limited Testing:** The reliability was only tested at three sites. More facilities could be included in reliability testing. • Not all data elements required for measure calculation were tested. • **Denominator Exclusions:** There are concerns that the denominator exclusions are for the database manager to decide, which lacks standardization. • **Appropriate Use of Capnography:** The interrater reliability should look at whether the reported data and the audited data were in agreement, not whether the use of capnography was “appropriate”. • **Validity:** There are concerns about the interpretation of the survey results for face-value validity, the details on face validity assessment, and the total number of randomly selected participants. • There are also concerns about bias as face validity was done with current participants of the program, not an independent group of experts. • It also does not discuss the reasons behind a high percentage of respondents do not agree.

CBE #4045– Waveform Capnography in Ventilated Patients: Percent of patient transport contacts with advanced airways in whom continuous waveform capnography was used, continued 4

Equity (optional) - Extent to which the measure can identify differences in care for certain patient populations, which can be used to advance health equity and reduce disparities in care.

Equity (n=21)	Strengths	Limitations
No Consensus 14% Met 14% Not Met, but Addressable 71% Not Met	• **Inclusivity:** The measure includes all patients who are transported with ventilators. • **Standard of Care:** There should be no variation in equity since this is a standard of care.	• **Equity Considerations:** Lack of equity considerations in the measure. It suggests that demographic data could be used to assess inequities.

CBE #4045– Waveform Capnography in Ventilated Patients: Percent of patient transport contacts with advanced airways in whom continuous waveform capnography was used, continued 5

Use and Usability - Extent to which potential audiences (e.g., consumers, purchasers, providers, and policymakers) are using or could use measure results for both accountability and performance improvement to achieve the goal of high quality, efficient health care for individuals or populations.

Use and Usability (n=21)	Strengths	Limitations
No Consensus 62% Met 33% Not Met, but Addressable 5% Not Met	• **Proven Usability:** The measure is already in use in many organizations, proving its usability. • **Standard of Care:** The measure is a standard of care in all settings. • **Existing Database**: The measure is currently in use within the GAMUT database since 2014 for internal QI, and QI with external benchmarking.	• **Lack of Patient and Family Input:** The absence of involvement of patient and family input might be leaving degrees of uncertainty that are not included in the approach towards understanding all the actors related to understanding, improving, and maintaining measure improvements. • **Lack of Details:** There is no clear description of planned uses within usability. In other sections, usability and use seem to be focused on consensus statements. • **Limited Information:** This measure has been in use for several years. However, there is very little information regarding the measure being used in the past.

Additional Measure Recommendations Discussion

Based on the measure discussions today, are there additional recommendations or solutions the developer can use to overcome any potential measure limitations?

Opportunity for Public Comment

Next Steps

Next Steps for Fall 2023

Meeting Summary

- Meeting summary will be posted to the E&M committee project page by February 26, 2024.

Appeals Period

- **Appeals Period:** February 26 – March 18
- Appeals committee will meet on March 27, 2024 to review eligible appeals. Please refer to the E&M Guidebook for more information about the appeals process.

Technical Report

- At the conclusion of the appeals period, a final technical report will be posted to the E&M Committee project page in April 2024.

Partnership for Quality Measurement
Powered by Battelle

Thank You!

Have questions? Contact us at
PQMsupport@battelle.org

Partnership for Quality Measurement

Powered by Battelle