Partnership for Quality Measurement
Powered by Battelle

National Consensus Development and Strategic Planning for Health Care Quality Measurement

Spring 2024 Cycle Endorsement and Maintenance (E&M) Meeting Discussion Guide

COST AND EFFICIENCY COMMITTEE

August 1, 2024

Prepared by:
Battelle
505 King Avenue, Columbus, Ohio 43201

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.

Table of Contents

List of Tables

List of Figures

Overview of Spring 2024 Measures for Review

For this measure review cycle, six measures were submitted to the Cost and Efficiency committee for endorsement consideration (Table 1). The measures focused on hospital visits after ambulatory surgical center (ASC) or hospital outpatient department procedures, hospitalizations for ambulatory-sensitive conditions, and readmission rates (Figure 1).

Table 1. Overview of Measures Under Endorsement Review

CBE Number	Measure Title	New/Maintenance	Developer/Steward
2539	Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy	Maintenance	Yale Center for Outcomes Research and Evaluation (Yale CORE)/Centers for Medicare & Medicaid Services (CMS)
3357	Facility-Level 7-Day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers	Maintenance	Yale CORE/CMS
3366	Hospital Visits After Urology Ambulatory Surgical Center Procedures	Maintenance	Yale CORE/CMS
3470	Hospital Visits after Orthopedic Ambulatory Surgical Center Procedures	Maintenance	Yale CORE/CMS
3495	Hospital-Wide 30-Day, All-Cause, Unplanned Readmission Rate (HWR) for the Merit-Based Incentive Payment System (MIPS) Eligible Clinician Groups	Maintenance	Yale CORE/CMS
4490	Hospitalizations for Ambulatory Care Sensitive Conditions among Home and Community Based Service (HCBS) Participants	New	The Lewin Group

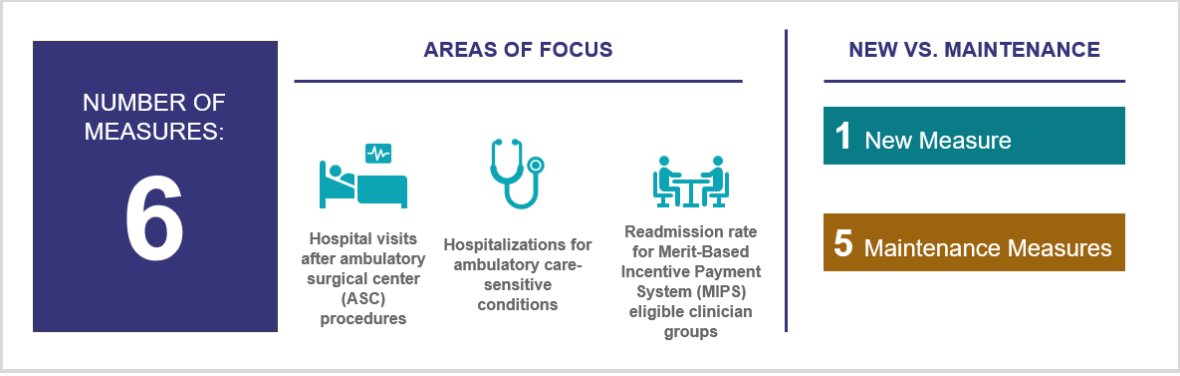

Figure 1. Spring 2024 Measures for Committee Review

Public Comment

Battelle accepts comments on measures under endorsement review through the Partnership for Quality Measurement (PQM) website and Public Comment Listening Sessions. For this evaluation cycle, the public comment period opened on May 16, 2024, and closed on June 14, 2024, and the Public Comment Listening Session was held on May 29, 2024.

Battelle received 27 public comments prior to the endorsement meeting. CBE #3357, #3366, and #2539 received 19 of the 27 comments, each having one to two supportive comments, one comment asking for clarification on patient age, and two comments expressing opposition for endorsement due to concerns with the risk adjustment model, reliability testing, and evidence of a business case. CBE #3495 received five comments: one supportive, one also questioning why accountable care organizations were not tested in the measure, and three comments opposing continued endorsement, criticizing the measure's risk adjustment, reliability testing, and inclusion criteria. Lastly, CBE #4490 received three comments, two of which were supportive of the measure and one asking for clarification on the measure title.

After the public comment period closed, developers/stewards had the opportunity to submit written responses to the public comments received. Summaries of the public comments and developer/steward responses are provided within the respective measure evaluation summaries of this discussion guide below.

Advisory Group Feedback

The Advisory Group was convened on June 6, 2024. Ten of 16 (60%) active Advisory Group members were in attendance to share feedback and ask questions regarding the measures under endorsement review. Developers/stewards of the respective measures were also in attendance and provided responses to the Advisory Group discussions. After the meeting, developers/stewards had the opportunity to submit additional written responses to Advisory Group member feedback and questions.

Summaries of the Advisory Group member discussions and developer/steward responses are provided within the respective measure evaluation summaries of this discussion guide below.

To support the review of the public comments and Advisory Group summaries, the number of comments or individuals that shared similar comments, feedback, and/or questions is represented as "a few" (2-3 individuals), "several" (4-6 individuals), and "many" (more than 6 individuals).

Measures Under Endorsement Review

CBE #2539: Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy [Yale CORE/CMS]

Measure Description: Facility-level risk-standardized rate of acute, unplanned hospital visits within 7 days of a colonoscopy procedure performed at a hospital outpatient department (HOPD) or ambulatory surgical center (ASC) among Medicare Fee-For-Service (FFS) patients aged 65 years and older. An unplanned hospital visit is defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission. The measure is calculated separately for ASCs and HOPDs.

Measure Status	
New or Maintenance: Maintenance Measure	**Used in An Accountability Application?** Yes - Public Reporting; Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
CBE Endorsement Status: Endorsed **Last Endorsement Review Cycle:** Spring 2020	**Proposed/Planned Use:** Hospital Outpatient Quality Reporting (Hospital OQR) Program, CMS; Ambulatory Surgical Center Quality Reporting (ASCQR) Program, CMS; Rural Emergency Hospital Quality Reporting (REHQR) Program, CMS

Measure Characteristics

Measure Type	Target Population(s)	Level of Analysis	Care Setting(s)
Outcome	FFS patients aged 65 years and older	Facility	Ambulatory Surgery Center; Hospital; Outpatient

Measure Overview
Rationale: The Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy measure (CBE #2539) (hereafter “Colonoscopy measure”) captures unplanned hospital visits 7 days after a colonoscopy procedure performed at an HOPD, or separately, at an ASC. The measure focuses on the outcome of unplanned hospital visits because this is a broad, patient-centered outcome that captures the full range of hospital visits resulting from adverse events or poor care coordination following the procedure. By providing HOPDs and ASCs with detailed information about patients who have an unplanned hospital visit, this measure supports quality improvement at facilities, and through public reporting of the measure allows for assessment and illumination of the variation in risk-adjusted hospital visits following colonoscopy.
Numerator: The outcome for this measure is all-cause, unplanned hospital visits within seven days of a qualifying outpatient colonoscopy. The measure defines a hospital visit as any emergency department (ED) visit, observation stay, or unplanned inpatient admission.

Measure Overview
Denominator: The target population for this measure includes low-risk colonoscopies performed in the outpatient setting (HOPD or ASC) for Medicare FFS patients aged 65 years and older. The measure is calculated separately for procedures performed at HOPDs vs ASCs and reported stratified by dual eligibility for HOPDs.
Exclusions: Colonoscopy Measure Exclusion Criteria: 1) Colonoscopies for patients who lack continuous enrollment in Medicare FFS Parts A and B in the 7 days after the procedure. Rationale: We exclude these patients to ensure full data availability for outcome assessment. 2) Colonoscopies that occur concurrently with high-risk upper gastrointestinal (GI) endoscopy procedures. Rationale: Patients undergoing concurrent high-risk upper GI endoscopy procedures, such as upper GI endoscopy procedures for the control of bleeding or treatment of esophageal varices, have a higher risk profile than typical colonoscopy patients. Therefore, these patients have a disproportionally higher risk for the outcome. 3) Colonoscopies for patients with a history of inflammatory bowel disease (IBD) or diagnosis of IBD at time of index colonoscopy or on the subsequent hospital visit outcome claim. Rationale: We exclude these patients because: • IBD is a chronic condition; patients with IBD undergo colonoscopy both for surveillance due to increased cancer risk and for evaluation of acute symptoms. IBD is likely to be coded as the primary diagnosis prompting the procedure irrespective of whether the patients are undergoing a screening procedure or a diagnostic procedure in the setting of an acute exacerbation of IBD. Therefore, we may not be able to adequately risk-adjust for these patients, as we cannot identify relatively well versus acutely unwell patients among visits coded as IBD. • Our aim is to capture hospital visits which reflect the quality of care. Admissions for acutely ill IBD patients who are evaluated with an outpatient colonoscopy and are subsequently admitted for medical treatment of an IBD flare do not reflect the quality of the colonoscopy. During measure development (see the 2014 Facility 7-day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy Measure Technical Report) we found that more than one-third of IBD patients admitted to the hospital with colonoscopy had a discharge diagnosis of IBD, indicating their admission was for medical treatment of their IBD. We therefore excluded this group so that providers who treat a disproportionate number of IBD patients will not be disadvantaged in the measure. • A post-index diagnosis of IBD, which represents a very small fraction of cases (less than 0.5% of the cohort) in the measure population, indicates that the condition was likely present at the time of the index colonoscopy but not coded. 4) Colonoscopies for patients with a history of diverticulitis or diagnosis of diverticulitis at time of index colonoscopy or on the subsequent hospital visit outcome claim. Rationale: We exclude these patients because: • It is unclear what the health status is of patients coded with a history or current diagnosis of diverticulitis, making it difficult to fully risk adjust for patients' health. Colonoscopies performed on patients with a history or current diagnosis of diverticulitis are likely to be coded as diverticulitis as the primary diagnosis irrespective of whether the patients are undergoing a screening procedure or a diagnostic procedure (i.e., are acutely unwell with active disease). Furthermore, the codes for diverticulitis and diverticulosis may not be consistently used; patients with diverticulosis may be erroneously coded as diverticulitis. Therefore, we may not be able to adequately risk adjust as we cannot identify relatively well versus acutely unwell patients among visits coded as diverticulitis. • Admissions for acutely ill patients with a history or current diagnosis of diverticulitis who are evaluated with an outpatient colonoscopy and are subsequently admitted for medical treatment of do not reflect the quality of the colonoscopy. During measure development (see the Facility 7-day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy Measure Technical Report) we found that more than one-quarter of patients with a history or current diagnosis of diverticulitis admitted to the hospital post colonoscopy had a discharge diagnosis of diverticulitis, indicating they were admitted for medical treatment of the condition. These admissions are likely unrelated to the quality of the colonoscopy. We

Measure Overview

therefore excluded this group so that providers who treat a disproportionate number of diverticulitis patients will not be disadvantaged in the measure.
• A post-index diagnosis of diverticulitis, which represents a very small fraction of cases (less than 0.5% of the cohort) in the measure population, indicates that the condition was likely present at the time of the index colonoscopy but not coded.
5) Colonoscopies followed by a subsequent outpatient colonoscopy procedure within 7 days.
Rationale: In these situations, the two colonoscopies are considered part of a single episode of care, for which the subsequent colonoscopy is considered the index procedure.

In addition, for colonoscopies performed at HOPDs, we exclude:
6) Colonoscopies that occur on the same day and at the same hospital as an emergency department (ED) visit that is billed on a different claim than the index colonoscopy, unless the ED visit has a diagnosis indicative of a complication of care.
Rationale: It is unclear whether the colonoscopy or ED visit occurred first. If the ED visit is coded with a diagnosis indicative of a complication of care, the measure assumes the ED visit occurred after the colonoscopy procedure and is counted in the measure. It is unlikely that a patient would experience an ED visit for an acute diagnosis at 1 facility and then travel to another facility for a routine colonoscopy on the same day. Accordingly, ED visits billed on the same day as a colonoscopy but at a different facility are included because they likely represent a routine procedure followed by a complication of care.

7) Colonoscopies that are billed on the same hospital claim as an ED visit and that occur on the same calendar day, unless the ED visit has a diagnosis indicative of a complication of care.
Rationale: In these situations, it is not possible to use claims data to determine whether the colonoscopy was the cause of, subsequent to, or during the ED visit. However, if the ED visit is coded with a diagnosis for a complication, the assumption is that it occurred after the colonoscopy procedure.
8) Colonoscopies that are billed on the same hospital outpatient claim and that occur after the ED visit.
Rationale: In these situations, we assume that the colonoscopy was subsequent to the ED visit and may not represent a routine colonoscopy procedure. Timing of the ED visits is determined using revenue center dates from the outpatient claim.
9) Colonoscopies that are billed on the same hospital outpatient claim as an observation stay.
Rationale: In these situations, it is not possible to use claims data to determine whether the colonoscopy was the cause of, subsequent to, or during the observation stay.

<u>1.15c Denominator Exclusions Details</u>
1) Colonoscopies for patients who lack continuous enrollment in Medicare FFS Parts A and B in the 7 days after the procedure. Lack of continuous enrollment in Medicare FFS for 7 days after the procedure is determined by patient enrollment status in FFS Parts A and B using the Medicare Enrollment Database. The enrollment indicators must be appropriately marked for the month(s) which fall within 7 days of the procedure date.
2) Colonoscopies that occur concurrently with high-risk upper GI endoscopy procedures. The list of the CPT codes for the upper GI endoscopy procedures identified as “high-risk” are in attached Data Dictionary, sheet “Colonos Exclusions”
3) Colonoscopies for patients with a history of IBD or diagnosis of IBD at time of index colonoscopy or on the subsequent hospital visit outcome claim. The ICD-10-CM codes that define IBD are in the attached Data Dictionary, sheet “Colonos Exclusions.”
4) Colonoscopies for patients with a history of diverticulitis or diagnosis of diverticulitis at time of index colonoscopy or on the subsequent hospital visit outcome claim. The ICD-10-CM codes that define diverticulitis are in the attached Data Dictionary, sheet “Colonos Exclusions.”
5) Colonoscopies followed by a subsequent outpatient colonoscopy procedure within 7 days. For cases in which a colonoscopy is followed by another colonoscopy within 7 days, the measure will use the subsequent colonoscopy as the index colonoscopy.

<u>The following are in addition to those above, but only for HOPDs:</u>

Measure Overview

6) Colonoscopies that occur on the same day and at the same hospital as an ED visit that is billed on a separate claim than the index colonoscopy, unless the ED visit has a diagnosis indicative of a complication of care. The billing and revenue center codes that define ED visits are in the attached Data Dictionary, sheet “Colonos Outcome ED Obs Stay.” The same facility is defined as having the same CMS Certification Number (CCN). Complications of care codes are shown in tab “Colonos Exclusions ED CoC” include the CCS categories such as “Complications of surgical procedures or medical care,” “Adverse effects of medical care,” and others.
7) Colonoscopies that are billed on the same hospital claim as an ED visit and that occur on the same calendar day, unless the ED visit has a diagnosis indicative of a complication of care.
The billing and revenue center codes that define ED visits are in the attached Data Dictionary, sheet “Colonos Outcome ED Obs Stay.” Complications of care codes are shown in tab “Colonos Exclusions ED CoC” include the CCS categories such as “Complications of surgical procedures or medical care,” “Adverse effects of medical care,” and others.
8) Colonoscopies that are billed on the same hospital outpatient claim and that occur after the ED visit.
The billing and revenue center codes that define ED visits are in the attached Data Dictionary, sheet “Colonos Outcome ED Obs Stay.”

9) Colonoscopies that are billed on the same hospital outpatient claim as an observation stay.

The billing and revenue center codes that define observation stays are in the attached Data Dictionary, sheet “Colonos Outcome ED Obs Stay.”

Measure is Risk-Adjusted and/or Stratified:
The risk-adjustment model includes 15 patient-level variables, including age, concomitant upper GI endoscopy, polypectomy during procedure, and 12 comorbidity variables obtained from inpatient, outpatient, and physician claims 12 months prior to index procedure. The measure does not include certain diagnoses that occur only at the time of the colonoscopy procedure toward risk adjustment because these diagnoses may represent complications of care.

Logic Model

Summary: The conceptual model for colonoscopy quality, depicted below, illustrates the pathway through which facilities can influence the outcome of post-procedural hospital visits. For example, the model identifies that patient-level factors, such as comorbidities and other risk factors (such as age), increase the risk of unplanned hospital visits. Better management of the risks associated with these comorbidities may be a potential avenue for facilities to reduce unplanned hospital visits. Provider-level factors (technical quality of the procedure, post-procedure provider accessibility), and facility-level factors (such as pre- and post-discharge patient communication, other post-procedural processes) may also contribute to the risk of unplanned hospital visits. Therefore, facilities may have opportunities to lower their unplanned hospital visit rates through quality improvement efforts focused on these factors. Examples of interventions that have been shown in the literature to reduce complications, and therefore decrease post-procedure acute care utilization, are discussed in Section 1.10, Section 2.2, and Section 6.2.1.

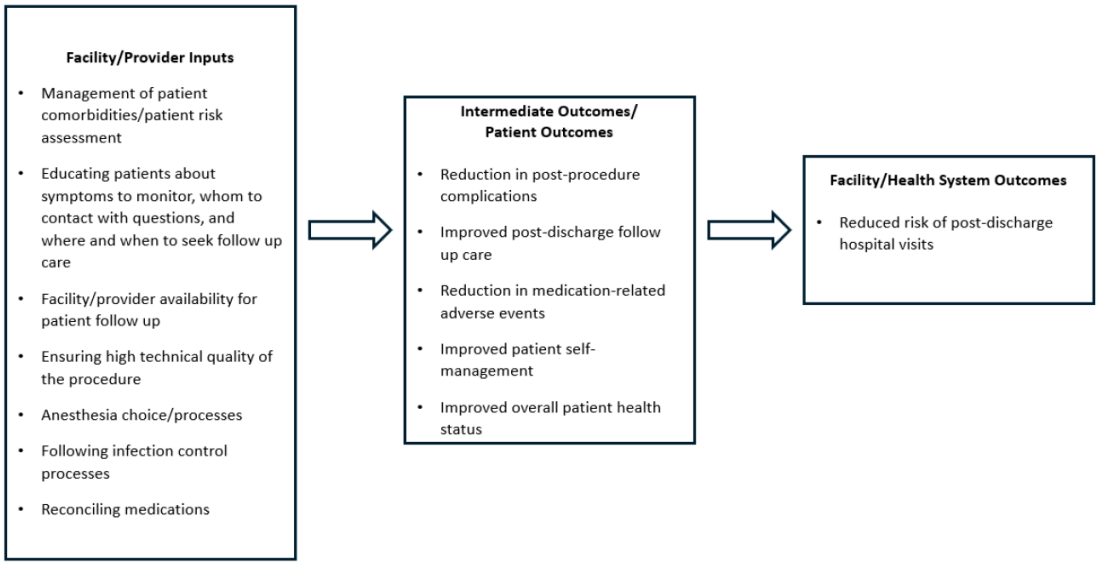

Figure 2. CBE #2539 Logic Model

Measure Evaluation Summary: CBE #2539

Importance
Staff Preliminary Rating: Met
Importance: The developer cites evidence from the primary literature interventions that the accountable entity can implement to improve hospitalizations, such as choice of sedation, improving low-volume providers and procedures, and proper management of patient medications. The developer further suggests facilities can implement a system of quality improvement and use CMS facility-level and claims-based reports, which include the principal diagnosis on admission, to tie quality improvement efforts to complications that are occurring. Data from July 1, 2020, to December 31, 2022, show a distribution of scores for Hospital Outpatient Departments (HOPDs) and Ambulatory Surgery Centers (ASCs), with mean scores of 13.18 and 9.92, respectively, against national averages. Approximately 30% of HOPDs and ASCs scored higher than their national averages. While there is no direct patient feedback on the measure's meaningfulness, the developer posits that both patients and providers benefit from measuring hospital visits post-colonoscopy, as it is a comprehensive indicator of patient outcomes. Evidence suggests that the quality of information provided to patients correlates with their experiences and concerns regarding colonoscopy, particularly noting that patients with complications rated the information they received lower than those without complications. The developer argues that reporting these measures to providers will encourage them to engage in quality improvement initiatives.

Feasibility
Staff Preliminary Rating: Met
Feasibility: The developer did not conduct a feasibility assessment of missing data for this measure, as it used a 100% sample of paid Medicare claims. The developer further states that because this is a claims-based measure, there is no added burden of reporting as all required data elements are routinely generated. The measure does not include any proprietary information.

Reliability	
Staff Preliminary Rating: Met	
Testing Level:	Accountable Entity Level
Testing Method:	Reliability testing was conducted using the **signal-to-noise method** on 2.5 years of data (7/20-12/22) and results reported show: For HOPDs, entity-level reliability was tested across **3,270 entities** and was **>0.6** for about **75%** of the entities. (The estimated **lower quartile was 0.555.)** For ASCs, entity-level reliability was tested across **1948 entities** and was **>0.6** for more than **90%** of the entities. (The estimated average reliability for the **2nd decile was 0.627**.)

Reliability: The measure is well-defined. Reliability was assessed at the entity level. Reliability statistics are above the established thresholds for at least 70% of the entities across HOPDs and ASCs.

Validity	
Staff Preliminary Rating: Not met but addressable	
Testing Level:	Accountable Entity Level
Testing Conducted:	Yes, validity testing was conducted using three methods: **(1) face validity, (2) association of measure scores with volume, and (3) validity of the outcome based on an analysis in claims data of the reasons why patients experienced a hospital visit (complications based on ICD-10 codes).**
Validity: Face validity results were strong, with 86.7% of a technical expert panel confirming the measure's validity and usefulness. Regarding the volume-outcome relationship, the developer reported a statistically significant correlation of -0.224 for HOPDs and -0.132 for ASCs, consistent with their hypothesis. However, it is unclear, based on the evidence provided, how facility procedural volume supports the validity of the measure, including the quality construct. Additionally, the developer provided supportive analysis of the outcome by providing ICD-10 codes associated with unplanned hospital visits post-procedure. The developer conducted statistical risk adjustment, selecting factors significantly correlated to the outcome but excluded social risk factors such as dual eligibility (DE) and the Area Deprivation Index (ADI) due to minimal impact. The model showed good discrimination with a C-statistic of 0.699.	

Equity	
Staff Preliminary Rating: Met	
Equity considered:	Yes
Equity: The developer reported that for HOPDs, patients with DE have an unadjusted hospital visit rate of 23.8%, compared with 12.6% for patients without DE. However, patients with high ADI have unadjusted hospital visit rates of 17.0% vs. 12.7% for patients without high ADI. This pattern was also similar for colonoscopies received at ASCs. The developer also looked at within facility and across facility rates for these subpopulations, finding that hospitals have worse outcomes for their DE patients compared with their non-DE patients. Looking across hospitals, the developer reports that only one facility had outcomes for DE patients that performed better than the national rate.	

Use & Usability	
Staff Preliminary Rating: Met	
Current or Planned Use:	Measure is currently used in the Hospital Outpatient Quality Reporting Program.
Use & Usability: The measure is currently used in the Hospital Outpatient Quality Reporting Program, Ambulatory Surgical Center Quality Reporting Program, and the Rural Emergency Hospital Quality Reporting Program. The developer cites evidence of interventions that can improve readmissions, as noted previously. The developer notes that stakeholders can submit questions about the measure through an online	

Use & Usability
tool and CMS responds to each question. The developer notes that these questions are considered in addition to literature reviews related to the measure and engagement with clinical experts for updating the measure. The developer also updates procedure codes used in the measure accordingly and adds that there have been no major substantive changes to the measure since its last endorsement in 2020. The developer reports that the national rate of hospital visits per 1,000 colonoscopies among HOPDs declined from 16.4 in 2018 reporting to 14.8 in 2019 reporting and 13.2 in 2023 reporting (current data). The developer reports no unexpected findings.

Public Comment[1]

Number of Comments Received: 5

Full text of developer/steward responses can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
One commenter shared a personal experience of needing to visit a hospital post-surgery and emphasized the importance of such follow-up visits, especially given the trend of quick patient turnover in hospitals.	Supportive	Thank you for your support of the post-procedure hospital visit measures [CBE 3470, 3357, 2539, and 3366]. You are correct that the outcome for these measures includes ED visits, observation stays, and inpatient admissions, 7 days after a procedure (which varies based on the measure).
One commenter suggested a modification to the current measure regarding the age range for Medicare Fee-for-Service patients. The commenter questions the rationale behind setting the age limit at 65 and older and proposes expanding the age range to include younger patients, potentially in their 50s. Although supportive of the measure, the commenter calls for a reevaluation of the age criteria to enhance the measure's effectiveness and relevance.	N/A	The measures are limited to Medicare patients because they are based on claims data (billing data) that are only nationally available and validated for Medicare patients. In the future we may be able to add Medicare Advantage patients, and there may be a future possibility of a Medicaid measure, but it is not possible for us to report nationally across all payers (private and public) due to the lack of a comprehensive, national all-payer database.
One commenter strongly opposed endorsement of this measure, citing significant flaws in its numerator, reliability, and risk adjustment methodology. The commenter argues that these	Non-supportive	We thank Dr. Miller for his comments. The Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy (hereafter, Colonoscopy) measure (CBE

[1] Comments, as submitted, can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
flaws could mislead patients about the quality of care at different ASCs.		#2539) is a reliable and valid measure that was initially endorsed in 2014 and re-endorsed in 2020 under largely the same criteria as the current Battelle process. The measure objectively meets all of Battelle's criteria for reliability and validity. We note that Dr. Miller submitted the same list of concerns for the HOPD Surgery measure, which was endorsed by these Cost & Efficiency Recommendation/Advisory Groups in the prior (Fall 2023) cycle. This Colonoscopy measure was developed using the same approach as the HOPD Surgery measure—it has the same 7-day unplanned hospital visit outcome, the same risk-adjustment approach, the same measure score calculation approach, and the same overall goal in improving patient care. The Colonoscopy measure is attributed to the facility and captures the outcome of hospital visits (ED visit, observation stay, or inpatient hospitalization) for colonoscopies performed in ASCs, and separately for those performed in HOPDs. The measure was developed with full transparency and public input, subject to several rounds of public comment (during measure development, rulemaking, and CBE endorsement), and broad stakeholder input. The current Battelle CBE criteria are largely the same as the prior National Quality Forum (NQF) criteria, and the measure has not changed substantially since first endorsed. Importantly, the measure's risk model continues to show good discrimination and calibration, which addresses many of the commenter's concerns about adequate risk adjustment. Yale/CORE's empiric approach to risk model development has ensured that patient-level variables that are significantly associated with the outcome are included in the model. Model testing shows that observed outcomes closely track predicted outcomes across a wide range of risk deciles. While we appreciate all the commenter's concerns, overall, the Colonoscopy measure is a reliable and valid measure of post-procedural outcomes and has support from the very surgeons that perform these procedures. To facilitate quality improvement, the facility receives, in addition to its measure score, detailed, claim-level information about each patient who met the inclusion criteria (including if they experienced a hospital visit, the type of visit, and the diagnosis associated with the visit). It is currently the only publicly

Comment Summary	Support Level	Summary of Developer Response
		reported measure that captures post-procedural outcomes for this high-volume procedure. Please see our full response to Dr. Miller's comment that addresses each of his concerns.
Two comments expressed concerns about the reliability of the measure, noting that the reliability score of 0.239 for hospital outpatient departments and 0.265 for ambulatory surgical centers with a minimum of 30 cases is significantly below the 0.7 threshold.	Non-supportive	According to Battelle's staff assessment, this measure does meet reliability criteria. Approximately 75% of facilities (with at least 30 cases) are above Battelle's threshold of 0.6 minimum reliability. Using 3 years of performance data, the median facility-level reliability score is 0.748 (IQR, 0.555-0.872) for HOPDs with at least 30 cases (the public reporting threshold) representing moderate reliability. At the current public reporting threshold, most facilities (about 75% of facilities with at least 30 cases) fall above the 0.6 minimum threshold stated in Battelle's current CBE guidebook. If CMS were to increase the case volume minimum so that all facilities exceeded this threshold, it would remove publicly available information from about 205 facilities that are currently publicly reported. We believe that median reliability of 0.6 (signal to noise) is sufficiently high for a facility-level publicly reported measure in a pay-for-reporting program. Increasing the minimum case volume for HOPDs has the tradeoff of removing important information available to the public on Care Compare.
One comment requested the committee re-evaluate the inclusion criteria for hospital visits in the measure's scoring and discuss the potential negative impact of including unrelated incidents on performance scores.	Non-supportive	According to Battelle's staff assessment, this measure does meet reliability criteria. Approximately 75% of facilities (with at least 30 cases) are above Battelle's threshold of 0.6 minimum reliability. The all-cause approach for the outcome for this measure is commonly used and accepted in quality measurement, and is the approach used in related measures in the outpatient setting (such as the Hospital Visits after Outpatient Surgery, CBE 2687, endorsed by the Cost & Efficiency Committee during the Fall 2023 CBE cycle); we have heard from stakeholders that alignment between similar measures is important. Furthermore, as shown in our submission (Table 7A), a complication is the most common reason for an unplanned post-procedural hospital visit, and the outcome window is only 7 days after the procedure, thus focusing on the time when we see (empirically) that highest volume of post-procedural hospital visits.

Advisory Group Feedback

Full text of developer/steward responses can be found on the PQM website.

Feedback/Questions	Summary of Developer Response
Similar Comments to #3357: Recognizing the comprehensive documentation within the submission, several committee members expressed similar concerns as noted with CBE #3357. Specifically with low event rates, one committee member questioned whether the measure has much discriminatory value between facilities, except for those in the fourth quartile, which identifies outliers.	Please refer to the developer's responses for CBE #3357. With respect to the low event rate, the developer noted that there are close to 4 million colonoscopies performed, with about 2 million in each setting over the 3-year period. It is important to monitor performance and give facilities feedback. The event rate is low, but the variation in event rates is broad. *Summary Response Received after the Advisory Group meeting:* Even with low outcome rates, we can show a meaningful distribution of measure scores as well as improvement over time. The colonoscopy measure is a case study of how the cycle of public reporting and specific facility feedback can yield broad improvements in performance if public reporting and facility feedback are sustained over time.
Colonoscopy Measures: One committee member asked whether this was the only colonoscopy measure that exists, as there may be better measures that look at the quality of the procedure. Another committee member noted that if this is the only measure, it doesn't mean it should be used if it's not a good measure.	The developer noted that they believe this is the only colonoscopy measure that is publicly reported and looks at the quality of colonoscopy. Therefore, this is a gap with respect to measurement. In addition, this is a low-burden measure, and the data are collected during the routine course of care. *Summary Response Received after the Advisory Group meeting:* There are no other publicly reported measures that capture outcomes related to colonoscopy. Therefore, this is a gap with respect to measurement. In addition, this is a low-burden measure, and the data are collected during the routine course of care.
Exclusions: A committee member asked: Why are irritable bowel syndrome (IBD) and diverticulitis excluded?	The developer noted that IBD and diverticulitis are relatively common issues, and they were identified by their technical expert panel as exclusions. *Summary Response Received after the Advisory Group meeting:* During measure development, we identified patients with IBD or diverticulitis as clinical situations that would present challenges with respect to risk adjustment and empirically determined that, based on their discharge diagnoses, it would be difficult to determine if their post-procedure hospital visits were planned or unplanned. Therefore, these patients are excluded from the measure.

Feedback/Questions	Summary of Developer Response
Denominator Clarification: One committee member asked about the units for the denominator.	The developer clarified that the denominator unit is per 1,000 visits. *Summary Response Received after the Advisory Group meeting:* The denominator is per 1,000 colonoscopies. Measure scores are expressed in units per 1,000 colonoscopies.
Stratification by Social Risk: In addition to the concerns expressed for CBE #3357, one committee member asked whether there is a consideration to stratify the measure by social risk factors, such as dual eligibility and ADI, as there may be less resources within the facility or community to support these patients. One committee member commented that because the hospital outpatient rates are higher than the ASCs, that could be due to areas of the country that do not have ASCs available and that the measure should consider the availability of these facilities within a given community. This may lead hospitals to be selective of lower-risk patients.	The developer noted that for hospital outpatient departments (HOPDs), the measure currently is stratified by dual eligibility, but they are not sure whether the measure will be stratified by ADI. Currently, this is being reported to facilities confidentially. CMS has been focusing on disparities for which there are enough data to support stratification. CMS recognized that there were disparities in the HOPD version of the measure and stratified by dual eligibility. Typically for the ASC measures, there aren't enough patients with social risk factors to support stratification. The developer added that they haven't performed any geographic analysis regarding the availability of ASCs but may consider looking into this if there are resources to do so. *Summary Response Received after the Advisory Group meeting:* For HOPDs, the related HOPD Surgery measure currently is stratified by dual eligibility and the information is being reported to facilities confidentially. CMS had planned on the same approach for the ASC measures, but as noted earlier, ASCs typically do not serve a high volume of patients with social risk, and empirically we have found that we cannot apply the stratification method to ASCs because there are not enough patients with social risk factors to support stratification. Regarding the hypothesis that ASC availability would impact outcome rates at HOPDs, we note that there are many confounders that would need to be considered in such an analysis, including the availability of high-quality providers, geographic differences in the regulation of ASCs by state, and the availability of/access to different settings that perform the same procedures (inpatient, HOPD, ASC, provider office settings), among others. We also note that ASC and HOPD rates cannot be compared because they are two different measures that are calculated separately.
Support for All-Cause: One committee member commented that they think it is a good approach to have an all-cause measure as these measures are claims based and the data can be noisy for a procedure-related measure.	The developer appreciated the comment. *Summary Response Received after the Advisory Group meeting:* We appreciate the support from the Advisory Group member.

Key Discussion Points:

- **All-Cause Outcome:** Does the committee have any additional concerns regarding the all-cause outcome, considering the developer's inclusion of the top 25 reasons for a hospital visit being related to the procedure?

- **Bias of the Outcome:** Does the committee have any concerns regarding the developer's response to the Advisory Group concerns regarding the lack of urgent care or office visits being captured in the measure?

- **Low Outcome Rates:** The developer responded to this Advisory Group concern that low outcome rates are mitigated by (1) minimum case volume thresholds for public reporting; (2) the statistical method that pulls measure scores for small facilities to the mean; and (3) that the measure is used in a pay-for-reporting program, not for pay-for-performance. Does the committee have any additional concerns?

- **Reliability:** Considering the developer's response to Advisory Group concerns and that greater than 70% of reliability estimates are above the expected value of 0.6, does the committee agree that the measure has sufficient reliability?

- **Validity:** Does the committee have concerns with the volume-outcome validity testing?

CBE #3357: Facility-Level 7-Day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers [Yale CORE/CMS]

Measure Description:

Facility-level risk-standardized ratio of acute, unplanned hospital visits within 7 days of a general surgery procedure performed at an ambulatory surgical center (ASC) among Medicare Fee-For-Service (FFS) patients aged 65 years and older. An unplanned hospital visit is defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission.

Measure Status	
New or Maintenance: Maintenance Measure	**Used in An Accountability Application?** Yes – Public Reporting, Quality Improvement with Benchmarking
CBE Endorsement Status: Endorsed **Last Endorsement Review Cycle:** Fall 2017	**Proposed/Planned Use:** Ambulatory Surgical Center Quality Reporting (ASCQR) Program (external benchmarking to multiple organizations)

Measure Characteristics

Measure Type	Target Population(s)	Level of Analysis	Care Setting(s)
Outcome	Medicare FFS patients aged 65 years and older, undergoing outpatient general surgery provedures in ASCs	Facility	Ambulatory Surgery Center

Measure Overview
Rationale: Several factors make unanticipated hospital visits a priority quality indicator. Because ASC providers are not aware of all post-surgical hospital visits that occur among their patients, reporting this outcome will help to illuminate problems that may not be currently visible (Zivanovic et al., 2020). In addition, the outcome of hospital visits is a broad, patient-centered outcome that reflects the full range of reasons leading to hospital use among patients undergoing same-day surgery. Public reporting of this outcome measure will provide ASCs with critical information and incentives to implement strategies to reduce unplanned hospital visits.
Numerator: The measure defines the outcome as any (one or more), all-cause, acute, unplanned hospital visit within seven days of an outpatient general surgery performed at an ASC; a hospital visit includes any emergency department (ED) visit, observation stay, or unplanned inpatient admission.
Denominator: The target population for this measure is Medicare FFS patients aged 65 years and older, undergoing selected outpatient general surgery procedures in ASCs that are within the scope of general surgery training. Specifically, the cohort of procedures includes the following types of procedures: abdominal, alimentary tract, breast, skin/soft tissue, wound, and varicose vein.
Exclusions: The measure excludes surgeries for patients without seven or more days of continuous enrollment in Medicare FFS Parts A and B after the surgery to ensure all patients have full data available for outcome assessment.

Measure Overview

Measure is Risk-Adjusted and/or Stratified:
In brief, the approach simultaneously models two levels (patient and facility) to account for the variance in patient outcomes within and between facilities (Normand & Shahian, 2007). At the patient level, the model adjusts the long-odds of hospital visits within seven days after the procedure for selected demographic, clinical, and procedure risk variables. The second level models the facility-specific intercepts as arising from a normal distribution. The facility intercept, or facility-specific effect, represents the ASC contribution to the risk of 7-day hospital visits, after accounting for patient risk and sample size, and can be inferred as a measure of quality. If there were no differences among ASCs, then after adjusting for patient risk, the facility intercepts would be identical across all ASCs.

Logic Model

Summary: The conceptual model for outpatient general surgery quality at an ASC, shown below, shows the pathway by which facilities can modify the outcome (all-cause, unplanned hospital visits within seven days of the orthopedic ASC procedure). For example, the model identifies that patient-level factors, such as comorbidities, increase the risk of unplanned hospitals visits. Better management of the risk associated with these comorbidities may be a potential avenue for facilities to reduce unplanned hospital visits. Provider-level factors (such as technical quality of the procedure, post-procedure provider accessibility), and facility-level factors (such as patient selection/risk assessment, pre- and post-discharge patient communication) may also contribute to the risk of unplanned hospital visits. Therefore, facilities may have opportunities to lower their unplanned hospital visit rates through quality improvement efforts focused on patient, provider, and facility factors.

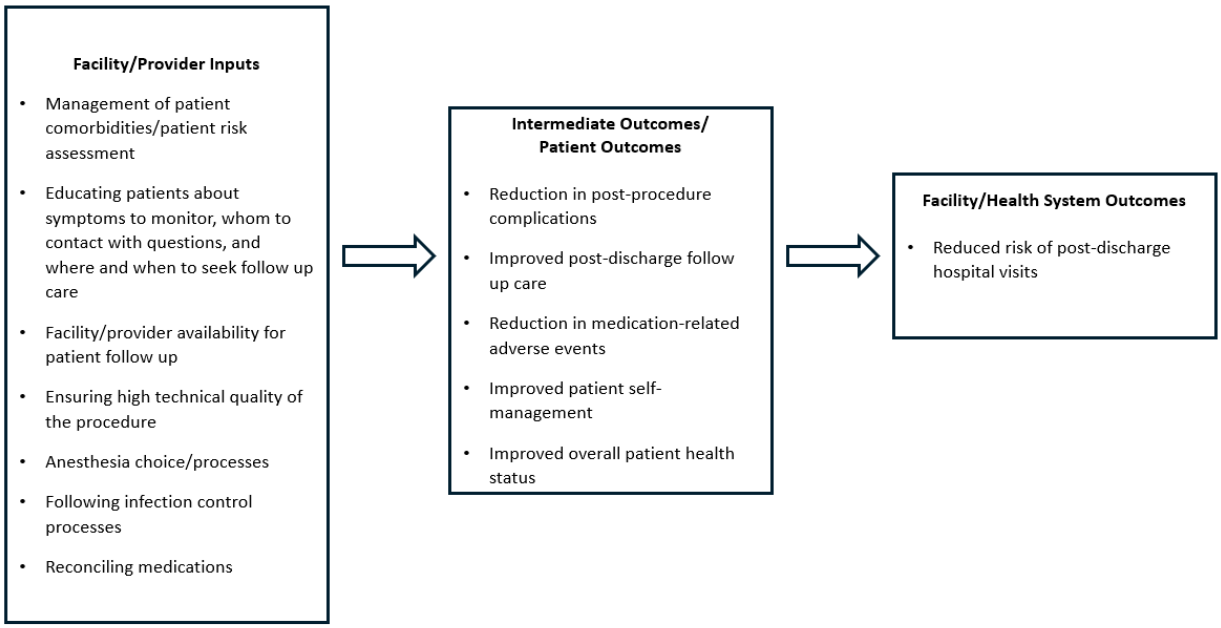

Figure 3. CBE #3357 Logic Model

Measure Evaluation Summary: CBE #3357

Importance
Staff Preliminary Rating: Met
Importance: This measure aims to reduce unplanned hospital visits post-surgery in ASCs. The developer cited evidence supporting interventions such as preoperative assessments, enhanced postoperative monitoring, and improved patient education to reduce hospitalizations. The developer suggests facilities adopt quality improvement systems and use CMS benchmarking reports to link improvements to specific complications. Using data from January 1, 2021, through December 31, 2022, the developer reported a distribution of scores with an overall score of 1.00 (national average) and a measure score range of 0.59-1.84, with a lower score being better. Despite lacking direct patient feedback about the meaningfulness of this measure focus, the developer argues it is necessary to inform ASC providers and patients of the outcome of hospital visits due to the increase in procedures in ASCs.

Feasibility
Staff Preliminary Rating: Met
Feasibility: The developer did not perform a feasibility assessment for missing data for this measure because it utilized a complete 100% sample of paid Medicare claims. Because this is a claims-based measure, the developer notes there is no added burden of reporting as all required data elements are routinely generated and there are no proprietary components of the measure.

Reliability	
Staff Preliminary Rating: Met	
Testing Level:	Accountable Entity Level
Testing Method:	Reliability testing was conducted using the **signal-to-noise method** and results reported show a median facility-level reliability score for ASCs as 0.690 (interquartile range [IQR], 0.555-0.824) for ASCs with at least 25 cases.
Reliability: The developer conducted accountable entity-level reliability by calculating signal-to-noise reliability scores on 2 years of data (1/21-12/22) for the risk-adjusted measure across 1,455 entities. Reliability statistics are above the established thresholds for at least 70% of the entities.	

Validity	
Staff Preliminary Rating: Not met but addressable	
Testing Level:	Accountable Entity Level

Validity	
Testing Conducted:	Validity testing was conducted using three methods: **(1) face validity during measure development, (2) validity through association with volume, and (3) validation of the outcome.**
Validity: Face validity results were strong, with 86.7% of a technical expert panel confirming the measure's validity and usefulness. Regarding the volume-outcome relationship, the developer hypothesized a weak-to-moderate negative relationship between facility volume and measure scores, supported by a statistically significant correlation of -0.207, aligning with prior literature showing better outcomes with higher facility volumes. However, it is unclear, based on the evidence provided, how facility procedural volume supports the validity of the measure, including the quality construct. Additionally, the developer provided supportive analysis of the outcome by providing ICD-10 codes associated with unplanned hospital visits within seven days of a qualifying surgery. The developer conducted statistical risk adjustment, selecting factors significantly correlated to the outcome but excluded social risk factors such as dual eligibility and the Area Deprivation Index due to minimal impact. The model showed good discrimination with a C-statistic of 0.699.	

Equity	
Staff Preliminary Rating: Not met but addressable	
Equity considered:	Yes
Equity: The developer did not identify whether the measure rates are different across different patient populations. Rather, the developer provided a rationale for why dual eligibility and Area Deprivation Index were not included in the risk model. For future submissions, the developer may consider whether the measure can identify differences in performance across different patient groups by their social and/or economic status.	

Use & Usability	
Staff Preliminary Rating: Not met but addressable	
Current or Planned Use:	Measure is currently used in the ASC Reporting Program.
Use & Usability: During its implementation, the measure had a "dry-run" period, allowing ASCs to familiarize themselves with the measure and ask questions through an online tool. Feedback from this period led to minor updates, such as adjustments to procedural codes and the admission algorithm, with no substantive updates since 2018. The developer has not reported any significant findings on improvement progress due to the ever-changing nature of procedures, making year-over-year performance assessments challenging. The developer notes that since the measure started publicly reporting in January 2024, there has not been sufficient time for facilities to fully implement changes and evaluate their impact on performance. The developer did not report any unexpected findings.	

Public Comment[2]

Number of Comments Received: 5

Full text of developer/steward responses can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
Two supportive comments shared personal experiences emphasizing the importance of this measure.	Supportive	Thank you for your support of the post-procedure hospital visit measures [CBE 3470, 3357, 2539, and 3366]. You are correct that the outcome for these measures includes ED visits, observation stays, and inpatient admissions, 7 days after a procedure (which varies based on the measure). We thank the commenter for the support of the measure and agree that the quality of postoperative care needs to be fine-tuned. This measure allows facilities to understand the root causes of any postoperative ED visit, observation, or inpatient stay and as such helps facilities identify opportunities for quality improvement.
Two comments opposed endorsement of this measure, criticizing its risk adjustment model, reliability, and lack of business case.	Non-supportive	We thank Dr. Miller for his comments. The Facility-Level 7-Day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers (hereafter, ASC General Surgery) measure (CBE #3357) is a reliable and valid measure that was endorsed previously in 2018 under largely the same criteria as the current Battelle process. The measure objectively meets all of Battelle's criteria for reliability and validity. We note that Dr. Miller submitted the same list of concerns for the HOPD Surgery measure, which was endorsed by this Cost & Efficiency Recommendation/Advisory Groups in the prior (Fall 2023) cycle. This ASC General Surgery measure was developed using the same approach as the HOPD Surgery measure—it has the same 7-day unplanned hospital visit outcome, the same risk-adjustment approach, the same measure score calculation approach, and the same overall goal in improving patient care. The ASC General Surgery measure is attributed to the facility and captures the outcome of hospital visits (ED visit, observation stay, or inpatient hospitalization) for general surgery procedures performed in

[2] Comments, as submitted, can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
		ASCs. It was first endorsed in 2018 by the NQF Surgery Committee. The measure was developed with full transparency and public input, subject to several rounds of public comment (during measure development, rulemaking, and CBE endorsement), and broad stakeholder input. The current Battelle CBE criteria are largely the same as the prior NQF criteria, and the measure has not changed substantially since first endorsed. Importantly, the measure's risk model continues to show good discrimination and calibration, which addresses many of the commenter's concerns about adequate risk adjustment. Yale/CORE's empiric approach to risk model development has ensured that patient-level variables that are significantly associated with the outcome are included in the model. Model testing shows that observed outcomes closely track predicted outcomes across a wide range of risk deciles. While we appreciate all the commenter's concerns, overall, the ASC General Surgery measure is a reliable and valid measure of post-procedural outcomes and has support from the very surgeons that perform these procedures. To facilitate quality improvement, the facility receives, in addition to its measure score, detailed, claim-level information about each patient who met the inclusion criteria (including if they experienced a hospital visit, the type of visit, and the diagnosis associated with the visit). It is currently the only publicly reported measure that captures post-surgical outcomes for these procedures in a health care system that spends billions of dollars each year on outpatient procedures and where procedural volume is shifting to the outpatient setting. Please see our full response to Dr. Miller's comment that addresses each of his concerns. According to Battelle's staff assessment, this measure does meet reliability criteria. Using 2 years of performance data, the median facility-level reliability score for ASCs is 0.690 (IQR, 0.555-0.824) for ASCs with at least 25 cases (the public reporting threshold) representing moderate reliability. At the current public reporting threshold, most facilities (about 75% of facilities with at least 25 cases) fall above the 0.6 minimum threshold stated in Battelle's current Endorsement & Maintenance Guidebook.
One comment questioned the rationale behind setting the age limit at 65 and older and	N/A	The measures are limited to Medicare patients because they are based on claims data (billing data) that are only nationally available

Comment Summary	Support Level	Summary of Developer Response
proposes expanding the age range to include younger patients, potentially in their 50s.		and validated for Medicare patients. In the future, we may be able to add Medicare Advantage patients, and there may be a future possibility of a Medicaid measure, but it is not possible for us to report nationally across all payers (private and public) due to the lack of a comprehensive, national all-payer database.

Advisory Group Feedback

Full text of developer/steward responses can be found on the PQM website.

Feedback/Questions	Summary of Developer Response
Concern with All-Cause Specification and Low Absolute Rates: A few committee members raised concern with the all-cause specification of the measure, expressing that while most reasons for why people may come back to the hospital could be related to the surgery, because the absolute rate of the visits is low (one committee member stated 2%), an unrelated visit could penalize a particular facility. In addition to low absolute rates, a committee member commented that the variation across those rates is low.	With respect to the all-cause outcome, the developer submitted the top reasons for a hospital visit following a procedure, which is Table 5 in the submission packet. Most of the reasons for the hospital visit are complications from the procedure or from lack of the appropriate care follow-up after the procedure. That is why outcome is all-cause. The developer noted that this measure has the lowest outcome rate compared to CBE #3366 and CBE #3470, adding that this measure is still important, and variation still exists. The measure adds value, as ambulatory surgical centers (ASCs) have their rates compared to other ASCs and are benchmarked against other geographic regions, and they also receive the claims-based reports. These are detailed reports that list every single patient, including what the procedure was; whether there was an inpatient stay, observation stay, or emergency department (ED) visit; and the principal diagnosis associated with the stay or visit. This information allows the ASC to identify the root cause of any problems that may be leading to those visits. Lastly, the developer added that more and more procedures are being moved to the ASC setting. CMS has an ASC-covered procedure list that changes over time as new procedures are continually added. Therefore, this measure will possibly capture additional procedures. No other publicly available measures capture outcome rates for these ASC procedures and are at the claims detail level for the facility. *Summary Response Received after the Advisory Group meeting:* The all-cause outcome is supported by the short outcome time window (7 days after the procedure, when most procedural-related complications occur) and evidence submitted by the developer that most post-procedural hospital visits are due to a

Feedback/Questions	Summary of Developer Response
	complication. Low outcome rates are mitigated by (1) minimum case volume thresholds for public reporting; (2) the statistical method that pulls measure scores for small facilities to the mean; and (3) that the measure is used in a pay-for-reporting program, not for pay-for-performance. As higher-risk procedures continue to migrate to the outpatient setting, this measure (and the related urology and orthopedic measures) retains its importance in monitoring performance in the ASC setting, where collectively these publicly reported ASC measures assess outcomes for the millions of procedures that are performed at ASCs annually.
Time Window Post-Surgery: A committee member raised a question as to why a 7-day window is used, rather than something like 72 hours, especially when considering the all-cause nature of the measure? In addition, the same committee member questioned whether there was a rate of false positives beyond the 72 hours, i.e., events that are not surgical related and therefore could distort the value that's being attributed to the facility?	The developer noted that in the methodology report submitted with the measure, they show a time course of hospital visits, ED visits, observation stays, and inpatient admissions after a procedure compared to baseline. The developer noted that there is a steep number of hospital visits immediately after the procedure, which is probably within 72 hours, and then it hits baseline at about 7 days. In the combination of the empiric data and the developer's technical expert panel feedback, they landed on the 7 days. The developer noted that it does not have any data about what would be considered a false positive and hasn't performed a detailed analysis looking at the variation in hospital visits between 72 hours and the rest of 7 days. However, looking at the actual list of complications, most of them are directly tied to the procedure. *Summary Response Received after the Advisory Group meeting:* The 7-day outcome window is supported empirically; the time course for post-procedural hospital visits was assessed during measure development and the decision was made to use 7 days based on that empiric data.
Low Reliability for Smaller Volume Providers: A few committee members noted that the reliability is low, specifically for the lower three deciles, which could be attributed to the low volume of procedures performed, and this is concerning for these providers. One committee member commented that even a threshold of 0.7 is too low to be acceptable. One committee member suggested having a multi-year scatter plot for the individual facilities to see how stable these results are year over year would provide a good indication of how stable (i.e., reliable) the measure is. One committee member noted that this measure could be good for looking at outliers. A few committee members noted that	The developer noted that the requirements for acceptable reliability have changed over time from the National Quality Forum and the Scientific Methods Panel to the Battelle standard of a minimum of 0.6 (for accountable entity-level testing). Recognizing that this is a new threshold, the developer noted that about 94% of procedures in this measure fall above that threshold. With respect to a year-over-year analysis, the developer commented that they haven't looked at that and they could. They noted a concern with ASCs, which is that these physician groups can make themselves out to be an ASC so they can bill for a facility fee. There is an uncertainty with respect to the turnover of surgeons, so you will see more of the surgeon signal compared to what you might see at the hospital level.

Feedback/Questions	Summary of Developer Response
if you're using this predicted-to-expected approach and a hierarchical model, it's going to be affected by a low volume. The concern is that there is shrinkage to the mean, in which poor-performing facilities aren't classified as poor.	*Summary Response Received after the Advisory Group meeting:* The statistical method used to calculate measure scores, in addition to the public reporting volume cutoff, ensures that facilities that are small and/or have low outcome rates are not unfairly characterized. Year-over-year comparisons are hampered by overlapping performance periods and cohort changes as additional types of procedures are approved for payment in the ASC setting.
Potential Bias in the Outcome: A few committee members raised concern regarding the potential bias with the outcome, because the measure does not capture people who may be seen at an urgent care facility or a physician office or return to the ambulatory surgical center due to complications. In addition, one committee member noted that many large hospitals are building urgent care centers within their system to shunt people from emergency departments (EDs) to urgent care. Thus, the measure's results may be somewhat distorted.	The developer noted that when this measure was developed, urgent care wasn't as prevalent as it is now, but they still consider urgent care and physician office visits as ambulatory care. That would be what is incentivized. If a person has a procedure done and they have a problem, the incentive is to seek outpatient care, such as physician offices, so the person doesn't have to go for acute care at a hospital. The developer noted that there are different types of urgent care, especially those urgent cares for people who can't get a primary care appointment and their primary care doctor refers them to urgent care within their system. The developer added that this type of urgent care would still be considered an office visit, and this is why the measure is incentivizing non-hospital-based ambulatory care. *Summary Response Received after the Advisory Group meeting:* The accountable entity for this measure is the ASC and the goal of the measure is to improve care at the ASC (before, during, and after the procedure) and to minimize the use of acute care hospital visits. ASCs and their staff should encourage post-surgical care in non-acute/non-hospital care settings, which could include the ASC itself or other ambulatory settings (clinic visit, urgent care, etc.).
Mortality Exclusion: A committee member raised concern that because mortality is excluded, if a person dies after the surgery, then that isn't counted as a problem.	The developer noted that this measure includes very low-risk, elective procedures, compared to inpatient procedures. The mortality rates are extremely low, lower than the outcome rate. Thus, the developer decided to not include mortality. *Summary Response Received after the Advisory Group meeting:* Mortality from these lower-risk procedures performed in relatively healthy patients is exceedingly low (2.5 hundredths of a percent) and the low outcome rate would likely complicate comparisons of quality across facilities.
Unintended Consequence: One committee member stated that because of the low rates, this measure might have a perverse incentive to reduce the use of	The developer stated that it does not believe the measure causes any perverse incentive to not perform procedures at ASCs. Rather, the rates of procedures at ASCs are increasing over time for lots of different reasons, such as convenience and cost.

Feedback/Questions	Summary of Developer Response
ambulatory surgery and therefore raise costs for the health care system.	The developer added that surgeons have a different view of this measure, as they want to track their performance. The developer noted that when this measure was reviewed by the National Quality Forum Surgery committee, a discussion point with the surgeons was that a complication or issue arising to the severity level requiring an ED visit or hospital-level visit was more concerning and more appropriate to capture compared to something that might be able to be addressed via an office visit or urgent care. *Summary Response Received after the Advisory Group meeting:* Low rates of post-procedural hospital visits are unlikely to decrease the use of ASCs. Objectively, the use of ASCs (as well as the number of ASCs) has been increasing.
Grouping of the Outcome: One committee member raised concern that the measure is grouping ED visits, observation stays, and inpatient admissions into one bucket. From a health care cost perspective, this blurs the severity of care, as one patient could stay an hour and then be released, while another could stay for days.	As noted above, ASCs receive detailed claims-based reports that have every single patient listed, including what the procedure was; whether there was an inpatient stay, observation stay, or ED visit; and the principal diagnosis associated with the stay or visit. This allows the ASC to identify the root causes of any problems that may lead to those visits. In addition, this measure is a binary (yes/no) outcome, as opposed to a length-of-stay count, like in the inpatient setting. There currently are measures that count actual days, but this has not been applied to the settings for this measure. This is because these are typically low-risk procedures. The purpose of the measure is to capture acute hospital utilization and not determine which complication was more severe than the other. It is meant to give feedback to facilities to improve those rates. *Summary Response Received after the Advisory Group meeting:* Separating the outcome into its components (ED visit, inpatient admission, observation stay) would be problematic because it would reduce the outcome rate in a setting where outcome rates are already low. In addition, capturing all components in one outcome reduces the opportunities for gaming. Facilities receive detailed reports that include specific information for each patient that indicates if the post-procedural outcome was an ED visit, inpatient admission, or observation stay, in addition to the principal discharge diagnosis, which allows ASCs to use this information to pinpoint areas for quality improvement.
Equity: A few committee members commented on equity and the inclusion of social risk factors. One committee member (patient partner) asked whether this measure should include socioeconomic status (SES) to	The developer noted that the measure doesn't create a bias towards people accessing these procedures or ASCs. This is an access issue, which is more upstream of the measure. The developer added that they looked at the proportion of patients with social risk factors at the facility level and found that the median was 0

Feedback/Questions	Summary of Developer Response
be used as a learning tool for addressing health issues related to people with low SES. Another committee member added that this measure is biased toward the middle class. Another committee member asked for more information as to how social risk factors were considered in the measure. Lastly, a committee member noted that the graphs of estimates of social risk factors included, with and without, are on the 45% line with no skew associated with the percentage among high-SES risk patients.	for dual eligibility and high Area Deprivation Index (ADI), which were the two of the social risk factors considered. There just aren't very many patients with social risk factors in the patient population in the analysis conducted. *Summary Response Received after the Advisory Group meeting:* Compared with HOPDs, ASCs serve a very low proportion of patients with social risk factors. This ASC measure, however, does not create disparities nor does it capture disparities in care. Measure testing with and without social risk factors in the risk model shows little impact of including these variables, possibly due to the very low proportion of patients with social risk served by ASCs.
Attribution: One committee member asked: How does this measure ensure penalties are assigned to ASCs and not unrelated hospitals?	The developer noted that they look at the claims, which show where the procedure was performed. They then can link the outcome with the actual procedure itself. *Summary Response Received after the Advisory Group meeting:* Claims data include a “place of service” code that allows us to identify procedures performed at ASCs.

Key Discussion Points:

- **All-Cause Outcome:** Does the committee have any additional concerns regarding the all-cause outcome, considering the developer’s inclusion of the top 25 reasons for a hospital visit being related to the procedure?

- **Bias of the Outcome:** Does the committee have any concerns regarding the developer’s response to the Advisory Group concerns regarding the lack of urgent care or office visits being captured in the measure?

- **Low Outcome Rates:** The developer responded to this Advisory Group concern that low outcome rates are mitigated by (1) minimum case volume thresholds for public reporting; (2) the statistical method that pulls measure scores for small facilities to the mean; and (3) that the measure is used in a pay-for-reporting program, not for pay-for-performance. Does the committee have any additional concerns?

- **Reliability:** Considering the developer’s response to Advisory Group concerns and that greater than 70% of reliability estimates are above the expected value of 0.6, does the committee agree that the measure has sufficient reliability?

- **Validity:** Does the committee have concerns with the volume-outcome validity testing?

- **Use and Usability:** Does the committee agree that the measure results can be used by ASCs to improve their scores over time?

CBE #3366: Hospital Visits After Urology Ambulatory Surgical Center Procedures [Yale CORE/CMS]

Measure Description:

Facility-level risk-standardized rate of acute, unplanned hospital visits within 7 days of a urology procedure performed at an ambulatory surgical center (ASC) among Medicare Fee-For-Service (FFS) patients aged 65 years and older. An unplanned hospital visit is defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission.

Measure Status	
New or Maintenance: Maintenance Measure	**Used in An Accountability Application?** Yes - Public Reporting; Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
CBE Endorsement Status: Endorsed **Last Endorsement Review Cycle:** Fall 2018	**Proposed/Planned Use:** Ambulatory Surgical Center Quality Reporting (ASCQR) Program

Measure Characteristics

Measure Type	Target Population(s)	Level of Analysis	Care Setting(s)
Outcome	Medicare FFS patients, aged 65 years and older, who have undergone a urology procedure in ASCs	Facility	Ambulatory Surgery Center

Measure Overview
Rationale: A hospital visit after outpatient surgery is unexpected, and many of the reasons for such hospital visits are preventable. Hospital visits following an ambulatory surgery vary from 0.5% to 9.0%, based on the type of surgery, outcome measured, and timeframe for measurement after surgery (Bongiovanni et al., 2021). Hospital visits can occur due to a range of potentially preventable adverse events including uncontrolled pain, urinary retention, surgical site infection, bleeding, septicemia, and venous thromboembolism. Patients also frequently report minor adverse events -- for example, uncontrolled pain, nausea, and vomiting -- that may result in unplanned acute care visits following surgery (Owens et al., 2014; Bongiovanni et al., 2021). Several factors make unanticipated hospital visits a priority quality indicator. Because ASC providers may not be aware of all post-surgical hospital visits that occur among their patients, reporting this outcome will help to illuminate problems that may not be currently visible (Zivanovic et al., 2020). In addition, the outcome of hospital visits is a broad, patient-centered outcome that reflects the full range of reasons leading to hospital use among patients undergoing same-day surgery. Public reporting of the ASC Urology measure will provide ASCs with critical information and incentives to implement strategies to reduce unplanned hospital visits.
Numerator: The measure defines the outcome as any (one or more), all-cause, unplanned hospital visit within seven days of an outpatient urology surgery; a hospital visit includes any emergency department (ED) visit, observation stay, or unplanned inpatient admission occurring within seven days after the ASC procedure.

Measure Overview
Denominator: The target population for this measure is Medicare FFS patients aged 65 years and older undergoing outpatient urology surgeries, typically performed by a urologist, at ASCs.
Exclusions: The measure excludes surgeries for patients without 7 or more days of continuous enrollment in Medicare FFS Parts A and B after the surgery. The measure excludes these patients to ensure all patients have full data available for outcome assessment.
Measure is Risk-Adjusted and/or Stratified: At the patient level, the model adjusts the log-odds of hospital visits within 7 days after the procedure for selected demographic, clinical, and procedure risk variables. The second level models the facility-specific intercepts as arising from a normal distribution. The facility intercept, or facility-specific effect, represents the ASC contribution to the risk of 7-day hospital visits, after accounting for patient risk and sample size, and can be inferred as a measure of quality. If there were no differences among ASCs, then after adjusting for patient risk, the facility intercepts would be identical across all ASCs.

Logic Model
Summary: The conceptual model for outpatient urologic surgery quality at an ASC, shown below, shows the pathway by which facilities can modify the outcome (all-cause, unplanned hospital visits within 7 days of the urologic ASC procedure). For example, the model identifies that patient-level factors, such as comorbidities, increase the risk of unplanned hospitals visits. Better management of the risk associated with these comorbidities may be a potential avenue for facilities to improve patients' post-procedure health status and reduce unplanned hospital visits. Provider-level factors (technical quality of the procedure, post-procedure provider accessibility), and facility-level factors (such as patient selection/risk assessment, pre- and post-discharge patient communication) may also contribute to the risk of unplanned hospital visits. Therefore, facilities may have opportunities to lower their unplanned hospital visit rates through quality-improvement efforts focused on patient, provider, and facility factors.

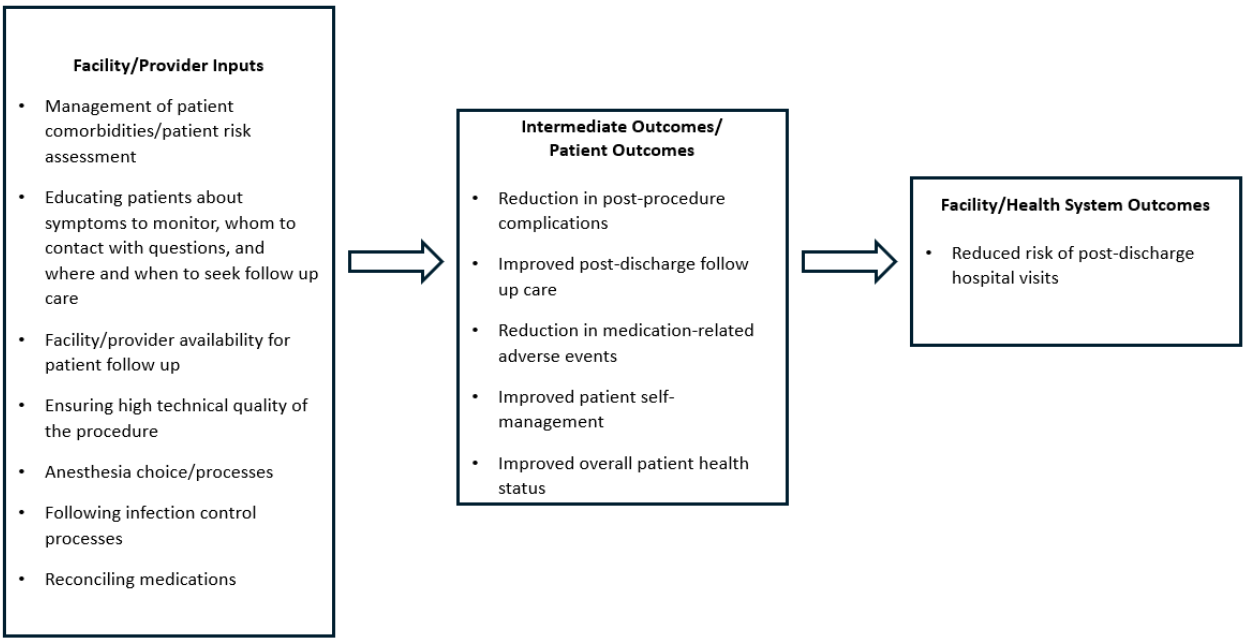

Figure 4. CBE #3366 Logic Model

Measure Evaluation Summary: CBE #3366

Importance
Staff Preliminary Rating: Met
Importance: The developer cites evidence of interventions that ASCs can implement to improve hospitalizations post-surgery, such as preoperative assessments, enhancing postoperative monitoring, and improving patient education. The developer suggests facilities can implement a quality improvement system and utilize both CMS facility-based and claims-based reports to enhance quality efforts. Data from January 1, 2021, to December 31, 2022, show an average facility-level measure score of 5.16, with about 30% of ASCs scoring above this average; lower scores indicate better performance.

Feasibility
Staff Preliminary Rating: Met
Feasibility: The developer did not perform a feasibility assessment for missing data for this measure because it utilized a complete 100% sample of paid Medicare claims. Because this is a claims-based measure, the developer notes there is no added burden of reporting as all required data elements are routinely generated, and there are no proprietary components of the measure.

Reliability	
Staff Preliminary Rating: Met	
Testing Level:	Accountable Entity Level
Testing Method:	Reliability testing was conducted using **signal-to-noise method,** and results reported an overall reliability estimate of 0.706 with a decile range of 0.422-0.949. Roughly 20% of ASCs have a signal-to-noise estimate less than 0.6.
Reliability: The measure is well-defined and precisely specified. The developer conducted signal-to-noise of the measure score with a minimum procedure volume of 35 (which is currently the public reporting cutoff). The developer reports an overall reliability estimate of 0.706 with a decile range of 0.422-0.949. Roughly 20% of ASCs have a signal-to-noise estimate of less than 0.6.	

Validity	
Staff Preliminary Rating: Not met but addressable	
Testing Level:	Accountable Entity Level
Testing Conducted:	Yes, validity testing was conducted using three methods: **(1) face validity during measure development, (2) validity through association with volume, and (3) validation of the outcome**.

Validity
Validity: Face validity results were strong, with 86% of a technical expert panel confirming the measure's validity and usefulness. Regarding the volume-outcome relationship, the developer hypothesized a weak-to-moderate negative relationship. However, the validity of this association is questionable, as it was only significant for facilities performing over 200 procedures and did not clearly demonstrate the ability to distinguish between facilities of varying quality. Additionally, the relevance of procedural volume to the measure's validity is unclear, as it was not included in the measure's logic model and its impact on quality is not explicitly defined. The developer also performed statistical risk adjustment using factors significantly correlated with the outcome and explored but ultimately excluded social risk factors such as dual eligibility and the Area Deprivation Index due to their minimal impact. The measure's discrimination was considered good, with a C-statistic of 0.615.

Equity	
Staff Preliminary Rating: Not met but addressable	
Equity considered:	Yes
Equity: The developer reports that ASCs do not serve a high proportion of patients with social risk factors, with the median facility proportion of patients with the dual eligibility variable being 0% (count of 0 patients) and for the high ADI, 1% (count of 1 patient). However, the developer did not identify whether the measure rates are different across other patient populations or how this measure may help to reduce disparities. Rather, for this section, the developer provided a rationale for why dual eligibility and Area Deprivation Index were not included in the risk model. For future submissions, the developer may consider whether the measure can identify differences in performance across different patient groups by their social and/or economic status.	

Use & Usability	
Staff Preliminary Rating: Not met but addressable	
Current or Planned Use:	Measure is currently used in the ASC Reporting Program.
Use & Usability: The measure is currently used in the ASC Reporting Program. The developer cites evidence of interventions that can improve hospital visits, as noted previously. The developer summarized that stakeholders may submit questions about the measure to CMS via an online tool. The developer notes that since re-endorsement in 2020, there have been no major reports or issues with the measure. The developer did not report any findings on the progress on improvement since the procedures are “ever-changing,” making it challenging to accurately assess year-over-year performance. The developer reports no unexpected findings.	

Public Comment[3]

Number of Comments Received: 4

Full text of developer/steward responses can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
One comment shared personal experiences of needing to visit a hospital post-surgery and emphasized the importance of such follow-up visits.	Supportive	Thank you for your support of the post-procedure hospital visit measures [CBE 3470, 3357, 2539, and 3366]. You are correct that the outcome for these measures includes ED visits, observation stays, and inpatient admissions, 7 days after a procedure (which varies based on the measure).
Two comments opposed endorsement of this measure, criticizing its risk adjustment model, low reliability, and a lack of business case.	Non-supportive	We thank Dr. Miller for his comments. Hospital Visits After Urology Ambulatory Surgical Center Procedures (hereafter, ASC Urology) measure (CBE #3366) is a reliable and valid measure that was endorsed previously in 2019 under largely the same criteria as the current Battelle process. The measure objectively meets all of Battelle's criteria for reliability and validity. We note that Dr. Miller submitted the same list of concerns for the HOPD Surgery measure, which was endorsed by this Cost & Efficiency Recommendation/Advisory Groups in the prior (Fall 2023) cycle. This ASC Urology measure was developed using the same approach as the HOPD Surgery measure—it has the same 7-day unplanned hospital visit outcome, the same risk-adjustment approach, the same measure score calculation approach, and the same overall goal in improving patient care. The ASC Urology measure is attributed to the facility and captures the outcome of hospital visits (ED visit, observation stay, or inpatient hospitalization) for urologic procedures performed in ASCs. It was first endorsed in 2019 by the NQF Admissions and Readmissions Committee. The measure was developed with full transparency and public input, subject to several rounds of public comment (during measure development, rulemaking, and CBE endorsement), and broad stakeholder input. The current Battelle CBE criteria are largely the same as the prior NQF criteria, and the measure has not changed substantially since first endorsed. Importantly, the measure's risk

[3] Comments, as submitted, can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
		model continues to show good discrimination and calibration, which addresses many of the commenter's concerns about adequate risk adjustment. Yale/CORE's empiric approach to risk model development has ensured that patient-level variables that are significantly associated with the outcome are included in the model. Model testing shows that observed outcomes closely track predicted outcomes across a wide range of risk deciles. While we appreciate all the commenter's concerns, overall, the ASC Urology measure is a reliable and valid measure of post-procedural outcomes and has support from the very surgeons that perform these procedures. To facilitate quality improvement, the facility receives, in addition to its measure score, detailed, claim-level information about each patient who met the inclusion criteria (including if they experienced a hospital visit, the type of visit, and the diagnosis associated with the visit). It is currently the only publicly reported measure that captures post-surgical outcomes for these procedures in a health care system that spends billions of dollars each year on outpatient procedures and where procedural volume is shifting to the outpatient setting. Please see our full response to Dr. Miller's comment that addresses each of his concerns. Using 2 years of performance data, the median facility-level reliability score is 0.720 (IQR, 0.573-0.849) for ASCs with at least 35 procedures (the public reporting threshold) representing moderate reliability. At the current public reporting threshold, most facilities (about 75% of facilities with at least 35 cases) fall above the 0.6 minimum threshold stated in Battelle's current CBE guidebook. If CMS were to increase the case volume minimum so that all facilities exceeded this threshold, it would remove publicly available information from about 25% of facilities that are currently publicly reported. We believe that median reliability of 0.6 (signal to noise) is sufficiently high for a facility-level publicly reported measure in a pay-for-reporting program. Increasing the minimum case volume has the tradeoff of removing more than half of facilities from reporting to the public.
One comment questioned the rationale behind setting the age limit at 65 and older and	N/A	The measures are limited to Medicare patients because they are based on claims data (billing data) that are only nationally available and validated for Medicare patients. In the future we may be able to

Comment Summary	Support Level	Summary of Developer Response
proposes expanding the age range to include younger patients, potentially in their 50s.		add Medicare Advantage patients, and there may be a future possibility of a Medicaid measure, but it is not possible for us to report nationally across all payers (private and public) due to the lack of a comprehensive, national all-payer database.

Advisory Group Feedback

Full text of developer/steward responses can be found on the PQM website.

Feedback/Questions	Summary of Developer Response
Similar Comments to #3357: Several committee members expressed the same concerns as noted with CBE #3357.	Please refer to the developer's responses for CBE #3357. *Summary Response Received after the Advisory Group meeting:* The all-cause outcome is supported by the short outcome time window (7 days after the procedure, when most procedural-related complications occur) and evidence submitted by the developer that most post-procedural hospital visits are due to a complication. Low outcome rates are mitigated by (1) minimum case volume thresholds for public reporting; (2) the statistical method that pulls measure scores for small facilities to the mean; and (3) that the measure is used in a pay-for-reporting program, not for pay-for-performance. As higher-risk procedures continue to migrate to the outpatient setting, this measure retains its importance in monitoring performance in the ASC setting, where collectively these publicly reported ASC measures assess outcomes for the millions of procedures that are performed at ASCs annually.
Higher Complication Rates: One committee member noted that there is a higher complication rate seen with this measure (4-6%). The committee member expressed feeling more comfortable with this measure compared to CBE #3357. In addition, the committee member noted that the first 25 complications on the list for this measure cover 50% of the admissions. The committee member said they thought this looks like a good measure, as those complications look to be related to the urology procedure.	The developer did not have a response to this comment during the meeting. *Summary Response Received after the Advisory Group meeting:* The unadjusted and adjusted outcome rates for the urology measure are higher than for the general surgery measure, which is expected (and consistent with the literature) due to the nature of the procedures performed within these cohorts. Together, the suite of ASC measures (general surgery, orthopedic, urology) captures this important outcome for procedures increasingly performed in the ASC setting. Hospitals receive claims-level information for all three measures to support quality improvement, and in all cases, there are no other outcome measures that capture this important and patient-sensitive outcome measure.

Feedback/Questions	Summary of Developer Response
Question Regarding the Dry Run: A committee member stated that the measure went through a "dry run" in 2018 to help educate measure entities. Currently there are no other publicly available quality reports related to these elective procedures. The committee member agreed the measure could underscore a measurement gap, but without understanding if there is an issue, the committee member wanted to know why time and money should be spent to measure this?	The developer clarified that the dry run was intended to provide facilities with their information before it went public. These measures first go through a dry run and then are publicly reported. This gives the facilities an idea of what's happening before the information is made public. The developer added that the measurement gap during the dry run for this measure is like what it is now. With some of the other measures that have been out there longer, there has been improvement. For instance, when looking at CBE #2539 (colonoscopy), CBE #3470 (orthopedic), CBE #3357 (urology), and CBE #3366 (general surgery), the colonoscopy measure has been out there the longest and there has been improvement over time. The remaining three measures have not been publicly reported for as long. The developer further noted that the intent for these measures was to capture all the procedures in ASCs but not in one measure. Thus, there is the orthopedic measure, the urology measure, colonoscopy, and the general surgery measure, which captures the rest. The general surgery measure has the lowest-risk procedures. *Summary Response Received after the Advisory Group meeting:* Quality and measurement gaps are established before measure development (as was the case for this measure). A dry run (also known as confidential reporting) comes just as the final measure is getting ready for public reporting and allows facilities to see and understand their results prior to public reporting.

Key Discussion Points:

- **All-Cause Outcome:** Does the committee have any additional concerns regarding the all-cause outcome, considering the developer's inclusion of the top reasons for a hospital visit being related to the procedure?

- **Bias of the Outcome:** Does the committee have any concerns regarding the developer's response to the Advisory Group concerns regarding the lack of urgent care or office visits being captured in the measure?

- **Low Outcome Rates:** The developer responded to this Advisory Group concern that low outcome rates are mitigated by (1) minimum case volume thresholds for public reporting; (2) the statistical method that pulls measure scores for small facilities to the mean; and (3) that the measure is used in a pay-for-reporting program, not for pay-for-performance. Does the committee have any additional concerns?

- **Reliability:** Considering the developer's response to Advisory Group concerns and that greater than 70% of reliability estimates are above the expected value of 0.6, does the committee agree that the measure has sufficient reliability?

- **Validity:** Does the committee have concerns with the volume-outcome validity testing?
- **Use and Usability:** Does the committee agree that the measure results can be used by ASCs to improve their scores over time?

CBE #3470: Hospital Visits after Orthopedic Ambulatory Surgical Center Procedures [Yale CORE/CMS]

Measure Description:

This measure was developed to improve the quality of care delivered to patients undergoing orthopedic procedures in an ambulatory surgical center (ASC). To assess quality, the measure calculates the risk-standardized rate of acute, unplanned hospital visits within seven days of qualified orthopedic surgeries or procedures performed at an ASC among Medicare fee-for-service (FFS) patients aged 65 years and older. An unplanned hospital visit is defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission.

Measure Status	
New or Maintenance: Maintenance	**Used in An Accountability Application?** Yes • Public Reporting • Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
CBE Endorsement Status: Endorsed **Last Endorsement Review Cycle:** Fall 2018	**Proposed/Planned Use:** Ambulatory Surgical Center Quality Reporting (ASCQR) Program

Measure Characteristics

Measure Type	Target Population(s)	Level of Analysis	Care Setting(s)
Outcome	Medicare FFS patients aged 65 years and older who have undergone an orthopedic procedure at an ASC	Facility	Ambulatory Surgery Center

Measure Overview
Rationale: The rationale for public reporting of the ASC Orthopedic Surgery remains imperative, given the significant growth in ASC utilization; in 2018, there were more than 23 million ambulatory surgeries performed at ASCs in the United States (Young et al., 2021). As ASCs increasingly become the preferred choice for lower-risk surgeries, including orthopedic procedures, evaluating postoperative outcomes, such as unplanned hospital visits, becomes pivotal for ensuring quality care and patient safety. This rising trend in orthopedic procedures at ASCs underscores the importance of incentivizing ASCs to address preventable complications and acute care needs, thus fostering continuous quality improvement in outpatient surgical care (Lopez et. al., 2021).
Numerator: The measure defines the outcome as any (one or more), all-cause, unplanned hospital visit within seven days of an outpatient orthopedic procedure; a hospital visit includes any emergency department (ED) visit, observation stay, or unplanned inpatient admission occurring within seven days after the ASC procedure.

Denominator: The target population for this measure is Medicare FFS patients aged 65 years and older undergoing selected outpatient orthopedic surgeries, typically performed by an orthopedist, at ASCs.
Exclusions: The measure excludes surgeries for patients without 7 or more days of continuous enrollment in Medicare FFS Parts A and B after the surgery. The measure excludes these patients to ensure all patients have full data available for outcome assessment.
Measure is Risk-Adjusted and/or Stratified: In brief, the approach simultaneously models two levels (patient and facility) to account for the variance in patient outcomes within and between facilities (Normand & Shahian, 2007). At the patient level, the model adjusts the log-odds of hospital visits within 7 days after the procedure for selected demographic, clinical, and procedure risk variables. The second level models the facility-specific intercepts as arising from a normal distribution. The facility intercept, or facility-specific effect, represents the ASC contribution to the risk of 7-day hospital visits, after accounting for patient risk and sample size, and can be inferred as a measure of quality. If there were no differences among ASCs, then after adjusting for patient risk, the facility intercepts would be identical across all ASCs.

Logic Model
Summary: The conceptual model for outpatient orthopedic surgery quality at an ASC, shown below, shows the pathway by which facilities can modify the outcome (all-cause, unplanned hospital visits within 7 days of the orthopedic ASC procedure). For example, the model identifies that patient level factors, such as comorbidities, increase the risk of unplanned hospitals visits (Bongiovanni et al., 2021). Better management of the risk associated with these comorbidities may be a potential avenue for facilities to reduce unplanned hospital visits. Provider-level factors (technical quality of the procedure, post-procedure provider accessibility), and facility-level factors (such as patient selection/risk assessment, pre- and post-discharge patient communication) may also contribute to the risk of unplanned hospital visits. Therefore, facilities may have opportunities to lower their unplanned hospital visit rates through quality improvement efforts focused on patient, provider, and facility factors.

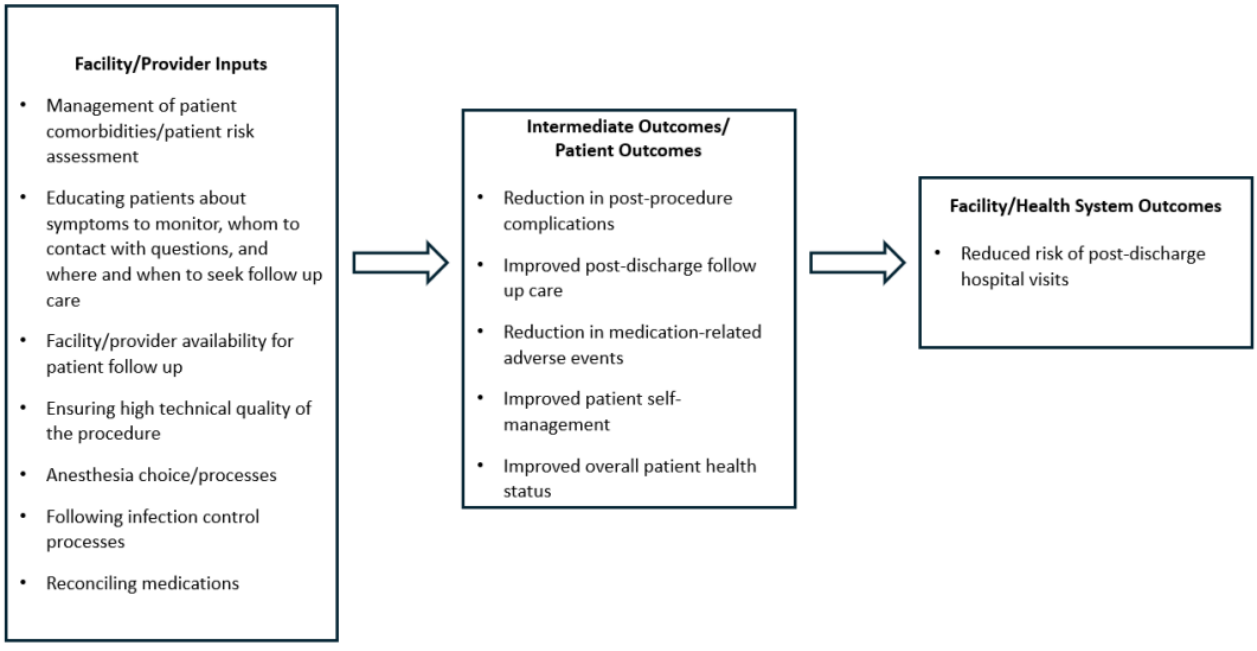

Figure 5. CBE #3470 Logic Model

Measure Evaluation Summary: CBE #3470

Importance
Staff Preliminary Rating: Met
Importance: The developer cites evidence of interventions ASCs can implement to improve hospitalizations, including preoperative assessments, enhancing postoperative monitoring, and improving patient education. The developer further suggests facilities can implement a system of quality improvement and use CMS facility-based and claims-based reports for quality improvement efforts. Using data from January 1, 2021, through December 31, 2022, the developer reported a distribution of scores with an overall score of 2.20 (mean facility-level rate) and a decile range of 1.88-2.62. Further, the developer cites evidence in which patients expressed concern over postoperative complications.

Feasibility
Staff Preliminary Rating: Met
Feasibility: The developer did not conduct a feasibility assessment of missing data for this measure, as it used a 100% sample of paid Medicare claims. The developer further states that because this is a claims-based measure, there is no added burden of reporting as all required data elements are routinely generated. The measure does not include any proprietary information.

Reliability	
Staff Preliminary Rating: Met	
Testing Level:	Accountable Entity Level
Testing Method:	Reliability testing was conducted using the **signal-to-noise method** calculated on 2 years of data (1/21-12/22) for the risk-adjusted measure across 1,754 entities and results show >0.6 for about 75% of the entities.
Reliability: The measure is well defined. Reliability was assessed at the entity level. Reliability statistics are above 0.6 for about 75% of the entities.	

Validity	
Staff Preliminary Rating: Not met but addressable	
Testing Level:	Accountable Entity Level
Testing Conducted:	Yes, validity testing was conducted using three methods: **(1) face validity during measure development, (2) validity through association with volume, and (3) validation of the outcome.**
Validity: The developer established face validity through a technical expert panel (TEP), where 92.3% of respondents affirmed the measure's usefulness and potential to improve the quality of care. The measure's outcome validity was supported by a claims-based analysis showing that the top 25 principal diagnosis codes associated with the outcome often indicated complications from the procedure. The developer conducted	

Validity
an association analysis between facility volume and the outcome, hypothesizing a weak-to-moderate negative correlation. The results were statistically significant, at -0.133, suggesting that higher procedural volumes might correlate with better outcomes. However, the relevance of procedural volume to the measure's validity remains unclear, as it was not part of the measure's logic model and its impact on quality is not well-defined. The developer conducted statistical risk adjustment and explored social risk factors such as dual eligibility and the Area Deprivation Index, which were not included in the final models due to minimal impact. The model demonstrated good discrimination with a C-statistic of 0.675.

Equity	
Staff Preliminary Rating: Not met but addressable	
Equity considered:	Yes
Equity: The developer reports that ASCs do not serve a high proportion of patients with social risk factors, with the median facility proportion of patients with the dual eligibility variable being 1% (count of 1 patient) and for the high ADI, 2% (count of 1 patient). The developer did not identify whether the measure rates are different across other patient populations or how this measure may help to reduce disparities. For future submissions, the developer may consider whether the measure can identify differences in performance across different patient groups by their social and/or economic status.	

Use & Usability	
Staff Preliminary Rating: Not met but addressable	
Current or Planned Use:	Measure is currently used in the ASC Reporting Program.
Use & Usability: The developer cites evidence of interventions that can improve hospital visits, as noted previously. The developer further suggests facilities can implement a system of quality improvement and use CMS benchmarking reports, which include the principal diagnosis on admission, to tie quality improvement efforts to complications that are occurring. The developer summarized the feedback approach for this measure, noting that since re-endorsement in 2020 there have been no major reports or issues with the measure. The developer did not report any findings on the progress on improvement because the procedures are “ever-changing,” making it challenging to accurately assess year-over-year performance. The developer reports no unexpected findings.	

Public Comment[4]

Number of Comments Received: 5

Full text of developer/steward responses can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
Two commenters expressed personal experiences and enthusiasm for the measure's potential to lead to better outcomes for patients undergoing similar procedures.	Supportive	Thank you for your support of this measure that captures hospital visits (ED visit, observation stay, inpatient admissions) following an orthopedic procedure performed at an ASC, which includes knee replacement surgery. Thank you for your support of the post-procedure hospital visit measures [CBE 3470, 3357, 2539, and 3366]. You are correct that the outcome for these measures includes ED visits, observation stays, and inpatient admissions, 7 days after a procedure (which varies based on the measure).
Two commenters opposed endorsement of this measure, criticizing its risk-adjustment model, low reliability, and a lack of business case.	Non-supportive	We thank Dr. Miller for his comments. Hospital Visits after Orthopedic Ambulatory Surgical Center Procedures (hereafter, ASC Orthopedic Surgery) measure (CBE #3470) is a reliable and valid measure that was endorsed previously in 2019 under largely the same criteria as the current Battelle process. The measure objectively meets all of Battelle's criteria for reliability and validity. We note that Dr. Miller submitted the same list of concerns for the HOPD Surgery measure, which was endorsed by this Cost & Efficiency Recommendation/Advisory Groups in the prior (Fall 2023) cycle. This ASC Orthopedic Surgery measure was developed using the same approach as the HOPD Surgery measure—it has the same 7-day unplanned hospital visit outcome, the same risk-adjustment approach, the same measure score calculation approach, and the same overall goal in improving patient care. The ASC Orthopedic Surgery measure is attributed to the facility and captures the outcome of hospital visits (ED visit, observation stay, or inpatient hospitalization) for orthopedic procedures performed in ASCs. It was first endorsed in 2019 by the NQF Admissions and

[4] Comments, as submitted, can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
		Readmissions Committee. The measure was developed with full transparency and public input, subject to several rounds of public comment (during measure development, rulemaking, and CBE endorsement), and broad stakeholder input. The current Battelle CBE criteria are largely the same as the prior NQF criteria, and the measure has not changed substantially since first endorsed. Importantly, the measure's risk model continues to show good discrimination and calibration which addresses many of the commenter's concerns about adequate risk adjustment. Yale/CORE's empiric approach to risk model development has ensured that patient-level variables that are significantly associated with the outcome are included in the model. Model testing shows that observed outcomes closely track predicted outcomes across a wide range of risk deciles. According to Battelle's staff assessment, this measure does meet reliability criteria. Using 2 years of performance data, the median facility-level reliability score is 0.759 (IQR, 0.605-0.869) for ASCs with at least 35 procedures (the public reporting threshold) representing moderate reliability. At the current public reporting threshold, most facilities (about 75% of facilities with at least 35 cases) fall above the 0.6 minimum threshold stated in Battelle's current CBE guidebook.
One commenter questioned the rationale behind setting the age limit at 65 and older and proposes expanding the age range to include younger patients, potentially in their 50s.	N/A	The measures are limited to Medicare patients because they are based on claims data (billing data) that are only nationally available and validated for Medicare patients. In the future we may be able to add Medicare Advantage patients, and there may be a future possibility of a Medicaid measure, but it is not possible for us to report nationally across all payers (private and public) due to the lack of a comprehensive, national all-payer database.

Advisory Group Feedback

Full text of developer/steward responses can be found on the PQM website.

Feedback/Questions	Summary of Developer Response
Similar Comments to #3357: Several committee members expressed the same concerns as noted with CBE #3357.	Please refer to the developer's responses for CBE #3357. *Summary Response Received after the Advisory Group meeting:*

	Please refer to the developer's responses for CBE #3357. The orthopedic measure cohort shift (the addition of total hip and total knee arthroplasty) underscores the importance of measures that will track outcomes for procedures that will migrate to the outpatient/ASC setting in the future.
Stratification by Type of Procedure: One committee member commented that because more orthopedic procedures are moving to the ambulatory setting and if a facility has a high complication rate, some of the smaller, lower-risk procedures will be hidden amongst some of the higher-risk procedures. Thus, there might be some value to stratifying based upon procedure risk.	The developer agreed with this point, noting that right now, there may not be enough data to be able to stratify, but hopefully in the future there will be. There is a hip and knee orthopedic measure, which is more focused on particular complications such as infections and mechanical malfunction. However, once the volume of procedures is high enough, it will be possible to stratify by those procedures. *Summary Response Received after the Advisory Group meeting:* Yale/CORE, the developer, agrees that it would be ideal to be able to measure outcomes for individual procedures or smaller groups of procedures; however, given that the volume-specific procedures are likely not yet large enough to split them off into a separate measure, they are currently grouped at a higher level (such as orthopedic, urology).
Clinical Grouper: One committee member asked for clarification on whether the measure would be updated to version 24 or version 28 of the CMS hierarchical condition categories.	The developer noted that they update every year, and there is a lag. The developer will confirm that they have updated to version 24. *Summary Response Received after the Advisory Group meeting:* Yale/CORE confirms that we used version 24 of CMS's Hierarchical Condition Categories for the version of the measure that was submitted for endorsement maintenance.

Key Discussion Points:

- **All-Cause Outcome:** Does the committee have any additional concerns regarding the all-cause outcome, considering the developer's inclusion of the top reasons for a hospital visit being related to the procedure?

- **Bias of the Outcome:** Does the committee have any concerns regarding the developer's response to the Advisory Group concerns regarding the lack of urgent care or office visits being captured in the measure?

- **Low Outcome Rates:** The developer responded to this Advisory Group concern that low outcome rates are mitigated by (1) minimum case volume thresholds for public reporting; (2) the statistical method that pulls measure scores for small facilities to the mean; and (3) that the measure is used in a pay-for-reporting program, not for pay-for-performance. Does the committee have any additional concerns?

- **Reliability:** Considering the developer's response to Advisory Group concerns and that greater than 70% of reliability estimates are above the expected value of 0.6, does the committee agree that the measure has sufficient reliability?

- **Validity:** Does the committee have concerns with the volume-outcome validity testing?

- **Use and Usability:** Does the committee agree that the measure results can be used by ASCs to improve their scores over time?

CBE #3495: Hospital-Wide 30-Day, All-Cause, Unplanned Readmission Rate (HWR) for the Merit-Based Incentive Payment System (MIPS) Eligible Clinician Groups [Yale CORE/CMS]

Measure Description:

This measure is a re-specified version of the hospital-level measure, "Hospital-Wide All-Cause, Unplanned Readmission Measure" (NQF #1789), which was developed for patients who are 65 years or older, are enrolled in Fee-for-Service (FFS) Medicare and are hospitalized in non-federal hospitals. This re-specified measure attributes hospital-wide index admissions to up to three participating MIPS Eligible Clinician Groups ("providers"), rather than to hospitals. It assesses each provider's rate of 30-day readmission, which is defined as unplanned, all-cause readmission within 30 days of hospital discharge for any eligible condition. The measure reports a single summary risk adjusted readmission rate (RARR), derived from the volume-weighted results of five different models, one for each of the following specialty cohorts based on groups of discharge condition categories or procedure categories: surgery/gynecology; general medicine; cardiorespiratory; cardiovascular; and neurology.

Measure Status	
New or Maintenance: Maintenance	**Used in An Accountability Application?** Yes • Public Reporting • Payment Program • Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
CBE Endorsement Status: Endorsed **Last Endorsement Review Cycle:** Fall 2019	**Proposed/Planned Use:** Merit-based Incentive Payment System (MIPS), part of the Quality Payment Program (QPP).

Measure Characteristics

Measure Type	Target Population(s)	Level of Analysis	Care Setting(s)
Outcome	Patients 65 and older enrolled in FFS Medicare and Hospitalized in non-Federal Hospitals	Clinician: Group/Practice	Clinician Office/Clinic; Hospital: Inpatient

Measure Overview
Rationale: The Hospital-Wide 30-Day, All-Cause, Unplanned Readmission Rate for the Merit-Based Incentive Payment System (MIPS) Eligible Clinician Group (MIPS HWR) measure addresses unplanned readmissions at the clinician group level for Medicare Fee-For-Service (FFS) beneficiaries aged 65 or older. The measure is risk adjusted and based on administrative claims. This measure is a re-specified version of the hospital-level measure, Hospital-Wide All-Cause, Unplanned Readmission (consensus-based entity [CBE] #1789) and related to the Hybrid Hospital-Wide All-Cause, Unplanned Readmission measure (CBE #2879e) that is under review in this same endorsement cycle (Spring 2024). The MIPS

Measure Overview
HWR measure has the same cohort, outcome, and claims-based risk variables as CBE #2879e and promotes a systems-level approach by clinicians and a focus on high-risk conditions, such as chronic obstructive pulmonary disease (COPD) and heart failure.
Numerator: The outcome for this measure is any unplanned readmission to a non-federal, short-stay, acute-care or critical access hospital within 30 days of discharge from an eligible index admission. Planned readmissions are not counted in the outcome. In the case of multiple readmissions during the 30-day period, only one of the readmissions is counted for the outcome. If a patient is readmitted to the same hospital on the same calendar day of discharge for the same condition as the index admission, the measure considers the patient to have had one single continuous admission (that is, one index admission). However, if the condition is different from the index admission, this is considered a readmission in the measure. The measure attributes the outcome (readmission) to up to three clinician groups to account for the reality that multiple healthcare roles can influence readmissions. The following three types of clinician groups are included in the multiple attribution approach: Discharge Clinician Group, Primary Inpatient Care Provider Group, and Outpatient Primary Care Physician Group.
Denominator: Eligible index admissions include acute care hospitalizations for Medicare Fee-for-Service (FFS) beneficiaries age 65 or older at non-federal, short-stay, acute-care, or critical access hospitals that were discharged during the performance period. Beneficiaries must have been enrolled in Medicare FFS Part A for the 12 months prior to the date of admission and 30 days after discharge, discharged alive, and not transferred to another acute care facility. Admissions for all principal diagnoses are included unless identified as having a reason for exclusion. A hospitalization that counts as a readmission for a prior stay may also count as a new index admission if it meets the criteria for an index admission.
Exclusions: From the cohort we exclude admissions for patients who were: 1. Discharged against medical advice. 2. Hospitalized in a prospective payment system (PPS)-exempt cancer hospital. 3. Hospitalized primarily for medical treatment of cancer. 4. Hospitalized primarily for a psychiatric disease. 5. Hospitalized for “rehabilitation care or fitting of prostheses and adjustment devices” (CCS 254). 6. Not able to be attributed to a clinician group. 7. Not continuously enrolled in Medicare FFS Part A or B for at least 30 days following discharge from the index admission. 8. With a principal or a secondary diagnosis code of COVID-19 coded as present on admission on the index admission claim.
Measure is Risk-Adjusted and/or Stratified: To harmonize with the existing hospital-level HWR measure (CBE #1789) the same claims-based risk factors were adopted for this clinician-group measure. As described below, we first considered adjustment for clinical conditions and then examined additional risk imparted by social risk factors after the potential for greater disease burden is included in the risk model. We believe that this is consistent with the evidence that people who experience greater social risk are more likely to have more disease burden compared with those who do not. We describe both approaches below. Please see Section 1.18 for equations used to derive the measure score that include risk adjustment.

Logic Model
Summary: The MIPS HWR measure attributed to MIPS-eligible clinician groups is an adaptation of a publicly reported measure attributed to hospitals (a version of which, the Hybrid HWR measure, is going through CBE endorsement in this same cycle; CBE #2879e). The goal of the clinician-group measure (MIPS HWR) is to improve patient outcomes by providing patients and clinicians with information about

Logic Model
clinician-group level, risk-standardized readmission rates of unplanned, all-cause readmission after admission for any eligible condition within 30 days of hospital discharge. Complex and critical aspects of hospital care, such as communication between providers, prevention of, and response to, complications, patient safety and coordinated transitions to the outpatient environment, all contribute to patient outcomes. This readmission measure was developed to identify clinician groups whose performance is better or worse than would be expected based on their patient case-mix, and therefore promote quality improvement and better inform consumers about care quality

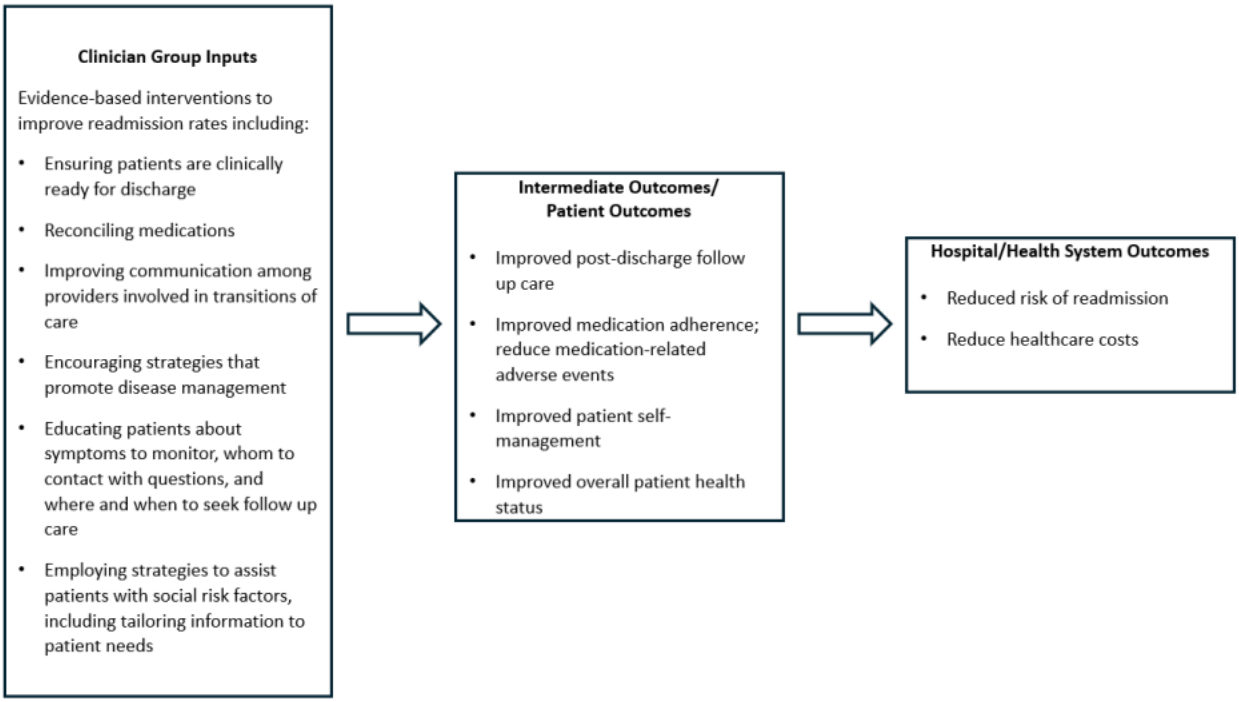

Figure 6. CBE #3495 Logic Model

Measure Evaluation Summary: CBE #3495

Importance
Staff Preliminary Rating: Met
Importance: The developer cites evidence from the primary literature in which hospital interventions that improve readmission rates, such as medication reconciliation ensuring patient readiness at discharge, improving communication between providers involved in the transition of care, and others. The developer notes that several hospital strategies can be used by groups of clinicians. Using data from January 1, 2021, through December 31, 2021, the developer reported a distribution of scores with an overall rate of 15.32% and a decile range of 13.62-17.41%. For a measure in which fewer readmissions is better, this distribution shows less-than-optimal performance for unplanned readmission, including slight variation in performance amongst providers. The developer cites interviews with patients and caregivers, who noted that readmissions resulted in confusion, frustration, and suffering.

Feasibility
Staff Preliminary Rating: Met
Feasibility: The developer did not conduct a feasibility assessment of an assessment of missing data for this measure, as it used a 100% sample of paid Medicare claims. The developer further states that because this is a claims-based measure, there is no added burden of reporting as all required data elements are routinely generated. The measure does not include any proprietary information.

Reliability	
Staff Preliminary Rating: Met	
Testing Level:	Accountable Entity Level
Testing Method:	Reliability testing was conducted using **signal-to-noise method.** Minimum reliability values are above **0.6** for all clinician groups with at least **200 patients** except for the surgical cohort, which has a mean reliability of **0.84** and a minimum of **0.49**.
Reliability: The measure is well-defined. Reliability was assessed at the entity level. Reliability statistics are above 0.6 for all entities with more than 200 patients per year, with the exception of the surgical cohort. Reliability rating is "Met" for entities with at least 200 patients. It should be specifically stated that the measure be only reported for entities with at least 200 patients.	

Validity
Staff Preliminary Rating: Met

Validity	
Testing Level:	Accountable Entity Level
Testing Conducted:	Yes, validity testing was conducted through **empirical validity testing** and by systematic assessment of measure **face validity** via a technical expert panel (TEP) of national experts and stakeholder organizations.
Validity: The developer conducted accountable entity-level validity testing by correlating the measure score to the CMS Hospital Overall Star Ratings and Hospital Star Ratings readmission group scores. The developer hypothesized a weak-to-moderate negative relationship, indicating that improved performance on this measure would result in improved Star rating scores. The developer reported weak and negative correlations, as expected. The developer also convened a TEP for face validity testing. Twelve out of 17 (70%) of TEP members agreed that the MIPS HWR measure scores are valid, useful, and can provide information for improving the quality of care. The developer also conducted statistical risk adjustment for each cohort, including patient-level risk factors that were relevant to the cohort based on the developer's conceptual model. The developer also explored social risk factors, such as dual eligibility and the Area Deprivation Index. The developer did not include these in the final models due to the minimal impact these social risk factors have on the measure scores. The developer reported C-statistics ranging from 0.63-0.68, which indicate good model discrimination.	

Equity	
Staff Preliminary Rating: Not met but addressable	
Equity considered:	Yes
Equity: The developer identified differences in unadjusted patient-level outcomes in patients with dual eligibility and Area Deprivation Index, but it is unclear if these are statistically significant. The developer mentions that these social risk factors were not included in the final risk model, due to the very weak association between measure scores and the proportion of patients with social risk among clinician groups. Because the social risk factors are not included in the model, the developer may consider doing significance testing on the differences of scores for cohorts with and without the social risk factor.	

Use & Usability	
Staff Preliminary Rating: Not met but addressable	
Current or Planned Use:	Measure is currently used in the MIPS program.
Use & Usability: The developer cites evidence of interventions that can improve readmissions, as noted previously. The developer describes various mechanisms in which feedback on the measure was obtained, including a developer-convened TEP, a public comment period, consultation from experts, and through CMS's QPP Q&A process. The developer noted that feedback received from its TEP informed a modification to the attribution approach by adopting a multiple attribution approach. The developer states that it does not have access to data that would span a time period for sufficiently comparing performance. The developer provides some results of the hospital-level readmission measure, demonstrating that improvement with this measure means that reductions in readmissions are attainable. The developer reports no unexpected findings.	

Public Comment[5]

Number of Comments Received: 5

Full text of developer/steward responses can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
Two commenters expressed support for the measure and highlighted the financial and professional impact that clinicians and health care systems might face due to unplanned readmissions.	Supportive	During measure development, CORE carefully evaluated and empirically tested six different attribution approaches and presented them to a national group of clinicians and patients (the technical expert panel). Our empirical evaluation of the selected attribution methods was comprised of analyses that allowed us to understand the implications of each approach regarding feasibility, validity, reliability, and sample size. The TEP strongly supported attributing readmissions to more than one type of clinician, and TEP input was the impetus for selecting the multiple attribution approach (to the primary inpatient clinician, discharging clinician, and outpatient clinician). We note that this measure is calculated at the physician group level, not at the individual clinician level.
One commenter strongly opposed endorsement of this measure, arguing that it is neither valid nor reliable for evaluating the quality or efficiency of care provided. The commenter outlined several significant issues with the measure's numerator, attribution methodology, and risk adjustment methodology, suggesting that these flaws could mislead patients about the quality of care provided by their physicians and unfairly penalize clinicians financially.	Non-supportive	The Hospital-Wide 30-Day, All-Cause, Unplanned Readmission Rate (HWR) for the Merit-Based Incentive Payment System (MIPS) Eligible Clinician Groups (hereafter, MIPS HWR) measure objectively meets all of Battelle's endorsement criteria for reliability and validity; Battelle's Staff Assessment notes the measure has "met" the reliability and validity criteria. In addition, the MIPS HWR measure was endorsed by NQF in 2020 under largely the same criteria as the current Battelle process. Please see Yale/CORE's full response for specific answers to concerns about the unplanned readmission outcome (well-established and widely used for decades), attribution (strongly supported by the national TEP, comprised of clinicians and consumers), risk adjustment (risk model approach and testing support validity of the model and the measure), and reliability (which meets Battelle's standards, according to Battelle's own Staff Assessment).
One commenter asked why this measure did not provide evidence, performance gap analysis, and reliability and validity testing for	Non-supportive	The ACO measure is a separate measure with a different CBE number—CBE #1789. There are three HWR measures, each under a different CBE number, due to attribution differences (the MIPS version

[5] Comments, as submitted, can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
accountable care organizations (ACOs) that are part of MIPS.		is attributed to three different types of providers), and differences in risk adjustment.
One commenter expressed significant concerns regarding the evidence cited to support the measure and its low reliability scores. The commenter urged the committee to reevaluate the measure's design and implementation, particularly focusing on a clear and evidence-based link between clinician group actions and readmission outcomes; reliability of the measure to meet or exceed the minimum standard; and risk adjustments to account for social determinants of health that may impact readmission rates.	Non-supportive	During measure development, CORE carefully evaluated and empirically tested six different attribution approaches and presented them to a national group of clinicians and patients (the technical expert panel). Our empirical evaluation of the selected attribution methods for each test measure was comprised of analyses that allowed us to understand the implications of each approach regarding feasibility, validity, reliability, and sample size. The TEP strongly supported attributing readmissions to more than one type of clinician and their input drove the selection of the multiple attribution approach (to the primary inpatient, discharging, and outpatient clinician). Regarding overall attribution to clinician groups vs. hospitals, our full response summarizes and cites studies showing that timely follow-up following a hospitalization is tied to lower readmission rates and that higher numbers of readmission-related activities are associated with lower readmission rates in primary care practices. The attending and discharging clinicians are ultimately responsible and accountable for the in-hospital care (management of patient comorbidities for example) and the discharge processes related to the risk of readmission.

Advisory Group Feedback

Full text of developer/steward responses can be found on the PQM website.

Feedback/Questions	Summary of Developer Response
Misattribution: A committee member expressed concern about the lack of a robust classification error assessment to ensure fair performance assessments. The committee member expressed that the measure may inaccurately categorize clinicians into performance deciles; this could lead to unfair payment adjustments that do not reflect the clinicians' quality of care. Additionally, the committee member noted that the measure misattributes readmissions, often assigning them to clinicians who have no control of the cause of readmission. The committee member supports holding physicians accountable but criticized the measure for not achieving this effectively. Several committee members echoed these sentiments.	The developer clarified that the measure applies to clinician groups rather than individual physicians, noting that a group must have at least 200 admissions in a year for outcomes to be attributed to the group. Regarding classification error, the developer mentioned they would discuss with their analysts what could be done. *Summary Response Received after the Advisory Group meeting:*

Feedback/Questions	Summary of Developer Response
	Reliability and classification: This measure meets Battelle's reliability thresholds (Battelle's Staff Assessment rated the measure as "met" for reliability), and these thresholds ensure acceptable minimum reliability for provider groups with at least 200 cases. Higher reliability reduces the likelihood of misclassification and in their work on misclassification, Adams et al., recommend direct measurement of reliability (https://www.nejm.org/doi/full/10.1056/NEJMsa0906323), which is what Battelle requires and what we have submitted with our endorsement maintenance submission. In addition, we currently only have access to 1 year of performance data, so we are not able to compare changes in year-over-year classification this time. We note, however, that changes in classification can represent differences in quality. Attribution: During measure development we identified and tested at least six attribution approaches (please see Appendix D of the MIPS HWR methodology report, which outlines the options that were tested and why the approaches were adopted or discarded). Because care is team-based, TEP members felt that the three-provider attribution approach was the most relevant to the way care is delivered and the fairest to providers. Each provider (the primary inpatient, discharging, and outpatient) was seen by the TEP as responsible for the care that could impact readmission. We note that this measure is now only calculated at the clinician group level, not for individual providers.
Addressing Social Risk Factors: Several committee members raised concerns about socioeconomic data and its significance in accurately measuring quality of care, independent of CMS's payment adjustments in MIPS. The committee members stressed that the measure should accurately reflect the quality of care delivered by physician groups and differentiate between factors within and outside their control, further highlighting the importance of the measure's construction beyond its application in payment programs. A committee member questioned the developer on the systemic issues concerning people with low socioeconomic resources experiencing worse outcomes. The question was whether this measure elucidates the reasons behind these disparities to effectively address and rectify them.	The developer elaborated on the analysis conducted regarding social risk factors, mentioning that they used dual eligibility and high ADI for their assessments. The findings indicated no strong correlation between social risk factors and measure scores, suggesting that these factors do not significantly impact the scores. The developer clarified that this measure is not stratified by social risk factors; however, the hospital-level measure does include such stratifications, which are reported confidentially to hospitals and may include factors such as dual eligibility and possibly the ADI. The developer explained that stratification at the hospital level helps explain how patients with and without social risk factors are treated within the same hospital and how hospitals compare in terms of treating these patients.

Feedback/Questions	Summary of Developer Response
	Within the MIPS program, adjustments to providers' payments are based on the proportion of patients with social risk factors, which is a common practice in CMS programs to accommodate providers treating patients with disparities. *Summary Response Received after the Advisory Group meeting:* Our empiric findings suggest that there is no strong impact of social risk factors on measure scores, suggesting that clinical variables in the risk model largely account for patients' social risk; the measure is not adjusted for social risk factors. In alignment with recommendations from the Office of the Assistant Secretary for Planning and Evaluation (ASPE), rather than adjust quality measures for social risk, CMS adjusts for social risk at the payment level within MIPS, where providers can earn MIPS bonus points if they care for patients with social risk factors. So that CMS and the health care system can better understand the relationship between social risk and readmission, the hospital version of the HWR measure is stratified by social risk factors and confidentially reported to hospitals as stratified.
Support for the Measure: A committee member referenced a decrease in observed readmission rates across subgroups over time, interpreting these as consistent, albeit small, improvements as an indicator of the measure's effectiveness in encouraging groups to reduce readmissions.	The developer appreciated the comment. *Summary Response Received after the Advisory Group meeting:* Yale/CORE does not have access to more than 1 year of data, so we are unable to perform an analysis to characterize any potential improvement. We therefore provided evidence for the potential for improvement in this measure by demonstrating that the hospital version of the HWR measure has shown improvement between the 2016/2017 performance year and the 2021/2022 performance year.
Mortality and Validity Implications: A committee member raised a concern about the absence of mortality assessments in the measure, noting that a hospital could have a lower readmission rate simply because it has a higher mortality rate. This gap in measurement could misrepresent a hospital's performance based on readmission rates alone without considering mortality outcomes. A committee member emphasized the importance of moving beyond face validity to assess hospital performance, specifically addressing the relationship between 30-day mortality and readmission rates. The	The developer acknowledged the issue of competing risks, particularly in the context of a short 30-day window used for assessing readmissions, where the percentage of patients dying within this period is relatively low. The developer noted they are exploring ways to address these competing risks, including coding adjustments related to end-of-life care such as DNR (Do Not Resuscitate) and palliative care. The developer confirmed that competing risks do not majorly compromise the measure's effectiveness.

Feedback/Questions	Summary of Developer Response
committee member expressed concern about patients dying who might otherwise be readmitted; therefore, analyzing the correlation between mortality and readmission could provide valuable insights.	The developer acknowledged the suggestion to analyze the correlation between mortality and readmissions. *Summary Response Received after the Advisory Group meeting:* The developer acknowledged the issue of competing risks, particularly in the context of a short 30-day window used for assessing readmissions, where the percentage of patients dying within this period is relatively low. The developer noted that they are exploring ways to address these competing risks, including coding adjustments related to end-of-life care such as DNR (Do Not Resuscitate) and palliative care. The developer confirmed that competing risks do not majorly compromise the measure's effectiveness.

Key Discussion Points:

- **Attribution:** Does the committee have any additional concerns regarding the measure's attribution, considering the developer's response?

- **Mortality:** Does the committee have any concerns regarding the absence of mortality assessments within the measure, given the developer's response regarding competing risks?

- **Use and Usability:** Does the committee agree that the measure results can be used by clinician groups to improve their scores over time?

CBE #4490: Hospitalizations for Ambulatory Care Sensitive Conditions among Home and Community Based Service (HCBS) Participants [The Lewin Group/CMS]

Measure Description:

For Medicaid HCBS participants aged 18 years and older, this measure calculates the state level observed and risk-adjusted rates of hospital admissions for ambulatory care sensitive conditions, including select behavioral health conditions, per 1,000 participants for chronic and acute ambulatory care sensitive conditions. This measure has three rates reported for potentially avoidable acute inpatient hospital admissions: chronic conditions composite; acute conditions composite; and chronic and acute conditions composite.

Measure Status	
New or Maintenance: New	**Used in An Accountability Application?** No • Public Reporting • Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
CBE Endorsement Status: N/A **Last Endorsement Review Cycle:** N/A	**Proposed/Planned Use:** N/A

Measure Characteristics

Measure Type	Target Population(s)	Level of Analysis	Care Setting(s)
Outcome	Medicaid HCBS participants aged 18 years and older	Population Geographic Ares (State)	Hospital: Inpatient;Other (Home and community-based services)

Measure Overview
Rationale: Ambulatory care sensitive conditions are a set of common chronic and acute conditions that can be treated effectively in ambulatory-care settings to prevent or minimize complications. Variations in the rate of potentially avoidable hospitalizations by state, race or ethnicity, and income indicate that performance gaps at the health-system level exist (Agency for Healthcare Research and Quality, 2018). The cost for treatment of ambulatory care sensitive conditions varies greatly across care settings. On average, the cost of treatment for an ambulatory care sensitive condition in an emergency department setting is twice that of an ambulatory care setting; inpatient care costs are much higher, averaging about four times the cost of emergency-department care (Galarraga et al., 2015). Effective management of ambulatory care sensitive conditions in outpatient and ambulatory care settings can help lower overall healthcare costs and avoid preventable hospitalizations.
Numerator: The number of potentially avoidable acute inpatient hospital admissions ambulatory care sensitive conditions include:

Measure Overview
Chronic conditions rate: Diabetes short-term complications, diabetes long-term complications, low-extremity amputation, chronic obstructive pulmonary disease (COPD), persistent asthma, hypertension, heart failure, Parkinson disease, renal disease, and seizure disorder. *Acute conditions rate:* Acute bronchitis, acute heart failure, constipation, dehydration, falls, pneumonia, complicated urinary tract infection, ketoacidosis (with or without coma), malnutrition, cellulitis, sepsis, and pressure ulcers. *Total rate:* Sum of acute and chronic composites for acute inpatient hospital admissions.
Denominator: Adults receiving Medicaid HCBS, aged 18 years and older, within each state.

Exclusions: The following statuses are excluded from the population eligible for inclusion in the denominator:
 	
• Acute hospital transfers;
• Hospice; and
• Hospitalizations for obstetric conditions.
 	
In addition, condition-specific exclusions (i.e., diagnoses that would be included within the conditions identified in the numerator) are removed from the measure numerator. These include:
 	
Chronic Conditions:
 	
• A procedure code for lower extremity amputation AND any diagnosis for diabetes.
  • Excluding any discharge with a diagnosis for traumatic amputation of the lower extremity or toe amputation procedure.
• Primary diagnosis of chronic obstructive pulmonary disease.
  • Excluding any discharge with a diagnosis for cystic fibrosis and anomalies of the respiratory system.
• Primary diagnosis of persistent asthma.
  • Excluding any discharge with a diagnosis for cystic fibrosis and anomalies of the respiratory system.
• Primary diagnosis of chronic obstructive pulmonary disease.
  • Excluding any discharge with a diagnosis for cystic fibrosis and anomalies of the respiratory system.
• Primary diagnosis of heart failure.
  • Excluding any discharges with a cardiac procedure.
• Primary diagnosis of hypertension.
  • Excluding any discharge with a cardiac procedure or diagnosis of State I–IV kidney disease with a dialysis procedure.
 	
Acute Conditions:
 	
• Primary diagnosis of pneumonia.
  • Excluding any discharge with a diagnosis of sickle cell anemia, hemoglobin S disease or procedure or diagnosis for immunocompromised state.
• Primary diagnosis of complicated urinary tract infection.
  • Excluding any discharge with a diagnosis of kidney or urinary tract disorder or procedure or diagnosis for immunocompromised state.

Measure Overview

• Primary diagnosis of cellulitis.
  • Excluding any discharge with a procedure or diagnosis for immunocompromised state.
• Primary diagnosis of pressure ulcer.
  • Excluding any discharge with a procedure or diagnosis for immunocompromised status.

Measure is Risk-Adjusted and/or Stratified: A total of 4,040,676 participants were identified in the analytic population; 3,960,577 participants were included in the testing population for risk adjustment. **Exhibit 3**, in the performance gap attachment, shows the distribution based on various characteristics (i.e., age, gender, race or ethnicity, dually enrolled status, rurality, history of prior hospitalization, HCC, and submitting state). For risk-adjustment purposes, the measure developer removed participants with unknown gender (N=36), unknown rurality status (N=44,519, 1 percent), or residing in states or territories (N=35,621, 1 percent) that were removed due to data quality issues.

Logic Model
Summary: Evidence indicates that there are approximately 3.5 million potentially avoidable hospital admissions every year, which account for $33.7 billion in aggregate hospital costs (McDermott & Jiang, 2020). The HCBS ACSC measure will help monitor rates of avoidable hospitalizations for ambulatory care sensitive conditions among HCBS participants. Appropriate primary care may also prevent the development or worsening of various chronic conditions and prevent individuals from returning to emergency or inpatient care settings for treatment. Continuity of care improvement efforts, such as increasing the average primary care visits to an optimal rate (generally three or four visits, annually, depending on health status and condition-specific needs), has been shown to reduce the risk of hospitalizations for ambulatory care sensitive conditions (Kao et al., 2019). Reduced hospital admissions and effective care coordination have the potential to contribute to healthcare cost savings as well as improve quality of care.

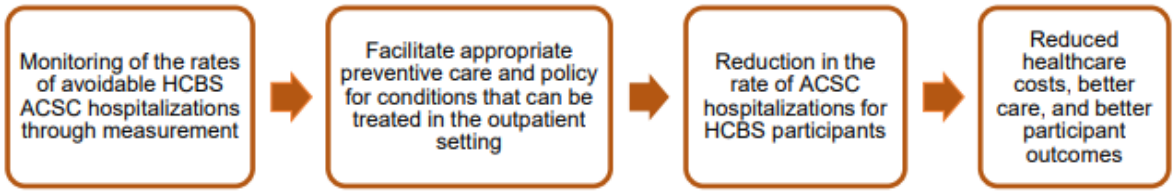

Figure 7. CBE #4490 Logic Model

Measure Evaluation Summary: CBE #4490

Importance
Staff Preliminary Rating: Met
Importance: The developer cites studies demonstrating the cost impact of potentially avoidable hospital admissions. The developer also notes that persons receiving Medicaid may be more likely to experience these admissions. Empirical data demonstrated a significant opportunity for improvement (e.g., a 30% reduction or 73,000 fewer admissions).

Feasibility
Staff Preliminary Rating: Met
Feasibility: The measure is based on administrative (claims) data. Overall, the measure is based on readily available data with minimal burden for data collection and reporting.

Reliability	
Staff Preliminary Rating: Not met but addressable	
Testing Level:	Accountable Entity Level
Testing Method:	**Signal-to-noise reliability** was conducted on **2019 data** for chronic, acute, and total admissions for both the observed rates and the expected. Reliability for observed and risk-adjusted rates is **>0.6** for all measures and all states.
Reliability: The measure is well-defined. However, reliability was assessed for the observed rates and expected rates only. It's not possible to determine the reliability of the ratio of observed to expected based on the reliability of the individual values. Reliability should be tested on the measure as it is intended to be reported.	

Validity	
Staff Preliminary Rating: Not met but addressable	
Testing Level:	Accountable Entity Level
Testing Conducted:	Yes, validity testing was conducted using **face validity,** and results reported generally supported the proposition that the measure's eligible population should include all adults receiving Medicaid HCBS.
Validity: The face validity results were equivocal. The accountable entity is the state, so the claim that needs substantiation is whether there are differences in state actions between better- and worse-performing states. Overall, the developer should provide a more robust logic model, literature review, and face validity process. The developer should also support the validity claim at the accountable entity level (state).	

Equity	
Staff Preliminary Rating: Met	
Equity considered:	Yes
Equity: Empirical data demonstrate differences in performance across gender and race. The submission might benefit from a stronger conceptual rationale for equity differences across states.	

Use & Usability	
Staff Preliminary Rating: Met	
Current or Planned Use:	Planned use in Public Reporting programs and Quality Improvement with Benchmarking
Use & Usability: For a new measure, the suggestion is that there is reason to accept the claim that no barriers exist to implementing the proposed improvements. However, that claim should be substantiated.	

Public Comment[6]

Number of Comments Received: 3

Full text of developer/steward responses can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
One comment expressed support for the measure but also expressed confusion on the title of the measure. The commenter asked for clarification on the term "Hospitalizations for ambulatory care-sensitive conditions."	Supportive	• For this measure, CMS defines “ambulatory care-sensitive conditions” as diagnoses of acute or chronic health issues that lead to potentially preventable hospitalizations if they are not treated or managed effectively in the outpatient setting (via primary or secondary care). • CMS will consider the interpretation of the measure name prior to measure implementation to ensure it is interpretable in plain language. Thank you for the feedback.
Two comments expressed support for the measure and suggested that there might be gaps in HCBS quality reporting, identifying areas for improvement, and enhancing overall	Supportive	Thank you for the feedback.

[6] Comments, as submitted, can be found on the PQM website.

Comment Summary	Support Level	Summary of Developer Response
health outcomes that this measure could help address.		

Advisory Group Feedback

Full text of developer/steward responses can be found on the PQM website.

Feedback/Questions	Summary of Developer Response
Feasibility Concerns: A few committee members raised data quality concerns. Members referenced the use of data from 2018 and 2019 in the measure specification and testing and ongoing problems with Medicaid data, raising doubts about the measure's feasibility for consistent reporting. A committee member asked about the 2–3-year data lag, and how this measure will be implemented?	The developer noted that the testing utilized data from 2021, deliberately omitting data from 2020 due to the public health emergency, which skewed the representation of care provided during that year. The developer mentioned challenges in accurately capturing data for the HCBS population and shared a plan to re-evaluate the measure using 2022 data. In response to the data lag, the developer highlighted the delays in processing national claims data for Medicare and Medicaid. While Medicare data processing has improved from 18 months to about 12 months, Medicaid data processing takes longer, approximately 2 to 2.5 years. The developer stated that the delays represent a systemic issue in using national claims data. *Summary Response Received after the Advisory Group meeting:* Administrative claims data from 2018 and 2019 represent the best picture for hospitalizations for ambulatory care-sensitive conditions when accounting for variations in standards of care during the COVID-19 public health emergency and capturing a full year of data (as 12 months of 2022 claims for Medicare and Medicaid populations are not yet available). Newer data will be used to calculate state-level performance once they are available later this year. Thank you for the feedback.
Variation in HCBS: A few members questioned the definition of HCBS used in the measure, noting it was not clearly specified. The concern was that variations in the definition and availability of HCBS across states could lead to discrepancies in access rates. This variability	Regarding the definition of HCBS, the developer emphasized that while there are differences in how states define and administer HCBS, the core populations generally include individuals who need assistance with activities of daily living and those with developmental disabilities. This measure only applies to those who are receiving

Feedback/Questions	Summary of Developer Response
could result in different outcomes in the measure due to diverse service offerings and payment rates by states. Additionally, a few committee members highlighted the range of hospitalization rates across states, noting that when risk adjustment models are applied, the variation in predicted versus actual hospitalizations narrows, and states with initially high rates of admission appear more favorable in the actual-to-predicted ratio. The question posed was whether the measure could be considered meaningful or reliable for comparing state performances given the substantial variation in HCBS programs and the moderate performance of the risk adjustment model.	services, so the number of people potentially eligible but not receiving services does not impact the measure. However, the variation in services offered by each state could influence the measure. Responding to the concern about variability and comparability of HCBS across states, the developer mentioned the introduction of the access rule, which aims to standardize and mandate reporting in the future. The developer also noted that states have discretion in how they implement their programs and which components they include. The developer expressed optimism that this new requirement would enhance the overall quality of care within Medicaid and specifically for HCBS. *Summary Response Received after the Advisory Group meeting:* Through the Ensuring Access to Medicaid Services rule, Medicaid is making substantial improvements in standardization of measurement within and across states. Future results for CBE #4490 will reflect implementation of the rule and provide a more stable definition of HCBS. Thank you for the feedback.

Key Discussion Points:

- **Feasibility:** Considering the concerns expressed regarding Medicaid data availability and the developer's response, to what extent does the committee agree or have additional concerns?

- **Reliability:** Reliability was assessed for the observed rates and expected rates only. The committee should consider this and any additional evidence from the developer.

- **Variation of HCBS:** Considering the concerns related to variation of HCBS services and the developer's response, to what extent does the committee have concerns regarding measure's potential impact and interpretation of performance?