Partnership for Quality Measurement
Powered by Battelle

National Consensus Development and Strategic Planning for Health Care Quality Measurement

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

COST AND EFFICIENCY COMMITTEE

February 10, 2025

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.se

Table of Contents

List of Tables

List of Figures

Overview of Fall 2024 Measures for Review

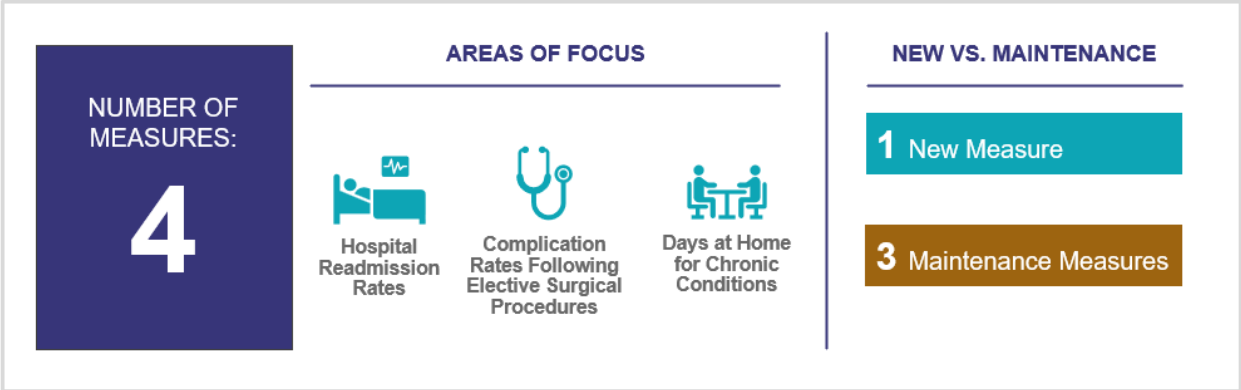
During this measure review cycle, developers and stewards submitted four measures to the Cost and Efficiency committee for endorsement consideration (Table 1). The measures focused on hospital readmission rates, complication rates following elective surgical procedures, and home care for chronic conditions (Figure 1).

Table 1. Overview of Measures Under Endorsement Review

CBE Number	Measure Title	New/Maintenance	Developer/Steward
1550	Hospital-Level Risk-Standardized Complication rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA)	Maintenance	Yale Center for Outcome Research and Evaluation/Centers for Medicare & Medicaid Services (CMS)
1891*	Hospital 30-Day, All-Cause, Risk-Standardized Readmission Rate (RSRR) Following Chronic Obstructive Pulmonary Disease (COPD) Hospitalization	Maintenance	Yale CORE/CMS
2879e	Hybrid Hospital-Wide Readmission (HWR) Measure with Claims and Electronic Health Record Data	Maintenance	Yale CORE/CMS
4555	Days at Home for Patients with Complex, Chronic Conditions	New	Yale CORE/CMS

** This measure was withdrawn from the Fall 2024 cycle by the measure steward on December 17, 2024, and will be deferred to the Fall 2025 cycle. As public comment and Advisory Group feedback were received before the withdrawal, the information is included below for transparency. However, this measure will not be discussed at the Recommendation Group Endorsement Meeting in February 2025.*

Figure 1. Fall 2024 Measures for Committee Review



Public Comment

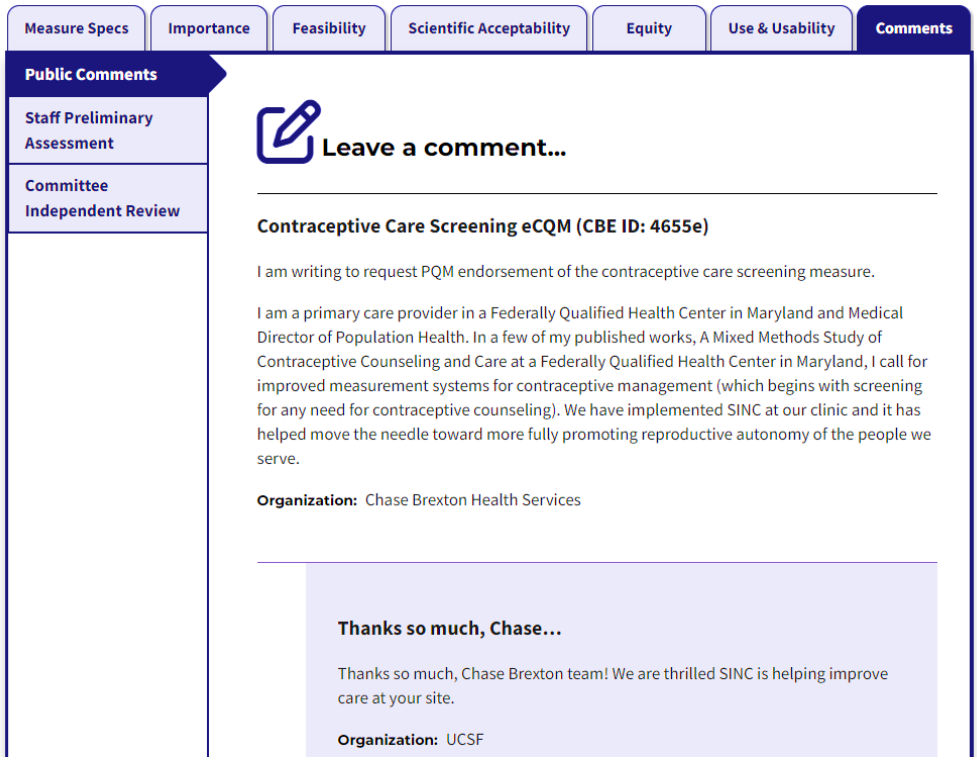
Battelle accepts comments on measures under endorsement review through the Partnership for Quality Measurement (PQM) website and Public Comment Listening Sessions. In this

evaluation cycle, the public comment period opened on November 15, 2024, and closed on December 16, 2024. Battelle held a Public Comment Listening Session on November 21, 2024.

After the public comment period closed, developers/stewards had the opportunity to respond to public comments on the measure page in the Submission Tool and Repository Measure Database (STAR). To view the public comments and responses, go to the "Comments" tab and select "Public Comments" in the left navigation pane(Figure 2). Each comment has a bold heading followed by the body of the comment. Developer responses, if any, appear as shaded replies beneath the comments. Note that developers are not obligated to respond to public comments. Lastly, the measure evaluation summaries below contain the number of public comments received for each measure

Lastly, the measure evaluation summaries below contain the number of public comments received.

Figure 2. Viewing Public Comments and Developer Responses



Advisory Group Feedback

The Advisory Group convened on December 3, 2024. Twenty of 29 (69%) active Advisory Group members attended to share feedback and ask questions regarding the measures under endorsement review. Developers/stewards of the respective measures also attended and provided responses to the Advisory Group questions. After the meeting, developers/stewards had the opportunity to submit additional written responses to Advisory Group member feedback and questions (Appendix A).

The measure evaluation summaries in this discussion guide contain overviews of the Advisory Group member discussions and developer/steward responses.

To support the review of the public comments and Advisory Group summaries, the number of comments received or number of individuals who shared similar comments, feedback, and/or questions is represented as “a few” (two to three individuals), “several” (four to six individuals), and “many” (more than six individuals). This discussion guide also employs four key categories—Supportive, Dissenting, Mixed, and Probing—to structure and enhance the Recommendation Group discussion.

- **Supportive**: This includes views and comments that express agreement, encouragement, or reinforcement of the measure.
- **Dissenting**: This captures opinions that disagree with or oppose what has been stated about the measure or what has been provided within the measure submission.
- **Mixed**: This category encompasses feedback that contains both supportive and dissenting elements.
- **Probing**: This involves questions or comments that seek to explore, clarify, or delve deeper into aspects of the measure.

Measures Under Endorsement Review

CBE #2879e: Hybrid Hospital-Wide Readmission (HWR) Measure with Claims and Electronic Health Record Data [Yale CORE/CMS]

Specifications

Measure Description: *Hybrid Hospital-Wide Readmission (HWR) Measure with Claims and Electronic Health Record Data measures facility-level risk-standardized rate of readmission (RSRR) within 30 days of discharge from an inpatient admission, among Medicare Fee-For-Service (FFS) and Medicare Advantage (MA) patients aged 65 years and older.*

Index admissions are divided into five groups based on their reason for hospitalization (e.g., surgery/gynecology, general medicine, cardiorespiratory, cardiovascular, and neurology); the final measure score (a single risk-standardized readmission rate) is calculated from the results of these five different groups, modeled separately. Variables from administrative claims and electronic health records are used for risk adjustment.

Staff Preliminary Assessment Rating[1]

Importance: Met

Rationale: This maintenance measure includes a cohort of patients admitted for a wide range of diagnoses and is intended to monitor and support quality improvement efforts for reducing 30-day readmission rates. It includes Core Clinical Data Elements (CCDEs) to improve risk adjustment and shows variation in hospital readmission rates. This measure, supported by a logic model, depicts hospital processes such as effective communication and coordinated transitions to improved health outcomes and reduced readmissions, highlighting the need for better care coordination as evidenced by patient and caregiver experiences.

Feasibility: Met

Rationale: For this eCQM all necessary data elements are consistently collected in electronic format. The CCDE used for risk are structured within EHR systems and include data transmitted from other electronic systems or derived from clinician assessments, coded using standards like RXNORM or SNOMED. These elements are captured during routine care, ensuring they do not disrupt workflow, and are complemented by readily available claims data. The feasibility of extracting these elements has been validated through extensive testing across multiple health systems, with ongoing adjustments to reporting requirements by CMS to address concerns about reporting burdens and enhance data capture capabilities in hospitals.

Reliability: Met

Rationale: The results demonstrate sufficient reliability at the accountable entity level.

Validity: Not Met but Addressable

Rationale: The developer conducted person- or episode-level validity testing and found varying agreement between EHR-based and chart-abstracted data, while accountable-entity validity testing showed expected correlations. Statistical risk adjustment methods used are appropriate

[1] Located under the “Comments” tab, then “Staff Preliminary Assessment.”

and demonstrate variation in the prevalence of risk factors across measured entities, contribute to unique variation in the outcome, and show the impact of risk adjustment for providers at high or low extremes of risk. The model performance is acceptable. Going forward, additional studies that either rule-out potential confounding (in addition to risk-adjustment) or describe features of potential mechanisms will strengthen causal claims.

Equity: Not Met but Addressable

Rationale: The developer assessed the impact of DE status and high ADI on hospital readmission measures, finding that while unadjusted rates were higher for these social risk factors, their inclusion in the risk-adjustment model had minimal impact due to high correlation coefficients near 0.999. Despite these findings, the results do not show how the scores differ across these social risk factors, which may be a future consideration for this domain.

Use & Usability: Not Met but Addressable

Rationale: The current hybrid HWR measure, including only Medicare FFS admissions, is active in the HIQR program, with plans to incorporate Medicare MA admissions. The developer provides evidence of interventions that can lead to improvements in reducing readmission rates such as enhanced discharge communication, medication counseling, and structured post-discharge care. CMS actively collects and incorporates feedback through Quality Net and annual updates, leading to significant measure enhancements including the integration of MA patients. Although no unintended impacts have been identified to patients, hospitals expressed challenges with EHR data integration and standardization, particularly in meeting IQR reporting thresholds. CMS is addressing these challenges through ongoing hospital feedback sessions, with potential protocol adjustments considered for future reporting cycles.

Public Comment

Number of Comments Received During the Public Comment Period: 4

Comments and their responses from measure developers can be found on the measure page under the "Comments" tab (*Figure 2*).

Advisory Group Feedback

Feedback/Questions	Summary of Developer Response
Feasibility, Quality Reporting Document Architecture (QRDA) File Submission, and Fast Healthcare Interoperability (FHIR): A few Advisory Group members had questions about the feasibility of the measure. An Advisory Group member asked for clarity on whether there have been measure feasibility issues with implementation and collection of electronic health record data. Another Advisory Group member indicated that they like the concept of combining an eCQM with the claims-based measure but they were concerned about the feasibility of different EHR systems being able to provide the required code set from reading the specifications. Because not all EHRs support QRDA extracts, they asked if there was any flexibility or other code sets available for this measure. Another Advisory Group member noted that any EHR certified to 170.315(c)(1) should be able to generate a QRDA. An Advisory Group member asked how CMS programs handle the measure if certain hospitals are unable to submit QRDA files (e.g., do they use the non-hybrid measure or does CMS suppress results for hospitals unable to submit these data)? Another Advisory Group member also asked if the developer considered a more modernized ability to transmit data using United States Core Data for Interoperability (USCDI) code set allowing for FHIR application programming interface (API) functionality.	The measure has been implemented, for voluntary reporting, starting with hospitalizations from 2022 through 2023. Some hospitals have struggled to meet reporting thresholds that CMS has outlined on the programmatic side and there have been ongoing discussions and changes made by CMS to give hospitals more time to have a higher level of reporting of some of the EHR information. Certified EHR systems can generate QRDA files for submission, and they have not found that to be a problem. The 2015 edition of certified electronic health record technology (CEHRT) systems can produce a QRDA file for submission. The ability to submit QRDA files is relevant across many CMS programs. Hospital participation in IQR is up to the individual facility. In the 2024 Outpatient Prospective Payment System (OPPS) Rule, CMS finalized 2 additional years of voluntary reporting for the portion of these specifications derived from EHR data to give hospitals more time to prepare for mandatory reporting. ± For hybrid measures and eCQMs, there are considerations for incorporating more USCDI elements and moving toward a FHIR platform. The Quality Data Model (QDM) and QRDA files are currently used but they will be ready when CMS makes that final determination to switch to FHIR.
Definition of Hybrid Measure: An Advisory Group member asked what it means for this to be a hybrid measure.	The measure uses two data sources: EHR data for risk adjustment and claims data to define the measure population and the outcome. Certain information that can only be obtained from the EHR (e.g., labs) provides a better picture of the severity of patients' conditions when they arrive at the hospital initially. The use of EHR data for risk

Feedback/Questions	Summary of Developer Response
	adjustment helps to balance the playing field and the case mix of patients at different hospitals when they arrive at the hospital.
Measure Results for Medicare Advantage: A few Advisory Group members asked for additional information on measure results for MA patients. An Advisory Group member inquired whether any testing had been conducted to assess whether hospitals with a higher or lower proportion of patients with Medicare Advantage have advantages or disadvantages. Another Advisory Group member expressed their happiness that the measure includes MA patients. They asked the developer to speak generally about the testing that was done regarding performance for MA patients. An Advisory Group member expressed concern that MA patients may be more likely to be put in observation status (compared with FFS patients) and asked if the developer has seen this.	The developer analyzed MA data and found they could be utilized in this measure. Findings showed some differences in risk factors for MA patients, but when those were adjusted in the risk model, readmission outcomes were very similar but slightly lower in the MA population. The developer referred Advisory Group members to their publication: Incorporating Medicare Advantage Admissions Into the CMS Hospital-Wide Readmission Measure. In this cohort study, adding MA admissions to the HWR measure was associated with improved measure reliability and precision and enabled the inclusion of more hospitals and beneficiaries. After MA admissions were included, the performance quintile changed for one in three hospitals, with the greatest shifts among hospitals with a high percentage of MA admissions. Regarding the likelihood of MA patients being put in observation status, the developer empirically found that observation rates for MA vs. FFS patients are similar. For example, for pneumonia readmissions, 5.69% of MA admissions (within 30 days following an index admission) were put in observation status, compared with 4.39% of FFS patients. Readmissions and observation stays are not mutually exclusive in the data; some of the observation stays will become inpatient admissions so even if MA admissions were more likely to begin as observation stays, the rate of inpatient admissions is even more similar between FFS and MA. Results were similar for other readmissions (e.g., acute myocardial infarction, heart failure).[±] The developer ran this analysis using the condition-specific readmission measures rather than HWR because to run this analysis they needed access to outpatient claims (observation stays and ED visits) that come from another source of data (the HWR measure only uses inpatient claims).
Use of Clinical Variables: A few Advisory Group members commented on the use of clinical variables. An Advisory Group member inquired about the utility of adding clinical variables to the measure. They noted that the versions of the measure with and without the clinical variables are highly correlated but the impact of	This measure is proof of concept and that there is interest and support from CMS to have more measures move in this hybrid direction. Stakeholder feedback indicated that having more detailed information on patients directly from the EHR tells a better picture of the status of

Feedback/Questions	Summary of Developer Response
adding clinical variables is marginal. They asked whether this measure was more of a proof of concept and if there might be potential future value to add clinical variables. Another Advisory Group member stated that if seven of 15 CCDEs are missing during the first 2 days of the admission process, the hospitalization episode is excluded from the measure. They indicated that CCDEs are basic data elements and if they are missing from the medical record that indicates poor quality of care. The Advisory Group member recommended removing this exclusion, stating that it should not be prioritized as a missing data concern but be potentially indicative of higher likelihood of readmission.	that patient and there is potential to add more clinical data elements into the measure specifications. While there are limitations to CCDE, the developer conducted analyses and received stakeholder feedback to include variables that are feasible to collect and reflect accurate measurements for those patients. While the developer agrees that these CCDEs are routinely collected on adult inpatients and reflect a standard of care, there may be many reasons for missing CCDEs on matched patients. The CCDE collection timestamp will be expanded for 2028 reporting to be the first reported during the hospital encounter, which may decrease missingness. Even if data are missing for certain patients, the measure aims to include those patients to the extent statistically reasonable so that those patient outcomes may be used in assessing hospital quality.±
Patient Involvement: A few Advisory Group members asked about patient feedback groups regarding hybrid measures and what the developers heard from patients.	The developer interviewed patients and caregivers for a technical expert panel (TEP) related to readmissions; patients and caregivers shared their stories of frustration, confusion, and suffering, as they or their loved ones faced unexpected returns to the hospital after discharge. In the interviews, patients cited experiences such as return to the hospital following exacerbation of a condition caused by changes in medication after discharge, returns to the hospital due to infection after an inpatient procedure, and other signs of poor coordination of care including insufficient communication from providers and hospital staff. Hospital readmission is disruptive to patients and caregivers and puts patients at additional risk of hospital-acquired infections and complications. Readmissions are also a major source of patient and family stress and may contribute substantially to loss of functional ability and independence, particularly in older patients (Covinsky et al., 2003). Covinsky KE, Palmer RM, Fortinsky RH *et al*. Loss of independence in activities of daily living in older adults hospitalized with medical

Feedback/Questions	Summary of Developer Response
	illnesses: *Increased vulnerability with age*. J Am Geriatr Soc 2003; 51: 451–458.
Matching Patients to Data: An Advisory Group member asked for more information about how patients are matched from the different data sets. For example, is the claims data the source of truth for the denominator patient population? They also asked for clarification on how the developer ensures successful patient matching of EHR data with the claims data to achieve valid measure results and accurate performance rates.	The four linking variables (patient identifier [Medicare Beneficiary Identifier], admission date, discharge date, and hospital identifier) are available in both the claims and EHR data, which are used to link patients. The hospitals are asked to submit both data sources for all their eligible discharges. The developer stated that the source of truth for the denominator is the claims file. From the policy side, CMS sets the threshold for the percentage of linking variables hospitals must submit to receive payment for participation in the program. The measure can be calculated with a certain amount of missing data using their imputation method (i.e. if there are CCDE missing, impute for the median value reported for that CCDE assuming a typical patient with a normal status). If there is too much missing data, that hospital would be excluded.
Correlation Between Hybrid HWR and Condition-Specific Readmission Measures: An Advisory Group member asked how this Hybrid HWR measure is correlated to the six disease-specific readmission measures.	While there is overlap, the measure populations are quite different. The Hybrid HWR measure has a broad measure population compared to the smaller, more direct measures that have more pointed patient populations. For validity testing (Table 10 in attachments), the developer compared the HWR measure against the Star Ratings readmission group score. The correlation coefficient was -0.5. The Star Ratings group score includes all the readmission measures that are implemented in CMS programs, as well as some other measures like the outpatient return to hospital visit after elective surgery measure. This is a strong correlation.
Low Readmission Rates: An Advisory Group member asked about whether hospitals can "game" the measure. Specifically, hospitals may avoid readmissions by putting patients in observation or sending them to the ED. The Advisory Group member asked if the developer knows the observation and ED visit rates among the hospitals with low 30-day readmission rates.	The developer considered ways they might be able to obtain that information. They suggested taking the existing EDAC measures, because there is no hospital-wide EDAC measure, and comparing it with the present measure.
Missing Data: An Advisory Group member asked what happens if data from one of the measure cohorts are missing. If a hospital is eliminated, is there a standard sample number being used to validate the testing?	While hospitals should submit data for all available cohorts, the minimum requirement for a hospital to receive a measure score is submission of data for patients in at least one specialty cohort. If a hospital does not submit any patients for a given cohort, a SRR for

Feedback/Questions	Summary of Developer Response
	that cohort would not be calculated. The hospital's RSRR would then be calculated based only on those cohorts reported with patients.±
Calibration Plot: An Advisory Group member requested the calibration plot comparing MA to Medicare FFS patients.	The calibration plots in Figures 1-2 of the attachments show similar performance for both MA and Medicare FFS patients. Additionally, the developer explored different modeling approaches and ultimately decided to use the combined cohort (FFS plus MA) without adjusting for FFS because the C-statistics did not change significantly, and adding an indicator term for FFS may inappropriately give an advantage to hospitals treating a greater proportion of MA beneficiaries.±
Risk Adjustment: A few Advisory Group members asked for more information related to risk adjustment. An Advisory Group member indicated that the new risk-adjustment approach merits some discussion and asked for confirmation on using CCDEs. They noted that the C-statistic for risk adjustment model of the medicine cohort meets their threshold but is still lower than other categories of patients. They inquired whether this is sufficient. An Advisory Group member stated that distribution of risk-adjusted rates is relatively narrow (10th-90th percentile is 13.8-15.0) and asked how meaningful the measure remains given this finding. An Advisory Group member called for a discussion of decision not to include social determinants in risk adjustment or stratification. They indicated the discussion of pathways is theoretical and does not fully explore this issue. They stated that research consistently shows minority-serving hospitals performing more poorly but does not address why this is the case.	C-statistics and model overfitting for all cohorts, including Medicine, were in acceptable ranges for a readmission measure. The Hybrid Voluntary Reporting dataset, the variation in RSRR between 10th-90th percentiles was 13.63% to 15.93%. However, this dataset is limited to hospitals that elect to participate in voluntary reporting (usually large, academic, teaching, non-critical-access hospitals), in which less variation in score would be expected. In the Claims-Only HWR (Medicare FFS + MA) dataset, which is nationally representative, the variation in RSRR between 10th-90th percentiles was 14.12% to 19.3%. There is meaningful variation in performance across hospitals: the-worst performing facility (RSRR 11.60%) is performing about 89% worse than the median (6.11%), while the best-performing facility (RSRR 1.52%) is performing 75% better than the median. The results for social risk factor testing examining high ADI and DE indicate that scores calculated with and without each social risk factor were highly correlated (near 1.0) and mean differences in measure scores were 0.031 and 0.055 percentage points, respectively. Odds ratios near 1.0 in each cohort for both datasets indicate no difference in the likelihood of readmission for patients with high ADI and DE. Additionally, risk adjustment for these factors may mask meaningful differences in the treatment of vulnerable populations. Testing for stratification of risk factors in the hybrid measure is ongoing.±

± The developer's full written response can be found in Appendix A.

Key Themes from Advisory Group Feedback, Public Comment, and Staff Assessments

Discussion Categories	Key Themes	Source of Comment	Summary of Comments
Dissenting	Measure Specifications	Public Comment	The Center for Healthcare Quality and Payment Reform criticizes the exclusion of certain post-discharge complications and deaths, while the Society of Hospital Medicine argues that the 30-day readmission window is too long and suggests a shorter window for more actionable feedback.
	Usability and Feasibility	Advisory Group; Staff Assessment; Public Comment	A few Advisory Group members had questions about the feasibility of the measure, specifically related to its implementation and collection of EHR data. Hospitals expressed challenges with EHR data integration and standardization, particularly in meeting IQR reporting thresholds. The American Medical Association, the Center for Healthcare Payment Report, and the Federation of American Hospitals noted significant burden and challenges with EHR data collection and alignment with clinical workflows.
	Validity	Staff Assessment; Public Comment	Additional studies that either rule-out potential confounding (in addition to risk-adjustment) or describe features of potential mechanisms will strengthen causal claims. Public comments also raised concerns about the overlap of all-cause readmissions with condition-specific measures, potentially leading to double counting in federal programs.
	Risk Adjustment	Advisory Group; Public Comment	A few Advisory Group members requested additional information on the approach to risk adjustment with particular interest in the use of social determinants and the meaningfulness of the measure based on the risk-adjusted rates. A public comment expressed that the risk adjustment is inadequate, as it does not account for socioeconomic factors.
Probing	Equity	Staff Assessment	For this optional domain, the developer evaluated social risk factors DE and ADI, finding minimal impact on adjusted measure scores, leading to the decision not to adjust for these factors, though the lack of stratification by social risk factors limits insights into outcome disparities.

CBE #1550: Hospital-Level, Risk-Standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA) [Yale CORE/CMS]

Specifications

Measure Description: *The measure estimates a hospital-level risk-standardized complication rate (RSCR) associated with elective primary THA and/or TKA procedures for Medicare patients (Fee-for-Service [FFS] and Medicare Advantage [MA]) aged 65 and older. The outcome (complication) is defined as any one of the specified complications occurring from the date of index admission to up to 90 days after the index admission. Complications are counted in the measure only if they occur during the index hospital admission or during a readmission. The complication outcome is a dichotomous (yes/no) outcome; if a patient experiences one or more of these complications in the applicable time period, the complication outcome for that patient is counted in the measure as a "yes."*

Staff Preliminary Assessment Rating

Importance: Not Met but Addressable

Rationale: This maintenance measure is supported by a logic model that clearly depicts hospital activities that can lead to desired outcomes, one of which is mitigating complications associated with THA/TKA procedures. The developer further summarizes evidence of how this measure focus area is meaningful to patients. The measure is supported by extensive literature on the need to improve care quality in this area. However, with half of patients are in the higher performance deciles, this implies that there is less variation in care quality among the providers being measured, and most are performing well.

Feasibility: Met

Rationale: The measure utilizes electronic data sources (e.g., claims), automatically generated within routine care, requires no additional data collection, and is non-proprietary, enhancing its accessibility and applicability.

Reliability: Not Met but Addressable

Rationale: There are potential issues with the accuracy of the results using the current reliability metrics. However, the identified limitations are deemed addressable, as the developer may consider performing additional reliability testing such as split-half reliability. By addressing this issue, there is potential to enhance the reliability.

To address these reliability concerns, the developer may consider refining the reliability testing methods or analytic approaches to improve the reliability assessment of the final measure.

Validity: Met

Rationale: For validity, the developer conducted both face and empiric validity testing. Face validity was assessed by a technical expert panel of 12 members from diverse backgrounds, with 91.7% affirming the measure's ability to differentiate hospital quality. Empiric testing involved correlating the measure score with hospital procedure volume, supported by research indicating that higher volumes correlate with better outcomes due to more experienced staff and standardized processes. The results showed a significant negative correlation between hospital volume and complication rates, particularly in hospitals with over 200 procedures, validating the

measure's effectiveness and the hypothesis that higher volumes lead to fewer complications. The risk-adjustment methods used are appropriate with an acceptable model performance.

Equity: Not Met but Addressable

Rationale: The analysis of the THA/TKA Complications measure showed that including social risk factors such as DE and high ADI does not significantly change the scores, indicating that existing risk model variables sufficiently account for these risks. However, the analysis did not explore how scores vary across different social risk factors, such as race, which is known to affect complication rates, suggesting an area for future consideration for this domain.

Use & Usability: Not Met but Addressable

Rationale: For maintenance, the current Medicare FFS-only measure is actively used in the HIQR and hospital value-based purchasing (HVBP) programs, with plans to incorporate Medicare MA patients. There is a clear feedback approach. However, it is unclear what, if any, feedback received led to any changes in the measure specifications. The developer also provides evidence of actions hospitals can take to improve measure performance, such as preoperative optimization, advanced surgical techniques, and rigorous postoperative care. Hospitals can also benefit from structured care models such as the Comprehensive Care for Joint Replacement (CJR) Model. Lastly, with recent changes to the measure cohort, the developer provides evidence of improvements in the FFS-only version of the measure over time and reports no unexpected findings due to the measure's use.

Public Comment

Number of Comments Received During the Public Comment Period: 2

Comments and their responses from measure developers can be found on the measure page under the "Comments" tab (*Figure 2*).

Advisory Group Feedback

Feedback/Questions	Summary of Developer Response
Risk Adjustment: A few Advisory Group members had questions regarding risk adjustment. An Advisory Group member noted that CMS decided not to adjust this quality measure for social risk factors and expressed their concerns about this approach. An Advisory Group member noted that the calibration of dual eligible patients looks less reliable than overall.	This new version of the measure, which includes Medicare Advantage, is not yet implemented (Hospital Value-Based Purchasing Program [HVPP]). Similar to measures reviewed by the Cost and Efficiency Committee during the Spring 2024 cycle, this measure is not adjusted. While it is not adjusted at the quality measure level, the payment programs may adjust the measure. For example, the HVPP is adding an adjustment for dual eligibility in the future so hospitals would not be penalized for treating a high proportion of patients with social risk factors. CMS does not adjust for social risk factors because they want to highlight differences in outcomes. Adjusting for these factors would obscure the true performance of facilities treating patients with those risk factors, making it harder to identify and address their needs. They do not wish to penalize providers already facing financial difficulties. Therefore, at the payment level, adjustments are made based on the proportion of patients with social risk factors or other variables chosen by the payment program. This way, disparities are visible and payments to the facilities are not affected. The developer assessed the impact of risk factors on this measure by correlating the measure scores when calculated with and without the risk factors. The correlation was very high, suggesting an adjustment at the quality measure level was not necessary. Regarding calibration, the calibration plots shown in Figure 4 shows that dual eligibility causes an overlap between the observed and predicted rates. This may be noise because there are not many dual-eligible patients. Referencing Table 10, the developer stated the median prevalence of patients with social risk factors for dual eligibility

Feedback/Questions	Summary of Developer Response
	is only 5%, compared to 20-23% for other measures. This suggests a problem upstream such that patients with social risk factors, who are using dual eligibility do not have access to these procedures.
Other Care Settings: An Advisory Group member asked for clarification as to whether complication rates in other care settings, such as urgent care or outpatient care, would be counted.	The developer has outpatient measures that also only look at a particular setting where the outcome occurred. In this case, they are looking only at an inpatient readmission and if a person has THA/TKA complications. Urgent care or visits to the clinician are not captured, as CMS is trying to incentivize outpatient care. The goal is for any follow-up visits to occur at the clinic level, not at the hospital due to increased cost and increased risk.
Stratification: An Advisory Group member asked if the measure is stratified to understand differences.	Though many of the other readmission measures are stratified, this measure is not.
Change in Care Settings: An Advisory Group member indicated knee and hip are off the Medicare FFS inpatient-only list for hospitals. Most of these procedures are outpatient in a hospital or an outpatient center. They asked if there is a need for this measure, because the setting of care is shifting to outpatient.	If patients are receiving the procedure inpatient, they are at higher risk. It is important to keep this measure as long as the volume (i.e., number of patients) is reasonable. The measure has a high-enough volume with good reliability, even across only 2 years. In addition, the measure is expanding to include Medicare Advantage beneficiaries. For the outpatient setting, there are other measures that capture this outcome, although not isolated to hip/knee. The developer has a hospital outpatient surgery measure and an ambulatory surgery center (ASC) orthopedic surgery measure.
Range of Variation: A few Advisory Group members shared comments about variation in performance. An Advisory Group member noted that that the reliability of the measure looks acceptable; however, they raised concern about the range of variation, noting that the performance gap looked fairly small. Another Advisory Group member noted that the original measure was developed with 2010 to 2013 data. The mean performance score was 3.3, then dropped and plateaued for several performance periods, and with the most recent data, has increased. They asked if the developer could use the measure over time longitudinally and compare rates. They also asked if the increase is more about the addition of Medicare Advantage patients or random variation.	In the section on performance gap, the developer found that there has been improvement in the Medicare fee-for-service only measure over time. Performance has improved, but it is also narrowing so the range got smaller. While there is less variation, there is opportunity for improvement. Some outliers are still an issue. The range of variation in the measure score shows a meaningful quality gap, given that the outcome rate is relatively low; small improvements in the outcome can be clinically meaningful. The performance gap is still wide enough in particular, because improvement is happening. These analyses are based on measure scores from an earlier version of the measure that did not yet capture outcomes from additional mechanical complication codes that have resulted in an increase (i.e., worsening) of measure scores due to the capture of additional complications. This, together with the new expansion of the cohort to include Medicare Advantage patients, justifies the continued implementation of this measure.±

Feedback/Questions	Summary of Developer Response
	The addition of mechanical complications makes it challenging to accurately show improvement results over time. Therefore, they truncated the comparison to before they added those mechanical complications. In a few years, they will be able to track over time.
Medicare Advantage and Dual-Eligible Patients: An Advisory Group member expressed concerns regarding the population of dual-eligible and Medicare Advantage patients. They indicated that these patients have been identified as high utilizers of health care resources and receive a lot of additional supports compared to FFS patients.	Regarding utilization, dual-eligible patients utilize more services, particularly post-acute care. The developer tested the impact of social risk factors and did not observe significant effects.
Mortality: An Advisory Group member asked for clarification about death being included as a one of the complications for the measure.	Death is an outcome for the complications measure.

± The developer's full written response can be found in Appendix A.

Key Themes from Advisory Group Feedback, Public Comment, and Staff Assessments

Discussion Categories	Key Themes	Source of Comment	Summary of Comments
Dissenting	Range of Variation	Advisory Group Staff Assessment	A few Advisory Group members had comments related to the range of variation in performance scores, noting that the performance gap is narrowing. The staff assessment also recognized that roughly half of the overall population fall in the first three deciles, which have the highest performance scores. This implies that there is less variation in care quality among the providers being measured.
	Limited Scope	Public Comment	The Center for Healthcare Quality and Payment Reform raised concern that the measure is limited in scope, as it only looks at inpatient hip or knee surgeries, excluding any outpatient and observation-treated complications.
	Reliability Testing	Staff Assessment; Public Comment	Current reliability metrics for RSRR may inaccurately reflect its true reliability, despite over 70% of entities meeting the 0.6 threshold. Addressing this issue may involve additional testing, such as split-half reliability. The American Medical Association recommends a minimum case threshold of over 25 individuals to ensure a reliability standard of at least 0.7. The Center for Healthcare Quality and Payment Reform commented that the results were poor due to small case volumes, which could lead to incorrect hospital classifications.
	Risk Adjustment	Advisory Group; Public Comment	A few Advisory Group members had questions related to the decision not to adjust for social risk factors and the calibration of dual-eligible patients. The Center for Healthcare Quality and Payment Reform

Discussion Categories	Key Themes	Source of Comment	Summary of Comments
			raised concern with the lack of social risk factor adjustment and that the measure inappropriately adjusts for Medicare Advantage beneficiaries.
	Usability – Feedback Mechanism	Staff Assessment	Although there is a process for collecting and considering feedback, the developer does not mention whether feedback received led to any changes in the measure specifications.
Probing	Equity	Staff Assessment	For this optional domain, the developer evaluated social risk factors DE and ADI, finding minimal impact on adjusted measure scores, leading to the decision not to adjust for these factors, though the lack of stratification by social risk factors limits insights into outcome disparities.

CBE #4555: Days at Home for Patients with Complex, Chronic Conditions [Yale CORE/CMS]

Specifications

Measure Description: *This is an ACO-level measure of days at home or in community settings (that is, not in acute care such as inpatient hospital or emergent care settings or post-acute skilled nursing) among adult Medicare Fee-for-Service (FFS) beneficiaries with complex, chronic conditions who are attributed to ACOs participating in the ACO REACH model. The measure includes risk adjustment for differences in patient mix across ACOs, with an additional adjustment based on patients' risk of death. A policy-based nursing home adjustment that accounts for patients' risk of transitioning to a long-term nursing home is also applied to incentivize community-based care. The performance period is one calendar year.*

Staff Preliminary Assessment Rating

Importance: Met

Rationale: This new measure aims to incentivize accountable care organizations (ACOs) to reduce excessive care in acute and post-acute settings, such as skilled nursing facilities (SNFs), by increasing the number of days patients spend at home, which has been linked to improved clinical outcomes and cost savings. It encourages ACOs to enhance care coordination and deliver timely primary and end-of-life care services, which are expected to reduce acute care needs, thus improving clinical outcomes and reducing costs. For maintenance review, the developer should consider providing additional evidence to support whether 12 days is a meaningful gap in care. What impact on health care costs and/or patient outcomes occurs with differences across deciles.

Feasibility: Met

Rationale: The measure uses routinely submitted claims data to define its cohort, risk-adjustment variables, and outcomes, maintaining its feasibility and non-proprietary nature without additional financial or administrative burdens on providers.

Reliability: Met

Rationale: The results demonstrate sufficient reliability at the accountable entity level.

Validity: Not Met but Addressable

Rationale: A new measure should at least include an argument for the validity of the data elements used in the measure specification. The data are largely from administrative or well-studied data sources, so the omission may be less problematic. Additionally, going forward, additional studies that either rule-out potential confounding (in addition to risk-adjustment) or describe features of potential mechanisms will strengthen causal claims.

The risk adjustment methods used are appropriate and demonstrate variation in the prevalence of risk factors across measured entities, contribute to unique variation in the outcome, and show the impact of risk adjustment for providers at high or low extremes of risk. The model performance is acceptable.

Equity: Met

Rationale: The developer's analysis within the ACO REACH (Realizing Equity, Access, and Community Health) Model utilized the Days in Care model and the CMS Disparity Method to address potential disparities based on Medicare-Medicaid dual-eligible status, ADI, and race. The findings revealed disparities, particularly for dual-eligible patients who experienced more days in care compared to their non-dual eligible counterparts, while non-white patients had slightly fewer days in care than white patients, and high ADI patients had more days in care than those from low ADI areas.

Use & Usability: Met

Rationale: This new measure is currently utilized within the CMS Innovation Center's ACO REACH Model. Studies referenced by the developer suggest that ACOs' investments in home-based care, preventive services, and post-acute follow-up can significantly reduce the need and costs associated with more intensive acute services. Although the measure was only implemented in 2023 and lacks sufficient data for trend analysis, no unexpected findings or concerns have been reported, and ACOs benefit from multiple feedback channels including webinars and a dedicated helpdesk.

Public Comment

Number of Comments Received During the Public Comment Period: 1

Comments and their responses from measure developers can be found on the measure page under the "Comments" tab (*Figure 2*).

Advisory Group Feedback

Feedback/Questions	Summary of Developer Response
Measure Clarifications: Several Advisory Group members had measure clarification questions: • *Measure Intent:* A few Advisory Group members requested an explanation of the measure, asking what it is really capturing and its overall intent. • *Measure Specifications:* A few Advisory Group members asked for clarification on the measure's denominator and numerator. One member noted that the patients seem aligned with the ACO REACH Model and asked if it includes fee-for-service patients A and B but not Medicare Advantage. They questioned whether the measure includes only facilities within the ACO or any hospital/facility. They also asked if a patient in the ACO REACH Model who receives care in a different state would be included in the denominator. Another member requested clarification on the numerator.	Consistent with CMS goals, consumers of care, patients, family members, and providers would like to increase use of care for patients in their home setting and in their community setting by reducing the number of days that they spend in care. To capture this, the measure counts the number of days a patient spends in a particular acute or post-acute care setting. Thus, "days at home" are defined as those days when a beneficiary is alive and not in an acute or post-acute care setting. This includes being admitted to the hospital, having an emergency department visit or an observation stay, and/or being seen at an inpatient rehabilitation facility, an inpatient psychiatric facility, a long-term care hospital, or a skilled nursing facility. A day in care, which is what the outcome or the numerator of the measure is, includes a planned or an unplanned acute care episode. Any acute or post-acute care visit a patient has within the ACO counts as 1 day in care so that there is no overlap in care. For example, if a patient is in the emergency department for 1 day, then they are admitted to the hospital on that same day, they still counted as having just 1 day in care. There are two numerator exclusions to count of days in care; a patient enrolled in hospice care and inpatient admissions related to obstetrics. The first day at home would be anything outside of the exclusions. Outpatient visits of any kind, even if performed by a provider in a hospital setting, are not counted as "days in care" but rather "days at home." The measure is intended to incentivize ACOs to facilitate and coordinate home- and community-based care to lower acute care admissions, reduce utilization and cost associated with acute care admissions for patients and providers, and improve patients' quality of

Feedback/Questions	Summary of Developer Response
	life by influencing care that allows them to spend more of their days at home. ± With respect to the measure specifications, the denominator for this measure does not include Medicare Advantage patients. The denominator includes only patients in Medicare FFS parts A and B. The denominator is for patients of ACOs that are specifically aligned to the ACO REACH Model. Alignment to ACOs is about determining which patients' outcomes each measured entity (i.e., ACO) is attributed to (that is, to which ACO a particular patient's outcomes during the measure period would be counted). The outcome itself does not depend on a "day in care" being at the aligned ACO (that is, an admission during the year to any hospital counts the same whether it is part of an ACO or not).±
Defining Home: An Advisory Group member asked the developer to define "home," specifically for patients or beneficiaries who live in nursing homes and assisted living facilities.	Being in an assisted living facility or a group home does count as a day at home, assuming that the patient is in their home and not in another care setting for any part of the day. For a patient who resides in a long-term nursing home, assuming they are not receiving any other skilled care, that is considered their place of residence or their home. The measure adjusts for an ACO's risk of transitioning their patient to a nursing home during the measurement period. The goal of this adjustment is to encourage ACOs to seek and facilitate home- and community-based care options without automatically moving a patient to a nursing home. Once a patient resides in a nursing home, that location is considered their place of residence and will count toward the total "days at home"; any other acute care admissions that meet the numerator criteria will still apply the same way to count that patient's "days in care."±
Visibility on Which Providers Are Part of An ACO: An Advisory Group patient participant asked how they would know if their hospital was a member of the ACO REACH program. They also asked if Medicare Advantage is part of an ACO.	Individuals can visit the CMS Innovation Center website for information on what ACOs are currently aligned to the model. As of their test data, 99 ACOs were participating. When they last checked there were just over 120 ACOs participating. The model is relatively new, having launched in 2021.

Feedback/Questions	Summary of Developer Response
	In response to the question about Medicare Advantage, an Advisory Group member noted that Medicare Advantage is not included in ACOs, as it is an insurance plan model, whereas an ACO is a health care delivery model, made up of various types of providers. The developer did not have a follow-up to this comment.
Validity Testing: An Advisory Group member asked if the developer had looked at the Star Ratings for validity testing or other markers of quality in the post-acute care settings. If so, how much does that correlate with the ACO's performance on a measure such as this one (CBE #4555)?	This is the first year of testing they have done with model data. The validity testing supports the measure, specifically for the relationship between the all-cause unplanned admissions and the risk-standardized all-cause readmission. The correlations between those were significant, because as admissions and readmissions increase, days at home decrease. The validity testing was guided by the other quality measures that are available specifically for ACOs participating in the REACH Model.
Limitation on Home Care: An Advisory Group member from a certified rural health clinic mentioned that there are specific guidelines on when they can visit patients at home. They send patient navigators to support patients in their homes. This should prompt discussions with Medicare about allowing providers more flexibility to visit patients at home, helping to keep them there instead of requiring hospital visits.	The measure aims to incentivize ACOs and groups of providers to utilize community resources. This way, patients can receive necessary care from a home health aide in their homes, rather than being admitted to the hospital for similar services.
Variation: An Advisory Group member indicated that the 10th-90th percentile variation of measure performance is low: 315 to 327 days. They requested more discussion of variation across ACOs and whether this measure captures significant variation.	The adjusted measure score represents days at home per person-year (unit of measurement that represents one person being observed or at risk for one year); the variation in performance between 10th and 90th percentile is 12 days per person-year, which is not small variation considering this represents a count of days at home for each patient attributed to an ACO REACH Model. The marginal value of an additional day in care is very high, and such a reduction in days in care is mathematically equivalent to the same increase in days at home. ±
Support for Measure: A few Advisory Group members expressed support for the measure, highlighting its importance to patients and communities. One member noted that the measure aligns well with the goals of population health, which focus on keeping people healthy and reducing the need for acute care.	N/A
Calibration: An Advisory Group member indicated the transition to nursing home and mortality models looks well calibrated and show	The deviance R-square is like that of other similar measures or studies. They indicated that because the intent is to account for patient

Feedback/Questions	Summary of Developer Response
acceptable C-statistics. They stated that the count models have low deviance R-square, and despite comments that these are typical of other endorsed measures, they would normally worry about noisiness of the measure. The split-sample correlation is very high, however, suggesting the measure is capturing something across providers.	risk factors, they are not trying to maximize the predictive ability of the model. The developer expects that some variation in the outcomes will be explained by the care delivered, which is what they intend to assess. ±

± The developer's full written response can be found in Appendix A.

Key Themes from Advisory Group Feedback, Public Comment, and Staff Assessments

Discussion Categories	Key Themes	Source of Comment	Summary of Comments
Supportive	Importance	Advisory Group	A few Advisory Group members shared their support of the measure as it addresses a critical area of concern for patients and their communities.
Dissenting	Measure Variation	Advisory Group Staff Assessment	An Advisory Group member indicated that the 10th-90th percentile variation of measure performance is low. The staff assessment noted that as a new measure, performance gap results are not required. However, it is unclear what quality impact would be with respect to a difference of roughly 12 days (between decile 1 and 10).
	Reliability	Public Comment	The Center for Healthcare Quality and Payment Reform raised concern with the reliability testing relies on a single split sample, without evaluating misclassification probability for individual ACOs.
	Validity	Staff Assessment	As a new measure submission, person- or encounter-level validity evidence was not provided. The data are largely from administrative or well-studied data sources, so the omission of this evidence may be less problematic. Further, the minimal variation in performance may not support a causal association between the entity and the measure focus.
	Risk Adjustment	Public Comment	The Center for Healthcare Quality and Payment Reform commented that the measure's adjustments for mortality and nursing home transition risk are inappropriate and potentially misleading, while the risk adjustment model is criticized for poor fit and lack of a comprehensive assessment.
	Limited Applicability	Public Comment	The Center for Healthcare Quality and Payment Reform noted that the measure is limited in applicability due to excluding individuals with low-risk scores and those in Medicare Advantage. The measure only includes Medicare beneficiaries attributed to REACH ACOs.

Appendix A

Following the Advisory Group meeting, developers/stewards had the opportunity to provide further written responses to feedback and questions from Advisory Group members. An abridged summary of these additional responses is presented in the discussion guide tables. The complete responses from developers/stewards, edited by Battelle staff for clarity and grammatical correctness, are included below.

CBE #1891: Full Responses Written by the Developer

Feedback/Question	Full Developer Response
Observation Status and ED Visits: An Advisory Group member noted that a challenge with readmission measures is variation in how different hospitals classify patient care (inpatient vs. observation status). Hospitals might treat patients the same but classify them differently, leading to inconsistencies in readmission reporting. They inquired if the developer had considered broadening the care setting from inpatient to include observations to address this issue. Another Advisory Group member noted that Medicare Advantage patients may be more likely to be put in observation status compared with Medicare FFS patients.	While the COPD Readmission measure captures only readmissions, CMS also has implemented measures that capture all post-discharge acute hospital-based care: the so-called "Excess Days in Acute Care (EDAC)" measures. The COPD EDAC measure, while in development, has not yet been implemented. We note that the measure has a Medicare Advantage indicator that adjusts for Medicare Advantage status. We also note that (using CY2022 data) the observed (unadjusted) readmission and observation rates for MA vs. FFS admissions are similar (readmission 18.04% vs. 18.16%, and observation 7.6% vs 6%, respectively).
Measure Target Population: An Advisory Group member inquired as to whether other lung disorders with similar symptoms, specifically bronchiectasis, are included in the measure.	While bronchiectasis can co-occur with COPD, it is a diagnosis distinct from COPD with different causes, symptoms, and treatments. For example, bronchiectasis is usually caused by infections or complications after a lung transplant, while COPD is primarily caused by smoking or exposure to lung irritants. In addition, in terms of symptoms, bronchiectasis causes cough, sputum production, and recurrent respiratory infections. COPD causes chronic respiratory symptoms, such as dyspnea, cough, expectoration, and/or exacerbations. Finally, bronchiectasis is treated with antibiotics to cure infections, while COPD is treated with anti-inflammatory drugs, such as inhaled corticosteroids. For these reasons, bronchiectasis is not included in the COPD Readmission cohort. The aim of the COPD Readmission measure is to capture COPD-related readmissions (e.g., exacerbations triggered by airflow obstruction and systemic inflammation specific to COPD), and therefore including bronchiectasis would impair the measure's

Feedback/Question	Full Developer Response
	specificity. Readmissions related to bronchiectasis are, however, captured by the Hospital-Wide Readmission measure, which consists of five clinical specialty cohorts; bronchiectasis falls within the "Medicine" cohort. Hospitals can identify these admissions because they receive patient-level data that identifies the diagnosis associated with the index admission in addition to the diagnosis associated with the readmission.
Performance Gap: An Advisory Group member indicated that the distribution of RSRRs is narrow. The min-max range is 7-8 percentage points, but that may reflect very extreme values. The interquartile range is narrow (18.1-18.9), and there is not a 10^{th}-90^{th} or 5^{th}-95^{th} range that excludes the most extreme values. They asked if there is enough variation in performance for measurement.	As we discuss in the "Usability" section, COPD readmission rates have improved. As shown in Table 14 (see "All Figures and Tables COPD Readmission" attachment), when comparing the distribution of risk-standardized scores from 2012-2015 to the most recent version of the FFS-only measure, hospital-level COPD RSRRs are lower across the entire distribution in the 2020-2023 performance period vs. 2012-2015 period. Given this improvement, we surmise that continued improvement is possible. Furthermore, the current measure now includes, for the first time, both Medicare FFS and MA patients, allowing CMS to track performance across a wider patient population. With this new measure, we see a meaningful difference in performance across hospitals. The best-performing hospital is performing 14% better than the median, and the worst-performing hospital is performing at about 32% worse than the median, showing evidence of a performance gap. We provide further evidence of variation by calculating and interpreting the median odds ratio (Merlo et al., 2006). The median odds ratio, in this context, calculates the odds of the outcome (readmission) if the same patient were treated at a higher-risk hospital compared with a lower-risk hospital. The median odds ratio for the COPD Readmission measure is 1.30. A value of 1.30 indicates that a patient has a 30% increase in the odds of readmission if they were initially admitted to a hospital for COPD at a higher-risk facility compared to a lower-risk facility. **References:** Merlo J, Chaix B, Ohlsson H, Beckman A, Johnell K, Hjerpe P, Råstam L, Larsen K. (2006) A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health, 60(4):2907.
Reliability Concerns: An Advisory Group member noted that reliability looks very low. Specifically, the signal-to-noise is never higher than 0.423, interhospital variance is only 0.02, and it seems to be measuring noise.	While we agree that the reliability is lower than ideal, we believe the measure is providing meaningful information. First, as noted in the prior response, we are seeing improvement over time, in performance (see "All Figures and Tables COPD Readmission" attachment). Second, some of the issues around reliability can be attributed to the greatly reduced admission volume that we observed at the start of the COVID-19 pandemic and which still persists. For example, admission volume (without removal of patients with a COVID diagnosis) was about 60% lower in April 2023

Feedback/Question	Full Developer Response
	compared with April 2019. Some research suggests that this decline was associated with a lower burden of non-COVID respiratory viral infections (So et al., 2021). Alternative explanations include changes in coding practices, changes in air quality, and changes in patient-level risk factors, such as smoking. It is possible that COPD volume will return to closer to pre-pandemic levels in future years, and we will be monitoring COPD admission volume annually. There are also other possible mitigation approaches to low reliability, including increasing the minimum sample size (currently 25) and increasing the number of years of data used to calculate the measure (currently 2 years). **References:** So JY, O'Hara NN, Kenaa B, Williams JG, deBorja CL, Slejko JF, Zafari Z, Sokolow M, Zimand P, Deming M, Marx J, Pollak AN, Reed RM. Population Decline in COPD Admissions During the COVID-19 Pandemic Associated with Lower Burden of Community Respiratory Viral Infections. Am J Med. 2021 Oct;134(10):1252-1259.e3. doi: 10.1016/j.amjmed.2021.05.008. Epub 2021 Jun 12. PMID: 34126098; PMCID: PMC8196237.

CBE #2879e: Full Responses Written by the Developer

Feedback/Question	Full Developer Response
Feasibility, QRDA File Submission, and FHIR: An Advisory Group member asked how CMS programs handle the measure if certain hospitals are unable to submit QRDA files (e.g., do they use the non-hybrid measure or does CMS suppress results for hospitals unable to submit these data)?	The Hybrid HWR measure is used in the Inpatient Quality Reporting program. There are several eCQMs in CMS programs that require QRDA files. As such, a hospital's ability to submit QRDA file is relevant to not only the hybrid measures but all eCQMs. The 2015 Edition CEHRT systems are able to produce a QRDA file for submission. If an entity is using a non-certified EHR system and is unable to generate a QRDA file, they have the option to partner with a vendor that is a clinical data registry that is certified through Drummond Group to compile their data and generate QRDAs for submission on their behalf.
Use of Clinical Variables: An Advisory Group member stated that if seven of 15 CCDEs are missing during the first 2 days of the admission process, the hospitalization episode is excluded from the measure. They indicated that CCDEs are basic data elements and if they are missing from the medical record that indicates poor quality of care. The Advisory Group member recommended removing this exclusion, stating that it should not	While we agree that these CCDEs are routinely collected on adult inpatients, and reflect a standard of care, there may be many reasons for missing CCDEs on matched patients: for example, if the lab or vital was outside the CCDE-collection window of 24 hours before to 2-24 hours after inpatient admission. This timestamp will be expanded for 2028 reporting to be the first reported during the hospital encounter, even if outside the 24-hour window. This criterion is used so that patients with less than the 13

Feedback/Question	Full Developer Response
be prioritized as a missing data concern but be potentially indicative of higher likelihood of readmission.	required CCDEs may still be included in the cohort, as statistical experts informed us that imputation should not be used on more than 40% missing data. Even if data are missing for certain patients, the measure aims to include those patients to the extent statistically reasonable so that those patient outcomes may be used in assessing hospital quality.
Calibration Plot: An Advisory Group member requested the calibration plot comparing MA to Medicare FFS patients.	Please see calibration plots below (Figures A1-A2), which show similar performance for both MA and Medicare FFS patients. Additionally, we explored different modeling approaches and ultimately decided to use the combined cohort (FFS plus MA) without adjusting for FFS because the C-statistics did not change significantly and adding an indicator term for FFS may inappropriately give an advantage to hospitals treating a greater proportion of MA beneficiaries.
Risk Adjustment: An Advisory Group member called for a discussion of decision not to include social determinants in risk adjustment or stratification. They indicated the discussion of pathways is theoretical and does not fully explore this issue. They stated that research consistently shows minority-serving hospitals performing more poorly but does not address why this is the case.	We conducted social risk factor testing for this measure for both the hybrid dataset and the claims only (with MA and Medicare FFS). Results show odds ratios near 1.0 in each cohort for both datasets, indicating no difference in the likelihood of readmission for patients with high ADI and DE. Additionally, measure scores calculated with and without each social risk factor were highly correlated (near 1.0) and mean differences in measure scores were 0.031 and 0.055 percentage points, respectively, for the claims-only dataset. Additionally, risk adjustment for these factors may mask meaningful differences in the treatment of vulnerable populations. We acknowledge that the previous claims-based HWR measure stratified by social risk and that current testing for stratification in the hybrid HWR measure is ongoing.

Figure A1. Calibration Plots: Deciles of Predicted v Observed Risk-Standardized Readmission Rate (RSRR) per Cohort (Medicare Fee-for-Service)

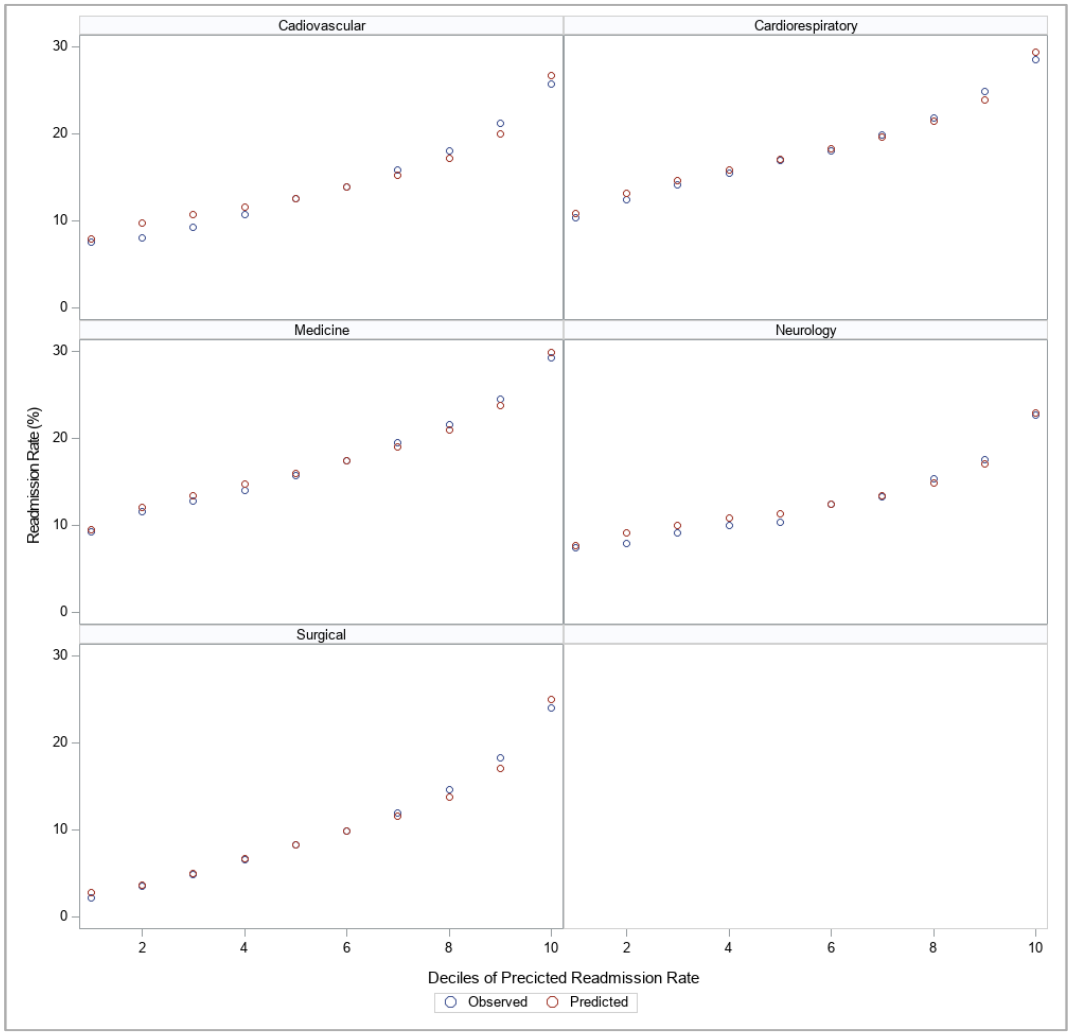
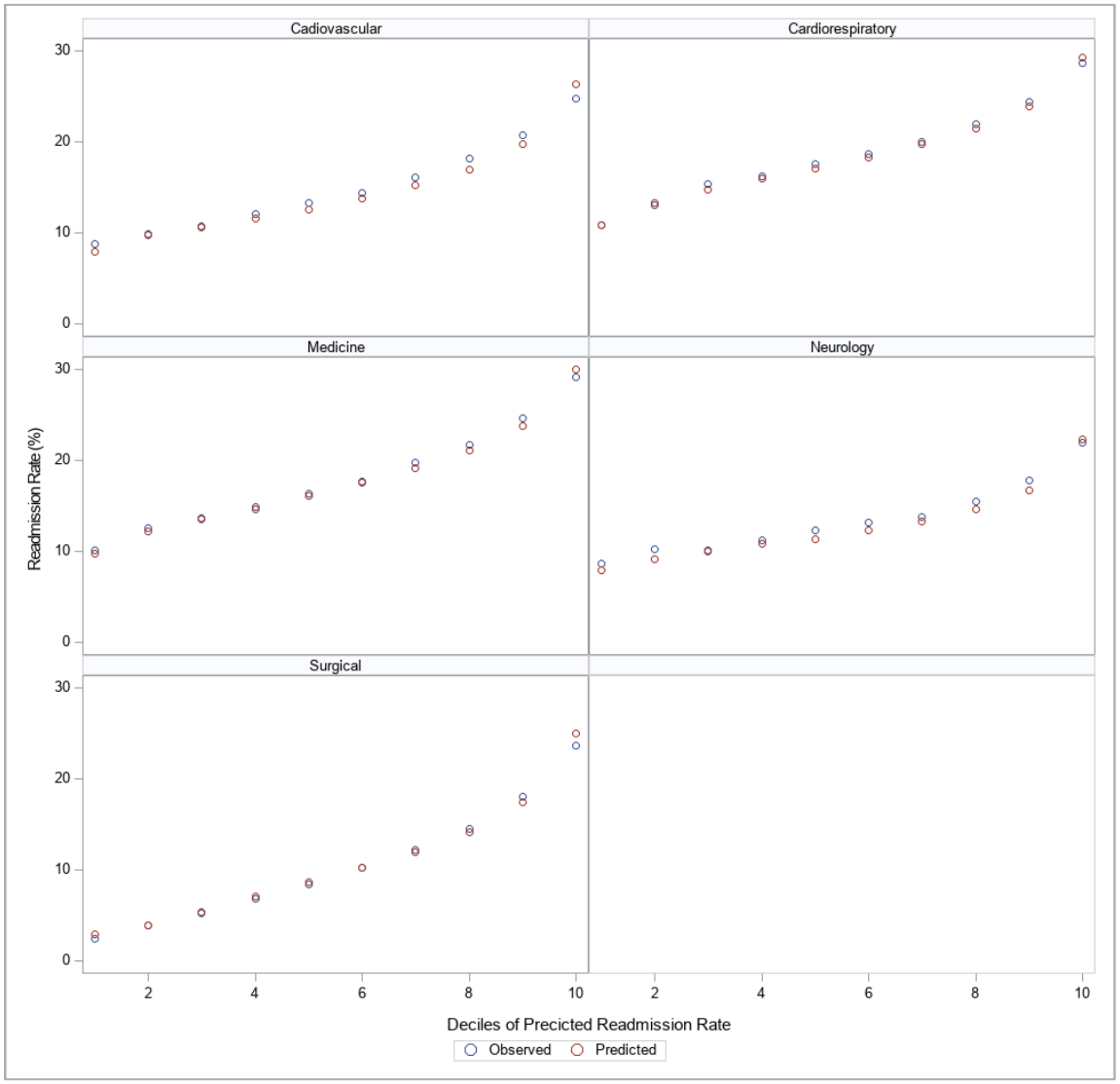


Figure A3. Calibration Plots: Deciles of Predicted v Observed Risk-Standardized Readmission Rate (RSRR) per Cohort (Medicare Advantage)



CBE #1550: Full Responses Written by the Developer

Feedback/Question	Full Developer Response
Range of Variation: A few Advisory Group members shared comments about variation in performance. An Advisory Group member noted that that the reliability of the measure looks acceptable; however, they raised concern about the range of variation, noting that the performance gap looked fairly small.	As noted during the meeting, the range of variation in the measure score shows a meaningful quality gap, given that the outcome rate is relatively low; small improvements in the outcome can be clinically meaningful. In addition, we have shown (with the FFS-only measure) that measure scores have improved over time. These analyses are based on measure scores from an earlier version of the measure that did not yet capture outcomes from additional mechanical complication codes that have resulted in an increase (worsening) of measure scores due to the capture of additional complications. This, together with the (new) expansion of the cohort to include MA patients justifies the continued implementation of this measure that captures inpatient procedures that are typically performed on patients with higher levels of comorbidities. As shown in Table 1 (see "All Figures and Tables THA TKA Complications" attachment), RSCRs of the MA+FFS measure range from 1.47% to 8.79%; the median is 3.50%; the 10th percentile is 2.91% and the 90th percentile is 4.42%. The best performer (1.47%) has a risk-standardized complication rate that is 72% better than the median (3.50%); the worst performer (8.79%) has a risk-standardized complication rate that is 2.5 times or 151% worse than the median (3.50%). We provide further evidence of variation by calculating and interpreting the median odds ratio (Merlo et al., 2006). The median odds ratio, in this context, calculates the odds of the outcome (complications) if the same patient were treated at a higher-risk hospital compared with a lower-risk hospital. The mean odds ratio for the THA/TKA Complications measure is 2.21, indicating that a patient has 2.2 times the odds (or 121% greater risk) of a complication if they were initially admitted to a hospital for THA/TKA procedure at a high-risk facility compared to a lower-risk facility, indicating meaningful variation in performance. **References:** Merlo J, Chaix B, Ohlsson H, Beckman A, Johnell K, Hjerpe P, Råstam L, Larsen K. (2006) A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health, 60(4):2907.

CBE #4555: Full Responses Written by the Developer

Feedback/Question	Full Developer Response
Variation: The transition to nursing home and mortality models looks well calibrated and shows acceptable C-statistics. The count models have low deviance R-square, and despite comments that these are typical of other endorsed measures, I would normally worry about noisiness of the measure. The split-sample correlation is very high, however, suggesting the measure is capturing something across providers.	The deviance R-square is similar to other similar measures or studies. Because the intent is to account for patient risk factors, we are not trying to maximize the predictive ability of the model. The outcomes will be affected by the care received by the patients, and ACO plays a significant role in the care of patients. We expect that some variation in the outcomes will be explained by the care delivered, which is what we intend to assess. We do not include variables that may be predictive of the outcomes but reflect the care provided by ACO.
Variation: The 10th-90th percentile variation of measure performance is low: 315 to 327 days. Would like to see more discussion of variation across ACOs and whether this captures serious variation that should be in the quality measurement portfolio.	The adjusted measure score represents days at home per person-year; the variation in performance between 10th and 90th percentile is 12 days per person-year, which is not small variation considering this represents a count of days at home for each patient attributed to an ACO REACH model. From another perspective, a difference of 12 days at home is equivalent to a difference of 12 days not in care – in other words, a patient at a 10th percentile ACO could expect to spend 12 more days in care than they would at the 90th percentile ACO. From our research and our engagement with patients and stakeholders, we understand that the marginal value of an additional day in care is very high, and such a reduction in days in care is mathematically equivalent to the same increase in days at home. In addition, ACOs are typically large organizations with hundreds to thousands of aligned patients, so even superficially small differences in score at the per-patient level can correspond to many days difference in total utilization across an ACO's entire population. Finally, we note that the outcome definition does only include acute hospital visits, inpatient admissions, and ED/observation stays; even for patients in this higher-risk cohort, the average number of such "days in care" per patient is fairly low to begin with.
Measure Clarifications Several Advisory Group members had measure clarification questions: • *Measure Intent:* A few Advisory Group members requested an explanation of the measure, asking what it is really capturing and its overall intent. • *Measure Specifications:* A few Advisory Group members asked for clarification on the measure's denominator and numerator. One member noted that the patients seem aligned	The Days at Home measure counts "days in care" to be those that are planned and unplanned acute care episodes including: acute care hospital admissions, emergency department visits, observation stays, inpatient rehabilitation facilities, inpatient psychiatric facilities, long-term care hospitals, and skilled nursing facilities. Any other visit such as an outpatient doctor/clinic visit, outpatient procedure, home health visit, assisted living facility or group home, residential psychiatric and substance abuse facilities, telehealth services, and unskilled nursing care is considered a "day at home." To reiterate, certain outpatient visits may physically occur at a hospital, but if billed as an outpatient visit, they will not count as a "day in care."

Feedback/Question	Full Developer Response
with the ACO REACH Model and asked if it includes fee-for-service patients A and B but not Medicare Advantage. They questioned whether the measure includes only facilities within the ACO or any hospital/facility. They also asked if a patient in the ACO REACH Model who receives care in a different state would be included in the denominator. Another member requested clarification on the numerator.	CORE team would like to clarify that alignment to ACOs is about determining which patients' outcomes each measured entity (the ACO) is attributed to. That is, to which ACO a particular patient's outcomes during the measure period would be counted. The outcome itself does not depend on a "day in care" being at the aligned ACO (that is, an admission during the year to any hospital counts the same whether it is part of an ACO or not). CORE also would like to clarify that the Days at Home measure was developed for a context with prospective alignment (in ACO REACH, patients are aligned to ACOs prior to the start of a performance year), so each ACO knows exactly which patients it is responsible for prospectively.
Defining Home: An Advisory Group member asked the developer to define "home," specifically for patients or beneficiaries who live in nursing homes and assisted living facilities.	The measure adjusts for an ACO's risk of transitioning their patient to a nursing home during the measurement period. The goal of this is to encourage ACOs to seek and facilitate home- and community-based care options, before automatically seeking to move a patient to a nursing home. Patients generally prefer remaining in their home or community to transitioning to a nursing home; the adjustment is in place to avoid incentivizing premature transitions to NH and to create greater incentives for ACOs to pursue better measure performance through home- and community-based care. Without this adjustment in place, a possible unintended consequence is that ACOs could be incentivized to more aggressively transition their patients to a nursing home, where daily unskilled nursing care could reduce the need for skilled or inpatient care and thus improve the Days at Home score. At times, transitioning a patient to a nursing home is necessary and appropriate for an individual, so we still expect some transitions to occur during the measurement period; once that patient has transitioned to a nursing home, that becomes considered their place of residence and their "days in care" count will continue as outlined by the numerator criteria.