
CBE ID

3502e

Title

Hybrid Hospital-Wide (All-Condition, All-Procedure) Risk-Standardized Mortality Measure with Claims and Electronic Health Record Data

Project

Management of Acute Events, Chronic Disease, Surgery, and Behavioral Health

Endorsement Status

Endorsed with Conditions

E&M Committee Rationale/Justification

When the measure returns for maintenance (5 years), the measure developer should have:

- Explored actionability of the measure with reporting entities (e.g., through qualitative data collection).

Is Under Review

No

Next Maintenance Cycle

Fall 2029

Previous Endorsement Cycle

Fall 2024

Steward

Centers for Medicare & Medicaid Services

1.0 New or Maintenance

Maintenance

1.1 Measure Structure

Single Measure

1.3 Electronic Clinical Quality Measure (eCQM)

Yes

1.6 Measure Description

Hybrid Hospital-Wide (All-Condition, All-Procedure) Risk-Standardized Mortality Measure with Claims and Electronic Health Record Data measure estimates a hospital-level 30-day risk-standardized mortality rate (RSMR), defined as death from any cause within 30 days after the index admission date for Medicare fee-for-service and Medicare Advantage patients who are

between the ages of 65 and 94.

Index admissions are assigned to one of 15 clinically cohesive and mutually exclusive divisions: six surgical divisions and nine non-surgical divisions, based on the reason for hospitalization. The surgical divisions are: Surgical Cancer (includes a surgical procedure and a principal discharge diagnosis code of cancer), Cardiothoracic Surgery, General Surgery, Neurosurgery, Orthopedic Surgery, and Other Surgical Procedures. The non-surgical divisions are: Cancer, Cardiac, Gastrointestinal, Infectious Disease, Neurology, Orthopedic, Pulmonary, Renal, Other Conditions. The final measure score (a single risk-standardized mortality rate) is calculated from the results of these 15 different divisions, modeled separately. Variables from administrative claims and electronic health records are used for risk adjustment.

1.7 Composite Measure

No

1.7 Measure Type

Outcome

1.8 Level of Analysis

Facility

1.9 Care Setting

Hospital: Inpatient

1.10 Measure Rationale

The goal of this measure is to improve patient outcomes by providing patients, physicians, hospitals, and policy makers with information about hospital-level, risk-standardized mortality rates following hospitalization for a range of medical conditions and surgical procedures. Measurement of patient outcomes allows for a broad view of quality of care that encompasses more than what can be captured by individual process-of-care measures. Complex and critical aspects of 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 but are difficult to measure by individual process measures. This measure was developed to identify institutions whose performance is better or worse than would be expected based on their patient case mix, and therefore promote hospital quality improvement and better inform consumers about care quality. While the broad measure score (hospital-wide mortality) provides a big picture view of hospital performance, the more granular division-level results can support the targeting of service-line quality improvement.

Mortality is a significant outcome that is meaningful to patients and providers. For the majority of Medicare beneficiaries admitted to acute care hospitals in the US, the goal is to avoid short-term mortality. According to recent internal analyses, from July 2018 to June 2019, there were about 6.9 million inpatient admissions among Medicare Fee for Service (FFS) and Medicare Advantage beneficiaries between the ages of 65 and 94 at about 4,700 U.S. hospitals. The observed mean hospital 30-day mortality rate was 6.12%. The range of hospital-level mortality scores for the HWM measure was 1.5%-11.6%.

1.11 Measure Webpage

<https://qualitynet.cms.gov/inpatient/measures/hybrid/methodology>

1.13 Data Dictionary

Not attached. I attest that all information will be provided where codes and/or value sets are needed (1.14a - 1.15c).

1.13a Attach Data Dictionary

[Hybrid HWM_Data Dictionary.xlsx](#)

1.14 Numerator

The measure outcome is death from any cause within 30 days of the admission date of the index admission. The numerator is a binary variable (1=yes/0=no) that indicates whether the patient died within 30 days of the index admission date.

1.14a Numerator Details

The measure outcome is death from any cause within 30 days of the admission date of the index admission.

1.15 Denominator

The index cohort includes all inpatient admissions for patients aged 65-94 years old that meet the following inclusion criteria: Enrolled in Medicare FFS/MA for one year prior to the index admission and during the month of the start date of the admission Aged between 65 and 94 Not transferred in from another acute care facility Not primarily treated for non-acute care (psychiatric care or rehabilitation) · Not enrolled in Medicare hospice 12 months prior to, on day of index admission, or within first 2 days of admission Not primarily treated for cancer and enrolled in hospice at anytime during admission stay Not with metastatic cancer · Not with select diagnoses for which hospitals have limited ability to influence survival Admissions meeting the following exclusion criteria are then removed from the denominator, to produce the final measure cohort: Discharged against medical advice Inconsistent or unknown vital status Primarily treated for crush injury, burns, intracranial injury, spinal cord injury, skull and face fracture, or open wounds of head, neck, and trunk With a principal or secondary diagnosis of COVID-19 With a principal diagnosis not assigned to any of the 15 divisions Hospitalizations not randomly selected (one index admission/patient/year) With an admission in a low volume CCS, defined as less than 100 admissions Admissions missing more than 5 of the 10 CCDE *For patients with multiple admissions, the measure selects only one admission, at random, for inclusion. There is no practical statistical modeling approach that can account or adjust for the complex relationship between the number of admissions and risk of mortality in the context of a hospital-wide mortality measure. Random selection ensures that providers are not penalized for a “last” admission during the measurement period; selecting the last admission would not be as accurate a reflection of the risk of death as random selection, as the last admission is inherently associated with a higher mortality risk. Random selection is also used in CMS’s condition-specific mortality measures. Note that random selection reduces the number of admissions, but does not exclude any patients from the measure. The cohort is defined using ICD-10 Clinical Modification codes identified in Medicare Part A Inpatient claims data. Assigning admissions to clinical divisions; (AHRQ) Clinical

Classifications System (CCS). There are a total of about 300 mutually exclusive AHRQ condition categories, most of which are single, homogenous diseases such as pneumonia or acute myocardial infarction. Some are aggregates of conditions, such as “other bacterial infections”. There are about 230 mutually exclusive procedure categories. Using the AHRQ CCS procedure and condition categories, the measure assigns each index hospitalization to one of 15 mutually exclusive divisions. The divisions were created based upon clinical coherence, consistency of mortality risk, adequate patient and hospital case volume for stable results reporting, and input from clinicians, patients, and patient caregivers on usability. The measure first assigns admissions with qualifying AHRQ procedure categories to one of six surgery divisions by identifying a defining surgical procedure. The defining surgical procedure is identified using the following algorithm: 1) if a patient only has one major surgical procedure then that procedure is the defining surgical procedure; 2) if a patient has more than one major surgical procedure, the first dated procedure performed during the index admission is the defining surgical procedure; 3) if there is more than one major surgical procedure on that earliest date, the procedure with the highest mortality rate is the defining surgical procedure. These divisions include admissions likely cared for by surgical teams. The surgical divisions are: Surgical Cancer (see note below), Cardiothoracic Surgery, General Surgery, Neurosurgery, Orthopedic Surgery, and Other Surgical Procedures. Note: For the Surgical Cancer division, any admission that includes a surgical procedure and a principal discharge diagnosis code of cancer is assigned to the Surgical Cancer division. The measure then assigns the remaining admissions into one of the nine non-surgical divisions based on the AHRQ diagnostic CCS of the principal discharge diagnosis. The non-surgical divisions are: Cancer, Cardiac, Gastrointestinal, Infectious Disease, Neurology, Orthopedic, Pulmonary, Renal, Other Conditions. The full list of the specific diagnosis and procedure AHRQ CCS categories used to define the divisions are attached in the Data Dictionary. Please see attached figure Hybrid HWM Flow Diagram of Inclusion and Exclusion Criteria and Division Assignment for the Index Admission.

1.15a Denominator Details

Please see the attached data dictionary that includes the details for the definition of each clinical division.

The index cohort includes all inpatient admissions for patients aged 65-94 years old that meet all of the following criteria:

- Enrolled in Medicare FFS/MA for one year prior to the index admission and during the month of the start date of the admission
- Aged between 65 and 94
- Not transferred in from another acute care facility
- Not primarily treated for non-acute care (psychiatric care or rehabilitation)
- Not enrolled in Medicare hospice 12 months prior to, on day of index admission, or within first 2 days of admission
- Not primarily treated for cancer and enrolled in hospice at any time during admission stay
- Not with metastatic cancer
- Not with select diagnoses for which hospitals have limited ability to influence survival

Assigning admissions to clinical divisions; (AHRQ) Clinical Classifications System (CCS).

There are a total of about 300 mutually exclusive AHRQ condition categories, most of which are single, homogenous diseases such as pneumonia or acute myocardial infarction. Some are aggregates of conditions, such as “other bacterial infections.” There are about 230 mutually exclusive procedure categories. Using the AHRQ CCS procedure and condition categories, the measure assigns each index hospitalization to one of 15 mutually exclusive divisions. The divisions were created based upon clinical coherence, consistency of mortality risk, adequate patient and hospital case volume for stable results reporting, and input from clinicians, patients, and patient caregivers on usability. Please see [Figure 1](#) for a flow chart that shows how admissions are assigned to divisions. The measure first assigns admissions with qualifying AHRQ procedure categories to one of six surgery divisions by identifying a defining surgical procedure. The defining surgical procedure is identified using the following algorithm: 1) if a patient only has one major surgical procedure then that procedure is the defining surgical procedure; 2) if a patient has more than one major surgical procedure, the first dated procedure performed during the index admission is the defining surgical procedure; 3) if there is more than one major surgical procedure on that earliest date, the procedure with the highest mortality rate is the defining surgical procedure. These divisions include admissions likely cared for by surgical teams. The surgical divisions are: Surgical Cancer (see note below), Cardiothoracic Surgery, General Surgery, Neurosurgery, Orthopedic Surgery, and Other Surgical Procedures. Note: For the Surgical Cancer division, any admission that includes a surgical procedure and a principal discharge diagnosis code of cancer is assigned to the Surgical Cancer division. The measure then assigns the remaining admissions into one of the nine non-surgical divisions based on the AHRQ diagnostic CCS of the principal discharge diagnosis. The non-surgical divisions are: Cancer, Cardiac, Gastrointestinal, Infectious Disease, Neurology, Orthopedic, Pulmonary, Renal, Other Conditions. The full list of the specific diagnosis and procedure AHRQ CCS categories used to define the divisions are attached in the Data Dictionary.

1.15b Denominator Exclusions

Admissions meeting any of the following criteria are then removed from the denominator, to produce the final measure cohort:

- Discharged against medical advice
- Inconsistent or unknown vital status
- Primarily treated for crush injury, burns, intracranial injury, spinal cord injury, skull and face fracture, or open wounds of head, neck, and trunk
- With a principal or secondary POA diagnosis of COVID-19
- With a principal diagnosis not assigned to any of the 15 divisions
- Hospitalizations not randomly selected (one index admission/patient/year)
- With an admission in a low volume CCS, defined as less than 100 admissions
- Admissions missing more than 5 of the 10 CCDEs

1.15c Denominator Exclusions Details

This measure excludes index admissions for patients:

- **Discharged against medical advice**

- Rationale: Providers did not have the opportunity to deliver full care and prepare the patient for discharge.
- **With inconsistent or unknown vital status (from claims data) or other unreliable claims data.**
 - Rationale: The measure does not include stays for patients where the admission date is after the date of death, or where the date of death occurs before the date of discharge but the patient was discharged alive because these are likely errors in the data.
- **Primarily treated for crush injury, burns, intracranial injury, spinal cord injury, skull and face fracture, or open wounds of head, neck, and trunk**
 - Rationale: Even though a hospital likely can influence the outcome of some of these conditions, in many cases death events are not a signal of poor quality of care when patients present with these conditions. These conditions are also infrequent events that are unlikely to be uniformly distributed across hospitals.
- **With a principal or secondary diagnosis of COVID-19**
 - Rationale: Patients with a primary or secondary diagnosis of COVID-19 are excluded from the measure cohort in response to the COVID-19 Public Health Emergency.
- **With a principal diagnosis not assigned to any of the 15 divisions**
- **Hospitalizations not randomly selected (one index admission/patient/year)***
 - Rationale: One admission is selected at random to ensure that providers are not penalized for a “last” admission during the measurement period; selecting the last admission would not be as accurate a reflection of the risk of death as random selection, as the last admission is inherently associated with a higher mortality risk.
- **With an admission in a low volume CCS, defined as less than 100 admissions**
 - Rationale: To calculate a stable and precise risk model, there are a minimum number of admissions that are needed. In addition, a minimum number of admissions and/or outcome events are required to inform grouping admissions into larger categories. These admissions present challenges to both accurate risk prediction and coherent risk grouping and are therefore excluded.
- **Admissions missing more than 5 of the 10 CCDEs**
 - Rationale: Patients for whom a large portion of CCDE is missing are excluded from the measure as their status upon hospital arrival would not be complete.

*For patients with multiple admissions, the measure selects only one admission, at random, for inclusion. There is no practical statistical modeling approach that can account or adjust for the complex relationship between the number of admissions and risk of mortality in the context of a hospital-wide mortality measure. Random selection ensures that providers are not penalized for a “last” admission during the measurement period; selecting the last admission would not be as accurate a reflection of the risk of death as random selection, as the last admission is inherently associated with a higher mortality risk. Random selection is also used in CMS’s condition-specific mortality measures. Note that random selection reduces the number of admissions but does not exclude any patients from the measure. The cohort is defined using ICD-10 Clinical Modification codes identified in Medicare Part A Inpatient claims data.

1.15d Age Group

Older Adults (65 years and older), Other

1.15e Age Range in Years

Older Adults (between 65 and 94)

1.16 Type of Score

Rate/proportion

1.17 Measure Score Interpretation

Better performance = Lower score

1.18 Calculation of Measure Score

Below we provide the individual steps to calculate the measure score:

Define Cohort

1. Create 15 mutually exclusive service-line divisions using groups of related conditions or procedures. See the data dictionary for divisions (Tables 6 and 7) and the inclusion/exclusion indicators (Tables 1-5). Reassign CCS groupings that define the clinical division to facilitate improved risk adjustment (see Section 4.4.2 for details). The specifications for each clinical division are defined in the data dictionary (Table 11. CSS Modifications).

2. Apply the inclusion/exclusion criteria to construct the measure cohort:

- Identify discharges meeting the inclusion criteria described in the denominator section above and assign to one of 15 divisions. Eligible discharges are from July 1 of one year to June 30 of the following year for any respective year.

Define outcome

3. Derive the measure outcome of 30-day mortality, by creating a binary flag that indicates whether the patient died within 30 days of the index admission date. (This is sourced from the date of death in the Medicare Enrollment Database.)

Define risk variables

4. Use patients' historical and index admission claims data, as well as CCDE values to create risk-adjustment variables. (Note: risk variables from claims are based on secondary diagnoses from inpatient claims for one year prior to and including the index admission, if present on admission)

Measure score calculation

5. For each specialty cohort group, estimate a separate hierarchical logistic regression model (HGLM) to produce a standardized mortality ratio (SMR), calculated as the ratio of the "predicted" mortality to the "expected" mortality at a given hospital. The HGLM is adjusted for age, selected clinical covariates, and a hospital-specific effect. Details about the risk-adjustment model can be found in the original measure development methodology report:

<https://www.qualitynet.org/inpatient/measures/mortality/methodology>.

6. Pool each specialty cohort SMRs for each hospital using a volume-weighted geometric mean to create a hospital-wide SMR (or RSMR). Calculations can be found attached and posted at:
<https://www.qualitynet.org/inpatient/measures/mortality/methodology>.

7. Use statistical bootstrapping to construct a 95% confidence interval estimate for each facility's RSMR. For more information about the measure methodology, please see the most recent Hybrid HWM Comprehensive Methodology Report attached and posted here:
<https://www.qualitynet.org/inpatient/measures/mortality/methodology>.

1.18a Attach measure score calculation diagram

[Measure Calculation Formulas-Hybrid HWM.pdf](#)

1.19 Measure Stratification Details

This measure is not stratified.

1.20 Types of Data Sources

Administrative Data, Claims Data, Electronic Health Records

1.25 Data Source Details

The components of this HWM measure, as specified in this CBE submission, are comprised of data from the following sources:

Cohort: Medicare fee-for-service claims and Medicare Advantage encounters; Medicare enrollment data.

Outcome: Medicare enrollment data

Risk adjustment: Medicare fee-for-service claims, Medicare Advantage encounters, supplemented with EHR data (core clinical data elements, or CCDE).

Feasibility of data collection is addressed in Section 3.1, "Feasibility".

Additional information on the data sources for this CBE submission can be found in Section 4.1 "Data and Samples"

1.26 Minimum Sample Size

There is no minimum sample size for the calculation of this measure.

2.1 Attach Logic Model

[HHWM Logic Model.pdf](#)

2.2 Evidence of Measure Importance

Death is a finite event, easy to measure accurately, and easily understood by patients and providers. For the majority of Medicare beneficiaries admitted to acute care hospitals in the US, the goal is to avoid short-term mortality. By measuring Hospital-Wide Mortality (HWM), CMS can ensure that efforts to reduce other outcomes, such as readmissions and resource utilization, are not resulting in unintended consequences. Specifically, this HWM measure complements the existing CMS Hospital-Wide All-Cause Risk-Standardized Readmission Measure (CBE ID: 2879) to allow assessment of trends in hospital performance for both outcomes, similar to other complementary pairs of readmission and mortality measures for specific conditions and procedures. Further, the HWM measure provides CMS with annually updated performance estimates for a larger proportion of the nation's hospitals, allowing significant performance outliers to be identified.

A hospital-wide mortality measure captures a large cohort of patients admitted for a wide range of diagnoses, and also illuminates a broad range of performance among hospitals. According to internal analyses, from July 2018 to June 2019, there were about 6.9 million inpatient admissions among Medicare FFS and Medicare Advantage beneficiaries between the ages of 65 and 94 at about 4,700 US hospitals. This comprehensive cohort is inclusive of patients not currently captured by existing condition-specific mortality measures and provides stakeholders with a broad quality signal (in addition to more granular data to support quality improvement). In addition to capturing an expansive cohort of patients, variation in hospital-level mortality rates demonstrates a quality gap: hospital-level mortality rates range widely, from 1.5% to 11.6% (using data from July 2018-June 2019). The average hospital-level, risk standardized 30-day mortality rate was 6.12%.

For many conditions and diagnoses, evidence supports that optimal medical care reduces mortality. For example, strategies that have been shown to be effective at reducing mortality include the adoption of evidence-based processes, such as those that prevent central line infections¹ and surgical-site infections,² early identification and management of sepsis,³ and use of evidence-based guidelines for treatment of heart failure,⁴ among others. In addition, deployment of Rapid Response Teams⁵ to attend to patients at the first sign of clinical decline, identification of high-risk patients on admission and increase nursing care and physician contact accordingly,⁶ and standardization of patient handoffs to avoid miscommunication or gaps in care⁷ have been effective at reducing errors and reducing mortality.

Some of the evidence-based recommendations apply to specific diagnoses. While condition- and procedure-specific initiatives to reduce mortality may broadly impact mortality rates across other conditions and procedures, there is likely more to be gained by a measure of hospital-wide mortality that can inform and encourage quality improvement efforts for patients not currently captured by existing CMS mortality measures. For example, a 2017 study of a standardized, inter-hospital transfer tool found that in-hospital mortality decreased for transferred patients following implementation of a one-page handover containing information critical for immediate patient care.⁸ Another study found that regionalization of cardiac hospitals increased transfer time for PCI and reduced 7-day mortality for STEMI patients.⁹

Finally, there is evidence that a hospital's organizational culture is linked to key measures of hospital quality performance.¹⁰ Since these cultural and leadership qualities affect the entire

hospital, a hospital-wide mortality measure may provide important incentives for hospitals to not only examine their care processes and improve care for individual conditions but may also provide incentives to encourage care transformation and improve overall organizational culture.

References:

1. Buetti, N., Marschall, J., Drees, M., Fakih, M. G., Hadaway, L., Maragakis, L. L., Monsees, E., Novosad, S., O'Grady, N. P., Rupp, M. E., Wolf, J., Yokoe, D., & Mermel, L. A. (2022). Strategies to prevent central line-associated bloodstream infections in acute-care hospitals: 2022 Update. *Infection control and hospital epidemiology*, 43(5), 553-569.
2. Seidelman, J. L., Mantyh, C. R., & Anderson, D. J. (2023). Surgical Site Infection Prevention: A Review. *JAMA*, 329(3), 244-252.
3. Rhee, C., Strich, J. R., Chiotos, K., Classen, D. C., Cosgrove, S. E., Greeno, R., Heil, E. L., Kadri, S. S., Kalil, A. C., Gilbert, D. N., Masur, H., Septimus, E. J., Sweeney, D. A., Terry, A., Winslow, D. L., Yealy, D. M., & Klompas, M. (2024). Improving Sepsis Outcomes in the Era of Pay-for-Performance and Electronic Quality Measures: A Joint IDSA/ACEP/PIDS/SHEA/SHM/SIDP Position Paper. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 78(3), 505-513.
4. Heidenreich, P., Bozkurt, B., Aguilar, D. et al. 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *JACC*. 2022 May, 79 (17) 1757-1780.
5. Hall KK, Lim A, Gale B. The Use of Rapid Response Teams to Reduce Failure to Rescue Events: A Systematic Review. *J Patient Saf*. 2020 Sep;16(3S Suppl 1):S3-S7..
6. Mann KD, Good NM, Fatehi F, Khanna S, Campbell V, Conway R, Sullivan C, Staib A, Joyce C, Cook D. Predicting Patient Deterioration: A Review of Tools in the Digital Hospital Setting. *J Med Internet Res*. 2021 Sep 30;23(9):e28209.
7. Starmer, A. J., Spector, N. D., Srivastava, R., West, D. C., Rosenbluth, G., Allen, A. D., Noble, E. L., Tse, L. L., Dalal, A. K., Keohane, C. A., Lipsitz, S. R., Rothschild, J. M., Wien, M. F., Yoon, C. S., Zigmont, K. R., Wilson, K. M., O'Toole, J. K., Solan, L. G., Aylor, M., Bismilla, Z., ... I-PASS Study Group (2014). Changes in medical errors after implementation of a handoff program. *The New England journal of medicine*, 371(19), 1803-1812.
8. Theobald CN, Choma NN, Ehrenfeld JM, Russ S, Kripalani S. Effect of a Handover Tool on Efficiency of Care and Mortality for Interhospital Transfers. *Journal of Hosp Medicine* 2017; 12(1):23-28.
9. Shen YC, Krumholz H, Hsia RY. Association of Cardiac Care Regionalization With Access, Treatment, and Mortality Among Patients With ST-Segment Elevation Myocardial Infarction. *Circ Cardiovasc Qual Outcomes*. 2021;14(3):e007195.
10. Linnander E, McNatt Z, Boehmer K, Cherlin E, Bradley E, Curry L. Changing hospital organisational culture for improved patient outcomes: developing and implementing the leadership saves lives intervention. *BMJ Qual Saf*. 2021 Jun;30(6):475-483. doi: 10.1136/bmjqs-2019-010734. Epub 2020 Jul 16.

2.4 Performance Gap

We refer readers to Section 1.18 for information on how performance scores are calculated.

As described in section 4.1.2, we provide results using a nationally representative dataset that includes both FFS and MA admissions and claims-based risk adjustment (but without the EHR-based data elements [CCDE] for enhanced case-mix risk adjustment), and separately, results from 2024 Voluntary Reporting (representing the measure as currently implemented, without MA admissions but with both the claims-based and clinical data elements from the EHR [CCDE] to enhance risk adjustment).

We characterize the degree of variation by reporting the distribution of RSMRs.

Measure Score Distribution

The distribution of measure scores from the Claims-Only HWM (Medicare FFS + MA) dataset and Hybrid HWM 2024 Voluntary Reporting dataset is shown below in Tables 1-3 and histograms are shown in Figures 3 and 4 (in the Table and Figures attachment).

There is wide variation in measure scores in the national dataset (Claims Only HWM [Medicare FFS + MA]): RSMRs for the 4,743 hospitals in the dataset range from 1.52% to 11.60% with a mean of 6.12% (standard deviation, 0.76%); the 25th percentile is 5.67% and the 75th percentile is 6.56% (Table 1). There is meaningful variation in performance across hospitals: the worst performing facility (RSMR 11.60%) is performing about 89% worse than the median (6.11%), while the best performing facility (RSMR 1.52%) is performing 75% better than the median.

As expected, we see less variation in the 1,136 hospitals within the Hybrid HWM 2024 Voluntary Reporting dataset. We see less variation in this dataset due to the voluntary nature of public reporting, where we expect that better performers may be more likely to choose to report. As shown in Table 2, RSMRs ranged from 2.16% to 6.17%, with a mean of 3.94% (standard deviation, 0.62%). The 25th percentile was 3.56% and the 75th percentile was 4.33%.

Table 1. Performance Scores by Decile

	Performance Gap												
	Overall	Minimum	Decile_1	Decile_2	Decile_3	Decile_4	Decile_5	Decile_6	Decile_7	Decile_8	Decile_9	Decile_10	Maximum
Mean Performance Score	6.34%	1.52%	5.26%	5.77%	6.00%	6.15%	6.27%	6.37%	6.52%	6.69%	6.90%	7.55%	11.60%
N of Entities	4,743	1	474	474	475	474	474	475	474	475	474	474	1
N of Persons													
/ Encounters	6,883,980	387	1,552,824	977,525	780,570	589,718	400,071	489,024	484,391	486,414	528,924	594,519	1,113
/ Episodes													

2.6 Meaningfulness to Target Population

With this hospital-wide mortality measure, hospitals can identify areas for improvement in care, therefore improving patient outcomes and 30-day mortality rates. In addition, the public reporting of hospital-level performance provides transparency/accountability to other stakeholders, including patients.

Overall, the Hybrid Hospital-Wide All-Cause Mortality measure is an important indicator of the quality of care provided by hospitals. 30-day mortality rates (that is, death within 30-days following hospital admission) may point to gaps in clinical practices, patient care protocols, and the overall delivery of healthcare services. By monitoring these rates, hospitals can identify areas

needing improvement and implement specific interventions that improve intermediate and overall outcomes (survival), an, an outcome that is meaningful to patients. This measure offers insights into a hospital's performance and informs patients, healthcare providers, and policymakers about the standard of care that currently exists.

This measure was developed with stakeholder input, including the input of patients. We had many discussions with stakeholders, including the patients on our TEP, and a patient workgroup, in which we heard support for the concept of measuring hospital-wide mortality after discharge and a strong desire for a measure that offers patients and providers meaningful, detailed, and statistically valid performance data.

References:

1. Medicare Payment Advisory Commission. (2005). *Report to the Congress: Medicare payment policy*. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs...

3.1 Contributions Towards Closing Care Gaps

We weigh social risk factor adjustment using a comprehensive approach that evaluates the following:

- Well-supported conceptual model for influence of social risk factors on measure outcome (detailed below);
- Feasibility of testing meaningful social risk factors in available data: and
- Empiric testing of social risk factors.

In the attachment, we summarize the findings of the literature review and conceptual pathways by which social risk factors may influence risk of the outcome. Our conceptualization of the pathways by which patients' social risk factors affect the outcome is informed by the literature cited below and IMPACT Act-funded work by the National Academy of Science, Engineering and Medicine (NASEM) and the Department of Health and Human Services Assistant Secretary for Policy and Evaluation (ASPE 2016; ASPE 2020).

Causal Pathways for Social Risk Variable Selection

There is a large body of literature linking various social risk factors to worse health status and higher mortality over a lifetime.^{2, 7, 19, 33} Although some recent literature evaluates the relationship between patient social risk factors and the mortality outcome, few studies directly address causal pathways or examine the role of the hospital in these pathways.^{5, 11, 18, 23, 24} Moreover, the current literature examines a wide range of conditions and risk variables with no clear consensus on which risk factors demonstrate the strongest relationship with mortality.

The social risk factors that have been examined in the literature can be categorized into three domains: (1) patient-level variables, (2) neighborhood/community-level variables, and (3) hospital-level variables.

Patient-level variables describe characteristics of individual patients and include the patient's income or education level.⁸ Neighborhood/community-level variables use information from sources such as the ACS as either a proxy for individual patient-level data or to measure environmental factors. Studies using these variables use one dimensional measures such as median household income or composite measures such as the Area Deprivation Index (ADI).^{20, 28, 31} Some of these variables may include the local availability of clinical providers.¹³⁻¹⁴ Hospital-level variables measure attributes of the hospital which may be related to patient risk. Examples of hospital-level variables used in studies are ZIP code characteristics aggregated to the hospital level or the proportion of Medicaid patients served in the hospital.^{11, 16-17}

The conceptual relationship, or potential causal pathways by which these possible social risk factors influence the risk of readmission following an acute illness or major surgery, like the factors themselves, are varied and complex. There are at least four potential pathways that are important to consider:

1. **Patients with social risk factors may have worse health at the time of hospital admission.** Patients who have lower income/education/literacy or unstable housing may have a worse general health status and may present for their hospitalization or procedure with a greater severity of underlying illness. These social risk factors, which are characterized by patient-level or neighborhood/community-level (as proxy for patient-level) variables, may contribute to worse health status at admission due to competing priorities (restrictions based on job), lack of access to care (geographic, cultural, or financial), or lack of health insurance. Given that these risk factors all lead to worse general health status, this causal pathway should be largely accounted for by current clinical risk-adjustment.
2. **Patients with social risk factors often receive care at lower quality hospitals.** Patients of lower income, lower education, or unstable housing have inequitable access to high quality facilities, in part, because such facilities are less likely to be found in geographic areas with large populations of poor patients. Thus, patients with low income are more likely to be seen in lower quality hospitals, which can explain increased risk of mortality following hospitalization.
3. **Patients with social risk factors may receive differential care within a hospital.** The third major pathway by which social risk factors may contribute to mortality risk is that patients may not receive equivalent care within a facility. For example, patients with social risk factors such as lower education may require differentiated care (e.g. provision of lower literacy information – that they do not receive).
4. **Patients with social risk factors may experience worse health outcomes beyond the control of the health care system.** Some social risk factors, such as income or wealth, may affect the likelihood of mortality without directly affecting health status at admission or the quality of care received during the hospital stay. For instance, while a hospital may make appropriate care decisions and provide tailored care and education, a lower-income patient may have a worse outcome post-discharge due to competing financial priorities which don't allow for adequate recuperation or access to needed treatments, or a lack of access to care outside of the hospital.

Although we analytically aim to separate these pathways to the extent possible, we acknowledge that risk factors often act on multiple pathways, and as such, individual pathways can be complex to distinguish analytically. Further, some social risk factors, despite having a strong conceptual

relationship with worse outcomes, may not have statistically meaningful effects on the risk model. They also have different implications on the decision to risk adjust or not.

Based on this model and that the Area Deprivation Index (ADI) and dual-eligibility variables aim to capture the social risk factors that are likely to influence these pathways (income, education, housing, and community factors) - the following social risk variables were considered for risk-adjustment:

- Dual-eligible status
 - Dual eligibility for Medicare and Medicaid is available at the patient level in the Medicare Master Beneficiary Summary File. The eligibility threshold for over 65-year-old Medicare patients considers both income and assets. For the dual-eligible (DE) indicator, there is a body of literature demonstrating differential health care and health outcomes among beneficiaries.²⁷
- Area Deprivation index (ADI): The ADI, initially developed by Health Resources & Services Administration (HRSA), is based on 17 measures across four domains: income, education, employment, and housing quality.^{20, 31}

The 17 components are listed below:

- Population aged ≥ 25 y with < 9 y of education, %
- Population aged ≥ 25 y with at least a high school diploma, %
- Employed persons aged ≥ 16 y in white collar occupations, %
- Median family income, \$
- Income disparity
- Median home value, \$
- Median gross rent, \$
- Median monthly mortgage, \$
- Owner occupied housing units, % (home ownership rate)
- Civilian labor force population aged ≥ 16 y unemployed, % (unemployment rate)
- Families below poverty level, %
- Population below 150% of the poverty threshold, %
- Single parent households with children aged < 18 y, %
- Households without a motor vehicle, %
- Households without a telephone, %
- Occupied housing units without complete plumbing, % (log)
- Households with more than 1 person per room, % (crowding)

ADI scores were derived using beneficiary's 9-digit ZIP Code of residence, which is obtained from the Medicare Enrollment Database, and is linked to 2017-2021 US Census/American Community Survey (ACS) data. In accordance with the ADI developers' methodology, an ADI score is calculated for the census block group corresponding to the beneficiary's 9-digit ZIP Code using 17 weighted Census indicators. Raw ADI scores were then transformed into a national percentile ranking ranging from 1 to 100, with lower scores indicating lower levels of disadvantage and higher scores indicating higher levels of disadvantage. Percentile thresholds established by the ADI developers were then applied to ADI percentile to dichotomize neighborhoods into more disadvantaged (high ADI areas=ranking equal to or greater than 85) or less disadvantaged areas (Low ADI areas= ranking of less than 85).

References

1. Blum, A. B., N. N. Egorova, E. A. Sosunov, A. C. Gelijns, E. DuPree, A. J. Moskowitz, A. D. Federman, D. D. Ascheim and S. Keyhani. "Impact of Socioeconomic Status Measures on Hospital Profiling in New York City." *Circ Cardiovasc Qual Outcomes* 7, no. 3 (2014): 391-7.
2. Brodish P.H., Hakes J.K. "Quantifying the individual-level association between income and mortality risk in the United States using the National Longitudinal Mortality Study." *Soc. Sci. Med.*, 170 (2016), pp. 180-187, 10.1016.
3. Buntin MB, Ayanian JZ. Social Risk Factors and Equity in Medicare Payment. *New England Journal of Medicine*. 2017;376(6):507-510.
4. Calvillo-King L, Arnold D, Eubank KJ, et al. Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review. *Journal of general internal medicine*. 2013;28(2):269-282.
5. Chang W-C, Kaul P, Westerhout C M, Graham M. M., Armstrong Paul W., "Effects of Socioeconomic Status on Mortality after Acute Myocardial Infarction." *The American Journal of Medicine*. 2007; 120(1): 33-39
6. Committee on Accounting for Socioeconomic Status in Medicare Payment Programs; Board on Population Health and Public Health Practice; Board on Health Care Services; Institute of Medicine; National Academies of Sciences, Engineering, and Medicine. *Accounting for Social Risk Factors in Medicare Payment: Identifying Social Risk Factors*. Washington (DC): National Academies Press (US); 2016 Jan 12.
(<https://www.ncbi.nlm.nih.gov/books/NBK338754/doi:10.17226/21858>)
7. Demakakos P, Biddulph JP, Bobak M, Marmot MG (2016a) Wealth and mortality at older ages: a prospective cohort study. *J Epidemiol Community Health* 70:346-353.
8. Department of Health and Human Services, Office of the Assistant Secretary of Planning and Evaluation. *Report to Congress: Social Risk Factors and Performance under Medicare's Value-based Payment Programs*. December 21, 2016.
(<https://aspe.hhs.gov/pdf-report/report-congress-social-risk-factors-and...>).
9. Eapen ZJ, McCoy LA, Fonarow GC, Yancy CW, Miranda ML, Peterson ED, Califf RM, Hernandez AF. Utility of socioeconomic status in predicting 30-day outcomes after heart failure hospitalization. *Circ Heart Fail*. May 2015; 8(3):473-80.
10. Foraker, R. E., K. M. Rose, C. M. Suchindran, P. P. Chang, A. M. McNeill and W. D. Rosamond. "Socioeconomic Status, Medicaid Coverage, Clinical Comorbidity, and Rehospitalization or Death after an Incident Heart Failure Hospitalization: Atherosclerosis Risk in Communities Cohort (1987 to 2004)." *Circ Heart Fail* 4, no. 3 (2011): 308-16.
11. Gopaldas R R, Chu D., "Predictors of surgical mortality and discharge status after coronary artery bypass grafting in patients 80 years and older." *The American Journal of Surgery*. 2009; 198(5): 633-638
12. Gilman M, Adams EK, Hockenberry JM, Wilson IB, Milstein AS, Becker ER. California safety-net hospitals likely to be penalized by ACA value, readmission, and meaningful-use programs. *Health Affairs (Millwood)*. Aug 2014; 33(8):1314-22.
13. Herrin J, Kenward K, Joshi MS, Audet AM, Hines SJ. Assessing Community Quality of Health Care. *Health Serv Res*. 2016 Feb;51(1):98-116. doi: 10.1111/1475-6773.12322. Epub 2015 Jun 11. PMID: 26096649; PMCID: PMC4722214.
14. Herrin J, St Andre J, Kenward K, Joshi MS, Audet AM, Hines SC. Community factors and hospital readmission rates. *Health Serv Res*. 2015 Feb;50(1):20-39. doi: 10.1111/1475-6773.12177. Epub 2014 Apr 9. PMID: 24712374; PMCID: PMC4319869.

15. Hamadi H, Moody L, Apatu E, Vossos H, Tafili A, Spaulding A. Impact of hospitals' Referral Region racial and ethnic diversity on 30-day readmission rates of older adults. *J Community Hosp Intern Med Perspect*. 2019;9(3):181-188.
16. Imran A, Rawal MD, Botre N, Patil A. Improving and Promoting Social Determinants of Health at a System Level. *Jt Comm J Qual Patient Saf*. 2022;48(8):376-384. Jha AK, Orav EJ, Epstein AM. Low-quality, high-cost hospitals, mainly in South, care for sharply higher shares of elderly black, Hispanic, and medicaid patients. *Health affairs* 2011; 30:1904-11.
17. Joynt KE, Jha AK. Characteristics of hospitals receiving penalties under the Hospital Readmissions Reduction Program. *JAMA*. Jan 23 2013; 309(4):342-3.
18. Kim C, Diez A V, Diez Roux T, Hofer P, Nallamotheu B K, Bernstein S J, Rogers M, "Area socioeconomic status and mortality after coronary artery bypass graft surgery: The role of hospital volume." *Clinical Investigation Outcomes, Health Policy, and Managed Care*. 2007; 154(2): 385-390
19. Kim D. The associations between US state and local social spending, income inequality, and individual all-cause and cause-specific mortality: The National Longitudinal Mortality Study. *Prev. Med*. 2015;84:62-68. doi: 10.1016/j.ypmed.2015.11.013.
20. Kind AJH, Buckingham W. Making Neighborhood Disadvantage Metrics Accessible: The Neighborhood Atlas. *New England Journal of Medicine*, 2018. 378: 2456-2458. DOI: 10.1056/NEJMp1802313. PMCID: PMC6051533. AND University of Wisconsin School of Medicine Public Health. 2023 Area Deprivation Index v4.0. Downloaded from <https://www.neighborhoodatlas.medicine.wisc.edu/>.
21. Kind, A. J., S. Jencks, J. Brock, M. Yu, C. Bartels, W. Ehlenbach, C. Greenberg and M. Smith. "Neighborhood Socioeconomic Disadvantage and 30-Day Rehospitalization: A Retrospective Cohort Study." *Ann Intern Med* 161, no. 11 (2014): 765-74.
22. Krumholz HM, Brindis RG, Brush JE, et al. 2006. Standards for Statistical Models Used for Public Reporting of Health Outcomes: An American Heart Association Scientific Statement From the Quality of Care and Outcomes Research Interdisciplinary Writing Group: Cosponsored by the Council on Epidemiology and Prevention and the Stroke Council Endorsed by the American College of Cardiology Foundation. *Circulation* 113: 456-462.
23. LaPar D J, Bhamidipati C M, et al. "Primary Payer Status Affects Mortality for Major Surgical Operations." *Annals of Surgery*. 2010; 252(3): 544-551
24. LaPar D J, Stukenborg G J, et al "Primary Payer Status Is Associated With and Resource Utilization for Coronary Artery Bypass Grafting." *Circulation*. 2012; 126:132-139
25. Lindenauer PK, Lagu T, Rothberg MB, et al. Income inequality and 30 day outcomes after acute myocardial infarction, heart failure, and pneumonia: retrospective cohort study. *BMJ*. 2013 Feb 14; 346:f521. doi: 10.1136/bmj.f521.
26. Normand S-LT, Shahian DM. 2007. Statistical and Clinical Aspects of Hospital Outcomes Profiling. *Stat Sci* 22 (2): 206-226. Reames BN, Birkmeyer NJ, Dimick JB, Ghaferi AA. Socioeconomic disparities in mortality after cancer surgery: failure to rescue. *JAMA surgery* 2014; 149:475-81.
27. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health & Human Services. Second Report to Congress on Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program. 2020. <https://aspe.hhs.gov/social-risk-factors-and-medicare-value-based-purch...>
28. Powell WR, Sheehy AM, Kind AJ. The area deprivation Index is the most scientifically validated social exposome tool available for policies advancing health equity. *Health Affairs Forefront*. 2023;

29. Regalbuto R, Maurer MS, Chapel D, Mendez J, Shaffer JA. Joint Commission requirements for discharge instructions in patients with heart failure: is understanding important for preventing readmissions? *Journal of cardiac failure*. 2014;20(9):641-649.
30. Shahian DM, Iezzoni LI, Meyer GS, Kirle L, Normand SL. Hospital-wide mortality as a quality metric: conceptual and methodological challenges. *American journal of medical quality: the official journal of the American College of Medical Quality*. 2012;27(2):112-123.
31. Singh, G. K. (2003). Area Deprivation and Widening Inequalities in US Mortality, 1969-1998. *American Journal of Public Health*, 93(7), 1137-1143.
<https://doi.org/10.2105/ajph.93.7.1137>
32. Trivedi AN, Nsa W, Hausmann LR, et al. Quality and equity of care in U.S. hospitals. *The New England journal of medicine* 2014; 371:2298-308.
33. van Oeffelen AA, Agyemang C, Bots ML, Stronks K, Koopman C, van Rossem L, Vaartjes I. The relation between socioeconomic status and short-term mortality after acute myocardial infarction persists in the elderly: results from a nationwide study. *Eur J Epidemiol*. 2012 Aug;27(8):605-13. doi: 10.1007/s10654-012-9700-z. Epub 2012 Jun 5.

Social Risk Factors Summary

While our testing results (see below, and in the attachment of figures and tables) show that patients with social risk factors (DE or high ADI) have higher unadjusted rates of the outcome, we find that the impact of each social risk factor on measure scores is minimal: measure scores calculated with and without each social risk factor are highly correlated (near 1), and differences between measure scores calculated with and without each social risk factor are near zero. Furthermore, we show that the distribution of measure scores across the quartiles of the hospital-proportion of each social risk factor overlap, demonstrating that it is possible for hospitals with a high proportion of patients with social risk factors to perform as well as hospitals with a lower proportion. Finally, our model calibration tests show that all 15 models are well calibrated for patients with each social risk factor (data not shown). These empiric results, together with the measure's use in a pay-for reporting (not pay for performance) program, support CMS's decision to not adjust the measure for social risk factors.

Analysis Approach

To provide a nationally representative set of testing results, social risk factor testing was conducted using the Claims-Only HWM (Medicare FFS + MA) (discharges July 1, 2018-June 30, 2019) dataset. We assessed the impact of social risk factors (High ADI and dual-eligible status) on the hybrid HWM measure by examining the following: prevalence of each social risk factor, association with the outcome, and impact on measure scores. We describe each analysis and the results below. We define the high ADI variable as an ADI score equal to or above 85.

Analysis #1: Variation in prevalence of the factor across measured entities (Medicare FFS and MA)

The prevalence of social risk factors at hospital-level in the HWM cohort varies across hospitals (Table 16 in the Tables and Figures attachment). In the Claims Only HWM (Medicare FFS and MA) dataset, the median percentage of dual-eligible patients was 14.2% (Interquartile Range

[IQR], 9.4%-21.4%) and the median percentage of patients with high ADI variable [score equal to or above 85] was 10.5% (IQR: 2.0%-26.8%).

Analysis #2: Observed outcome rates in patients with social risk factors

In the Claims Only HWM (Medicare FFS and MA) dataset, patient-level observed mortality rates were higher for dual-eligible patients (8.5%) compared with non-dual-eligible patients (5.9%) (Table 17 in the Tables and Figures attachment). The observed mortality rate for patients with the high ADI variable is higher (7.2%) compared to patients without the high ADI variable (6.2%).

Analysis #3: Impact of social risk factor on hospital-level measure scores

To determine the impact of adding social risk factors on measure scores, we compared correlation coefficients of measures scores calculated with and without the social risk factors in the model (Table 18, Figures 10 and 11 in the Tables and Figures attachment), and we compared differences in measure scores (Table 18). Hospitals' risk-standardized mortality rates (RSMRs) are highly correlated: the correlation coefficient of RSMRs between hospitals using the Claims-Based HWM (Medicare FFS and MA) dataset, calculated with and without the high ADI variable is 0.999 (Figure 10) and correlation coefficient between measure scores calculated with and without the dual-eligible variable is 0.999 (Figure 11). The median change in hospitals' RSMRs when adding either social risk factor is small (-0.0015 for high ADI and -0.0023 for dual-eligibility for the Claims-Only HWM (Medicare FFS and MA) dataset (Table 18) dataset.

4.1 Feasibility Assessment

As part of broader measure development, we originally tested the feasibility of electronic extraction of the EHR-based data elements used to enhance risk adjustment (the core clinical data elements or CCDEs). The CCDE are a set of data elements that are captured on most adults admitted to acute care hospitals, are easily extracted from EHRs, and can be used to risk adjust hospital outcome measures for a variety of conditions and procedures. Feasibility testing included: 1) identification of potentially feasible clinical data through qualitative assessment, 2) empirical feasibility testing of several clinical data elements electronically extracted from two large multi-facility health systems, and 3) validity testing of the CCDE at an additional health system. Results from these analyses show conceptual feasibility by a Technical Expert Panel (TEP), while empirical feasibility demonstrates consistent capture and match rate of CCDE from the EHRs. For more information on our initial feasibility testing conducted during measure development, please see the Hybrid HWM methodology report and 2013 Core Clinical Data Elements Technical Report attached to this form.

Prior to measure implementation, CMS received feedback through the FY2022 Inpatient Prospective Payment System Final Rule¹ indicating concerns about reporting burden, in terms of variation in readiness and eCQM reporting capabilities across hospitals. This concern was addressed by delaying implementation for several years after rule finalization by adding one rounds of Confidential Reporting (and two rounds of Confidential Reporting for Hybrid Hospital-Wide Readmission, which shares similar data elements and submission/ collection processes) to allow hospitals and their vendors additional time to upgrade IT systems, improve data mapping and other capabilities, and increase staff training for measure reporting. This eCQM reporting

cycle was delayed in comparison to reporting requirements for other Hospital IQR Program measures.

As also discussed in Section 6.2.3, since the Hybrid HWM measure has been implemented, we have received feedback from stakeholders, in particular through 2024 Voluntary Reporting of the measure. In their feedback, hospitals noted challenges in CCDE capture. For example, we found about 35% of admissions were missing the “Platelet” lab test variable for the Surgical Orthopedics division. This was due to the fact that hospitals had reported the value in a unit (femtoliters) that is unusable for measure calculation, due to standardization. To address this problem, an alternative standardization strategy has been implemented for 2025 Reporting so that all values can be used.

Also as described in Section 6.2.3, hospitals also provided feedback about challenges in meeting the IQR reporting threshold for submission of CCDE (within 24 hours before/after inpatient admission for 90% of discharges, and linking variable [used to merge EHR to claims data] for 95% of discharges) that are required to receive their Annual Payment Update. CMS was responsive to these comments and has proposed that the submission of CCDE remain voluntary for 2025 reporting.² Additionally, CORE (the measure developer) is updating the data collection approach (effective with the 2025 Annual Update Cycle) to expand the CCDE lookback period beyond the 24 hours prior to/after inpatient admission, to the first result captured during the hospital encounter. By increasing the window from which CCDE can be extracted, hospitals are likely to report CCDE for a higher percentage of discharges, improving their ability to meet the IQR submission percentage.

Finally, we also note that after initial feasibility testing of the CCDE during measure development we identified potential for barriers related to data collection for some data elements. For example, we found a lower capture rate for certain lab tests. Because of a lower capture rate among surgical patients, we finalized the specifications to utilize CCDE lab values in risk adjustment for surgical divisions only if optionally reported by hospitals; imputation of missing lab values was not performed. We refer readers to the Use/Usability section of this form for recent updates. Additionally, because of low capture rates of bicarbonate for all divisions, we expanded the Bicarbonate Lab Test value set to include carbon dioxide lab codes, which are often performed in lieu of bicarbonate lab tests. We refer readers to section 4.3.4 Validity Testing Results (Missing Data), where results show missingness for bicarbonate lab tests for 2024 Voluntary Reporting ranged from 5.70% to 13.09%, showing improvement for this data element from initial development testing.

Analyses around missing data are presented in Section 4.3.4.

The estimated costs of data collection are minimal, as this measure utilizes information from EHR systems already within the hospital. We estimate 12 hours for one employee to extract and submit patient files through the Quality Reporting Document Architecture (QRDA) Submission Portal, consistent with all eQMs. This measure is not intended to influence clinical workflow, as CCDE were selected by a Technical Expert Panel (TEP) because they are routinely captured on all adult inpatients, and data are submitted electronically (CCDE); other measure components, including other risk adjustment variables, numerator and denominator inclusion and exclusion are captured using Medicare inpatient claims.

References

1. Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long-Term Care Hospital Prospective Payment System and Policy Changes and Fiscal Year 2022 Rates; Quality Programs and Medicare Promoting Interoperability Program Requirements for Eligible Hospitals and Critical Access Hospitals; Changes to Medicaid Provider Enrollment; and Changes to the Medicare Shared Savings Program; Corrections. Published October 20, 2021.
<https://www.federalregister.gov/documents/2021/10/20/2021-22724/medicar...>
2. Medicare and Medicaid Programs: Hospital Outpatient Prospective Payment and Ambulatory Surgical Center Payment Systems; Quality Reporting Programs, Including the Hospital Inpatient Quality Reporting Program; Health and Safety Standards for Obstetrical Services in Hospitals and Critical Access Hospitals; Prior Authorization; Requests for Information; Medicaid and CHIP Continuous Eligibility; Medicaid Clinic Services Four Walls Exceptions; Individuals Currently or Formerly in Custody of Penal Authorities; Revision to Medicare Special Enrollment Period for Formerly Incarcerated Individuals; and All-Inclusive Rate Add-On Payment for High-Cost Drugs Provided by Indian Health Service and Tribal Facilities. Published July 10, 2024.
<https://www.cms.gov/medicare/payment/prospective-payment-systems/ambula...>

4.2 Attach Feasibility Scorecard

[eCQM-Feasibility-Scorecard_Hybrid HWM_Fall 2024.xlsx](#)

4.3 Feasibility Informed Final Measure

Based on results from testing shown below, the CCDE were selecting for the final specifications to be feasible to extract, and routinely collected for all adult inpatient EHRs.

To address CBE's requirement for feasibility in relation to data elements and measure logic, we reevaluated this measure against the feasibility domains (see attached Feasibility Scorecard). The results of feasibility assessment for the 10 data elements are below:

- The data elements are in a structured format within the EHR systems (scoring 1 for Availability),
- Some data elements were transmitted directly from other electronic systems into the EHR or resulted from clinician assessment or interpretation (scoring 1 for accuracy)
- This measure's data elements are coded using either RXNORM or SNOMED (scoring 1 for data standards)
- The data elements required for this measure (lab values, vital signs, referral orders, problem list entries) are captured during the course of care and do not impact workflow (scoring 1 for workflow)

4.4 Proprietary Information

Not a proprietary measure and no proprietary components

5.1.1 Data Used for Testing

Please see Table 4 in the Tables and Figures attachment.

5.1.2 Differences in Data

For the updated testing in this CBE endorsement submission, we provide results from two datasets. Each dataset is described in detail in Table 4, in the attachment.

- **HWM Claims-Only Dataset.** This dataset includes both FFS and MA admissions and was used to provide a national dataset that includes all Medicare beneficiaries. While this dataset includes all claims-based variables used for risk adjustment, it does not include the CCDE EHR elements; the addition of the EHR data elements in the CCDE provides risk adjustment supplemental to claims-based risk adjustment, therefore results derived from the HWM claims-only dataset provides a close approximation for nationally representative results. Measure scores calculated with and without the CCDE are highly correlated.
- **2024 Hybrid HWM Voluntary Reporting Dataset:** This dataset was used to provide information on the integration of EHR data elements (CCDE) for case-mix risk-adjustment that supplements the claims-based variables; this dataset includes both claims-based risk adjustment variables and EHR-based data elements (CCDE). This dataset, however, does not include MA admissions, as MA were not part of the measure specifications at the time of data collection; MA admissions will be added to the measure in 2026 Reporting per the Fiscal Year 2024 Inpatient Prospective Payment System Final Rule.

See dataset descriptions in Table 4 (see Tables and Figures attachment) for further details on each dataset. We note that we also include results derived from datasets from original measure development as this is the data that was used for risk variable selection.

5.1.3 Characteristics of Measured Entities

For this measure, hospitals are the measured entities. All non-federal, short-term acute care inpatient US hospitals (including territories) with Medicare fee-for-service (FFS) and Medicare Advantage (MA) beneficiaries aged between 65 and 94 years are included. For Data Element Validity and Feasibility testing, we present testing from the original development of this measure, using a historical dataset—21 Kaiser Permanente hospitals, that includes all-payer and all adults aged 18+. The number of measured entities varies by testing type: see Table 4 in the attachment.

5.1.4 Characteristics of Units of the Eligible Population

Please see Table 4 in the Tables and Figures attachment.

5.2.1 Level(s) of Reliability Testing Conducted

Person or encounter level (i.e., data element) (e.g., inter-abstractor reliability), Accountable entity level (i.e., measure score) (e.g., signal-to-noise analysis)

5.2.2 Method(s) of Reliability Testing

Data Element Reliability (Patient/Encounter Level)

Data element reliability for the EHR-based variables (CCDE) used in the Hybrid HWM measure has been established previously during the development of another measure (the Hybrid Hospital-Wide Readmission measure) (CBE ID: 2879). In this testing we used the capture rate to establish reliability. We refer readers to the attached 2013 Core Clinical Data Elements Technical Report (Version 1.1) for methodologic details.

Measure Score Reliability: Split-Sample

To ascertain measure score reliability, we calculated the intra-class correlation coefficient (ICC) using a split-sample (also known as the split-half) method in both the Claims-Only HWM (Medicare FFS + MA) (discharges July 1, 2022-June 30, 2023), and the Hybrid HWM 2024 Voluntary Reporting (July 1, 2022-June 30, 2023) datasets. We did not calculate signal-to-noise reliability for the overall measure score because the signal-to-noise calculation should be based on a statistical model;¹ the measure score (risk-standardized readmission rate or RSRR) for the HWM measure is a combined score that is not calculated from a single statistical model.

The reliability of a measurement is the degree to which repeated measurements of the same entity agree with each other. For measures of hospital performance, the measured entity is the hospital, and reliability is the extent to which repeated measurements of the same hospital give similar results. Accordingly, our approach to assessing reliability is to consider the extent to which assessments of a hospital using different but randomly selected subsets of patients produce similar measures of hospital performance. Hospital performance is measured once using a random subset of patients from a defined dataset from a measurement period, and then measured again using a second random subset exclusive of the first from the same measurement period, and the agreement of the two resulting performance measures compared across hospitals.²

For split-sample reliability of the measure, we randomly sampled half of patients within each hospital from a one-year measurement period, calculated the measure for each hospital, and repeated the calculation using the second half of patients. Thus, each hospital is measured twice, but each measurement is made using an entirely distinct set of patients. To the extent that the calculated measures of these two subsets agree, we have evidence that the measure is assessing an attribute of the hospital, not of the patients. As a metric of agreement, we calculated the intra-class correlation coefficient³. Specifically, we used the Claims-Only Hospital Wide Mortality (Medicare FFS + MA) and 2024 Hybrid Hospital-Wide Mortality Voluntary Reporting datasets, randomly split each into two approximately equal subsets of patients, and then calculated the RSRR for each hospital for each sample. The agreement of the two RSRRs was quantified for hospitals in each sample using the intra-class correlation as defined by ICC (2,1).³

Using two non-overlapping random samples provides a conservative estimate of the measure's reliability, compared with using two random, but potentially overlapping samples which would exaggerate the agreement. Moreover, because our final measure is derived using hierarchical logistic regression, and a known property of hierarchical logistic regression models is that smaller volume hospitals contribute less 'signal', a split sample using a single measurement period would introduce extra noise. This leads to an underestimate in the actual split-sample reliability that would be achieved if the measure were reported using the full measurement period, as evidenced by the Spearman Brown prophecy formula.⁴ We used this formula to estimate the reliability of the

measure if the whole cohort were used, based on an estimate from half the cohort.

References

1. Adams J, Mehrota, A, Thoman J, McGlynn, E. (2010). Physician cost profiling – reliability and risk of misclassification. *NEJM*, 362(11): 1014-1021.
2. Rousson V, Gasser T, Seifert B. "Assessing intrarater, interrater and test-retest reliability of continuous measurements," *Statistics in Medicine*, 2002, 21:3431-3446.
3. Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 1979, 86, 420-3428.
4. Spearman, Charles, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271-295.

5.2.3 Reliability Testing Results

Data Element Reliability—Clinical Data Element Capture

Data element reliability for the EHR-based variables (CCDE) used in the Hybrid HWM measure has been established previously during the development of another measure (the Hybrid Hospital-Wide Readmission measure) (CBE ID: 2879). Data element reliability testing showed a rate of capture of greater than 90% for all cohorts (except for Surgical/Gynecological, which does not use lab values). We refer readers to the attached 2013 Core Clinical Data Elements Technical Report (Version 1.1) for testing details. This data element testing was analyzed within the clinical groupings of the Hybrid Hospital-Wide Readmission measure, which has fewer clinical groupings than the Hybrid HWM measure. However, we would expect the results to be representative of the data elements that are used within the more granular clinical groupings of the mortality measure.

Measure Score Reliability Results

In the Claims-Only HWM [Medicare FFS and Medicare Advantage] dataset (Dataset 2), there were 4,247 hospitals in the development sample and 4,446 hospitals in the validation sample. The intraclass correlation between the two RSRRs for each sample was 0.736, which meets the current CBE threshold for split-sample reliability (0.6).

In the Hybrid HWM 2024 Voluntary Reporting dataset (Dataset 3), there were 1,039 hospitals in the development sample and 1,042 hospitals in the validation sample. The intraclass correlation between the two RSRRs for each sample was 0.784, which also meets the current CBE threshold for reliability (0.6).

We note that we did not complete Table 5 in this form because the split-sample reliability calculation results in a single statistic, not a distribution.

5.2.4 Interpretation of Reliability Results

Data Element Reliability

Based on prior testing of the CCDE for a related measure (the Hybrid Hospital-Wide Readmission

measure, CBE ID: 2879), we determined that the CCDE for the HWM measure are sufficiently reliable for use for enhancement of risk adjustment. Please also see CCDE data element validity testing (Section 4.3.3) that also satisfies the requirement for data element reliability testing.

Measure Score Reliability Results

The split-sample reliability score (using the Claims-Only HWM [Medicare FFS and Medicare Advantage] and Hybrid HWM 2024 Voluntary Reporting datasets) of 0.736 and 0.784, respectively, meet the current CBE threshold for split-sample reliability (0.6).¹

Reference:

1. Battelle (2023). Endorsement & Maintenance (E&M) Guidebook. Partnership for Quality Measurement. July 2024.

5.3.1 Level(s) of Validity Testing Conducted

Person or encounter level (i.e., data element) (e.g., sensitivity and specificity), Accountable entity level (i.e., measure score) (e.g., criterion validity)

5.3.3 Method(s) of Validity Testing

Data Element Validity Testing (CCDE)

Chart Abstraction:

We developed electronic specifications (e-specifications) using the Measure Authoring Tool (MAT) and analyzed extracted data from EHRs. We assessed the ability of hospitals to use the e-specifications to query and electronically extract CCDEs from the EHR, within 24 hours before or up to 24 hours after inpatient admission for labs; within 24 hours before or up to 2 hours after inpatient admission for vital signs, for all adult inpatient admissions occurring over the course of one year.

Validity testing assessed the accuracy of the electronically extracted CCDEs compared to the same CCDEs gathered through manual abstraction (from the EHR) in a subset of 368 charts identified in the data query in 3 hospitals that used Cerner as their EHR Vendor (Dataset 4), and 391 charts identified in the data query in data extracted from 1 hospital with 391 admissions that used GE Centricity as their clinical EHR (Dataset 5).

We calculated the number of admissions that needed to be randomly sampled from the EHR dataset and manually abstracted to yield a statistical margin of error (MOE) of 5% and a confidence level of 95% for the match rates between the two data sources. Sites then used an Access-based manual abstraction tool provided (along with training) to manually abstract the CCDEs from the random samples of the medical records identified through the EHR data query. The manual chart abstraction data is considered the “gold standard” for the purpose of this analysis.

We conducted validity testing on the critical EHR data elements in the Hybrid HWM measure. For each continuous data element, we were only interested in the case where the electronic

abstraction value exactly matched the manual abstraction value. We therefore only calculated the raw agreement rate between data from electronic and manual chart abstraction. For simple data values, we believe taking this approach, as compared to reporting statistical tests of accuracy, better reflects the concept of matching exact data values rather than calculated measure results. Therefore, we do not report statistical testing of the accuracy of the EHR derived data value as compared with the abstracted value. Instead, we counted only exact matches in the data value as well as the time and date stamp associated with that value when we calculated the match rate. The 95% confidence level was established based on the sample size and reflects the exact match rate using these criteria.

Missing Data

For the EHR data elements used in the measure's risk models, we anticipate that there will be some missing data. We examined rates of missing data using the Hybrid HWM 2024 Voluntary Reporting dataset (discharges July 1, 2022-June 30, 2023), which includes CCDE submitted by 1,162 hospitals during Confidential Reporting (Table 9 in the Tables and Figures attachment). We characterize CCDE as "missing" when: 1) the hospital did not report any data for that value, or 2) when the reported value is unusable in risk-adjustment (e.g. missing units or data not able to be standardized [string data, or values which cannot be converted to primary Unified Code for Units of Measure (UCUM) units without additional information]). For measure calculation, where CCDE values are missing or unusable, we impute the median value reported across all hospitals for that CCDE, to profile a "typical" patient.

Measure Score Validity

Face Validity

We systematically assessed the face validity of the Hybrid HWM measure score as an indicator of quality by confidentially soliciting the TEP members' agreement with the following statement via an online survey following the final TEP meeting: "The risk-standardized mortality rates obtained from the Hybrid Hospital-Wide Mortality Measure as specified can be used to distinguish between better and worse quality facilities." The survey offered participants response options on a six-point scale (1=Strongly Disagree, 2=Moderately Disagree, 3=Somewhat Disagree, 4=Somewhat Agree, 5= Moderately Agree, and 6=Strongly Agree).

Empiric Validity

To assess the construct validity of the HWM measure, we identified and assessed the measure's correlation with other measures with publicly reported data that we hypothesized would be related to mortality based on the evidence for similar or overlapping causal pathways (see the Logic Model in Section 2.1).

Because the Hybrid HWM measure is not nationally representative, due to lack of availability of EHR data from a nationally representative set of hospitals, we used the claims-only HWM measure

(the same measure but without EHR elements used to supplement claims-based risk adjustment) for validity testing. As we have previously shown that measure scores calculated with these two measures (hybrid, claims only) are highly correlated (>0.99), it is acceptable to use the claims-only measure for this analysis.

In order to test the validity of the HWM measure score, we examined whether better performance on the claims-only HWM measure was related to better performance for other relevant structural and outcome measures. However, there are multiple challenges associated with this approach:

1. There are many measures that use a variety of criteria to define a high performing hospital, including: adherence to core processes of care, complications and safety measures, and patient satisfaction.
2. Together with our Technical Workgroup, which consists of nationally recognized experts in measure development, as well as other measurement experts, we have concluded that there is no single recognized and accepted “gold standard” measure that specifically measures factors most relevant to such a broad measure as hospital-wide mortality. Our approach was to select three separate assessments against which we could compare the measure score with the hypothesis that a trend toward correlation with these external assessments would support a conclusion of high measure score validity.

After reviewing available measures, we selected the following three to use for validity testing.

1. Nurse-to-bed ratio: Several studies have found that higher levels of nurse staffing are associated with improved patient outcomes and lower mortality rates.¹⁻⁴ We used a nurse-to-bed ratio calculated using two fields from the American Hospital Association’s (AHA) annual survey. The AHA surveys all hospitals in the United States and the response rate averages 85–95 percent annually [5], covering about 6,000 hospitals.⁵ Staffing is measured as the numbers of full-time and part-time RNs, and LPNs. Within the American Hospital Associations annual survey from 2019, we used the fields “FTEN” and “HOSPD”, which are self-reported fields that are defined in the AHA data dictionary as the number of reported full-time registered nurse and number of hospital beds.

2. Hospital Star Rating Mortality Group Score: CMS’s Overall Hospital Star Rating assesses hospitals’ overall performance (expressed on CMS’ Care Compare tool on Care Compare, graphically as stars) based on a weighted average of group scores from five different domains of quality (mortality, readmissions, safety, patient experience, timely and effective care). The Mortality Group is comprised of the mortality measures that are publicly reported on the care compare tool. The Mortality Group Score is calculated using a simple average of the scores for the individual measures. The Mortality Group score is on a higher-is-better scale. We used mortality group scores from 4,581 hospitals from the 2021 release. The full methodology for the Overall Hospital Star Rating can be found at:

<https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic...>

3. Overall Hospital Star Rating: CMS’s Overall Hospital Star Rating assesses hospitals’ overall performance based on a weighted average of group scores (the Summary Score) from five different domains of quality (mortality, readmissions, safety, patient experience, timely and effective care). For the validity testing presented in this testing form, we used hospital’s Summary

Score from 4,544 hospitals from the 2021 release that includes results from measures with dates of data that align with the HWM testing dataset. The full methodology for the Overall Hospital Star Rating can be found at

<https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=OnetPublic...>

We examined the relationship of performance on the claims-only measure scores (RSMR) with each of the three external measures of hospital quality. For the external measures, the comparison was against performance within quartiles for Nurse-to-bed ratio, Star Ratings Mortality Group score, or Overall Hospital Star Ratings Summary Score. We then hypothesized the strength and the direction of the relationship for each measure (Table 6 in the Tables and Figures attachment). For the HWM measure, a lower measures score means better performance, therefore for comparator measures where better performance is hypothesized to be related to be better performance on HWM (such as nurse-to-bed ratio), the direction of the association is shown as “negative.”

We then examined the relationship of performance of the HWM measure scores (RSMRs) with each of these external measures of hospital quality as measured by Pearson correlation coefficients (Table 7 in the Tables and Figures attachment). We also compared performance of the HWM measure within quartiles of the comparator measures (Figures 5-7). For purposes of this testing, we used the Claims-Based HWM (Medicare FFS and MA) dataset (discharges July 1, 2018, through June 30, 2019) as it is a national sample, similar to the Star Ratings.

References:

1. Aiken LH, Clarke SF, Sloane DM, Sochalski J, Silber JH. Hospital Nurse Staffing and Patient Mortality, Nurse Burnout, and Job Dissatisfaction. *Journal of the American Medical Association*. 2002;288(16):1987-93.
2. American Hospital Association (AHA) The AHA Annual Survey Database Fiscal Year 1997 Documentation. Chicago: Health Forum; 1999.
3. Griffiths P, Ball J, Murrells T, et al Registered nurse, healthcare support worker, medical staffing levels and mortality in English hospital trusts: a cross-sectional study. *BMJ Open* 2016;6:e008751. doi: 10.1136/bmjopen-2015-008751
4. Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse-Staffing Levels and The Quality of Care in Hospitals. *New England Journal of Medicine*. 2002;346:1719-22.
5. Needleman J., Buerhaus P., Pankratz V.S., Leibson C.L., Stevens S.R., Harris M. Nurse staffing and inpatient hospital mortality. *N. Engl. J. Med*. 2011;364:1037-1045. doi: 10.1056/NEJMsa1001025. Centers for Medicare & Medicaid Services (CMS). Hospitals - Overall hospital quality star rating | Provider Data Catalog. Data.CMS.gov. <https://data.cms.gov/provider-data/topics/hospitals/overall-hospital-qu...>;

5.3.4 Validity Testing Results

Summary

Our validity testing results, described below, provide strong evidence for the validity of the EHR based data elements (CCDE) (based on comparison of data elements through chart abstraction and analysis of missing data), and validity of the measure score, shown through construct validity,

and face validity.

Validity of EHR Data Elements

We note that all candidate CCDE were previously identified by a Technical Expert Panel to be routinely collected on all adult inpatients, in order to guide clinical decision making. See the 2013 Core Clinical Data Elements Technical Report (Version 1.1) attached for details.

Of the candidate variables, chart abstraction for validity testing was done in Dataset 4 (Electronically extracted data from 3 hospitals with Cerner as their EHR) and Dataset 5 (Data element validity dataset- from one hospital using GE Centricity as their EHR vendor). [Table 7](#) in the Tables and Figures attachment demonstrates the comparison between electronic and manual abstraction of data in the two health systems. We found that the percent agreement between the EHR-based variables and chart-abstracted values ranged from 14.66% to 97.22%.

We note that a post-validation review of the code used by the hospital in Dataset 5 (one hospital with GE Centricity as EHR vendor), revealed that the hospital experienced a number of errors. The most significant of which was extracting data only within an incorrect two-hour window for laboratory test results (the correct window was 24 hours). Additionally, physical exam (vital signs) data were extracted based on the date/time that results were documented rather than the date/time the physical exams were performed, driving down the accuracy of these data. However, post-validation review of the code used by the hospital in Dataset 4 (three hospitals with Cerner as their EHR vendor) showed no such errors in the query executed. As a result, the match rate for Dataset 4 was much higher.

Our analysis of missing data (Table 8 in the Tables and Figures attachment) showed that reporting of CCDE varied across cohorts, from 1.02% missing for Systolic Blood Pressure in the Surgical Orthopedic division, to 20.60% for White Blood Cell Count in the Non-Surgical Cardiac Division.

Measure Validity Testing Results

Face Validity

A total of six of the eight TEP members completed the face validity survey. Of the six respondents, five respondents (83%) indicated that they somewhat, moderately, or strongly agreed and one somewhat disagreed with the following statement: "The risk-standardized mortality rates obtained from the Hybrid Hospital-Wide Mortality Measure as specified can be used to distinguish between better and worse quality facilities." Survey results from the TEP indicated high agreement (83%) regarding the face validity of the Hybrid HWM measure.

Empiric Measure Score Validity

Table 11 in the Tables and Figures in the attachment shows the results of the analyses examining associations between the HWM measure and both structural and quality metrics in the same

causal pathway, as described in Section 2.1. All analyses were performed with the Medicare FFS + MA claims-only dataset.

Nurse to bed ratio. As hypothesized, we found a weak but significant, negative association (-0.138, $p < .0001$) between the HWM measure scores and the nurse-to-bed ratio (Table 9 in the Tables and Figures attachment). We demonstrate this relationship graphically in [Figure 5](#) in box-whisker plots showing the distribution of the HWM measure RSMRs within each quartile of Nurse-to-bed Ratio. The blue diamonds represent the mean RSMRs of Nurse-to-bed ratio quartiles. The association between HWM RSMRs and Nurse-to-bed ratio suggests that hospitals with lower HWM RSMRs (better performance) are more likely to have higher Nurse-to-bed ratios (better performance), which is consistent with our logic model, and supported by the literature.

Mortality Group Score and Star Rating Summary Score: As hypothesized, the HWM measure score was moderately, negatively and significantly associated with both the Star Rating Standardized Mortality Group Score (-0.574, $p < .0001$), and the Summary Score (-0.324, $p < .0001$), meaning that higher scores (better performance) on the comparator measures were associated with lower scores (better performance) on the HWM measure. This is expected because the star ratings quality measures focus on, or contain a portion of, the same domain of quality as the HWM measure (mortality). Box plots (whisker plots) that visualize these relationships are shown below in Figures 5-7 (see attachment).

[Figure 6](#) shows the box-whisker plots of the HWM measure RSMRs within each quartile of Star Rating Mortality Group Score. The blue diamonds represent the mean HWM RSMRs within Star Rating Mortality Group score quartiles. The correlation between HWM RSMRs and Star Rating Mortality Group score is -0.574, which demonstrates that hospitals with lower HWM RSMRs (better performance) are more likely to have higher Star-Rating Mortality Group scores (better performance).

[Figure 7](#) shows the box-whisker plots of the HWM measure RSMRs within each quartile of the Star Rating Summary Score. The blue diamonds represent the mean RSMRs within Summary Score quartiles. The correlation between HWM RSMRs and Star Rating Summary Score is -0.324, which demonstrates that hospitals with lower HWM RSMRs (better performance) are more likely to have higher Star Rating Summary Scores (better performance).

5.3.5 Interpretation of Validity Results

Our data element, construct, and face validity testing results all support the validity of the Hybrid HWM Measure. We discuss each category of testing below.

Data Element Validity

Our chart abstraction results, presented here, show a high percent agreement for most variables. We note that the lower capture rate (see reliability section) and lower % agreement rate for the bicarbonate value have been addressed in measure updates that were made to the accepted value for this data element (see Section 3.1); we also note that this value is only used within the surgical divisions, where agreement was lowest, if reported by hospitals. The rate of missing values continues to be low for most data elements and not likely to introduce bias. We note that the

impact of missing values for the White Blood Cell laboratory test is very low. As we employ an imputation strategy for missing data, missingness is unlikely to have a meaningful impact on measure scores. We refer readers to section 4.3.4 Testing Results (Missing data) which shows improved missingness from development testing. We note that the missingness of the “platelet” variable in the Surgical Orthopedics division (about 34%) is not due to limitations in data capture, but rather due to an implementation strategy to count all unusable values as missing. For this particular CCDE, values were reported in a unit that is unusable for measure calculation, due to an inability to standardize across units. An alternative standardization strategy has been implemented for 2025 Reporting to account these values as useable for this variable. Please see Section 3.1 and Section 6.2.3 for more details. We note that for missing CCDE values, multiple imputation is used to impute a value based on the characteristics of the CCDE reported. To minimize any small potential for bias from CCDE values, we account for potential outlier values, using winsorization, as well as account for missing values in our risk models. It is expected that CCDE reporting will continue to improve in future years, when the CCDE lookback period is expanded beyond 24 hours, and as hospitals gain familiarity with the measures.

Measure Score and Face Validity

Face validity and measure score validity testing support measure score validity. We found associations between structural (nurse-to-bed ratio) and quality metrics (Star Rating Mortality Group Score, and Summary Score) that were significant, with the expected strength, and in the expected direction. TEP face validity voting results were strong and provide support for the validity of the measure.

5.3.2 Type of Accountable Entity Level Validity Testing Conducted (derived)

Empirical validity testing at the accountable entity-level (e.g., criterion validity, construct validity, known groups analysis), Systematic assessment of face validity of the measure’s performance score as an indicator of quality or resource use

5.4.1 Methods Used to Address Risk Factors

Statistical risk adjustment model with risk factors

5.4.2 Conceptual Model Rationale

This section addresses clinical risk variables; please see Section 5.1 for a discussion of social risk factors.

Approach to Variable Selection

Our approach to risk adjustment was tailored to and appropriate for a publicly reported outcome measure, as articulated in the American Heart Association (AHA) Scientific Statement, “Standards for Statistical Models Used for Public Reporting of Health Outcomes”.¹ The measure estimates hospital-level 30-day all-cause RSMRs using hierarchical logistic regression models. In brief, the approach simultaneously models data at the patient and hospital levels to account for variance in patient outcomes within and between hospitals.²

The approach to risk adjustment differs from usual claims-based measures. The original claims-

only HWM measure uses claims data to adjust for two aspects of risk: 1) case mix or how sick individual admitted patients are; and, 2) service mix or the proportion of admitted patients with various different principal discharge diagnoses.

The goal of the hybrid measure is to enhance risk adjustment using clinical data from the electronic health record (EHR). To select candidate variables for the Hybrid risk model, we began with the list of all administrative claims-based risk-adjustment variables included in the claims-only HWM measure, described below. We then added EHR-based risk variables, also described below.

Selecting Risk Variables

Candidate Comorbid Risk Variables

Our goal is to develop parsimonious models that include clinically relevant variables strongly associated with the risk of mortality in the 30 days following an index admission. For candidate variable selection, using the development sample we started with the CMS Condition Categories (CC)s grouper, used in previous CMS risk-standardized outcome measures, to group ICD-9 codes into comorbid risk adjustment variables.

To select candidate variables, a team of clinicians reviewed all CMS-CCs and combined some of these CMS-CCs into clinically coherent groups to ensure adequate case volume. Any combined CMS-CCs were combined using both clinical coherence and consistent direction of mortality risk prediction across the CMS-CC groups in the majority of the 15 divisions.

Potential Complications of Care During Hospitalization

Complications occurring during hospitalization are not comorbid illnesses and do not reflect the health status of patients upon presentation. In addition, they likely reflect hospital quality of care, and, for these reasons, should not be used for risk adjustment. Although adverse events occurring during hospitalization may increase the risk of mortality, including them as risk factors in a risk-adjusted model could lessen the measure's ability to characterize the quality of care delivered by hospitals. We have previously reviewed every CMS-CC and identified those which, if they were to occur only during the index hospitalization, are more likely than not to represent potential complications rather than pre-existing comorbidities. For example: fluid, electrolyte, or base disorders; sepsis; and acute liver failure are all examples of CMS-CCs that could potentially be complications of care.

For the claims-only HWM measure, we took a two-step approach to identifying complications of care. First, we searched the secondary diagnosis codes in the index admission claim for all patients in the measure and identified the presence of any ICD-9 code associated with a CMS-CC. If these codes appeared only in the index admission claim, we flagged them because they are potential complications of care. Next, we determined if these potential complications of care were associated with a "present on admission" code. Any potential complication of care with an associated "present on admission" code was kept in the risk model; any potential complication of care without an associated "present on admission" code was removed under the assumption that it represented a complication of care. In this way, we supplemented the existing approach to identifying potential complications of care used in CMS's publicly reported mortality measures by

incorporating “present on admission” codes. Our analyses demonstrate that a majority of hospitals currently use “present on admission” codes across a majority of conditions. Therefore, we felt that a combined approach to excluding complications of care from the risk model that used both the existing methodology and “present of admission” codes allow the measure to capture as many clinically appropriate risk variables as possible while simultaneously removing complications of care from the risk model.

Final Comorbid Risk Variable Selection

To inform variable selection, we used the development sample to create 500 bootstrap samples for each of the service-line divisions. (This analysis was performed prior to removing the divisions Other Non-Surgical Conditions and Other Surgical Procedures; therefore, this analysis was completed on 15 divisions. During ICD-10 re-specification, these divisions were added back to the measure.) For each sample, we ran a standard logistic regression model that included all candidate variables. The results were summarized to show the percentage of times that each of the candidate variables was significantly associated with 30-day mortality (at the $p \leq 0.05$ level) in the 500 bootstrap samples (for example, 90% would mean that the candidate variable was significant at $p \leq 0.05$ in 90% of the bootstrap samples). We also assessed the direction and magnitude of the regression coefficients.

We found that models containing all risk factors performed similarly to models containing a more limited set of “significant” risk factors, described below. We therefore used a fixed, common set of comorbidity variables in all of our models for simplicity and ease of implementation and analysis. We describe below the steps for variable selection.

- a. The CORE Project Team reviewed the bootstrapping results and decided to provisionally examine risk adjustment variables at or above a 90% cutoff in one of the 15 service-line division models (in other words, retain variables that were significant at the $p \leq 0.05$ level in at least 90% of the bootstrap samples for each division). We chose the 90% cutoff because this threshold has been used across other measures and produced a model with adequate discrimination.
- b. In order to develop a statistically robust and parsimonious set of comorbid risk variables, we then chose to limit the variables to those that met a 90% threshold in at least 13/15 divisions. This step resulted in the retention of 31 risk factors, including age and 19 comorbid risk variables. This resulted in C-statistics that did not change by more than 0.02 in any of the 15 divisions compared to models that contained all possible risk variables.

CCDE Risk Variable Selection

To select candidate clinical EHR variables, we began with the list of candidate CCDE variables shown below in Table 10 in the Tables and Figures attachment. The table includes all tested, candidate elements. Of those tested, ultimately ten variables were chosen for the final model.

First, we looked at how many admissions in our Clinical Hybrid Dataset were missing values for each CCDE. The non-surgical divisions had fewer than 10% of admissions that were missing values. However, in the surgical divisions, while vitals were missing in fewer than 10% of admissions, the laboratory result values were missing in 15% - 50% of admissions, depending upon division. For development purposes only, we imputed values for missing labs or vital signs,

as described below:

For all admissions missing any vital signs and for admissions within the non-surgical divisions missing any laboratory result values, we used multiple imputation (imposing limits to ensure the imputed values were within clinical possibilities) with 5 copies of data with different imputations based on a multi-normal distribution.

For admissions within the surgical divisions missing any laboratory results, we randomly imputed a value within the normal range for that lab. For the normal ranges, see Table 11 Candidate Clinical EHR Risk Variable (CCDE) Mortality Association Modelling Approaches below. Rationale: Surgical patients that are missing initial labs are most likely elective surgical admissions that had the labs collected within 30 days PRIOR TO ADMISSION. It is less likely that a patient with an extremely abnormal lab value would undergo an elective surgery without having the labs checked again on admission. This approach is for development purposes only.

Second, we selected which CCDE would be the most appropriate to include in the hybrid HWM measure. We approached risk variable selection from the perspective of ensuring a parsimonious list of clinical EHR variables that would minimize hospital burden to report the data and provide face validity from a clinical perspective.

Therefore, we first sought to ensure that each candidate variable was modeled in a clinically appropriate way. For example, the laboratory value sodium has a U-shaped predictive association with mortality: Normal sodium levels are associated with a low risk of mortality, while both abnormally high and abnormally low levels are associated with an increased risk of mortality. The association between each CCDE variable and mortality was reviewed by four clinicians and selected based on the best association. See Table 3 Candidate Clinical EHR Risk Variable (CCDE) Mortality Association Modelling Approaches for the approach used for each risk variable. In addition, we report the normal values used for imputing missing laboratory results within the surgical divisions.

Based upon this information, we selected a standard set of clinically coherent risk variables in order to ensure that each division-level risk model included key laboratory results and vital signs data (see final list of EHR risk variables above, in Section 2b3.1.1). As with prior hybrid measures that use EHR data in their risk model, we did not include risk variables if they were strongly correlated with another variable. For example, we selected systolic blood pressure but not diastolic blood pressure, as these variables were highly correlated and provide very similar risk prediction. Using a standard set of clinically selected variables produced improved c-statistics compared to the models based purely upon stepwise selection. We also tested allowing the risk variables to vary across the 15 divisions (using stepwise selection) but still forcing in clinical variables and found that the model discrimination (c-statistic) was very similar, in some cases identical, to using a standard set of variables. Therefore, we proceeded with a common set of 10 clinical risk variables plus age across all divisions for the remainder of the risk model development work. for

Service-Line Adjustment

We use the AHRQ CCS grouper to group all ICD-10 **principal discharge diagnoses** into clinically coherent categories (categories have been somewhat modified as described below). For

all AHRQ principal discharge diagnosis code CCSs with sufficient volume (CCSs with fewer than 100 admissions are excluded), we also included a discharge diagnosis-specific indicator in the model. This ensures that the principal discharge diagnosis for each patient is also included in the risk model, in addition to the 20 comorbid risk variables described above.

Rationale: Discharge diagnosis categories differ in their baseline mortality risks and hospitals will differ in their relative distribution of these discharge diagnosis categories (service mix) within each division.^{3,4,5} Therefore, adjusting for principal discharge diagnosis categories levels the playing field across hospitals with different service mixes. See the data dictionary for the CCSs (tabs HWM Non-SurgCohortDiv CCS and HWM SurgicalCohortDiv CCS) that comprise each of the divisions in this measure and Tables 13-14 for the parameter estimates for the CCS categories for each division.

CCS modifications: Note that in addition to using the AHRQ CCS grouper to define the CCS categories in each division (see section S.7 of the submission form), we made two types of modifications: (1) We modified selected CCS highly heterogenous CCS categories to create more homogenous CCS risk variable groups, and so increased the face validity of risk model, described below, and (2) we combined low-mortality CCSs (those with mortality rate of 1% or lower), also described below.

Heterogenous CCSs: In parallel with our approach during measure development in ICD-and in response to feedback from our TEP and Technical Workgroup, we addressed heterogeneity within specific AHRQ CCS groups where the risk of mortality varied significantly across the different ICD-10 diagnoses within the CCS. We calculated the correlation between mortality rates grouped by principal discharge diagnosis ICD-10 code within each CCS. We identified any CCS with an intra-class correlation (ICC) score >0.05 as having high heterogeneity. (The ICC is used in this context to identify heterogeneity of mortality risk across ICD-10 codes within the ICC. The value 0.05, or 5%, is a conventional threshold for accounting for between group heterogeneity.) To address the heterogeneity, three clinicians independently, and through consensus, modified the highly heterogeneous CCSs using clinically informed recategorizations, by either splitting the CCSs into more than one CCS, moving ICD-10 codes to more clinically coherent CCSs, or removing from inclusion ICD-10 codes where quality of care less likely impacts survival, and/or where there were a small number of patients. During ICD-10 respecification, we identified 44 highly heterogeneous CCSs and made modifications to 20 of them, as described in the data dictionary, tab “HWM_CCS_Modifications.”

Low-mortality CCSs: During initial measure development, the patient-level risk models for two divisions (the “Other” surgical and non-surgical divisions) did not converge due to the large number of CCS category codes in these divisions, and due to low mortality rates associated with some of the CCSs in these divisions (which are used for service-line risk adjustment). However, the TEP and Patient and Family Caregiver Workgroup had a strong interest in retaining these admissions (more than half a million admissions) in the measure. To address this issue, within each division, CCSs with low mortality rates (those less than or equal to 1%) are combined into one independent group, which reduces the total number of risk variables (CCS category codes) in the model.

References

1. Blum, A. B., N. N. Egorova, E. A. Sosunov, A. C. Gelijns, E. DuPree, A. J. Moskowitz, A. D. Federman, D. D. Ascheim and S. Keyhani. "Impact of Socioeconomic Status Measures on Hospital Profiling in New York City." *Circ Cardiovasc Qual Outcomes* 7, no. 3 (2014): 391-7.
2. Brodish P.H., Hakes J.K. "Quantifying the individual-level association between income and mortality risk in the United States using the National Longitudinal Mortality Study." *Soc. Sci. Med.*, 170 (2016), pp. 180-187, 10.1016.
3. Calvillo-King L, Arnold D, Eubank KJ, et al. Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review. *Journal of general internal medicine.* 2013;28(2):269-282.
4. Chang W-C, Kaul P, Westerhout C M, Graham M. M., Armstrong Paul W., "Effects of Socioeconomic Status on Mortality after Acute Myocardial Infarction." *The American Journal of Medicine.* 2007; 120(1): 33-39.
5. Demakakos P, Biddulph JP, Bobak M, Marmot MG (2016a) Wealth and mortality at older ages: a prospective cohort study. *J Epidemiol Community Health* 70:346-353.

5.4.2a Attach Conceptual Model

[4.4.2 Conceptual Model Rationale_Hybrid HWM.pdf](#)

5.4.3 Variable Distribution Across Measured Entities

We refer reader to risk variable frequencies for 15 divisions in the attached data dictionary.

We provide results from both the claims-only HWM dataset (MA + FFS cohort, claims-based risk adjustment), and the 2024 Voluntary Reporting dataset (FFS cohort plus CCDE enhanced risk adjustment) to demonstrate results using a national sample, and to include the CCDE variable.

CORE's Approach to Annual Model Validation

CORE's measures undergo an annual measure reevaluation process, which ensures that the risk-standardized models are continually assessed and remain valid, given possible changes in clinical practice and coding standards over time. Modifications made to measure cohorts, risk models, and outcomes are informed by review of the most recent literature related to measure conditions or outcomes, feedback from various stakeholders, and empirical analyses, including assessment of coding trends that reveal shifts in clinical practice or billing patterns. Input is solicited from a workgroup composed of up to 20 clinical and measure experts, inclusive of internal and external consultants and subcontractors.

For 2024 Voluntary Reporting of the Hybrid HWM measure, we:

- Updated the ICD-10 code-based specifications used in the measures. Specifically, we:
 - Incorporated the code changes that occurred in the FY 2023 version of the ICD-10-CM/PCS (effective with October 1, 2022+ discharges) and FY 2023 version of the Yale-modified v4.0 of the Agency for Healthcare Research and Quality (AHRQ) Healthcare Cost and Utilization Project (HCUP)'s beta version 2019.1 Clinical Classification Software (CCS) into the cohort definitions and risk models;

- Applied a modified version of the FY 2023 V24 CMS-Hierarchical Condition Category (HCC) crosswalk that is maintained by RTI International to the risk models; and
- Monitored code frequencies to identify any warranted specification changes due to possible changes in coding practices and patterns
- Evaluated the stability of the risk-adjustment model between initial development and voluntary reporting by examining the model variable frequencies, model coefficients, and the performance of the risk-adjustment model in each year.
- For each of the conditions, we assessed logistic regression model performance in terms of discriminant ability. We computed two summary statistics to assess model performance: the predictive ability and the area under the receiver operating characteristic (ROC) curve (c-statistic).

5.4.4 Risk/Case-Mix Adjustment Modeling and/or Stratification Results

Please see attached data dictionary (Tables 15 and 16) for the final variables for each of the 15 risk models with associated odds ratios, for both the Claims-Only HWM dataset and the Hybrid HWM 2024 Voluntary Reporting dataset.

5.4.5 Calibration and Discrimination

To assess model performance, we assessed model discrimination, calibration, and overfitting. To assess discrimination, we computed two discrimination statistics, the c-statistic and predictive ability. For all analyses, we provide results from both the claims-only (Medicare FFS and MA) and hybrid 2024 VR datasets.

The **c-statistic** is the probability that predicting the outcome is better than chance, which is a measure of how accurately a statistical model can distinguish between a patient with and without an outcome.

Predictive ability measures the ability to distinguish high-risk subjects from low-risk subjects; therefore, for a model with good predictive ability, we would expect to see a wide range in observed outcomes between the lowest and highest deciles of predicted outcomes. To calculate the predictive ability, we calculated the range of mean observed outcomes between the lowest and highest predicted deciles of outcome probabilities.

For assessments of **model calibration**, we provide calibration plots, with mean predicted and mean observed outcomes plotted against deciles of predicted outcomes. The closer the predicted outcomes are to the observed outcomes, the better calibrated the model is.

In addition, we provide an analysis of **overfitting**. Overfitting refers to the phenomenon in which a model accurately describes the relationship between predictive variables and outcome in the development dataset but fails to provide valid predictions in new patients. Estimated calibration values of γ_0 close to 0 and estimated values of γ_1 close 1 provide evidence of good calibration of the model.

Please see the attachment "Hybrid HWM All Tables and Figures" for the model testing results, which are described below.

Model Performance Results

Discrimination and Calibration

As shown in Table 12 and 13 (see attachment "Hybrid HWM All Tables and Figures"), across clinical divisions, c-statistics range from 0.743 to 0.896 in the Claims-Only HWM (Medicare FFS + MA) (discharges July 1, 2018-June 30, 2019) dataset and 0.749 to 0.901 in the Hybrid HWM 2024 Voluntary Reporting (discharges July 1, 2022-June 30, 2023) dataset, respectively.

Model testing shows a wide range of predictive ability (Table 12 and 13), and risk decile plots (Figures 8 and 9 in the attachment "Hybrid HWM All Tables and Figures") show that higher deciles of the predicted outcomes are associated with higher observed outcomes, demonstrating good calibration of the models across both datasets (Claims only, and HWM 2024 Voluntary Reporting).

Overfitting

Table 14 and 15 (see attachment "Hybrid HWM All Tables and Figures", page 13), show overfitting results, demonstrating that γ_0 in the validation samples is close to zero, and γ_1 is close to one across clinical divisions for the national dataset.

5.4.6 Interpretation of Risk/Case-mix Factor Findings

Interpreted together, our diagnostic results demonstrate the risk-adjustment model adequately controls for differences in patient characteristics.

Our model testing results provide evidence for good discrimination (c-statistics) across clinical divisions. Our calibration plots show that higher deciles of predicted outcomes are associated with higher observed outcomes, which show good calibration of the models. The models also show a wide range of predictive ability. The overfitting values for the national dataset show satisfactory calibration for each clinical division in the national dataset.

5.4.7 Final Approach to Address Risk Factors

Statistical risk adjustment model with risk factors

6.1.1 Current Status

In use

6.1.3 Current Use(s)

Public Reporting, Quality Improvement with Benchmarking (external benchmarking to multiple organizations), Quality Improvement (Internal to the specific organization), Other

6.1.3 Program Details

Name of the program and sponsor

Hospital inpatient quality reporting program (IQR), CMS

URL of the program

<https://qualitynet.cms.gov/inpatient>

Purpose of the program

The Hospital IQR, (Medicare Modernization Act [2003]) incentivizes hospitals to report data on inpatient care quality. Results help Medicare beneficiaries make informed healthcare decisions, and promote better quality of care.

Geographic area and percentage of accountable entities and patients included

The Hospital IQR program includes acute care hospitals across the nation with nearly 4,500 hospitals and 70 million Medicare Beneficiaries

Applicable level of analysis and care setting

The level of measurement is the facility; the setting is the Hospital Inpatient.

6.1.3a Other Current Use

While this re-specified measure with MA and FFS admissions is currently not in use, the prior hybrid HWM measure (with FFS admissions only) is in use for public reporting and for quality improvement.

6.2.1 Actions of Measured Entities to Improve Performance

As described in the logic model and in Section 2.1, there are specific, evidence-based actions that hospitals can take to reduce 30-day mortality rates among admitted patients. These interventions include conditions/procedure specific interventions, such as following guideline-based protocols for the care for patients with stroke or AMI, following evidence-based practices for reducing surgical-site infections, and following standardized care for ventilated patients (see Section 2.2 for details). In addition, as described in detail in Section 2.2, there are broader, hospital-wide approaches, such as the use of early warning systems for deteriorating patients, medication reconciliation, and good patient and family communication and coordination of care at discharge, that can contribute to lowering mortality rates.

Specifically for this measure, hospitals can also use the detailed reports shared by CMS to support quality improvement. These reports provide hospitals with their detailed measure results, discharge-level data, and state and national results at the division level, to help hospitals identify specific areas for improvement.

6.2.2 Feedback on Measure Performance

CMS receives feedback on all its measures through the publicly available Q&A tool on Quality Net. Through this tool, we have received, since the last submission, only basic questions about the measure, including the cohort definition, the outcome definition, and specific questions about a facility's data. We did not receive any suggestions for changes to the claims-based portion of this measure.

Additionally, the EHR portion of this measure goes through the Annual Updates Process (required for all eCQMs), which includes coding and logic review. Since 2023 Voluntary Reporting of the measure, we have also received suggestions from stakeholders regarding logic and coding updates through this process, as well as through JIRA.

6.2.3 Consideration of Measure Feedback

Major changes to the Hybrid HWM measure since it was last endorsed in 2019 include the addition of Medicare Advantage patients, which was finalized in the 2024 Inpatient Prospective Payment System (IPPS) Rule to be incorporated in the measure for discharges June 30, 2024-July 1-2025, for 2026 Reporting. ¹ For more information, see Appendix E of the Hybrid HWM comprehensive Methodology Report attached.

Additionally, the exclusion criteria for admissions in a low volume CCS, defined as less than 100 admissions is currently being reassessed for 2025 Reporting as part of updating the SAS pack for measure calculation in the initial testing data. This strategy was employed in order to address convergency issues and to calculate a stable and precise risk model within the context of small sample size. This statistical challenge no longer exists for 2025 Reporting, as the measure will have a national sample. The anticipated impact of updating this exclusion is minimal.

Minor measure updates to the EHR portion of the measure include:

- Annual digital quality measure maintenance, including coding, value set, and logic updates.
- Temporary COVID-related measure modifications:
 - Exclusion of patients with a principal or secondary diagnosis (POA) of COVID-19 from the measure cohort.
 - Risk-adjustment for patients with history of COVID-19.

CORE's annual cycle of measure reevaluation aims to make continuous improvement on the measure, and to be responsive to stakeholder input. Through stakeholder Q&A, developers have been made aware of implementation challenges faced by hospitals from 2024 Voluntary Reporting:

- Hospitals provided feedback regarding the topic of acceptable CCDE units for submission:
 - Hospitals provided feedback in having difficulty determining which units for CCDE are acceptable for submission and ultimately used for measure calculation. For values submitted with units that cannot be converted to the primary UCUM unit, the value is

set to missing. The current strategy for missing or unusable data is to substitute the median value reported for that CCDE, assuming a somewhat typical patient. Developers review units submitted by hospitals for data pre-processing each year, with the goal of including as many units as possible for measure calculation. We note limitations with units unable to be standardized to a common unit without additional lab values, and unusable data such as text/string data (e.g. “high- see Dr. John”), an ongoing challenge in the eCQM community.

- During 2024 Voluntary Reporting we discovered that hospitals were reporting values for the “Platelet” variable within only one of the 15 divisions (Surgical Orthopedics division) in the unit of femtoliters, which is a unit that is currently unusable for measure calculation because there is no formula to convert femtoliters to the appropriate platelet unit (/mm³) for standardization. The measure has been updated to accept femtoliters and impute an appropriate value for measure calculation for 2025 Reporting and future years.
- Hospitals provided feedback regarding submission of linking variables used to merge claims to EHR data, as well as IQR threshold requirements (which are not used towards measure calculation):
 - Hospitals expressed concern reaching IQR program requirements, in which hospitals must submit CCDE (within 24 hours before or up to 24 hours after inpatient admission for labs; within 24 hours before or up to 2 hours after inpatient admission for vital signs) for 90% of discharges and linking variable (used to merge EHR to claims data) for 95% of discharges in order to receive their Annual Payment Update. These comments were heard by CMS and the measure developer, and the proposal for the submission of CCDE to remain voluntary for 2025 reporting was included as a rider to the Outpatient Prospective Payment System Proposed Rule.² Additionally, an update to expand the CCDE lookback period from the 24 hours prior to/after inpatient admission to the start of the hospital encounter is being finalized through the 2025 Annual Updates Cycle. Stakeholders have noted that ED and observation stays have changed in the past several years, with longer ED and observation visit lengths of stay, making it more difficult to submit CCDE within the 24 hours. By increasing the window of time from which CCDE can be extracted, hospitals are likely to report CCDE for a higher percentage of discharges, increasing their ability to meet the IQR submission requirements.
 - Hospitals also expressed concern meeting reporting threshold laboratory requirements for surgical patients. For 2025 Reporting, the exclusion criteria, which removes patients with more than 5 of 10 CCDE missing from the cohort, will be updated to accommodate division-specific CCDE missingness. Additionally, surgical lab values will be removed from the risk-model.
 - Additionally, through stakeholder Q&A, hospitals voiced difficulty submitting a linking variable, Medicare Beneficiary Identifier (MBI), for Medicare Advantage patients. While MBI is available for patients for the claims portion of this measure, hospitals note its collection is not fully integrated into hospitals EHR programs. CMS and the measure developer are aware of this limitation for hospitals beginning with Reporting

Year 2026 and have been in contact with multiple hospitals to address this issue for future reporting.

Reference

1. Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long- Term Care Hospital Prospective Payment System and Policy Changes and Fiscal Year 2024 Rates; Quality Programs and Medicare Promoting Interoperability Program Requirements for Eligible Hospitals and Critical Access Hospitals; Rural Emergency Hospital and Physician-Owned Hospital Requirements; and Provider and Supplier Disclosure of Ownership; and Medicare Disproportionate Share Hospital (DSH) Payments: Counting Certain Days Associated with Section 1115 Demonstrations in the Medicaid Fraction. <https://www.govinfo.gov/content/pkg/FR-2023-08-28/pdf/2023-16252.pdf>
2. Medicare and Medicaid Programs: Hospital Outpatient Prospective Payment and Ambulatory Surgical Center Payment Systems; Quality Reporting Programs, Including the Hospital Inpatient Quality Reporting Program; Health and Safety Standards for Obstetrical Services in Hospitals and Critical Access Hospitals; Prior Authorization; Requests for Information; Medicaid and CHIP Continuous Eligibility; Medicaid Clinic Services Four Walls Exceptions; Individuals Currently or Formerly in Custody of Penal Authorities; Revision to Medicare Special Enrollment Period for Formerly Incarcerated Individuals; and All-Inclusive Rate Add-On Payment for High-Cost Drugs Provided by Indian Health Service and Tribal Facilities. Proposed on July 22, 2024. <https://www.federalregister.gov/documents/2024/07/22/2024-15087/medicar...>

6.2.4 Progress on Improvement

The Hybrid HWM measure has undergone one round of Voluntary Reporting in 2024, in which a small subset of hospitals participated (n= 1,125). As such, progress on improvements cannot be generalized due to limited sample, years for comparison, and the self-selecting nature of hospitals that participate in voluntary reporting (in which they typically score better on the mortality outcome).

6.2.5 Unexpected Findings

There were no unintended impacts during implementation of this measure on patients or in care delivered by hospitals. However, there were some challenges with respect to EHR data (CCDE) submission, and with CMS programmatic (IQR) data reporting threshold requirements (not used for measure calculation) as described in the Section 6.2.3.

7.1 Supplemental Attachment

[HHWM CBE Fall Cycle Supplemental Documents.zip](#)

Developer POC email

Lisa.Suter@yale.edu

Measure Developer POC

Lisa Suter
Yale/YNHH Center for Outcomes Research and Evaluation (CORE)
New Haven, CT
United States

The measure developer is different from the measure steward

Yes

Steward Address

Raquel Myers
Windsor Mill, MD
United States

Steward Organization

Centers for Medicare & Medicaid Services

Steward Organization URL

<https://www.cms.gov/medicare/quality/initiatives/hospital-quality-initiative/in...>

Steward POC email

raquel.myers@cms.hhs.gov