

# **Reevaluation of CMS Claims-Based Hospital Outcome Measures (Risk Variable Reselection) Technical Expert Panel Summary Report**

September 29, 2022

## **Prepared by:**

Yale New Haven Health Services Corporation (YNHHS) — Center for Outcomes Research and Evaluation (CORE)

This material was prepared by YNHHS/CORE under contract to the Centers for Medicare & Medicaid Services (CMS), an agency of the U.S. Department of Health and Human Services. The contents presented do not necessarily reflect CMS policy.

Reevaluation of CMS Claims-based Hospital Outcome Measures (Risk Variable Reselection)  
 Technical Expert Panel Summary Report ..... 3

- Background ..... 3
- Measure Reevaluation Team ..... 3
- TEP Composition ..... 4
- TEP Meetings ..... 6
- First TEP Meeting Overview ..... 6
  - Executive Summary of TEP Meeting 1 ..... 7
  - Detailed Summary of TEP Meeting 1 ..... 7
- Second TEP Meeting Overview ..... 9
  - Executive Summary of TEP Meeting 2 ..... 9
  - Detailed Summary of TEP Meeting 2 ..... 10
  - Input Received After Second TEP Meeting ..... 13
- Third TEP Meeting Overview ..... 13
  - Executive Summary of TEP Meeting 3 ..... 13
  - Detailed Summary of TEP Meeting 3 ..... 14
  - Input Received After Third TEP Meeting ..... 19
- Next Steps ..... 19
- Ongoing Reevaluation ..... 20
- Conclusion ..... 20
- Appendix A. CORE Reevaluation Team ..... 21
- Appendix B. TEP Meeting Schedule ..... 22

## **Center for Outcomes Research and Evaluation Project Team**

Kashika M. Sahay, PhD, MPH – Project Co-Lead  
Steven Spivack, PhD, MPH – Project Co-lead  
Zhenqiu Lin, PhD – Analytic Director  
Elizabeth Triche, PhD – Associate Director  
Shu-Xia Li, PhD – Associate Director  
Anna Sigler, MPH – Project Manager  
Jennifer Falcone, BA – Project Coordinator  
Yongfei Wang, MS – Statistician  
Si Zhou, MS – Statistician  
Emily Bean, MPH – Research Associate  
Karen Dorsey Sheares, MD, PhD – Subject Matter Expert  
Lisa Suter, MD – Contract Director

## **Acknowledgements**

This work is a collaborative effort, and the authors gratefully acknowledge the CORE patient and family engagement and X4 teams for their collaboration, prior TEP members and team members, and CMS for their contributions to this work.

# Reevaluation of CMS Claims-based Hospital Outcome Measures (Risk Variable Reselection) Technical Expert Panel Summary Report

## Background

The Centers for Medicare & Medicaid Services (CMS) has contracted with Yale New Haven Health Services Corporation/Center for Outcomes Research and Evaluation (CORE) to reevaluate CMS hospital claims-based outcome measures previously developed by CORE. As part of this project, CORE is working to determine if modifications can be made to improve the statistical model performance and face validity of the measures.

CORE is obtaining stakeholder input on this work and has assembled a national Technical Expert Panel (TEP) of clinicians, health services researchers, statisticians, patient advocates, and other stakeholders. The TEP is providing input to help shape considerations for methodological approaches to risk variable reselection (RVR).

This report reflects the work of the CORE measure reevaluation team and the TEP members, and summarizes the issues discussed, as well as feedback and recommendations received from the TEP over the span of three meetings. CORE will update the report to include feedback and recommendations from future meetings as they occur.

## Measure Reevaluation Team

The CORE measure reevaluation team consists of individuals with expertise in outcome measure development, health services research, clinical medicine, statistics, and measurement methodology. This work is led by Kashika M. Sahay, PhD, MPH, and Steven Spivack, PhD, MPH. Lisa G. Suter, MD, Contract Director and Professor of Medicine in the Section of Rheumatology, Department of Internal medicine at Yale University School of Medicine, and Elizabeth Triche, PhD, Associate Director, oversee the work. See [Table A1](#) in Appendix A for a list of all CORE measure reevaluation team members.

Raquel Myers, PhD, JD, MPH serves as the project's Contracting Officer Representative, providing ongoing input.

As part of reevaluation work, the risk model respecification team is actively working to update the risk-adjustment methodology for all 21 hospital claims-based outcome measures currently in public reporting (see Table 1 below). The goal of this work is to improve the validity of the measures and to address stakeholder concerns.

**Table 1. 21 Hospitals Claims Measures Currently in Public Reporting**

| Hospital Readmissions Reduction Program (HRRP)             | Hospital Value-Based Purchasing Program (HVBP)               | Inpatient Quality Reporting (IQR)                      |
|--|--|--|
| Acute Myocardial Infarction Readmission                    | Acute Myocardial Infarction Mortality                        | Hospital-Wide Readmission                              |
| Heart Failure Readmission                                  | Heart Failure Mortality                                      | Acute Myocardial Infarction Payment                    |
| Pneumonia Readmission                                      | Pneumonia Mortality  | Heart Failure Payment                                  |
| Chronic Obstructive Pulmonary Disease Readmission          | Chronic Obstructive Pulmonary Disease Mortality              | Pneumonia Payment                                      |
| Coronary Artery Bypass Graft Readmission                   | Coronary Artery Bypass Graft Mortality                       | Total Hip Arthroplasty/Total Knee Arthroplasty Payment |
| Total Hip Arthroplasty/Total Knee Arthroplasty Readmission | Total Hip Arthroplasty/Total Knee Arthroplasty Complications | Acute Myocardial Infarction Excess Days in Acute Care  |
|  |  | Heart Failure Excess Days in Acute Care                |
|  |  | Pneumonia Excess Days in Acute Care                    |
|  |  | Stroke Mortality                                       |

**TEP Composition**

In September 2019, CORE released a public call for nominations to convene the TEP. Potential TEP members were recruited via email to individuals, professional societies, and organizations recommended by the measure reevaluation team and stakeholder groups, email blasts sent to CMS email listservs, and through a posting on [CMS’ website](#).

At the time of the first TEP meeting in 2020, the TEP was comprised of 16 members; in 2021, 14 TEP members were able to reconvene for a second TEP meeting; and at the time of the third TEP meeting in 2022, 11 of 13 TEP members attended. Due to some TEP members no longer being able to participate, CORE added 3 additional TEP members, two of whom represented patients and families. Current participants are listed in [Table 2](#). The TEP is comprised of clinicians, health services researchers, statisticians, patients, patient advocates/caregivers, health insurance representatives, and hospital administrators. The role of the TEP is to provide input to CORE on key methodological and clinical decisions for the RVR task. The original appointment term for the TEP was from October 2019 through April 2020. CORE reconvened the TEP in late 2021 after a pause due to COVID-19. The current TEP term is from September 2022 – March 2023.

Responsibilities of TEP members include:

- Reviewing background materials provided by CORE prior to each meeting;
- Participating in TEP meetings held by webinar/teleconference; and
- Providing input on key clinical and methodological decisions.

**Table 2. TEP Roster — Member Name, Professional Role, Organization, and Location**

| Name, Credentials, and Professional Role   | Organizational Affiliation, City, State   |
|--|---|
| <b>Ann Borzecki, MD, MPH; <i>Attending Physician and Research Scientist</i></b>                                  | <ul style="list-style-type: none"> <li>• VA Bedford Healthcare System, Bedford, VA</li> <li>• Center for Healthcare Organization and Implementation Research, Bedford, MA</li> </ul>  |
| <b>Sarah Brinkman, MBA, MA, CPHQ; <i>Quality Program Manager</i></b>   | <ul style="list-style-type: none"> <li>• Stratis Health, Minneapolis, MN</li> </ul>   |
| <b>Steven Coffee, MA, EM CQSL, Colonel; <i>USAF Director, Military Personnel (patient)</i></b>                   | <ul style="list-style-type: none"> <li>• Headquarters US Cyber Command</li> <li>• Patients for Patient Safety, World Health Organization (WHO), Woodbridge, VA</li> </ul>   |
| <b>Michael Duan, MS; <i>Principal Data Scientist</i></b>   | <ul style="list-style-type: none"> <li>• Premier, Inc., Charlotte, NC</li> </ul>  |
| <b>Richard Dutton, MD, MBA; <i>Anesthesiologist, Adjunct Professor, Chief Quality Officer</i></b>                | <ul style="list-style-type: none"> <li>• Baylor University Medical Center</li> <li>• Texas A&amp;M University</li> <li>• US Anesthesia Partners, Dallas, TX</li> </ul>  |
| <b>Ryan Merkow, MD, MS; <i>Surgical Oncologist, Health Services and Outcomes Researcher, Faculty Scholar</i></b> | <ul style="list-style-type: none"> <li>• Northwestern University, Surgical Outcomes and Quality Improvement Center</li> <li>• American College of Surgeons, Division of Research and Optimal Patient Care, Chicago, IL</li> </ul> |
| <b>Matthias Cheung, RPh, PhD; <i>Adjunct Professor of Pharmacy, Medical Reviewer and Writer (patient)</i></b>    | <ul style="list-style-type: none"> <li>• University of the Pacific, Thomas J. Long School of Pharmacy, Stockton, CA</li> <li>• Eversana Life Science Services, LLC, Chicago, IL</li> </ul>  |
| <b>Shaifali Ray, MHA; <i>Principal (patient)</i></b>   | <ul style="list-style-type: none"> <li>• Seva Solutions Group, Wheaton, IL</li> </ul>   |
| <b>Sachin Shah, MD, MPH; <i>Physician Scientist, Member of Faculty</i></b>                                       | <ul style="list-style-type: none"> <li>• Massachusetts General Hospital</li> <li>• Harvard Medical School, Boston, MA</li> </ul>  |
| <b>Lynn Stillman, RN; <i>Program Manager Payment Innovation</i></b>  | <ul style="list-style-type: none"> <li>• Elevance Health (formerly Anthem Blue Cross/Blue Shield of New Hampshire), Bedford, NH</li> </ul>  |

| Name, Credentials, and Professional Role   | Organizational Affiliation, City, State   |
|--|---|
| <b>Mary Vaughan-Sarrazin, PhD; Associate Professor, Department of Internal Medicine, Director, Quantitative Unit of Health Services and Clinical Research Core, Investigator</b> | <ul style="list-style-type: none"> <li>• University of Iowa</li> <li>• Iowa City VA Medical Center, Iowa City, IA</li> </ul>  |
| <b>Thomas Webb, MBA, PhD candidate; Associate Vice President of Quality Analytics</b>  | <ul style="list-style-type: none"> <li>• Rush University Medical Center, Chicago, IL</li> </ul>   |
| <b>Bonnie Weiner, MD, MSEC, MBA; Cardiologist, Professor of Medicine, Director Interventional Cardiology Research, Chief Medical Officer, Senior Medical Director</b>            | <ul style="list-style-type: none"> <li>• University of Massachusetts Medical School</li> <li>• Worcester Medical Center</li> <li>• Accreditation of Cardiovascular Excellence</li> <li>• Avania, Harvard, MA</li> </ul> |

## TEP Meetings

CORE held its first TEP meeting on March 11, 2020 (TEP Meeting 1), its second on December 6, 2021 (TEP Meeting 2), and its third on September 29, 2022 (TEP Meeting 3). CORE anticipates holding additional meetings through 2023 (see [Appendix B](#) for the TEP meeting schedule). This report contains a summary of TEP Meeting 1 through TEP Meeting 3, as well as any feedback received from TEP members after the meeting.

TEP meetings follow a structured format. CORE presents key issues identified during measure reevaluation and a proposed approach to addressing them, and TEP members review, discuss, and advise on the issues.

## First TEP Meeting Overview

Prior to TEP Meeting 1, CORE provided TEP members with detailed meeting materials outlining CORE’s current approach to risk adjustment. Materials prepared for the meeting included:

- The slide deck for the meeting;
- The meeting agenda; and
- A backgrounder.

TEP members provided input on CORE’s current risk adjustment approach and shared thoughts on approaches to RVR for CMS’ 21 publicly reported outcome measures.

## Executive Summary of TEP Meeting 1

### Overview of Information Presented by CORE

CORE:

- Presented a high-level overview of risk adjustment;
- Introduced CORE's current approach to measure risk adjustment; and
- Presented a preview of TEP Meeting 2.

### Overview of TEP Feedback

The TEP provided feedback on CORE's risk adjustment approach. Specifically, the TEP's feedback included:

- Questions about CORE's risk adjustment approach, including which data sources are used, whether Medicare Advantage patients are included in the measure cohorts, and how CORE conducts vetting of clinical risk variables; and
- A concern about the conceptual approach to risk adjustment.

## Detailed Summary of TEP Meeting 1

CORE welcomed participants, introduced the CORE project team, presented the TEP confidentiality agreement, and conducted a roll call of meeting participants; 14 of 16 TEP members were in attendance. CORE noted they would reach out to TEP members not in attendance for their input.

CORE provided an overview of the RVR project and reviewed goals of the meeting, including an introduction to risk adjustment, CORE's current approach to risk adjustment, and obtaining the TEP's feedback on the best approaches for reselecting clinical risk variables for 21 claims-based hospital outcome measures currently in CMS public reporting programs.

### Introduction to Risk Adjustment and Current Risk Adjustment Approach

#### CORE Presentation to the TEP on Risk Adjustment

- CORE provided an overview on risk adjustment in the context of quality measurement, and CORE's processes for identifying candidate clinical and demographic risk factors, as well as their approach to considering social risk factors (SRFs) in measure risk-adjustment models.
- CORE noted that it compiles a list of clinical and demographic factors that have a conceptual relationship with the measure outcome and are independent of quality of care, generally compiled by examining peer-reviewed literature and receiving input from clinical experts. Clinical risk variables are defined using individual International Statistical Classification of Diseases, Tenth Revision (ICD-10) codes and Hierarchical Condition Categories (HCCs), which group ICD-10 codes into clinically similar groups.

- CORE reviewed the steps taken to test the clinical and demographic variables for inclusion in the risk model, and how the candidate risk variable list is narrowed, using analytic methods including stepwise logistic regression and bootstrapping. Final risk models are based on overall model performance, clinical relevance, and significance.
- CORE noted SRFs are considered, examining the conceptual and empirical relationship between the SRF and outcome, the SRF's accessibility in the data, and influence on hospital-level effects. Inclusion of SRFs is impacted by additional factors such as potential unintended consequences, existing CMS programmatic adjustments, alternative program-level approaches, and tradeoffs for each measure.

#### TEP Questions/Feedback and CORE responses

- TEP question: Among claims under consideration, both for the readmission outcome and the 12-month lookback period, are Medicare Part B claims included in addition to hospital claims?
  - CORE response: Only inpatient claims are used to identify the readmission outcome, but the 12-month lookback period for the condition- and procedure-specific measures uses both inpatient and outpatient claims to identify risk variables.
- TEP question: Does CORE include Medicare Advantage patients in the measure cohorts?
  - CORE response: Medicare Advantage patients are currently excluded from the measures because they do not have associated claims; CORE, under contract to CMS, is currently working on further analyses to investigate methods for potentially incorporating Medicare Advantage patients into measures moving forward.
- TEP question: How are risk variables clinically vetted?
  - CORE response: Regarding clinical vetting, engagement begins during measure development, including review of codes with TEPs, in-house clinical experts, and external clinicians and researchers. During measure reevaluation, CORE continues to consult with clinicians and researchers to vet newly released ICD-10 codes.
- TEP comment: The outlined approach to identifying variables based on conceptual association may be flawed; many variables screened based on bivariate associations may be mediators, rather than confounders, and would not be appropriate to include in risk models.
- TEP question: How are factors such as gender and race considered in the risk models, given the existing controversy over whether these patient-level factors should be considered genetic or demographic in nature?
  - CORE response: CORE seeks to include biological, clinical, demographic, and comorbid conditions in risk models and carefully reviews these variables individually on a measure-specific basis.
- One TEP member asked if CORE could walk through a step-by-step example of how the risk adjustment process is conducted.

- CORE noted they would present an example of the risk model development process at the next TEP meeting.

## Second TEP Meeting Overview

Prior to TEP Meeting 2, CORE provided the TEP members with detailed meeting materials outlining CORE's current approach to risk adjustment. Materials prepared for the meeting included:

- The slide deck for the meeting;
- The meeting agenda;
- A backgrounder with key definitions and terminology; and
- A preview video with an example of current risk adjustment strategies.

TEP members provided input on CORE's current risk adjustment approach, and shared thoughts on approaches to RVR for the hospital measures with a specific focus on frailty and data feature engineering. TEP members were given an example of Heart Failure Mortality to discuss, and engaged in several round-robin discussions about the overall approach. TEP members were generally positive about the data feature engineering approaches. They had some concerns about potential gaming and a need for reevaluation efforts to ensure that hospitals with different present-on-admission (POA) coding guidelines were not penalized.

With respect to frailty, some TEP members wondered if individual codes relating to frailty would already be captured and the model may be artificially overfit with an additional frailty term.

## Executive Summary of TEP Meeting 2

### Overview of Information Presented by CORE

CORE:

- Presented a high-level overview of timeline for risk model respecification over multiple years;
- Provided background on current risk variable selection approaches and potential innovations using code-based and condition category-based features;
- Solicited feedback with respect to the data feature engineering approach;
- Discussed empirical results relating to different operational definitions of frailty to consider different statistical and data-based approaches to capturing frailty; and
- Presented a summary of next steps for future TEP engagement.

### Overview of TEP Feedback

- The TEP provided feedback on CORE's data feature engineering approach and frailty examples and asked how the POA logic is applied to diagnosis codes. The TEP was interested in evaluating frailty data and wondered if an indicator variable comprised of

individual ICD-10 codes differs from including the same codes individually within the risk model.

- The TEP noted concerns about:
  - Correlation/association versus causation of model inputs;
  - Differences in hospital coding practices, including:
    - Under-coding
    - Variation in coding requirements for critical access hospitals that may have implications for their coding practices;
  - Predictive model may show a negative coefficient, but this does not correspond to a “true” reduction in risk (ex. hypertension has a negative coefficient); and
  - Conflation of causal inference and predictive value in the risk model: note that identifying causal predictors requires a whole different approach than risk model development and that negative predictors are just a feature of prediction models; trying to manipulate those and weed out certain predictors will only lead to a decrease in predictive power.
- The TEP offered suggestions, including:
  - Potentially removing risk variables with negative coefficients to avoid hospital gaming;
  - Consideration of a penalty in the model to deprioritize gameable codes (ex. palliative care);
  - Including clinical adjudication to support the accuracy of administrative data elements used for risk modeling;
  - Suggesting the use of bootstrapping for model estimates;
  - Concerns about definitions of frailty and how upcoding may favor high resource hospitals; and
  - Agreement with the overall approach to consider ICD-10 codes individually and as part of condition categories (CCs).

## Detailed Summary of TEP Meeting 2

CORE welcomed participants, introduced the CORE project team, presented the TEP confidentiality agreement, and conducted roll call of meeting participants; 12 of 14 TEP members were in attendance. CORE team members were introduced, and participants shared their credentials and interest in being on this TEP. The discussion was divided into two broad topics: clinical data feature engineering and risk adjustment for patient frailty, with round robin discussion at the end of each topic. After content discussion, the facilitator asked for TEP feedback on the logistics of the TEP meeting and the CORE team discussed next steps for TEP engagement.

### Data Feature Engineering

#### CORE Presentation to the TEP on Data Feature Engineering

- CORE described a hybrid approach to data feature engineering that includes individual ICD-10 codes and CCs to be input separately. A hybrid approach will allow CORE to

select a parsimonious set of important individual ICD-10 codes while grouping other codes in CCs and combined CCs; grouping codes with similar clinical and statistical characteristics may reduce the number of model inputs, while also potentially reducing the potential for gaming.

- CORE clarified that feature engineering at the grouper level includes separating the conditions noted on the index admission claims from those on the historical claims, splitting broad CCs to allow for highly predictive codes to influence model characteristics, and combining CCs where there is clinical and statistical similarity among included codes.
- CORE summarized the approach for risk model optimization, noting the pragmatic goal is to test and validate the model approach through an iterative process that includes identifying the most consistently predictive ICD-10 codes, examining the impact of the codes within the existing CCs, restructuring CCs based on these results, and testing model performance and face validity.
- CORE asked for feedback on the validity of this approach and any items that should be considered in greater detail.
- Individual code analysis demonstrates tradeoffs in the conceptual approach. CORE provided sample results based on a single performance period which included history code frequency, odds ratio range, and the mortality outcome rate range associated with each of these different approaches.

#### TEP Feedback/CORE Response on Data Feature Engineering

- Several TEP members expressed positive feedback for using individual codes in the reselection process.
- TEP members expressed some concerns about negative coefficients and the potential for gaming.
  - CORE response: The goal of quality improvement efforts is not to establish causation, but rather measure a quality signal between hospitals.
- One TEP member noted that an important goal of the overall approach should be to improve the model c-statistic.
- One TEP member noted critical access hospitals have variation in coding requirements especially with respect to POA codes. Thus, there is a need to stratify by hospital type to ensure coding practices do not unfairly penalize rural hospitals.
- One TEP member noted considering adding a penalty in the regression model to deprioritize gameable codes.
  - CORE response: CORE noted that payment algorithms are determined by CMS and cautioned that the scope of this discussion is risk adjustment; the purpose of risk adjustment is to adjust for differences in case mix, not to identify causation.
- One TEP member noted that having a consistent amount of historical data is important; another noted that individual codes are potentially subject to gaming and reevaluation activities should consider timing of reselection carefully.

- CORE response: CORE clarified 12 months before the index admission is included as history and reevaluation activities (including RVR) are scheduled to occur every three years.

## Frailty

### CORE Presentation to the TEP on Frailty

- CORE presented several options for operationalizing the complex construct of frailty as well as a conceptual model for why frailty should be considered.
- Options for operationalizing frailty include:
  - Using the combined CC approach, or ‘Marked Disability and Frailty,’ which incorporates the Healthcare Common Procedure Coding System codes and use of durable medical equipment, such as power wheelchairs and oxygen;
  - Including frailty-related ICD-10 codes shown in the literature to predict decline;
  - Using a count of relevant frailty codes in a patient’s history; and/or
  - Including a validated quantitative scale for frailty, or ‘Claims-based Frailty Index.’
- CORE asked TEP members to consider whether CORE should examine a scale-based definition or a code-based approach and if there are other options CORE might consider.

### TEP Feedback/CORE Responses on Frailty

- TEP feedback: ICD-10 codes that currently exist may lack functional status because the coders may not have the information needed for coding these details. Validation might be needed to ensure that the codes accurately reflect the patient’s functional status.
  - CORE response: CORE’s guiding principle for risk variable selection is based on a need to be able to consistently capture standardized selection data within a national database; that does not exist yet for objective measures of frailty.
- One TEP member noted that a frailty indicator from ICD-10 codes does not necessarily differ from having the ICD-10 codes as individual variables in the model.
- TEP feedback: CORE should clarify whether frailty will be measured using a ‘Deficit Accumulation Framework for Frailty,’ as opposed to alternative frailty definitions, which includes five specific domains and has not been translated into a claims-based definition.
  - CORE response: CORE used the Harvard validated claims-based frailty index (CFI) as an example of a validated claims-based scale for frailty, but CORE is not proposing implementing the CFI into the risk adjustment model.
- TEP feedback: It is possible individual codes stand by themselves to describe frailty. However, some codes could be grouped into an indicator variable when that makes sense, rather than trying to construct a composite rating score.
  - CORE response: CORE is looking at the potential additive effect of a separate frailty variable beyond the individual codes.
- TEP feedback: Disease registries such as cardiology often incorporate frailty and functional scales; CORE should consider the use of the functional scales more broadly, however availability of functional scales may vary significantly.

- CORE response: CORE acknowledges that it is challenging to incorporate functional status into a claims-based metric in a standardized way. CORE stated that several of the validated claims-based scales use the deficit accumulation framework.

### **Input Received After Second TEP Meeting**

One TEP member asked questions about the implications of risk adjustment for payment purposes. CORE responded via email that the risk adjustment is for calculating risk scores in a standardized way to compare hospitals across the country. CMS programs make decisions relating to payment through other mechanisms that are not the focus of the current project.

### **Third TEP Meeting Overview**

Prior to TEP Meeting 3, CORE provided TEP members with a TEP charter, updated TEP participant list, project background information, and results of empirical analyses.

During the meeting, the CORE team:

- Presented CORE’s approach to condition-based data feature selection using the risk model for Heart Failure Mortality;
- Presented considerations/framework for selecting risk model variables, operationalizing frailty, and addressing social risk factors using the Heart Failure Mortality model as an example; and
- Provided an overview of upcoming TEP interaction.

TEP members were asked for their thoughts and suggestions regarding:

- Not risk adjusting for “protective” individual codes with odds ratios less than one (OR<1);
- Adjusting for the do not resuscitate (DNR) ICD-10 code Z66;
- The proposal to adjust for frailty and their preferences for defining frailty, and
- patient and hospital-level social risk testing.

### **Executive Summary of TEP Meeting 3**

#### **Overview of Information Presented by CORE**

CORE:

- Presented CORE’s approach to condition-based data feature selection for the risk adjustment model for Heart Failure Mortality;
- Presented considerations/framework for selecting risk model variables, operationalizing frailty, and addressing social risk factors in the Heart Failure Mortality model; and
- Provided an overview of upcoming TEP engagement.

## Overview of TEP Feedback

The TEP provided feedback on not risk adjusting for “protective” individual codes with OR<1. Specifically, the TEP’s feedback included:

- The majority of TEP participants favored removing the protective codes with OR<1 from the risk model; and
- TEP participants noted the need to better understand what the protective codes are measuring and concern about the potential for unanticipated effects on hospital coding practices.

The TEP provided feedback on adjusting for the DNR ICD-10 code Z66. Specifically, the TEP’s feedback included:

- The majority of TEP participants favored adjusting for DNR POA, and some TEP participants favored adjusting for DNR regardless of POA; and
- TEP participants were skeptical about the accuracy and completeness of the POA status and noted different requirements for different types of hospitals.

The TEP provided feedback on the proposal to adjust for frailty and their preferences for defining frailty. The majority of TEP participants voted against including a frailty variable in the risk model, citing concerns about lack of statistical evidence, computational burden, and provider burden.

The TEP provided feedback on including hospital-level social risk in the model. Specifically, the TEP’s feedback included:

- The majority of TEP participants believed current data sources are inadequate to account for social risk; and
- The TEP suggested implicit biases, low health literacy, and access to care are key drivers of disparities, and hospitals with larger disparities need support to provide better outcomes rather than economic penalties.

## Detailed Summary of TEP Meeting 3

CORE welcomed participants, introduced the CORE project team and three CMS participants, and conducted a roll call of TEP participants; 11 of 13 TEP members were in attendance. CORE noted it would reach out to TEP members not in attendance for their input.

CORE provided background information and updates on measure reevaluation and risk model specification and reviewed goals of the meeting, including requesting TEP feedback on the use of individual codes in the risk model, including frailty in the risk model, and potentially assessing the effects of social risk in the risk model.

## Use of Individual Codes in the Risk Model

### CORE Presentation to the TEP on Use of Individual Codes in the Risk Model

- CORE explained the prior approach to risk adjustment, which focused on CCs and described the proposed approach that relies on individual ICD-10 codes and may account for frailty and social risk.
  - CORE noted they hoped to get input from the TEP on how to handle “protective” ICD-10 codes in the risk model, which were defined as risk variables with  $OR < 1$ .
  - CORE noted they were also interested in TEP input on including DNR as a candidate risk variable.
- CORE presented results of the individual ICD-10 codes model using the Heart Failure Mortality measure as an example.
  - CORE ran the model using the pre-COVID data from the 2021 performance period (2017–2019 data).
  - CORE shared the results of the individual codes analysis, which started with 568 ICD-10 codes meeting the frequency and statistical significance thresholds, finding 123 codes that were statistically significant in all 1000 bootstrap iterations, 71 of which were associated with higher odds of mortality.
    - There were 52 protective codes with  $OR < 1$ ;
    - The model was minimally affected by removing the protective codes;
    - The model with all statistically significant variables (123 codes plus age) had a C-statistic of 0.787; and
    - The model with only increased risk variables ( $OR > 1$ ), with 71 codes plus age, had a C-statistic of 0.771.
- CORE proposed only including risk variables with  $OR > 1$  in the model.
- CORE proposed adjusting for the DNR ICD-10 code Z66, rather than excluding these patients from quality measurement; a significant proportion (17%) of the Heart Failure Mortality cohort had a DNR code.

### TEP Feedback/CORE Responses on Use of Individual Codes in the Risk Model

- TEP participants were asked for their thoughts and suggestions regarding not adjusting for protective individual codes with  $OR < 1$ .
- Nine TEP participants responded to the polling question, **“Do you agree with dropping all individual codes with  $OR < 1$  in the individual codes approach?”**, with seven TEP participants (78%) responding affirmatively to drop the protective codes and two TEP participants citing opposition.
  - Five TEP participants cited improved face validity of the measure when protective codes are removed from the risk model.
  - Both TEP participants who opposed dropping protective codes were clinicians and their concerns included ignoring what we do not understand, removing some protective elements while we are rewarding protective activities in value-based purchasing, preferring a higher C-statistic (0.787 vs 0.771) when including

- protective codes, and potential reclassification of hospitals due to the methodological change.
- Several TEP participants noted the need to better understand what the protective codes are measuring, since these outcomes may be the result of appropriate medical management of patients.
  - Several TEP participants mentioned potential unintended effects on hospital coding practices.
  - TEP participants were asked for their thoughts and suggestions regarding adjusting for the DNR ICD-10 code Z66.
    - Six TEP participants who verbally responded to this question, supported including DNR POA in the risk model, and two TEP participants indicated potential support of including DNR in the model, regardless of POA status.
    - Several TEP participants noted concerns about the timing of the DNR. Even though DNR is not on the POA exempt list, there were concerns about the reliability (accuracy and completeness) of the POA status; one TEP participant noted critical access hospitals are not required to code POA.
    - Some TEP participants noted the presence of a DNR, regardless of the timing, impacts clinical decision making and captures severity of illness that is not easily captured otherwise.

### **Including Frailty in the Risk Model**

#### CORE Presentation to the TEP on Including Frailty in the Risk Model

- CORE presented information about the potential to include frailty in the risk model, and shared empirical results using Heart Failure Mortality as a case study for updating the risk modeling approach.
- CORE evaluated two approaches to define frailty:
  - Frailty as defined within the Accountable Care Organization (ACO) Multiple Chronic Conditions (MCC) measure (ACO MCC approach), a National Quality Forum endorsed binary measure that includes durable medical equipment and mobility issues, protein calorie malnutrition, and other claims-based frailty markers; and
  - A CFI, which uses a deficit accumulation index conceptual framework and can be defined as either a binary or categorical variable.
- CORE noted the results of these two approaches are not strongly correlated, and fewer patients are identified as frail when we use the ACO MCC definition (47.1%) versus the binary CFI definition (70.8%). Both frailty definitions have significant bivariate associations with the Heart Failure Mortality outcome, with observed mortality at 14.6% for patients identified as frail with the ACO MCC approach, and 11.2% for patients identified as frail with the binary CFI approach.
- CORE voiced support for the ACO MCC definition of frailty.
- CORE explained there is limited statistical benefit to adding a frailty variable in multivariate models, but it may make sense to add a frailty variable for face validity purposes.

## TEP Feedback/CORE Responses on Including Frailty in the Risk Model

- TEP participants were asked for their thoughts and suggestions regarding the proposal to adjust for frailty and their preferences for defining frailty.
- Eleven TEP participants responded to the poll question, **“Based on the empirical results presented, would you recommend adjusting for frailty based on the following options?”** and of the 11 TEP participant responses:
  - Five of eleven (45%) selected “No, frailty not captured well in codes”;
  - Four of eleven (36%) selected “Yes, ACO MCC”;
  - Two of eleven (18%) selected “Yes, CFI”; and
  - None (0%) selected “Yes, some other frailty code”.
- After a round-robin discussion, six of 11 (55%) TEP participants voted against including frailty in the model due to lack of statistical evidence to support it, as neither frailty marker improved the C-statistic. Additionally, TEP participants noted added computational and provider burden as reasons to exclude frailty.
- TEP participants who voted to include a frailty variable favored the ACO MCC definition of frailty.
- Additional frailty considerations noted by TEP participants included:
  - Other factors, such as lack of access, may affect the proportion of patients using durable medical equipment, which is significant since durable medical equipment use is a factor in both frailty definitions considered;
  - A suggestion CORE might explore a simpler model with fewer variables and a frailty indicator; and
  - Concerns that evaluating the frailty variable only in the Heart Failure Mortality risk model could overlook opportunities where there may be more impact for other quality measures/disease states, like readmission.

## **Including Social Risk in the Risk Model**

### CORE Presentation to the TEP on Including Social Risk in the Risk Model

- CORE presented information about including social risk variables in the risk model, including TEP survey results with feedback about social risk.
  - Twelve of the 13 (92%) TEP participants responded to the survey.
  - The focus was on social risk variables, data sources, and the approach to hospital-level testing.
  - Survey respondents identified social risk variables at both the person and community/neighborhood level.
  - Results suggested there are pros and cons to the various data sources, such as:
    - Incomplete, inaccurate, and unavailable data;
    - Influence of state laws, such as for dual eligibility;
    - Nuances in defining income, such as variation over time; and
    - Variability in documentation.
  - CORE noted it is likely they would need to construct a composite measure of social risk.

- CORE noted the social risk metrics being considered for empirical testing include:
  - Agency for Healthcare Research and Quality Socioeconomic Status Index;
  - Dual eligibility; and
  - Area Deprivation Index.
- CORE noted race is also included in the empirical results being presented.
- CORE presented the bivariate results for selected social risk variables.
  - These results, similar to the literature, show some seemingly counter-intuitive results, which could be due to people not accessing services or adverse outcomes outside the 30-day timeframe being evaluated.
  - CORE needs to consider if there are ways that social risk is already embedded in its models (ex. in coding practices).

#### TEP Feedback/CORE Responses on Including Social Risk in the Risk Model

- There was general agreement among TEP participants that current data sources are inadequate to account for social risk. Implicit biases, low health literacy, and access to care are key drivers of disparities, and hospitals with larger disparities need support to provide better outcomes rather than economic penalties. TEP members acknowledged that social risk testing is an important and complex construct in risk adjustment. They supported testing various social risk concepts at both the area and person level to try to better represent the construct, but acknowledged that risk adjustment may not be the most appropriate way to account for these differences.
- Overall, several specific concerns about the data available for these analyses were discussed, including:
  - Concerns about ascertainment;
  - Risk adjustment for social risk making disparities less visible;
  - Collection of the data reflecting assumptions of people collecting the data rather than responses of patients; and
  - Race data reported as White/non-White are inadequate to draw conclusions.
- Other measurement concerns were discussed, including:
  - Readmissions being affected differently than mortality; and
  - Patient selection in who receives hospital care leading to outcomes that are counterintuitive.
- Additional measurement/testing strategies to explore were proposed, including:
  - A TEP participant who uses Veterans Administration data suggested that Veteran Administration-specific metrics “service-connected disability” and “priority group” showed stronger associations with disparities than the typical variables for race/ethnicity;
  - Broader use of secondary diagnoses to capture severity of illness;
  - Consider controlling for the propensity for admission by race/ethnicity;
  - Conducting death certificate analyses to better understand overall mortality to supplement mortality data for the hospital setting; and
  - Using the new CMS structural measure for hospitals to document the social health needs of their patients (going into voluntary reporting in 2023 and

mandatory reporting in 2024), along with current datasets to better understand disparities.

- TEP comments not specific to the questions above included:
  - TEP participants suggested including lower frequency codes (ex. rare conditions/genetic diseases which did not meet the frequency threshold) but with high impact on mortality outcomes in the risk model, and one TEP participant suggested CMS' major complication or comorbidity list could be considered to filter the low frequency codes.
  - A TEP participant suggested CMS might consider signaling through rulemaking intent for future use of social risk ICD-10 codes in quality reporting to encourage more consistent use of the codes.
  - \*\*CORE note: the current model was restricted to clinical risk factors. Z-codes for social risk related variables were deliberately excluded. However, none of the Z-codes would have met frequency thresholds for inclusion in the Heart Failure Mortality cohort.

### **Input Received After Third TEP Meeting**

One TEP member provided additional feedback after the meeting, including support for the individual codes model, adjusting for DNR, and including frailty in the model. This member highlighted the relationship between Black and Brown communities having higher disease burden and possibly higher DNR status and noted that hospitals caring for greater proportions of Black and Brown patients shouldn't be disproportionately penalized because of their case burden. Regarding social risk factors, this TEP member also touched upon access to care (including access to transportation) and food insecurity (distance to a grocery store) as important social variables but acknowledged the lack of reliable data at this time to include in the model.

### **Next Steps**

The next steps for this TEP include moving forward with the RVR framework for mortality measures and applying it to other measures.

In the next phase of interaction, the TEP will have a deeper discussion about the application of the final framework to other measures, including readmission; participants can look forward to communication about their availability in mid-2023.

The RVR team will circulate this report to TEP members for feedback as well as brief CMS on the TEP discussion points. CORE will reconvene the TEP to review empirical results once the model has been tested.

Ongoing reevaluation efforts described in the next section will include implementing the approaches discussed with the TEP and sharing additional empirical results for TEP review.

## Ongoing Reevaluation

Ongoing reevaluation work will focus on applying the final framework to additional measures and testing it for different conditions and outcomes to then arrive at a respecified model for all measures under reevaluation. This includes:

- Assessing the clinical and face validity of final measure specifications;
- Testing model and measure performance (reliability and validity);
- Continuing to evaluate potential adjustments for social risk factors; and
- Applying the framework to more recent years of data, post-COVID-19.

## Conclusion

The TEP provided valuable feedback on the use of individual ICD-10 codes, including frailty, and assessing social risk in the risk model. The TEP was supportive of excluding protective ICD-10 codes with  $OR < 1$  and including DNR if it was POA in the model. The TEP was split in the decision to adjust for frailty and did not favor adjusting for race or social risk within the risk model due to concerns about the quality of available data to account for these factors. However, they did suggest exploring other data sources for social risk and assessing frailty and social risk for other measures more broadly. Overall, they felt the framework for assessing individual codes allowed for an understanding of the clinical risk factors.

## Appendix A. CORE Reevaluation Team

**Table A1. Center for Outcomes Research and Evaluation Risk Variable Reselection Reevaluation Team**

| Name                                  | Team Role                              |
|---------------------------------------|--|
| <b>Kashika M. Sahay</b> , PhD, MPH    | Project Co-lead                        |
| <b>Steven Spivack</b> , PhD, MPH      | Project Co-lead                        |
| <b>Zhenqiu Lin</b> , PhD              | Analytic Director                      |
| <b>Elizabeth Triche</b> , PhD         | Project Co-lead and Associate Director |
| <b>Shu-Xia Li</b> , PhD               | Project Co-lead and Associate Director |
| <b>Anna Sigler</b> , MPH              | Project Manager                        |
| <b>Jennifer Falcone</b> , BA          | Project Coordinator                    |
| <b>Yongfei Wang</b> , MS              | Statistician                           |
| <b>Si Zhou</b> , MS                   | Statistician                           |
| <b>Emily Bean</b> , MPH               | Research Associate                     |
| <b>Karen Dorsey Sheares</b> , MD, PhD | Subject Matter Expert                  |
| <b>Lisa Suter</b> , MD                | Project Director                       |

## Appendix B. TEP Meeting Schedule

CORE will engage and seek input from the TEP as they develop the measure through email communication and meetings:

1. **TEP Meeting 1:** Wednesday, March 11, 2020; 5:00 – 7:00 PM EST (Location: teleconference/webinar)
2. **TEP Meeting 2:** Monday, December 6, 2021; 2:00 – 4:00 PM EST (Location: teleconference/webinar)
3. **TEP Meeting 3:** September 29, 2022; 4:00 – 7:00 PM EST (Location: teleconference/webinar)
4. **TEP Meeting 4:** Middle/Late 2023 (Location: teleconference/webinar)