

CBE ID

4125

Title

Thirty-day Risk-Standardized Death Rate among Surgical Inpatients with Complications (Failure-to-Rescue)

Project

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

Endorsement Status

Endorsed with Conditions

E&M Committee Rationale/Justification

Perform additional reliability testing for endorsement review.

Is Under Review

No

Next Maintenance Cycle

Fall 2029

Previous Endorsement Cycle

Fall 2023

Initial Endorsement

Mon, 03/18/2024 - 13:29

Steward

Centers for Medicare & Medicaid Services

1.0 New or Maintenance

New

1.3 Electronic Clinical Quality Measure (eCQM)

No

1.6 Measure Description

Percentage of surgical inpatients who experienced a complication and then died within 30-days from the date of their first "operating room" procedure. Failure-to-rescue is defined as the probability of death given a postoperative complication.

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

N/A as this is not a paired measure.

Website URL not available; Final measure specifications for implementation will be made publicly available on CMS' appropriate quality website, once finalized through the CBE endorsement and CMS rulemaking processes.

1.11 Measure Webpage

<http://notavailable/seerationaleabove>

1.13 Data Dictionary

Attached

1.13a Attach Data Dictionary

[Specification_Failure-to-Rescue_Final.zip](#)

1.14 Numerator

Patients who died within 30 days from the date of their first "operating room" procedure, regardless of site of death.

1.14a Numerator Details

Number of verified deaths (STUS_CD=20) within 30 days from the date of the first eligible operating room procedure (**Table 1**), regardless of site of death, among discharges meeting the inclusion and exclusion rules for the denominator.

This measure uses submitted claims data and vital status data from the Medicare Beneficiary Summary File (or equivalent resources, such as a Vital Status File) to calculate the measure score. All data elements necessary to calculate this numerator are defined with the attached technical specifications.

1.15 Denominator

Patients aged 18 years and older admitted for certain procedures in the General Surgery, Orthopedic, or Cardiovascular Medicare Severity Diagnosis Related Groups (MS-DRGs) who were enrolled in the Medicare program and had a documented complication that was not present on admission. Documented complications include: cardiac events, congestive heart failure, hypotension or shock or hypovolemia, pulmonary embolus or deep vein thrombosis or phlebitis, cerebrovascular accident (CVA) or transient ischemic attack (TIA), coma, seizure, psychosis, nervous system complications, pneumonia or pneumonitis, pneumothorax/effusion, respiratory compromise or bronchospasm, internal organ damage or perforation, peritonitis, gastrointestinal bleed and blood loss, sepsis, deep wound infection or wound complication, renal dysfunction, gangrene/amputation, intestinal obstruction or ischemia, retained foreign body, pressure injury, orthopedic complication, hepatitis or jaundice, pancreatitis, necrosis of bone (thermal or aseptic), osteomyelitis, disseminated intravascular coagulation (DIC), pyelonephritis, or other postsurgical complication.

1.15a Denominator Details

DENOMINATOR OVERALL

Discharges for patients ages 18 through 89 years with any listed ICD-10-PCS procedure code for an operating room procedure (**Table 1**) and all of the following:

-Enrolled in the Medicare program

-Any admission type in which the earliest ICD-10-PCS code for an operating room procedure (**Table 1**) occurs within the qualifying period, starting three days prior to the date of admission and ending at the date of discharge

-Meet the inclusion and exclusion criteria for one of the denominator complication categories (**Tables 3-5**)

And meeting one of the following criteria:

-Eligible discharges assigned to the General Surgery, Orthopedic, or Cardiovascular Medicare Severity Diagnosis Related Groups (MS-DRGs: **Table 2**)

OR

-Eligible discharges assigned to the ECMO or Tracheostomy Medicare Severity Diagnosis Related Groups (**Table 2**; MS-DRGs 003 or 004), *and*

- with an MDC for diseases and disorders of the circulatory system; digestive system; hepatobiliary system and pancreas; musculoskeletal system and connective tissue; skin, subcutaneous tissue and breast; or endocrine, nutritional and metabolic diseases (**Table 2**; MDCs 05, 06, 07, 08, 09, 10), *and*

-with any listed ICD-10-PCS code for a procedure assignable to MS-DRG 003 or 004 (**Table 1**;

FTRPXCHGTOMSDRG003004P), that, in the absence of a code for ECMO (**Table 5**) or tracheostomy (**Table 6**), would assign the discharge to a denominator eligible MS-DRG (**Table 2**), and

-without any listed ICD-10-PCS procedure code for ECMO (**Table 5**), and

-without any listed ICD-10-PCS procedure code for tracheostomy (**Table 6**) occurring before or on the same day as the first non-tracheostomy operating room procedure

Denominator Category 1_Cardiac Event

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for cardiac event not present on admission (**Table 3**; FTR1CARDEVENTD) or any listed ICD-10-PCS procedure code for cardiac event (**Table 4**; FTR1CARDEVENTP) at least one day after the first qualifying operating room procedure (**Table 1**)

Denominator Category 2_Congestive Heart Failure

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for congestive heart failure not present on admission (**Table 3**; FTR2CHFD)

Denominator Category 3_Hypotension/Shock/Hypovolemia

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for hypotension, shock, or hypovolemia not present on admission (**Table 3**; FTR3SHOCKD)

Denominator Category 4_Pulmonary Embolus/Deep Vein Thrombosis/Phlebitis

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for pulmonary embolus, deep vein thrombosis or phlebitis not present on admission (**Table 3**; FTR4PEDVTPHD) or any listed ICD-10-PCS procedure code for pulmonary embolus, deep vein thrombosis or phlebitis (**Table 4**; FTR4PEDVTPHP) at least one day after the first qualifying operating room procedure (**Table 1**)

Denominator Category 5_Cerebrovascular Accident (CVA)/TIA

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for cerebrovascular accident or transient ischemic attack not present on admission (**Table 3**; FTR5CVAD)

Denominator Category 6_Coma

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for coma not present on admission (**Table 3**; FTR6COMAD)

Denominator Category 7_Seizure

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for seizure not present on admission (**Table 3**; FTR7SEIZD)

Denominator Category 8_Delirium/Psychosis

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for psychosis not present on admission (**Table 3**; FTR8PSYCHD)

Denominator Category 9_Nervous System Complications

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for nervous system complications not present on admission (**Table 3**; FTR9NERVSYS)

Denominator Category 10_Pneumonia/Pneumonitis

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for pneumonia or pneumonitis not present on admission (**Table 3**; FTR10PNEUMOD)

Denominator Category 11_Pneumothorax

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for pneumothorax not present on admission (**Table 3**; FTR11PTXD) or any listed ICD-10-PCS procedure code for pneumothorax (**Table 4**; FTR11PTXP) at least one day after the first qualifying operating room procedure (**Table 1**)

Denominator Category 12_Respiratory Compromise/Bronchospasm

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for respiratory compromise or bronchospasm not present on admission (**Table 3**;

FTR12RESPCOMPD) or any listed ICD-10-PCS procedure code for respiratory compromise/bronchospasm (**Table 4**; FTR12RESPCOMPP) at least one day after the first qualifying operating room procedure (**Table 1**)

Denominator Category 13_ Internal Organ Damage/Perforation

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for internal organ damage or perforation not present on admission (**Table 3**; FTR13ORGDAMD) or any listed ICD-10-PCS procedure code for internal organ damage or perforation (**Table 4**; FTR13ORGDAMP) at least one day after the first qualifying operating room procedure (**Table 1**)

Denominator Category 14_ Peritonitis

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for peritonitis not present on admission (**Table 3**; FTR14PERITD)

Denominator Category 15_ GI Bleed and Blood Loss

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for gastrointestinal bleeding or blood loss not present on admission (**Table 3**; FTR15GIBLEEDD)

Denominator Category 16_ Sepsis

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for sepsis not present on admission (**Table 3**; FTR16SEPSISD)

Denominator Category 17_ Deep Wound Infection/Wound Complication

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for deep wound infection or wound complication not present on admission (**Table 3**; FTR17WOUND) or any listed ICD-10-PCS procedure code for deep wound infection or wound complication (**Table 4**; FTR17WOUNDP) at least one day after the first qualifying operating room procedure (**Table 1**)

Denominator Category 18_ Renal Dysfunction

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for renal dysfunction not present on admission (**Table 3**; FTR18RENALD)

Denominator Category 19_Gangrene/Amputation

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for gangrene or amputation not present on admission (**Table 3**; FTR19GANGAMPD)

Denominator Category 20_Intestinal Obstruction/Ischemia

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for intestinal obstruction or ischemia not present on admission (**Table 3**; FTR20INTOBTISCHD)

Denominator Category 21_Foreign Body

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for foreign body not present on admission (**Table 3**; FTR21FORBODYD)

Denominator Category 22_Pressure Injury

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for pressure injury not present on admission (**Table 3**; FTR22PID)

Denominator Category 23_Orthopedic Complication

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for orthopedic complication not present on admission (**Table 3**; FTR23ORTHOCOMPD) or any listed ICD-10-PCS procedure code for orthopedic complication (**Table 4**; FTR23ORTHOCOMP) at least one day after the first qualifying operating room procedure (**Table 1**)

Denominator Category 24_Hepatitis/Jaundice

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for hepatitis or jaundice not present on admission (**Table 3**; FTR24HEPATD)

Denominator Category 25_Pancreatitis

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for pancreatitis not present on admission (**Table 3**; FTR25PANCD)

Denominator Category 26_Necrosis of Bone (Thermal or Aseptic)

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for necrosis of bone (thermal or aseptic) not present on admission (**Table 3**; FTR26NECBOND)

Denominator Category 27_Osteomyelitis

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for osteomyelitis not present on admission (**Table 3**; FTR27OSTEOMYD)

Denominator Category 28_Disseminated Intravascular Coagulation (DIC)

Denominator-eligible discharges with a secondary ICD-10-CM diagnosis code of disseminated intravascular coagulation (DIC) not present on admission (**Table 3**; FTR28DICD)

Denominator Category 29_Pyelonephritis

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for pyelonephritis not present on admission (**Table 3**; FTR29PYNEPHD)

Denominator Category 30_Postprocedural/Transfusion Complication

Denominator-eligible discharges with any listed secondary ICD-10-CM diagnosis code for postsurgical complication not present on admission (**Table 3**; FTR30POSTSURGD)

This measure uses submitted claims data to calculate the measure score. All data elements necessary to calculate this denominator are defined with the attached technical specifications.

1.15b Denominator Exclusions

DENOMINATOR OVERALL EXCLUSIONS (FOR ALL CATEGORIES)

Exclude discharges:

-Patients aged >90 years

-
- Admitted from a hospice facility (ADMSOUR = F)
 - Do not resuscitate (DNR) status (ICD-10-CM Z66) present on admission (POA)
 - Contradictory death information (reported date of death before admit date, death date before discharge date when patient was reportedly discharged alive, discharge disposition reported as died but enrollee has subsequent claims)
 - No qualifying "operating room" procedure (**Table 1**) with a reported date
 - First or only qualifying "operating room" procedure (**Table 1**) was outside appropriate time window for that claim (i.e., 4 or more days before the date of admission, or after the date of discharge)
 - With an ungroupable MS-DRG (DRG=999)
 - With missing discharge disposition (STUS_CD=missing), gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DGNS_CD1=missing)
 - Discharged against medical advice (DISP=7)

Denominator Exclusions Category 1_Cardiac Event

Exclude discharges:

- with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for cardiac event (**Table 3**; FTR1CARDEVENTD)

Denominator Exclusions Category 2_Congestive Heart Failure

Exclude discharges:

- with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for congestive heart failure (**Table 3**; FTR2CHFD)

Denominator Exclusions Category 3_Hypotension/Shock/Hypovolemia

Exclude discharges:

- with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for hypotension, shock, or hypovolemia (**Table 3**; FTR3SHOCKD)
- with any listed principal ICD-10-CM diagnosis code for trauma (Table 7)

Denominator Exclusions Category 4_Pulmonary Embolus/Deep Vein Thrombosis/Phlebitis

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for pulmonary embolus, deep vein thrombosis or phlebitis (**Table 3**; FTR5PEDVTPHD)

Denominator Exclusions Category 5_Cerebrovascular Accident (CVA)/TIA

-Exclude discharges: with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for stroke, cerebrovascular accident or transient ischemic attack (**Table 3**; FTR5CVAD)

Denominator Exclusions Category 6_Coma

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for coma (**Table 3**; FTR6COMAD)

Denominator Exclusions Category 7_Seizure

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for seizure (**Table 3**; FTR7SEIZD)

Denominator Exclusions Category 8_Delirium/Psychosis

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for psychosis (**Table 3**; FTR8PSYCHD)

Denominator Exclusions Category 9_Nervous System Complications

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for nervous system complications (**Table 3**; FTR9NERVSYSD)

Denominator Exclusions Category 10_Pneumonia/Pneumonitis

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for pneumonia or pneumonitis (**Table 3**; FTR10PNEUMOD)

Denominator Exclusions Category 11_Pneumothorax

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for pneumothorax (**Table 3**; FTR11PTXD)

Denominator Exclusions Category 12_Respiratory Compromise/Bronchospasm

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for respiratory compromise or bronchospasm (**Table 3**; FTR12RESPCOMP)

Denominator Exclusions Category 13_Internal Organ Damage/Perforation

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for internal organ damage or perforation (**Table 3**; FTR13ORGDAMD)

Denominator Exclusions Category 14_Peritonitis

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for peritonitis (**Table 3**; FTR14PERITD)

Denominator Exclusions Category 15_GI Bleed and Blood Loss

Exclude discharges:

- with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for GI bleeding or blood loss (**Table 3**; FTR15GIBLEEDD)
- with any listed principal ICD-10-CM diagnosis code for trauma (Table 7)

Denominator Exclusions Category 16_Sepsis

Exclude discharges:

- with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for sepsis (**Table 3**; FTR16SEPSISD)

Denominator Exclusions Category 17_Deep Wound Infection/Wound Complication

Exclude discharges:

- with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for deep wound infection or wound complication (**Table 3**; FTR17WOUNDDD)

Denominator Exclusions Category 18_Renal Dysfunction

Exclude discharges:

- with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for renal dysfunction (**Table 3**; FTR18RENALD)

Denominator Exclusions Category 19_Gangrene/Amputation

Exclude discharges:

- with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for gangrene or amputation (**Table 3**; FTR19GANGAMPD)

Denominator Exclusions Category 20_Intestinal Obstruction/Ischemia

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for intestinal obstruction or ischemia (**Table 3**; FTR20INTOBSTISCHD)

Denominator Exclusions Category 21_Foreign Body

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for foreign body (**Table 3**; FTR21FORBODYD)

Denominator Exclusions Category 22_Pressure Injury

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for pressure injury (**Table 3**; FTR22PID)

Denominator Exclusions Category 23_Orthopedic Complication

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for orthopedic complication (**Table 3**; FTR23ORTHOCOMPD)

Denominator Exclusions Category 24_Hepatitis/Jaundice

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for hepatitis or jaundice (**Table 3**; FTR24HEPATD)

Denominator Exclusions Category 25_Pancreatitis

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for pancreatitis (**Table 3**; FTR25PANCD)

Denominator Exclusions Category 26_Necrosis of the Bone (Thermal or Aseptic)

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for necrosis of the bone (thermal or aseptic) (**Table 3**; FTR26NECBOND)

Denominator Exclusions Category 27_Osteomyelitis

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for osteomyelitis (**Table 3**; FTR27OSTEOMYD)

Denominator Exclusions Category 28_Disseminated Intravascular Coagulation (DIC)

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for disseminated intravascular coagulation (DIC) (**Table 3**; FTR28DICD)

Denominator Exclusions Category 29_Pyelonephritis

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for pyelonephritis (**Table 3**; FTR29PYNEPHD)

Denominator Exclusions Category 30_Postprocedural/Transfusion Complication

Exclude discharges:

-with any listed principal ICD-10-CM diagnosis code (or secondary diagnosis present on admission) for postsurgical complication (**Table 3**; FTR30POSTSURGD)

This measure uses submitted claims data to calculate the measure score. All data elements necessary to identify denominator complications are defined with the attached technical specifications.

1.15c Denominator Exclusions Details

This measure uses submitted claims data, in combination with validated death data from the Medicare Beneficiary Summary File (or equivalent resources, such as a Vital Status File) to calculate the measure score. All data elements necessary to calculate these denominator exclusions are defined with the attached technical specifications.

1.16 Type of Score

Rate/proportion

1.17 Measure Score Interpretation

Better performance = Lower score

1.18 Calculation of Measure Score

See attached.

1.18a Attach measure score calculation diagram

[FTR measure flow.pdf](#)

1.19 Measure Stratification Details

Not applicable; this measure is not stratified.

1.20 Types of Data Sources

Claims Data

1.25 Data Source Details

Medicare inpatient claims data, including Medicare Inpatient Encounter (shadow billing) data for Medicare Advantage enrollees, in combination with validated death data from the Medicare Beneficiary Summary File or equivalent resources. CMS receives death information from a number of sources. The main sources CMS uses to develop its death information are Medicare claims data from the Medicare Common Working File (CWF), online date of death edits submitted by family members, and benefit information used to administer the Medicare program collected from the Railroad Retirement Board (RRB) and the Social Security Administration (SSA). Overall, over 99.9% of death days have been validated. As for other CMS 30-day mortality measures, the "Valid Date of Death Switch" is used to confirm that the exact day of death has been validated.

1.26 Minimum Sample Size

There is no minimum sample size to calculate the measure. If a hospital has fewer than 25 denominator eligible records, the hospital's mortality rate and interval estimates will not be publicly reported. This approach is consistent with other CMS 30-day mortality measures used in

public reporting programs.

2.1 Attach Logic Model

[FTR Logic Model and Tables.pdf](#)

2.2 Evidence of Measure Importance

The concept of “failure to rescue” (FTR) was originally developed by Jeffrey Silber and colleagues and adapted by Jack Needleman and colleagues. Over the past three decades, numerous studies have identified associations with multiple hospital characteristics and processes of care and rates of failure to rescue. The current measure is an updated and completely re-tested version of a previously CBE-endorsed measure #0353, “Failure to Rescue 30-day Mortality.” This measure was stewarded by Silber and colleagues at the Children’s Hospital of Philadelphia (CHOP) and used extensively by the research and quality improvement communities. CHOP allowed CBE endorsement to lapse in 2021.

Hospital Characteristics and Staffing

A series of seminal papers by Silber et al. and Needleman et al. established the relationship between several hospital characteristics and failure to rescue rates. Silber et al. (1992) examined 5,972 Medicare patients admitted for cholecystectomy and transurethral prostatectomy and found that failure to rescue was independent of severity of illness at admission, but was significantly associated with the presence of surgical housestaff and a lower percentage of board-certified anesthesiologists. The adverse occurrence rate was independent of these hospital characteristics. In a larger sample of 74,647 patients who underwent general surgical procedures in 1991-92, Silber et al. (1997) found lower failure to rescue rates at hospitals with high ratios of registered nurses to beds. Failure rates were strongly associated with risk adjusted mortality rates, as expected, but not with complication rates. Finally, among 16,673 patients admitted for coronary artery bypass surgery, failure to rescue rates were lower (whereas complication rates were higher) at hospitals with magnetic resonance imaging facilities, bone marrow transplantation units, or approved residency training programs (Silber et al., 1995). In a 2002 publication, Needleman and Buerhaus confirmed that higher registered nurse staffing (RN hours/adjusted patient day) and better nursing skill mix (RN hours/licensed nurse hours) were consistently associated with lower failure to rescue rates among major surgery patients from 799 hospitals in 11 states in 1997, even using administrative data to define complications. An increase from the 25th to the 75th percentile on these two measures of staffing was associated with 5.9% (95% CI, 1.5% to 10.2%) and 3.9% (95% CI, -1.1% to 8.8%) decreases, respectively, in the rate of failure-to-rescue among major surgery patients.

Other more recent individual studies have reported similar significant associations between failure to rescue and hospital characteristics, including nurse staffing levels (Aiken et al., 2011; Brooks Carthon et al., 2012; Ma et al., 2015; Silber et al., 2007), greater nurse education or advanced nurse skill mix (Kendall-Gallagher et al., 2011; Kutney-Lee et al., 2013; Silber et al.,

2007), hospital volume (Gonzalez et al., 2014; Silber et al., 2009), nursing (ANCC) magnet status (Kutney-Lee et al., 2015; McHugh et al., 2013), resident-to-bed ratio or teaching status (Silber et al., 2007, 2009).

Several systematic reviews have reported confirmatory findings. A 2015 systematic review by Johnston et al. including 42 studies (some of which are described previously) identified several hospital characteristics associated with delayed escalation of care and higher FTR rates, including lower hospital volume, lower nurse staffing, and non-teaching status. The review identified 3 studies that found that mortality rates increased in patients with delayed escalation of care (odds ratio ranging from 2.1 to 3.1) and one study reporting that delayed transfer to the intensive care unit (ICU) was associated with 20% higher mortality compared to rapid transfer. A systematic review by Bourgon Labelle (2019) identified 15 studies finding significant associations between nurse staffing levels and improved failure to rescue rates (both in-hospital and 30-day) among patients with postoperative cardiac events. The review also identified 6 studies finding that a higher proportion of nurses with baccalaureate degrees was also significantly associated with lower 30-day failure to rescue rates. A systematic review by Twigg et al. (2019) identified nine studies reporting significant associations between nursing skill mix and failure to rescue rates among adult patients in acute care settings. In a systematic review by Audet et al. (2018), six studies were identified that reported significant associations between nursing education and lower risk of failure to rescue. Twigg and colleagues also found that the association between nursing education and failure to rescue was stronger for surgical patients than for non-surgical patients. In a meta-analysis of three studies, Liao et al. (2016) concluded that a 10% increase in nurses with a bachelor's degree or above was associated with a 5% reduction in risk of failure to rescue (OR: 0.95; 95% CI, 0.94-0.97; $p < 0.001$).

References

1. Aiken LH, Cimiotti JP, Sloane DM, Smith HL, Flynn L, Neff DF. Effects of nurse staffing and nurse education on patient deaths in hospitals with different nurse work environments. *Med Care*. 2011;49(12):1047-1053.
2. Audet LA, Bourgault P, Rochefort CM. Associations between nurse education and experience and the risk of mortality and adverse events in acute care hospitals: A systematic review of observational studies. *Int J Nurs Stud*. 2018;80:128-146.
3. Bourgon Labelle J, Audet LA, Farand P, Rochefort CM. Are hospital nurse staffing practices associated with postoperative cardiac events and death? A systematic review. *PLoS One*. 2019;14(10):e0223979.
4. Brooks Carthon JM, Kutney-Lee A, Jarrín O, Sloane D, Aiken LH. Nurse staffing and postsurgical outcomes in black adults. *J Am Geriatr Soc*. 2012;60(6):1078-1084.
5. Gonzalez AA, Dimick JB, Birkmeyer JD, Ghaferi AA. Understanding the volume-outcome effect in cardiovascular surgery: the role of failure to rescue. *JAMA Surg*. 2014;149(2):119-123.
6. Johnston MJ, Arora S, King D, et al. A systematic review to identify the factors that affect failure to rescue and escalation of care in surgery. *Surgery*. 2015;157(4):752-763
7. Liao LM, Sun XY, Yu H, Li JW. The association of nurse educational preparation and patient

- outcomes: Systematic review and meta-analysis. *Nurse Educ Today*. 2016;42:9-16.
8. Kendall-Gallagher D, Aiken LH, Sloane DM, Cimiotti JP. Nurse specialty certification, inpatient mortality, and failure to rescue. *J Nurs Scholarsh*. 2011;43(2):188-194.
 9. Kutney-Lee A, Sloane DM, Aiken LH. An increase in the number of nurses with baccalaureate degrees is linked to lower rates of postsurgery mortality. *Health Aff (Millwood)*. 2013;32(3):579-586.
 10. Kutney-Lee A, Stimpfel AW, Sloane DM, Cimiotti JP, Quinn LW, Aiken LH. Changes in patient and nurse outcomes associated with magnet hospital recognition. *Med Care*. 2015;53(6):550-557.
 11. Ma C, McHugh MD, Aiken LH. Organization of Hospital Nursing and 30-Day Readmissions in Medicare Patients Undergoing Surgery. *Med Care*. 2015;53(1):65-70.
 12. McHugh MD, Kelly LA, Smith HL, Wu ES, Vanak JM, Aiken LH. Lower mortality in magnet hospitals. *Med Care*. 2013;51(5):382-388.
 13. Needleman J, Berghaus P, Mattke S, Stewart M, Zelevinsky K. Nurse-staffing levels and the quality of care in hospitals. *N Engl J Med*. 2002;346(22):1715-1722.
 14. Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. *Med Care*. 1992;30(7):615-29.
 15. Silber JH, Rosenbaum PR, Schwartz JS, Ross RN, Williams SV. Evaluation of the complication rate as a measure of quality of care in coronary artery bypass graft surgery. *JAMA*. 1995;274(4):317-323.<https://pubmed.ncbi.nlm.nih.gov/7609261/>
 16. Silber JH, Rosenbaum PR, Williams SV, Ross RN, Schwartz JS. The relationship between choice of outcome measure and hospital rank in general surgical procedures: implications for quality assessment. *Int J Qual Health Care*. 1997;9(3):193-200.
<https://pubmed.ncbi.nlm.nih.gov/9209916/>
 17. Silber JH, Romano PS, Rosen AK, Wang Y, Even-Shoshan O, Volpp KG. Failure-to-rescue: comparing definitions to measure quality of care. *Med Care*. 2007;45(10):918-925.
 18. Silber JH, Rosenbaum PR, Romano PS, et al. Hospital teaching intensity, patient race, and surgical outcomes. *Arch Surg*. 2009;144(2):113-121.
 19. Twigg DE, Kutzer Y, Jacob E, Seaman K. A quantitative systematic review of the association between nurse skill mix and nursing-sensitive patient outcomes in the acute care setting. *J Adv Nurs*. 2019;75(12):3404-3423.

Processes of Care

Studies also show that other processes of care can influence failure to rescue rates. Failure to rescue has been found to be associated with measures of a hospital's aggressiveness of care (defined as the level of resources or inpatient spending), with hospitals that treat patients more aggressively having better surgical mortality and failure to rescue rates (Kaestner, 2010; Silber, 2010). Three recent systematic reviews have examined the relationship between the use of various hospital-based interventions and the risk of failure to rescue.

A 2022 systematic review by Burke et. al. including 52 articles identified three critical stages that lead to failure to rescue - failure to recognize complications, failure to relay information regarding

complications, and failure to react in a timely and appropriate manner – and six types of interventions that can improve failure to rescue rates within healthcare organizations:

1. Staffing levels and education: Based on 14 studies (meta-analysis, retrospective cohort studies, cross-section studies, case-control studies, case reports, and a descriptive project), the authors found that FTR is highly sensitive to nursing care, specifically nurse-patient ratios, patient turnover and nurse staffing in non-ICU settings, staffing patterns, training and opportunity for simulation. For example, a cohort study demonstrated that after implementation of minimum nurse staffing levels in California, FTR rates decreased significantly more in California than in comparison states, with improvements of up to 32.9% ($P < 0.05$) in the final implementation period, across quartiles of baseline nurse staffing.

2. Detection, early warning signs (EWS) systems and checklists: Based on 8 studies (RCT, observational studies systematic review, cross-sectional studies and respect to follow up study), the authors observed the importance of early warning symptom detection protocols and timely and appropriate escalation. For example, a randomized controlled trial (RCT) demonstrated improved patient management (SWAT-M, $P < 0.001$) and nontechnical skills ($P = 0.043$) between baseline and final ward rounds, whereas the control group showed no improvement ($P = 0.571$ and $P = 0.809$, respectively). A small learning effect was seen with improvement in patient assessment (SWAT-A) in both groups ($P < 0.001$).

3. Surveillance, communication and electronic monitoring: Based on 6 studies (retrospective cohort study, cross-sectional study, observation of a pilot, perspective single blinds observational study and a retrospective observational study of her control), the authors underscore the importance of nursing communication and continuous monitoring. For example, a retrospective cohort study demonstrated that when nursing surveillance was performed at least 12 times a day, there was a significant ($P = 0.0058$) decrease in the odds of experiencing failure to rescue (OR = 0.52) compared with when surveillance was delivered an average of <12 times a day.

4. Medical emergency and rapid response teams (RRT): Based on 8 studies (cluster RCT, retrospective audit, cross-sectional survey, case control, retrospective observational, descriptive/competitive study, longitudinal study and interrupted times serious population base study), the authors observe that significant variation in the design and reporting of studies examining medical emergency teams (METs) and RRTs limits the ability to draw clear conclusions regarding effectiveness. For example, a cluster RCT demonstrated similar incidence of the composite primary outcome in the control and MET hospitals (5.86 versus 5.31 per 1000 admissions, $P = 0.640$), as well as of the individual secondary outcomes (cardiac arrests, 1.64 versus 1.31, $P = 0.736$; unplanned ICU admissions, 4.68 versus 4.19, $P = 0.599$; and unexpected deaths, 1.18 versus 1.06, $P = 0.752$). A reduction in the rate of cardiac arrests ($P = 0.003$) and unexpected deaths ($P = 0.01$) was seen from baseline to the study period for both groups

combined, suggesting an effect of study participation unrelated to the MET program.

5. Relaying information about complications: Based on 9 studies (cohort study, literature review of six studies, cross-sectional survey, multi center qualitative study, observational, perspective observational and observational, questionnaire-based), the authors conclude that interprofessional communication and nurse physician relationship are of paramount importance, and recommended use of SBAR as a communication tool. For example, one study involved prospective collection of predefined surgical critical events and communications, patient interviews, and sporadic clinical questioning of junior clinicians. The authors reported that of 80 critical patient events identified across four hospitals, 26 (33%) were not communicated to attending surgeons. Although residents felt that attending contact was unnecessary for safe patient care in 61 (76%) of these events, discussions with attending physicians changed management in 33% (18/54) of cases in which they occurred.

6. Reacting to a patient in a timely manner with the correct evidence-based management: Based on 3 studies (audit of single center two units, retrospective cohort with contemporaneous control group, and retrospective cohort), the authors found that timely and evidence-based interventions have a significant impact on patient outcomes; for example, timely administration of antibiotics to patients with sepsis.

A 2015 systematic review by Johnston et al. identified several interventions that can improve timely escalation of care, including new vital sign charts and improved documentation, escalation protocols, and communication tools. Four studies found that these interventions increased the number of escalation-of-care calls or physician communications regarding deterioration. One pre-post cohort study found that an escalation protocol led to a non-significant decrease in in-hospital cardiac arrests (3% vs. 9% pre-implementation) and a significant decrease in ICU admission rates (23% vs. 46% pre-implementation, $p < 0.001$). A second pre-post cohort study found that use of a new vital signs chart led to a non-significant decrease in in-hospital cardiac arrest (0.5% vs. 1.8% pre-implementation) and a significant decrease in mortality (0.6% vs. 2.6% pre-implementation).

In the recent *Making Healthcare Safer III* report, Hall et al. (2020a) examined two patient safety practices with the potential to impact failure to rescue rates - patient monitoring systems and rapid response teams. Of the 8 included studies examining the impact of patient monitoring systems, there was moderate but inconsistent evidence that systems with continuous monitoring lead to reductions in failure to rescue events. Hall et al. (2020a, 2020b) identified 10 studies (including 3 meta-analyses and 3 systematic reviews) examining the impact of rapid response teams (RRTs) on failure to rescue events. This systematic review found that the implementation of RRTs was associated with decreases in inpatient mortality and in-hospital cardiac arrest. Two of the three meta-analyses found that RRT implementation significantly decreased mortality rates among adult inpatients (pooled relative risk [RR] range, 0.87-0.88), while the third found no

difference in overall mortality (pooled RR, 0.92; 95% CI, 0.82-1.04). Three meta-analyses identified overall decreases in non-ICU cardiac arrest after RRT implementation (pooled RR range, 0.62-0.65). Hall et al. reported mixed results on the impact of RRT on ICU transfer rates - one meta-analysis including 10 studies found no association while one systematic review found that RRTs reduced unplanned ICU admissions.

References

1. Burke JR, Downey C, Almouadaris AM. Failure to Rescue Deteriorating Patients: A Systematic Review of Root Causes and Improvement Strategies. *J Patient Saf.* 2022;18(1):e140-e155.
2. Hall KK, Lim A, Gale B. Failure To Rescue. In: Hall KK, Shoemaker-Hunt S, Hoffman L, et al. Making Healthcare Safer III: A Critical Analysis of Existing and Emerging Patient Safety Practices [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2020a. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK555513/>
3. Hall KK, Lim A, Gale B. The Use of Rapid Response Teams to Reduce Failure to Rescue Events: A Systematic Review. *J Patient Saf.* 2020b;16(3S Suppl 1):S3-S7.
4. Johnston MJ, Arora S, King D, et al. A systematic review to identify the factors that affect failure to rescue and escalation of care in surgery. *Surgery.* 2015;157(4):752-763
5. Kaestner R, Silber JH. Evidence on the efficacy of inpatient spending on Medicare patients. *Milbank Q.* 2010;88(4):560-594.
6. Silber JH, Kaestner R, Even-Shoshan O, Wang Y, Bressler LJ. Aggressive treatment style and surgical outcomes. *Health Serv Res.* 2010;45(6 Pt 2):1872-1892.

2.3 Anticipated Impact

By using failure-to-rescue (FTR), a risk-standardized measure of death after an adverse occurrence, hospitals can identify opportunities to improve their quality of care. Hospitals and health care providers benefit from knowing not only their institution's mortality rate, but also their institution's ability to rescue patients after clinical deterioration. The measure is especially important if hospital resources needed for preventing complications are different from those needed for rescue. We anticipate that this measure will encourage hospitals to focus on early identification and rapid treatment of complications, thereby improving the overall quality of care. Failure to rescue measures have been repeatedly validated by their consistent association with nurse staffing, nursing skill mix, technological resources, rapid response systems, and other activities that improve early identification and prompt intervention when complications arise after surgery.

Performance Results from Beta Testing:

Risk-standardized rates show substantial variation in performance scores across the 2,055 eligible facilities with at least 25 qualifying denominator records. Specifically, the distribution of 30-day

Failure to Rescue risk-standardized death rates in our test data is as follows:

5th percentile: 0

15th percentile: 21.15

25th percentile: 29.33

35th percentile: 35.15

45th percentile: 40.35

55th percentile: 46.88

65th percentile: 53.35

75th percentile: 60.95

85th percentile: 71.64

95th percentile: 98.01

Median: 43.48

Mean: 46.62

This empirical analysis demonstrates considerable opportunity for improvement if facilities at the 75th percentile (60.95 risk-standardized deaths per 1,000 qualifying surgical cases) could move across the interquartile range to the 25th percentile (29.33 risk-standardized deaths per 1,000 surgical cases), which would represent a 50% decrease in the frequency of deaths after postoperative complications.

See **Table 1 logic model attachment** for a distribution of performance scores for the current CMS PSI 04 compared to the proposed measure. Compared with the current CMS PSI 04 measure that is used for public reporting, the proposed measure has a much higher minimum volume threshold (25 versus 3), covers over 8 times more denominator patients, and captures about 2.1 times more numerator events (deaths). The numerator increase is largely due to the application of this measure to both Medicare Advantage and FFS enrollees, as well as the inclusion of deaths after hospital discharge but within 30 days of the index operative procedure.

2.5 Health Care Quality Landscape

This measure has been designed and tested to replace CMS PSI 04, which is currently being used in the Hospital Inpatient Quality Reporting (HIQR) Program (00134-02-C-HIQR, formerly CBE

#0351). This redesign is intended to address stakeholder concerns about the existing PSI 04 measure, which include:

1. Complications sometimes develop BEFORE the index operation in PSI 04, even before transfer to the index hospital (I.e., the operation is part of an effort to “rescue” the patient).
2. The heterogenous cohort includes patients with very high-risk surgery (e.g., trauma surgery, burn surgery, organ transplants, intracranial hemorrhage) and very low-risk surgery (e.g., eye, ear, urolithiasis).
3. Mean length of stay and prevalence of early discharge to post-acute facilities vary across hospitals, causing bias in comparing performance.
4. PSI 04 appears to slightly disadvantage major referral centers, even after risk-adjustment.

The respecified FTR measure will create a more homogenous denominator population and capture post-discharge deaths within 30-days after the first denominator-qualifying operation. This redesign is intended to better align the measure with the previously CBE-endorsed measure of Failure-to-Rescue: 30-Day Mortality (CBE #0353, endorsed 2008, renewed 2012 and 2015, allowed to lapse 2021).

2.6 Meaningfulness to Target Population

Measures of failure-to-rescue among hospitalized surgical patients have been found to be useful by multiple stakeholders in the United States. For example, in the Fiscal Year (FY) 2022 Medicare Hospital Inpatient Prospective Payment System (IPPS) Proposed Rule (CMS-1752-P, April 2021), CMS proposed to retire PSI 04 from use in CMS programs. In response, CMS received many communications from patients, caregivers, patient advocacy organizations, employers and employer coalitions, and others. These communications clearly articulated the perceived value of CMS PSI 04 as a broad measure of postoperative mortality and hospitals’ skill at rescuing patients who experience complications. In response, CMS did not finalize the proposal to retire PSI 04 and invested in improving it in response to stakeholder feedback.

3.1 Contributions Towards Closing Care Gaps

Using data from all 2,907 hospitals in our test data set, we conducted a social disparities analysis and found:

-Hispanic patients have similar risk of Failure to Rescue (OR=0.93; 95% CI, 0.82-1.05) as non-Hispanic patients, after adjusting for age and other factors in the risk-adjustment model.

-Black patients (OR=0.96; 95% CI, 0.91-1.01) and patients of "other" race (OR=1.06; 95% CI, 0.94-1.20) have similar risk of Failure to Rescue as White patients, after adjusting for age and

other factors in the risk-adjustment model.

-Risk of Failure to Rescue is unrelated to sex, after adjusting for age and other factors in the risk-adjustment model. Sex was considered as a risk-adjustment feature but was found to have no marginal predictive value.

-Analyses of observed, expected, and risk-adjusted rates in all of the above patient cohorts confirm that the comorbidities, operative, and demographic factors in the risk-adjustment model account for some increased risk of Failure to Rescue among Black patients and patients of “other” race (average expected rate 5.81% and 5.71%, respectively, versus 4.18% among White patients), and that any residual bias is neither clinically nor statistically significant.

Empirical analyses confirm that this measure is neutral across social risk groups, including race, ethnicity, urbanicity/rurality, and sex, due to adjustment for all the patient characteristics described above. Age was explicitly included in the risk-adjustment model, so its effect was directly removed. These findings are as expected based on our conceptual model.

4.1 Feasibility Assessment

Because this measure is based on readily available administrative claims data, feasibility is not an issue. A similarly designed measure (CMS PSI 04) has been used by CMS for over a decade. No difficulties have been reported with respect to data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, or time and cost of data collection. Hospitals routinely generate and transmit claims in a timely manner for all Medicare beneficiaries.

4.3 Feasibility Informed Final Measure

No feasibility assessment was completed due to the reasons outlined above.

4.4 Proprietary Information

Not a proprietary measure and no proprietary components

4.4a Fees, Licensing, or Other Requirements

There are no fees associated with use of this claims-based measure. The measure specifications will be available upon request through the CMS QualityNet Help Desk.

5.1.1 Data Used for Testing

This measure was originally developed using data on Medicare FFS discharges from Inpatient Prospective Payment System (IPPS) hospitals, including hospitals in Maryland and excluding Veterans Administration hospitals, from for the period July 1, 2019, through December 31, 2019 and July 1, 2020 through June 30, 2021. Q1 and Q2 2020 data were excluded due to the blanket Extraordinary Circumstance Exception (ECE) for COVID-19. These data included roughly 12.4 million inpatient discharges from 3,357 hospitals where Medicare was the primary payer.

The measure was then tested on Medicare data from January 1, 2021 through June 30, 2022, including monthly inpatient claims files (Research Identifiable Files, or RIF) and Medicare Beneficiary Summary Files. These data included roughly 10.5 million discharges from 3,163 hospitals where Medicare was the primary payer. We used CMS+VA PSI v13.0 software to calculate the number of cases meeting the definition for the numerator and denominator for the current CMS PSI 04 measure and the proposed failure-to-rescue measure. We specifically evaluated the impact of changing the numerator definition from in-hospital death to 30-day death, with the 30-day window starting on the day of the first “operating room” procedure.

5.1.2 Differences in Data

Not applicable.

5.1.3 Characteristics of Measured Entities

Descriptive characteristics of the hospitals and Medicare FFS population included in testing are shown in **Tables 2 and 3 of the logic model attachment**.

5.1.4 Characteristics of Units of the Eligible Population

The test data set includes 417,054 inpatient encounters from 2,163 Medicare Inpatient Prospective Payment System hospitals, including Maryland hospitals. Of these hospitals, 2,055 met the minimum denominator threshold of 25 for reporting their Failure to Rescue rate. These hospitals are very diverse, representing all bed size categories, teaching status categories, nursing skill mix and staffing categories, and location (urban/rural) categories. Test hospitals are situated in all 50 US states and the District of Columbia.

5.2.1 Level(s) of Reliability Testing Conducted

Accountable entity level (i.e., measure score) (e.g., signal-to-noise analysis)

5.2.2 Method(s) of Reliability Testing

We applied split-half and test-retest approaches to estimate the reliability of this risk-adjusted measure at the accountable entity (hospital) level, using the intraclass correlation coefficient (ICC) as an estimator. As formulas are not allowed in the online form, see **logic model**

attachment pg. 9-10 for the methodology.

By design, hospital-level risk-adjusted outcome measures are centered around a global mean with an approximately normal distribution (allowing for the fact that the tails of the distribution may be augmented with hospitals that are true quality outliers). Because this ICC depends only on the ratio of between-hospital to within-hospital estimated variance components, and the relevant denominator for each hospital, we can estimate reliability as a function of the hospital's denominator size, using an application of the Spearman-Brown prophecy formula. We applied this methodology to hospital subsamples that were formed by randomly dividing the available year of patient data from each hospital into two, then executing the measure code separately on each split-half, to yield two estimates per hospital.

The higher the ICC, the greater the statistical reliability of the measure, and the greater the proportion of variation that can be attributed to systematic differences in performance across hospitals (i.e., signal as opposed to noise). We used the rubric established by Landis and Koch (1977) to interpret ICCs:

0 - 0.2: slight agreement

0.21 - 0.39: fair agreement

0.4 - 0.59: moderate agreement

0.6 - 0.79: substantial agreement

0.8 - 0.99: almost perfect agreement

1: perfect agreement

References

1. Dickens, William T. "Error components in grouped data: is it ever worth weighting?." *The Review of Economics and Statistics* (1990): 328-333.
2. Landis, J. Richard, and Gary G. Koch. "The measurement of observer agreement for categorical data." *biometrics* (1977): 159-174.
3. Spearman-Brown Prophecy Formula" in: Frey, B. (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vols. 1-4). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781506326139

5.2.3 Reliability Testing Results

Signal-to-noise reliability was estimated as an intraclass correlation coefficient based on a two-way mixed model with facility random effects (C,1), inflating the denominator and numerator for

each hospital from 18 months to a full 24-month reporting period. To further improve reliability for public reporting, we recommend empirical Bayesian shrinkage (i.e., smoothing) to reduce random noise, in accord with standard methods for all AHRQ and CMS Patient Safety Indicators. The smoothed rate is a weighted average of the reference population rate and the local (hospital) risk-adjusted rate. If the data from the individual hospital include many observations and provide a numerically stable estimate of the rate, then the smoothed rate will be very close to the risk-adjusted rate, and it will not be heavily influenced by the reference population rate. Conversely, the smoothed rate will be closer to the reference population rate if the hospital rate is based on a small number of observations and may not be numerically stable, especially from year to year. As a weighted average of the risk-adjusted rate and the rate observed in the reference population, the smoothed rate is calculated with a shrinkage estimator, as described in this report: <https://qualityindicators.ahrq.gov/Downloads/Resources/Publications/202...> .

Signal-to-noise reliability was estimated as an intraclass correlation coefficient based on a two-way mixed model with facility random effects (C,1).

- Minimum: 0.231
- 25th percentile: 0.388
- Median: 0.568
- 75th percentile: 0.738
- Maximum: 0.973

Please note the functionality of the decile table below was not working at the time of submission. As such, the decile information is included below for reference in the following format: Decile #/Reliability value/ # Entities/Total Persons

Overall/0.7039(mean)/2055/1087624

Minimum/0.2314/21/525

Decile 1/0.2571/205/15853

Decile 2/0.3248/206/21776

Decile 3/0.3879/206/29419

Decile 4/0.4671/206/40024

Decile 5/0.5379/205/53027

Decile 6/0.6016/205/69384

Decile 7/0.6697/205/92901

Decile 8/0.7384/206/129893

Decile 9/0.8106/205/199744

Decile 10/0.8861/206/435603

Maximum/0.9733/1/8099

5.2.4 Interpretation of Reliability Results

Failure to Rescue demonstrates moderate signal-to-noise reliability at most test facilities, based on a 24-month reporting period with both Medicare FFS and Medicare Advantage enrollees, as the mean and median ICC values equal 0.704 and 0.568, respectively. The comparable metrics for the currently reported version of CMS PSI 04, which is limited to Medicare fee-for-service patients, are 0.256 and 0.209, respectively, based on CMS+VA PSI v13 software applied to the 2023-reported performance period. The percentage of all eligible entities with reliability of at least 0.4 for Failure to Rescue is approximately 73% (based on a 24-month reporting period), versus 25% for the currently reported version of CMS PSI 04.

As with any 30-day mortality measure, reliability at the hospital level varies in accord with the size of the hospital and its eligible denominator. Minimum volume thresholds can be applied and adjusted, as needed, to address low reliability at low-volume hospitals. By regulation, the current minimum (denominator) volume threshold for all CMS 30-day risk-standardized mortality measures is 25. Overall, testing results showed that Failure to Rescue, as currently specified, can distinguish true performance across hospitals of typical size and volume.

Table 2. Accountable Entity Level Reliability Testing Results by Denominator, Target Population Size

Accountable Entity-Level Reliability Testing Results													
	Overall	Minimum	Decile_1	Decile_2	Decile_3	Decile_4	Decile_5	Decile_6	Decile_7	Decile_8	Decile_9	Decile_10	Maximum
Reliability	0.7039 (mean)	0.2314	0.2571	0.3248	0.3879	0.4671	0.5379	0.6016	0.6697	0.7384	0.8106	0.8861	0.9733
Mean Performance Score	2055	21	205	206	206	206	205	205	205	206	205	206	1
N of Entities	1087624	525	15853	21776	29419	40024	53027	69384	92901	129893	199744	435603	8099

5.3.1 Level(s) of Validity Testing Conducted

Accountable entity level (i.e., measure score) (e.g., criterion validity)

5.3.3 Method(s) of Validity Testing

Convergent validity refers to the degree to which multiple measures of a single underlying

concept are positively correlated with each other. To assess the convergent validity of the measure, we have compared the measure results with related measures of patient safety and outcomes. For this comparison, we drew on hospital-level quality measure results publicly available on data.Medicare.gov. Using Spearman rank correlation coefficients, we compared hospital-level failure-to-rescue rates with rates of risk-standardized 30-day readmission and mortality rates (e.g., hospital-wide unplanned all-cause readmissions), complications for hip/knee replacement patients and a composite measure of patient safety and adverse events. Correlations among these measures would support the validity of the failure-to-rescue measure because they measure a similar quality construct of patient safety. However, we do not expect strong correlations because patient safety is a complex construct, and these measures differ from the failure-to-rescue measure in terms of the populations and conditions being measured.

Known groups validity is a type of construct validity that focuses on a measure's ability to discriminate between groups of measured entities that are known to differ on the underlying latent construct. With respect to hospital quality and safety, prior research has demonstrated several "known groups" that can be identified from the available data:

- Hospital resident-to-bed ratio, stratified as major teaching/academic (at least 0.25 fulltime equivalent [FTE] residents per bed), minor teaching/academic (more than 0 but less than 0.25 FTE residents per bed), and non-teaching
- Hospital nurse-to-bed ratio, stratified as highly staffed (more than 2.0 FTE licensed nurses per bed), moderately staffed (1.0-2.0 nurses per bed), poorly staffed (less than 1.0 nurses per bed)
- Hospital nurse skill mix, estimated as the proportion of all nursing FTEs or nursing hours that are provided by registered nurses (versus licensed vocational/practical nurses), stratified as relatively low (less than 85%), medium (85-97.5%), and high (over 97.5%)
- Hospital urban/rural location.

We hypothesized that failure-to-rescue rates would be lower at major teaching hospitals, urban hospitals, and hospitals with high nurse staffing and skill mix than at non-teaching hospitals, rural hospitals, and hospitals with low nurse staffing and skill mix, respectively.

Face validity refers to the degree to which evidence, clinical judgement, and theory support the interpretations of a measure score. Face validity is an assessment by experts that determines the extent to which a measure, at face value, appears to reflect what it is intended to assess. To determine face validity, we obtained input from members of the TEP to determine whether they think the measure as specified will help inform consumers and help providers improve quality.

5.3.4 Validity Testing Results

Convergent validity was assessed using other measures of hospital quality that are used in Federal programs, focusing on measures that do not cover postoperative mortality. For all but one of these comparisons, the proposed measure demonstrates higher convergent validity than the current CMS PSI 04 measure (**Table 4 in the logic model attachment**). Of note, the Spearman rank correlation coefficient between this measure and the 30-day hospital-wide unplanned readmission measure was 0.229 ($p < 0.001$). These findings show the expected direction and strength. Hospitals with higher nurse staffing and skill mix tend to have lower death rates after serious postoperative complications. Hospitals that identify complications late or fail to treat them aggressively tend to have higher 30-day readmission rates and higher death rates after serious postoperative complications.

As shown in **Table 5 of the logic model attachment**, the data support these hypotheses for all “known groups” except rural/urban location. Full-time equivalent nurse-to-bed ratio was classified as <1; 1-2; or 2. Relative to the 496 hospitals with the lowest nurse staffing, the 1,266 hospitals with intermediate nurse staffing had an overall rate ratio of 0.98, and the 445 hospitals with the highest nurse staffing had an overall rate ratio of 0.84 ($p < 0.001$). Similar results were found for nursing skill mix; 872 hospitals with the highest ratios of RN-to-total nurse staffing had an overall rate ratio of 0.83 ($p < 0.001$), compared with the 328 hospitals with the lowest ratios.

Face validity results are as follows:

- 9 of 10 members (90%) voted “yes” that the measured outcome (rate of 30-day mortality among surgical inpatients with complications) provides a representation of relevant quality in a facility.
- 9 of 10 members (90%) voted “yes” that implementation of the measure in hospital inpatient quality reporting programs (in place of current PSI 04) is likely to lead to improve quality of care by reducing the frequency of failure to rescue.
- 5 of 5 members (100%) who are employed by a “measured entity” (i.e., employed or affiliated with hospital organizations) voted “yes” that the proposed measure is easy to understand and may be useful for decision-making.

The one member who disagreed felt that the proposed denominator expansion (adding patients who experience less serious complications after surgery) makes the measure less relevant to identifying hospitals’ performance in rescuing higher risk/serious cases. The member indicated that other CMS mortality measures address lower risk cases, while PSI 04 is unique in its focus on patients with a very high risk of death. In response, the team highlighted that there is only one current mortality measure that focuses on surgical cases and that measure is limited to CABG. This proposed expansion is bringing a new and broader population of surgical patients into the measurement sphere. These patients better represent “typical” surgical patients undergoing

bariatric surgery, orthopedic surgery, cancer surgery, colorectal surgery, etc. Only if patients with mild-to-moderate complications are brought into the denominator can we focus attention on preventing the progression of complications from mild to serious, which is the core of the failure-to-rescue concept. The improvements to this measure make it unique as a measure of surgical outcomes (failure-to-rescue) across a broad set of non-emergency procedures.

5.3.5 Interpretation of Validity Results

Systematic assessment of face validity of the performance measure score confirms that the score is believed to accurately reflect hospital performance with respect to postoperative care, and to distinguish good from poor performance. The only negative vote in the expert panel process was motivated by concern about modifying the denominator population, compared with the current CMS PSI 04 measure, by excluding certain high-risk patients such as multiple trauma, burns, and transplants, and refocusing the denominator population on general surgery, orthopedic surgery, and cardiovascular surgery. However, this change was motivated by over a decade of feedback from the user community and both public and private stakeholders.

Empirical testing results confirm that the proposed Failure to Rescue measure, which is designed to align with the prior CBE-endorsed measure #0353 (“Failure to Rescue 30-day Mortality”), has superior convergent validity and known groups validity compared with the measure currently used in CMS programs, CMS PSI 04. These properties are also consistent with the performance of #0353, as previously reported to the CBE.

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

There are established risk factors for failure to rescue, many of which are outside hospitals’ control (e.g., age, comorbidity burden). Risk factors for failure to rescue can be categorized into three groups - (1) patient risk factors for mortality within 30 days of surgery, such as age, comorbidities, or preoperative ‘do not resuscitate’ orders; (2) social risk factors that can influence patient risk, such as patient functional status, race/ethnicity, or socioeconomic status, and; (3) hospital factors, such as nurse and resident staffing, staff skill mix, hospital volume and technological resources. Patient attributes (demographics, comorbid conditions, clinical signs and symptoms, functional risk factors, and others) present at the start of care are integral components of the risk model, in that they directly influence the measured outcome and hospitals have less control. Care processes and intermediate factors (or mediators) can influence failure to rescue rates. These factors are largely within a hospital’s control and are therefore not considered as risk

factors. These process factors are summarized in the Importance section. Examples of models that have been included in published studies are included in **Table 6 of the logic model attachment**.

5.4.2a Attach Conceptual Model

[Graphic for FTR RA Conceptual Model.zip](#)

5.4.3 Variable Distribution Across Measured Entities

Because of the large number of measured entities (2,907 with at least one denominator record; 2,055 with at least 25 denominator records), we are unable to report descriptive statistics for the risk variables at the entity level. For additional details regarding the overall frequency of all risk factors (and risk factors that were considered but not selected for the final model), please refer to **Table 3 in the logic model attachment**. Mean age varies across measured entities from a minimum of 63.9 years to a maximum of 79.4 years, with 25th, 50th, and 75th percentile values of 72.0, 73.3, and 74.4 years, respectively. Mean values of the Elixhauser (AHRQ) Comorbidity Risk of Mortality Index vary across measured entities from a minimum of -5.5 to a maximum of 22.4, with 25th, 50th, and 75th percentile values of 4.0, 6.0, and 7.8, respectively. Finally, as a summary measure of variation, the expected rate of Failure to Rescue varies across measured entities from a minimum of 2.21 per 1,000 surgical cases to a maximum of 130.67 per 1,000 surgical cases, with 25th, 50th, and 75th percentile values of 34.79, 44.28, and 52.51, respectively.

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

The final risk-adjustment model was estimated using cluster-adjusted multivariable logistic regression to optimize calibration, after testing both logistic and probit link functions. The model was also estimated using a mixed-level logistic model with hospital random effects, but the results (including the confidence intervals surrounding parameter estimates) were virtually unchanged, compared with simpler form models. All risk factors were dichotomous (0/1) except for:

- age, which was tested in both piecewise linear and categorical forms;
- discharge quarter, which was tested as a set of dummy variables to capture secular trends in risk-standardized mortality over time (and unmeasured secular trends in case mix due to the post-pandemic backlog in elective surgery);
- Modified Diagnosis-Related Groups (MDRGs) representing aggregates of adjacent CMS MS-DRGs without comorbidities or complications, with comorbidities or complications, or with major comorbidities or complications, which were tested as a fully saturated set of dummy variables;
- AHRQ's default Clinical Classifications Software Refined (CCSR) for International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM)-codes, applied to the principal diagnosis on each record, which were tested as a fully saturated set of dummy variables; and
- Elixhauser Index for Risk of In-hospital Mortality, which was tested as a continuous variable.

MDRGs were used to adjust for the type of operation for which the patient was admitted (excluding tracheostomy, which often follows a period of postoperative respiratory failure). CCRs were used to adjust for the principal reason for the patient's admission to the hospital. The Elixhauser Index was used to adjust for the combined effect of multiple comorbidities, including comorbidities that were not sufficiently frequent or sufficiently impactful to be selected as independent risk factors.

All data came from the fields available on Medicare FFS claims and Medicare Advantage shadow claims (inpatient encounter records), including ICD-10-CM diagnosis codes for comorbidities present on admission, ICD-10-CM principal diagnosis codes, ICD-10-PCS procedure codes affecting the CMS MS-DRG assignment, hospital-reported source of admission (i.e., transfer from another hospital), and demographic fields for age, sex, and discharge year and quarter. Interactions between COVID-19 present on admission and discharge quarter were used to account for the changing impact of COVID-19 over time, as population immunity has improved and more effective treatments have become available. Two transfer variables were created to adjust for the possibility that patients transferred from one hospital to another for an operation may be at higher risk than patients who remain at the hospital where they presented, even after adjusting for other measured patient characteristics. One of these features is based on transfers reported by the receiving hospital, and the other is based on transfers identified from Medicare claims data even without reporting by the receiving hospital.

Guided by the conceptual model, we developed the baseline risk adjustment model for FTR using the following process.

1. Randomly partitioned the full denominator data into an 80% training set and a 20% hold-out (model performance or evaluation) test set.
2. Created contingency tables for all categorical features to identify any that had zero cells for either the positive or negative outcome. These features were not considered further due to anticipated model convergence problems (i.e., quasi-complete separation). For continuous variables, such as age, we ran locally weighted bivariate regressions (i.e., locally weighted scatterplot smoothing, or LOWESS) to understand the functional form of the relationship. This analysis confirmed that the risk of FTR was not linearly related to age, except for the limited age range between 70 and 90 years.
3. Fit one model using the least absolute shrinkage and selection operator (LASSO) on the training set using 10-fold cross-validation (CV). This step helped to assess model fit on the training set, while facilitating parameter tuning (e.g., the lambda regularization parameter in the cross-validation [CV]-based LASSO). We chose the final model where the regularization parameter (lambda) was set to λ_{1se} , i.e., "one-standard-error" (i.e., the largest lambda at which the mean squared error (MSE) is within one standard error of the minimum MSE.). This rule is standard practice for improving generalization, and its suitability was confirmed using the hold-out test set.

4. Given that Lasso was able to provide a robust solution, with consistent selection of the same 120 ± 5 features, we did not use other penalized regression approaches (e.g., Elastic Net).
5. The final risk-adjustment model was a cluster-adjusted logistic regression model. The model was estimated on the entire dataset using the set of features selected by Lasso through 10-fold cross-validation and testing on the hold-out test set.
6. The risk-adjustment model was also tested with additional social drivers of health variables (Medicaid insurance, Hispanic ethnicity, Race), considered individually and collectively.

References

1. T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning (Springer, 2001), vol. 1.

5.4.4a Attach Risk/Case-mix Adjustment Modeling and/or Stratification Specifications

[FTR RISK MODEL v2.xlsx](#)

5.4.5 Calibration and Discrimination

We summarize model performance using the following measures:

-Overall model discrimination as assessed by C-statistic. The C-statistic is the area under the receiver-operator curve (i.e., AUC) that measures the discriminative ability of a regression model across all levels of risk. It also describes the probability that a randomly selected patient who experienced a fall with injury had a higher expected value than a randomly selected patient who did not experience that event. The AUC was 0.816 in the holdout test set (based on Lasso) and 0.818 for the final logistic model. These values indicate strong discrimination performance, relative to a random classifier with $AUC=0.5$.

-The precision-recall (PR) curve and the area under the curve (AUPRC). The PR curve and AUPRC are less sensitive to data imbalance or class imbalance (i.e., very rare events) than the AUC. The AUPRC was 0.184 in the holdout test set (based on Lasso), indicating good prediction at the individual patient level relative to a random classifier with $AUPRC=0.043$.

-Model calibration was assessed across deciles of patient risk using Hosmer-Lemeshow plots. The deciles of risk are ten mutually exclusive groups containing equal numbers of discharges, ranging

from very low-risk patients (according to the model) to high-risk patients. We do not provide Hosmer-Lemeshow test statistics because, given the large sample size of our data, the null hypothesis is almost always rejected. Moreover, the plots provide more detail on model fit than the overall Hosmer-Lemeshow statistic. Because over 43% of events occurred in the highest-risk decile, and over 63% occurred in the highest-risk quintile, the decile analysis is statistically unstable. However, the analysis suggests overestimation of risk among low-risk patients in the bottom five deciles (i.e., observed-to-expected ratios of 0.64-0.84 among patients with death rates under 2%), but very accurate estimation among high-risk patients in the top five deciles (i.e., observed-to-expected ratios of 0.99-1.11 among patients with death rates over 2%). Alternative link functions are being tested to better account for the overestimation of risk among low-risk patients.

5.4.5a Attach Calibration and Discrimination Testing Results

[FTR PSI04 CALIBRATION AND DISCRIMINATION TESTING final.pdf](#)

5.4.6 Interpretation of Risk/Case-mix Factor Findings

See above.

5.4.7 Final Approach to Address Risk Factors

Statistical risk adjustment model with risk factors

6.1.2 Current or Planned Use(s)

Public Reporting

6.2.1 Actions of Measured Entities to Improve Performance

No facility is expected to have a zero rate for this measure because it targets rescue from severe conditions that have a non-zero death rate. In many cases, rescue procedures may be unsuccessful or the decision to discontinue them may be made. When treatment of a complication has been unsuccessful, providers and family members often decide to order “do not resuscitate” or “palliative care” or “comfort measures only”; these choices do not affect the measure because they are generally consequences of the patient’s clinical deterioration, not direct causes of it.

However, there are evidence-supported interventions that hospital can implement to improve timely identification of clinical deterioration and treatment of preventable complications, including improved nurse staffing, simulation training, standardized communication tools, electronic monitoring and/or warning systems, and rapid response systems.

Developer POC email

BColip@mathematica-mpr.com

Measure Developer POC

Brittany Colip
Mathematica
600 Alexander Park Ste 100
Princeton, NJ 08540
United States

Measured/accountable entity (reliability and/or validity) methodology and results (if available)

Measured entity (reliability and validity) methodology and results (if available)

Responder for Survey

Patient

Specify number of risk factors

126

The measure developer is different from the measure steward

Yes

Steward Address

Donta Henson
7500 Security Boulevard
Baltimore, MD 21244
United States

Steward Organization

Centers for Medicare & Medicaid Services

Steward Organization URL

<https://www.cms.gov/>

Steward POC email

DONTA.HENSON1@cms.hhs.gov