

CBE ID

4130e

Title

Hospital Harm - Postoperative Respiratory Failure

Project

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

Endorsement Status

Endorsed

Is Under Review

No

Next Maintenance Cycle

Fall 2029

Previous Endorsement Cycle

Fall 2023

Initial Endorsement

Mon, 03/18/2024 - 13:23

Steward

Centers for Medicare & Medicaid Services

1.0 New or Maintenance

New

1.3 Electronic Clinical Quality Measure (eCQM)

Yes

1.6 Measure Description

This electronic clinical quality measure (eCQM) assesses the proportion of elective inpatient hospitalizations for patients aged 18 years and older without an obstetrical condition who have a procedure resulting in postoperative respiratory failure (PRF).

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 this is not a paired measure.

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.com/seerationaleabove>

1.12a Attach MADiE Output

[HH-PRF-v0-0-141-QDM-5-6.zip](#)

1.13 Data Dictionary

Attached

1.13a Attach Data Dictionary

[Postoperative Respiratory Failure Value Set Directory 11 01 2023.zip](#)

1.14 Numerator

The numerator is elective inpatient hospitalizations for patients with postoperative respiratory failure (PRF) as evidenced by: Criterion A: Mechanical Ventilation (MV) initiated within 30 days after First operating room (OR) procedure, as evidenced by: A.1. Intubation that occurs outside of a procedural area and within 30 days after the end of the First OR procedure of the encounter or A.2. MV that occurs outside of a procedural area within 30 days after the end of the First OR procedure of the encounter and is preceded by a period of non-invasive oxygen therapy between the end of the OR procedure and the MV occurrence, and without a subsequent OR procedure between the non-invasive oxygen therapy and the MV occurrence or Criterion B: MV with a duration of more than 48 hours after the First OR procedure, as evidenced by: B.1. Extubation that occurs outside of a procedural area more than 48 hours after the end of an OR procedure and within 30 days after the end of the First OR procedure, and is not preceded by a period of non-invasive oxygen therapy or a subsequent OR procedure between the end of the OR procedure and the extubation occurrence or B.2 Mechanical ventilation that occurs between 48 and 72 hours after the end of an OR procedure and within 30 days after the end of the First OR procedure, and is not preceded by a non-invasive oxygen therapy or a subsequent OR procedure between the end of the OR procedure and the MV occurrence

1.14a Numerator Details

The numerator is elective inpatient hospitalizations for patients with postoperative respiratory failure (PRF) as evidenced by:

- Criterion A: Mechanical Ventilation (MV) initiated within 30 days after First operating room (OR) procedure, as evidenced by:
- A.1. Intubation that occurs outside of a procedural area and within 30 days after the end of the First OR procedure of the encounter,

or

- A.2. MV that occurs outside of a procedural area within 30 days after the end of the First OR procedure of the encounter and is preceded by a period of non-invasive oxygen therapy between the end of the OR procedure and the MV occurrence, and without a subsequent OR procedure between the non-invasive oxygen therapy and the MV occurrence

Or

- Criterion B: MV with a duration of more than 48 hours after the First OR procedure, as evidenced by:
- B.1. Extubation that occurs outside of a procedural area more than 48 hours after the end of an OR procedure and within 30 days after the end of the First OR procedure, and is not preceded by a period of non-invasive oxygen therapy or a subsequent OR procedure between the end of the OR procedure and the extubation occurrence

or

- B.2 Mechanical ventilation that occurs between 48 and 72 hours after the end of an OR procedure and within 30 days after the end of the First OR procedure, and is not preceded by a non-invasive oxygen therapy or a subsequent OR procedure between the end of the OR procedure and the MV occurrence

The time period for data collection is during an elective inpatient hospitalization, which is defined as beginning at hospital arrival including time in observation or outpatient surgery when the transition between these encounters (if they exist) and the inpatient encounter are within an hour or less of each other.

All data elements necessary to calculate this numerator are defined within value sets available in the Value Set Authority Center (VSAC) and listed below:

- Intubation procedures are represented by the value set Intubation (2.16.840.1.113762.1.4.1248.179)
- Procedural areas are represented by the value set Procedural Hospital Locations (2.16.840.1.113762.1.4.1248.216)
- Operating room (OR) procedures are represented by the value set General and Neuraxial Anesthesia (2.16.840.1.113762.1.4.1248.208)
- Non-invasive oxygen therapies are represented by the value sets Non Invasive Oxygen Therapy (2.16.840.1.113762.1.4.1248.213) and Non Invasive Oxygen Therapy by Nasal Cannula or Mask (2.16.840.1.113762.1.4.1248.209)
- Mechanical ventilation (MV) procedures are represented by the value set Mechanical

Ventilation (2.16.840.1.113762.1.4.1248.107)

- Extubation procedure is represented by the direct reference code "Removal of endotracheal tube (procedure)" (SNOMEDCT Code 271280005)

To access the value sets for the measure, please visit the Value Set Authority Center (VSAC), sponsored by the National Library of Medicine, at <https://vsac.nlm.nih.gov/>.

1.15 Denominator

Elective inpatient hospitalizations that end during the measurement period for patients aged 18 and older without an obstetrical condition and at least one surgical procedure was performed within the first 3 days of the encounter.

1.15a Denominator Details

This measure includes all elective inpatient hospitalizations that end during the measurement period for patients aged 18 and older without an obstetrical condition and at least one surgical procedure was performed within the first 3 days of the encounter, and all payers. Elective inpatient hospitalizations include time in observation or outpatient surgery when the transition between these encounters (if they exist) and the inpatient encounter are within an hour or less of each other.

Measurement period is one year. This measure is at the hospital-by-admission level.

All data elements necessary to calculate this denominator are defined within value sets available in the Value Set Authority Center (VSAC) and listed below:

- Elective inpatient encounters are represented using the value set of Elective Inpatient Encounter (2.16.840.1.113762.1.4.1248.85) and direct reference code "Elective (qualifier value)" (SNOMEDCT Code 103390000)
- Observation encounters are represented using the value set of Observation Services (2.16.840.1.113762.1.4.1111.143)
- Outpatient surgery encounters are represented using the value set of Outpatient Surgery Service (2.16.840.1.113762.1.4.1110.38)
- Obstetrical condition diagnoses are represented using the value set of Obstetrics and VTE Obstetrics (2.16.840.1.113762.1.4.1248.33)
- Surgical procedures are represented using the value set of General and Neuraxial Anesthesia (2.16.840.1.113762.1.4.1248.208)

To access the value sets for the measure, please visit the Value Set Authority Center (VSAC),

sponsored by the National Library of Medicine, at <https://vsac.nlm.nih.gov/>.

1.15b Denominator Exclusions

Inpatient hospitalizations for patients:

- Who have mechanical ventilation that starts more than one hour prior to the start of the first operating (OR) procedure
- With arterial partial pressure of oxygen (PaO₂) < 50 mmHg within 48 hours or less prior to the start of the first OR procedure
- With arterial partial pressure of carbon dioxide (PaCO₂) > 50 mmHg combined with an arterial pH < 7.30 within 48 hours or less prior to the start of the first OR procedure
- With a principal diagnosis for acute respiratory failure
- With a secondary diagnosis for acute respiratory failure present on admission
- With any diagnosis present on admission for the existence of a tracheostomy
- Where a tracheostomy is performed before or on the same day as the first OR procedure
- With any diagnosis for neuromuscular disorder or degenerative neurological disorder
- With any procedure for selected pharyngeal, nasal, oral, facial, or tracheal surgery involving significant risk of airway compromise likely to require prophylactic retention of the endotracheal tube for at least 48 hours

1.15c Denominator Exclusions Details

The time period for data collection is during an elective inpatient hospitalization, beginning at hospital arrival including time in observation or outpatient surgery when the transition between these encounters (if they exist) and the inpatient encounter are within an hour or less of each other.

All data elements necessary to calculate this numerator are defined within value sets available in the Value Set Authority Center (VSAC) and listed below:

- Mechanical ventilation (MV) procedures are represented by the value set Mechanical Ventilation (2.16.840.1.113762.1.4.1248.107)
- Operating room (OR) procedures are represented by the value set General and Neuraxial Anesthesia (2.16.840.1.113762.1.4.1248.208)
- Arterial partial pressure of oxygen (P_aO₂) laboratory tests are represented by the direct reference code "Oxygen [Partial pressure] in Arterial blood" (LOINC Code 2703-7)
- Arterial partial pressure of carbon dioxide (P_aCO₂) laboratory tests are represented by the direct reference code "Carbon dioxide [Partial pressure] in Arterial blood" (LOINC Code 2019-8)
- Arterial pH laboratory tests are represented by the direct reference code "pH of Arterial

blood" (LOINC Code 2744-1)

- Acute respiratory failure diagnoses are represented by the value set Acute Respiratory Failure (2.16.840.1.113762.1.4.1248.88)
- The present on admission indicators are represented by the value set Present on Admission or Clinically Undetermined (2.16.840.1.113762.1.4.1147.197)
- Tracheostomy diagnoses are represented by the value set Tracheostomy Diagnoses (2.16.840.1.113762.1.4.1248.89)
- Tracheostomy procedures are represented by the value set Tracheostomy Procedures (2.16.840.1.113762.1.4.1248.181)
- Neuromuscular disorder diagnoses are represented by the value set Neuromuscular Disorder (2.16.840.1.113762.1.4.1248.91)
- Degenerative neurological disorder diagnoses are represented by the value set Degenerative Neurological Disorder (2.16.840.1.113762.1.4.1248.92)
- Selected pharyngeal, nasal, oral, facial, or tracheal surgeries involving significant risk of airway compromise likely to require prophylactic retention of the endotracheal tube for at least 48 hours are represented by the value set Head and Neck Surgeries with High Risk Airway Compromise (2.16.840.1.113762.1.4.1248.183)

To access the value sets for the measure, please visit the Value Set Authority Center (VSAC), sponsored by the National Library of Medicine, at <https://vsac.nlm.nih.gov/>.

1.16 Type of Score

Rate/proportion

1.17 Measure Score Interpretation

Better performance = Lower score

1.18 Calculation of Measure Score

See attached diagram.

1.18a Attach measure score calculation diagram

[PRF_Measure Calculation Diagram 11 01 2023.pdf](#)

1.19 Measure Stratification Details

N/A; this measure is not stratified.

1.20 Types of Data Sources

Electronic Health Records

1.25 Data Source Details

Hospitals collect EHR data using certified electronic health record technology (CEHRT). The MAT

output, which includes the human readable and XML artifacts of the clinical quality language (CQL) for the measure are contained in the eCQM specifications attached. No additional tools are used for data collection for eCQMs.

1.26 Minimum Sample Size

Because this measure is limited to elective surgical encounters and is a rare event, a minimum volume threshold of at least 100 in the denominator is recommended to achieve a minimal ICC of approximately 0.2.

2.1 Attach Logic Model

[Hospital Harm - PRF Logic Model and Tables 11 01 2023.pdf](#)

2.2 Evidence of Measure Importance

PRF is the most common serious postoperative pulmonary complication (Arozullah et al., 2000; Canet et al., 2015; Gupta et al., 2011; and Kor et al., 2014). Postoperative pulmonary complications (PPCs) increase postoperative mortality, and health care costs (Mohanty et al., 2016; Miskovic, 2017). Some cases of PRF are potentially preventable with optimal care. Factors that might contribute include careful management of intra- and perioperative ventilator use and fluids, reducing surgical duration, using regional anesthesia, preventing wound infection, and optimizing pain control (Stocking et al, 2022; Encinosa et al, 2008; Zrelak , 2012). Mechanical ventilation administered invasively (via an endotracheal tube) is inherently unpleasant and resource-intensive, and virtually all sentient patients would prefer to avoid prolonged periods of it (i.e., over 48 hours, or re-initiation of mechanical ventilation after extubation), if possible. An eCQM-based Hospital Harm PRF measure would enable hospitals to assess harm reduction efforts and modify their quality improvement efforts more reliably. The measure would also help to identify hospitals that have persistently high PRF rates. The measure will ensure that PRF events are tracked and that hospitals are incentivized to reduce the incidence of PRF. The eCQM would also be able to identify cases from an all-payer population, as it would not be dependent upon claims-based ICD-10-CM coded data.

Some of the many causes of PRF are such clear precipitants that the occurrence or identification of PRF per se have received little focus in recent medical literature: e.g., respiratory depression from excessive opiate administration; multiple system organ failure from postoperative sepsis; or postoperative cardiac arrest. The available studies therefore have focused on less understood (but potentially causal) pathways and identified more subtle associations between specific intraoperative risk factors and PRF (Blum et al., 2013; Attaallah et al., 2019; Shalev et al., 2014; Hughes et al., 2010; and Chandler et al., 2020). Analyzing data on 50,367 patient admissions for common adult surgical procedures using an anesthesia information system between 2004 and 2009, Blum et al. identified intraoperative risk factors associated with subsequent development of acute respiratory distress syndrome (ARDS) among patients with similar preoperative risk: ventilator drive pressure (OR=1.17 per cm H₂O), fraction inspired oxygen (OR=1.02 per 0.01), erythrocyte transfusion (OR=5.36), and crystalloid intravenous fluid administration

(OR=1.37 per liter). The number of different anesthetics administered during the admission was associated with higher risk of ARDS (OR=1.37) (Blum et al., 2013).

Hughes et al. identified intraoperative risk factors for the postoperative development of ARDS among 89 patients admitted to the ICU with PRF. In this study, patients who received more than 20mL/kg/h fluid resuscitation in the operating room had a higher chance of developing ARDS than those who received less than 10mL/kg/h (OR=3.8, p=0.04). Those who received between 10 and 20mL/kg/h had a non-significant odds ratio of 2.4 (p=0.14) (Hughes et al., 2010).

In multivariable analysis of the National Surgical Quality Improvement Program (NSQIP) database of adult inpatients who underwent neurosurgery under general anesthesia (2005-2010), Shalev and co-authors found that operative time exceeding 3 hours was associated with increased risk of reintubation (OR 2.9; 95%CI 1.8-4.8) (Shalev et al., 2014). In a retrospective time-matched cohort study, Attaallah et al. found that operative-specific risk factors including ASA status, elective case type, and surgical duration were significantly associated with postoperative respiratory failure (Attaallah et al., 2019). A recent matched case-control study conducted across five academic medical centers (n=638) found greater intraoperative ventilator volume and pressure and 24-hour fluid balance to be potentially modifiable factors associated with PRF (Stocking et al., 2021 and 2022).

Two studies describe quality improvement interventions that resulted in decreased rates of acute respiratory failure (ARF) (Braddock et al., 2014; Cassidy et al., 2013). In a one-year, prospective cohort intervention study involving 13,743 patients in a large academic medical center, Braddock et al. found that, adjusting for patient characteristics, implementation of a multifaceted, microsystem intervention utilizing in situ simulation training (TRANSFORM) was associated with a significantly decreased rate of ARF (Braddock et al., 2014). Multivariable logistic regression showed reduced odds of ARF following the intervention (OR 0.58, 95% CI 0.35 to 0.96).

In a pre-post intervention study of 250 patients at an academic safety net hospital, Cassidy et al. found a trend towards fewer unplanned intubations following the I COUGH intervention, which emphasized incentive spirometry, coughing and deep breathing, oral care, patient and family education, head-of-bed elevation, and promoting mobilization (Cassidy et al., 2013). The incidence of unplanned intubations declined from 2.0% to 1.2% in the intervention group (p = 0.09) but remained relatively stable at comparable NSQIP hospitals (1.4% to 1.6%). Risk-adjusted NSQIP data showed that unplanned intubations fell from an observed-to-expected (OE) ratio of 2.10 (95% CI 1.42 to 2.98) before I COUGH to an OE ratio of 1.31 (95% CI, 0.87 to 1.97) after the intervention; however, the authors did not report the statistical significance of this difference.

A systematic review of incentive spirometry after upper abdominal surgery found no evidence that this intervention is effective in preventing pulmonary complications, including acute respiratory inadequacy (Guimaraes et al., 2009). However, another systematic review by Lawrence et al. evaluated all interventions to prevent postoperative pulmonary complications after non-cardiothoracic surgery. These authors identified good evidence suggesting that lung expansion therapy (for example, incentive spirometry, deep breathing exercises, and continuous positive airway pressure) reduces postoperative pulmonary risk after abdominal surgery and fair evidence suggesting that selective nasogastric tube decompression (i.e., minimization of their use) after abdominal surgery reduces risk. Fair evidence also suggests that short-acting neuromuscular blocking agents result in lower rates of residual neuromuscular blockade and may reduce risk for pulmonary complications (Lawrence et al., 2006).

Several studies found that PRF is associated with longer length of stay (Rahman et al., 2013; Gajdos et al., 2013; and Marda et al., 2013). In a multivariable analysis of National Inpatient Sample (NIS) data from 2002-2010, Rahman et al. found that length of stay was significantly longer for patients with PRF (median 8.0 days) compared to those without respiratory failure (median 4.0 days, $p < 0.0001$) (Rahman et al., 2013). Using NSQIP data, Gajdos et al. found that failure to wean from ventilator and reintubation were associated with longer postsurgical length of stay in all age groups compared with participants not having these complications (median length of stay ≥ 19 days with complications; $p < 0.001$) (Gajdos et al., 2013). In a smaller study ($n=178$), Marda et al. found that mean duration of intensive care unit (ICU) and hospital stay after surgery was significantly longer in patients who had PPCs, including respiratory failure, as compared to patients without PPCs (9.5 ± 14.8 days vs. 2.7 ± 1.8 days, [$p < 0.001$]; 22.6 ± 16.8 days vs. 7.6 ± 2.8 days [$p < 0.001$], respectively) (Marda et al., 2013).

Several studies also found that PRF is associated with higher 30-day readmission rates (Sabate et al., 2014; Rosen et al., 2013; and Lawson et al., 2013). In three studies included in a recent literature review by Sabate et al., the estimated increased costs in U.S. dollars associated with PRF ranged from \$5,983 to \$7,109 per procedure (for complications not requiring ventilation) to \$118,841 to \$120,579 (for complications requiring tracheostomy), in part due to more readmissions (Sabate et al., 2014). In a cross-sectional analysis of VA patient treatment files, including 1,807,488 index hospitalizations and 262,026 readmissions, Rosen et al. found that 30-day readmission rates after surgical hospitalizations with a PSI 11 event (17.8%) were significantly higher than after surgical hospitalizations without a PSI 11 event (9.9%) ($p < 0.0001$), with an adjusted odds ratio of 1.39 (95% CI 1.25 to 1.54) (Rosen et al., 2013). In a cohort study of NSQIP data from the American College of Surgeons (ACS) and Medicare inpatient claims ($n=90,932$), the rate of unplanned intubation within 30 days of an index procedure was significantly higher among patients with a 30-day readmission (4.1%) than among those without a 30-day readmission (1.8%, $p < 0.001$) (Lawson et al., 2013). Likewise, prolonged ventilation was more frequent among readmitted patients (4.4%) than among patients who were not readmitted (2.7%, $p < 0.001$). Bath et al. used Medicare data (MedPAR) from 2009 to 2012 and found that the odds of 30-day readmission among patients undergoing abdominal aortic aneurysm repair were increased among patients with postoperative respiratory failure (OR=1.44, $p < 0.0001$) (Bath et al., 2018).

Four different population-based studies have demonstrated that PRF is independently associated with mortality. Based on NIS data of morbidly obese patients who underwent bariatric surgery, Masoomi et al. found that patients who developed ARF had significantly greater in-hospital mortality than those who did not develop this complication (5.69% versus 0.04%, $p < 0.01$) (Masoomi et al., 2013). Based on an analysis of data from 165,600 senior patients undergoing non-emergent major general surgeries from the ACS NSQIP registry, Gajdos et al. found that reintubation had one of the highest failure-to-rescue rates among all postoperative complications (25.6%) (Gajdos et al., 2013). In multivariable analysis of 5,318 adults undergoing cardiothoracic surgery at a single institution, the risk of perioperative mortality was significantly increased among patients with a respiratory failure complication (OR 3.2, 95% CI 2.2 to 4.9) (Rahmanian et al., 2013). Gray et al. retrospectively examined 57,000 inpatient discharges at six hospitals between July 2012 and June 2014 and found that hospitalizations with a PSI 11 event were associated with an additional 3.78 hospital days, compared to hospitalizations without a PSI 11 event ($p < 0.001$), as well as a significantly increased risk of in-hospital mortality (OR=248.93; $p < 0.001$) (Gray et al., 2017). One small study ($n = 450$) of patients from the ACS NSQIP database undergoing thoracoabdominal aortic aneurysm (TAAA) repair did not find such an association between reintubation and mortality (Bensley et al., 2013).

While the recent literature does not identify many ways to prevent PRF, the evidence base is increasing and providers intuitively seek to minimize the occurrence of PRF through many of their routine practices. Adoption of this eCQM has the potential to improve the quality of care for surgical patients and, therefore, increase patient safety, which is a priority area identified by the National Quality Strategy (Rosen et al., 2013). This eCQM would fill a gap in measurement for the all-payer population. Additionally, with a systematic EHR-based patient safety measure in place, hospitals can more reliably assess harm reduction efforts and modify their efforts in near real-time. In this way, greater achievements in reducing postoperative respiratory failure and enhancing hospital performance on patient safety outcomes can be expected.

Please see tables 14-21 in the **logic model attachment** for clinical practice guidelines.

References:

1. Arozullah, A. M., Daley, J., Henderson, W. G., & Khuri, S. F. (2000). Multifactorial risk index for predicting postoperative respiratory failure in men after major noncardiac surgery. The National Veterans Administration Surgical Quality Improvement Program. *Annals of surgery*, 232(2), 242-253.
2. Attaallah A.F., Vallejo M.C., Elzamzamy O.M., Mueller M.G., Eller W.S. (2019). Perioperative risk factors for postoperative respiratory failure. *J Perioper Pract.* 29(3), 49-53.
3. Bensley R.P., Curran T., Hurks R., et al. (2013). Open repair of intact thoracoabdominal aortic aneurysms in the American College of Surgeons National Surgical Quality

- Improvement Program. *J Vasc Surg*, 58(4), 894-900.
4. Blum J.M., Stentz M.J., Dechert R., et al. (2013). Preoperative and intraoperative predictors of postoperative acute respiratory distress syndrome in a general surgical population. *Anesthesiology*. 118(1), 19-29.
 5. Cassidy M.R., Rosenkranz P., McCabe K., Rosen J.E., McAneny D. (2013) I COUGH: reducing postoperative pulmonary complications with a multidisciplinary patient care program. *JAMA surgery*. 148(8), 740-745.
 6. Canet, J., Sabaté, S., Mazo, V., Gallart, L., de Abreu, M. G., Belda, J., Langeron, O., Hoefft, A., Pelosi, P., & PERISCOPE group (2015). Development and validation of a score to predict postoperative respiratory failure in a multicentre European cohort: A prospective, observational study. *European journal of anaesthesiology*, 32(7), 458-470.
 7. Chandler D, Mosieri C, Kallurkar A, et al. (2020). Perioperative strategies for the reduction of postoperative pulmonary complications. *Best Pract Res Clin Anaesthesiol*, 34(2), 153-166.
 8. Bath J., Dombrovskiy V.Y., Vogel T.R. (2018). Impact of Patient Safety Indicators on readmission after abdominal aortic surgery. *J Vasc Nurs*. 36(4), 189-195.
 9. Braddock C.H., 3rd, Szaflarski N., Forsey L., Abel L., Hernandez-Boussard T., Morton J. (2014). The TRANSFORM Patient Safety Project: A Microsystem Approach to Improving Outcomes on Inpatient Units. *Journal of general internal medicine*.
 10. Encinosa, W. E., & Hellinger, F. J. (2008). The impact of medical errors on ninety-day costs and outcomes: an examination of surgical patients. *Health services research*, 43(6), 2067-2085.
 11. Gajdos C., Kile D., Hawn M.T., Finlayson E., Henderson W.G., Robinson T.N. (2013) Advancing age and 30-day adverse outcomes after nonemergent general surgeries. *Journal of the American Geriatrics Society*, 61(9), 1608-1614.
 12. Gupta, H., Gupta, P. K., Fang, X., Miller, W. J., Cemaj, S., Forse, R. A., & Morrow, L. E. (2011). Development and validation of a risk calculator predicting postoperative respiratory failure. *Chest*, 140(5), 1207-1215.
 13. Guimaraes M.M., El Dib R., Smith A.F., Matos D. (2009). Incentive spirometry for prevention of postoperative pulmonary complications in upper abdominal surgery. *The Cochrane database of systematic reviews*, (3).
 14. Gray D.M., 2nd, Hefner J.L., Nguyen M.C., Eiferman D., Moffatt-Bruce S.D. (2017). The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *Am J Med Qual*, 32(6), 583-590.
 15. Hughes C.G., Weavind L., Banerjee A., Mercaldo N.D., Schildcrout J.S., Pandharipande P.P. (2010). Intraoperative risk factors for acute respiratory distress syndrome in critically ill patients. *Anesthesia and analgesia*. 111(2), 464-467.
 16. Kor, D. J., Lingineni, R. K., Gajic, O., Park, P. K., Blum, J. M., Hou, P. C., Hoth, J. J., Anderson, H. L., 3rd, Bajwa, E. K., Bartz, R. R., Adesanya, A., Festic, E., Gong, M. N., Carter, R. E., & Talmor, D. S. (2014). Predicting risk of postoperative lung injury in high-risk surgical patients: a multicenter cohort study. *Anesthesiology*, 120(5), 1168-1181.
 17. Lawrence V.A., Cornell J.E., Smetana G.W. (2006). Strategies to reduce postoperative pulmonary complications after noncardiothoracic surgery: systematic review for the American College of Physicians. *Ann Intern Med*, 144(8), 596-608.
 18. Lawson E.H., Hall B.L., Louie R., et al. (2013). Association between occurrence of a postoperative complication and readmission: implications for quality improvement and cost savings. *Annals of surgery*, 258(1), 10-18.
 19. Marda M., Pandia M.P., Rath G.P., Bithal P.K., Dash H.H. (2013). Post-operative pulmonary

- complications in patients undergoing transoral odontoidectomy and posterior fixation for craniovertebral junction anomalies. *Journal of anaesthesiology, clinical pharmacology*, 29(2), 200-204.
20. Masoomi H., Reavis K.M., Smith B.R., Kim H., Stamos M.J., Nguyen N.T. (2013). Risk factors for acute respiratory failure in bariatric surgery: data from the Nationwide Inpatient Sample, 2006-2008. *Surg Obes Relat Dis*, 9(2), 277-281.
 21. Miskovic A., Lumb A.B. (2017) Postoperative pulmonary complications. *Br J Anaesth*, 118:317-334.
 22. Mohanty S., Rosenthal R.A., Russell M.M. et al. (2016). Optimal perioperative management of the geriatric patient: a best practices guideline from the American College of Surgeons NSQIP and the American Geriatrics Society. *J Am Coll Surg*, 222:930-947.
 23. Rosen, A. K., Loveland, S., Shin, M., Shwartz, M., Hanchate, A., Chen, Q., Kaafarani, H. M., & Borzecki, A. (2013). Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Medical care*, 51(1), 37-44.
 24. Rahman M., Neal D., Fargen K.M., Hoh B.L. (2013). Establishing standard performance measures for adult brain tumor patients: a Nationwide Inpatient Sample database study. *Neuro Oncol*. 15(11):1580-1588.
 25. Rahmanian P.B., Kroner A., Langebartels G., Ozel O., Wippermann J., Wahlers T. (2013). Impact of major non-cardiac complications on outcome following cardiac surgery procedures: logistic regression analysis in a very recent patient cohort. *Interactive cardiovascular and thoracic surgery*, 17(2), 319-326; discussion 326-317.
 26. Sabate S., Mazo V., Canet J. (2014). Predicting postoperative pulmonary complications: implications for outcomes and costs. *Case reports in anesthesiology*. 27(2), 201-209.
 27. Shalev D., Kamel H. (2014). Risk of Reintubation in Neurosurgical Patients. *Neurocritical care*.
 28. Stocking J. C, Drake C., Aldrich J. M., et al. (2021). Risk Factors Associated With Early Postoperative Respiratory Failure: A Matched Case-Control Study. *J Surg Res*. 261, 310-319.
 29. Stocking, J. C., Drake, C., Aldrich, J. M., Ong, M. K., Amin, A., Marmor, R. A., Godat, L., Cannesson, M., Gropper, M. A., Romano, P. S., Sandrock, C., Bime, C., Abraham, I., & Utter, G. H. (2022). Outcomes and risk factors for delayed-onset postoperative respiratory failure: a multi-center case-control study by the University of California Critical Care Research Collaborative (UC3RC). *BMC anesthesiology*, 22(1), 146.
 30. Zrelak, P. A., Utter, G. H., Sadeghi, B., Cuny, J., Baron, R., & Romano, P. S. (2012). Using the Agency for Healthcare Research and Quality patient safety indicators for targeting nursing quality improvement. *Journal of nursing care quality*, 27(2), 99-108.
 31. Zrelak, P. A., Utter, G. H., Sadeghi, B., Cuny, J., Baron, R., & Romano, P. S. (2012). Using the Agency for Healthcare Research and Quality patient safety indicators for targeting nursing quality improvement. *Journal of nursing care quality*, 27(2), 99-108.

2.3 Anticipated Impact

Postoperative respiratory failure (PRF), defined as unplanned endotracheal reintubation,

prolonged need for mechanical ventilation, or inadequate oxygenation and/or ventilation, is the most common serious postoperative pulmonary complication, with an incidence of up to 7.5% (the incidence of any postoperative pulmonary complication ranges from 10-40%) (Arozullah, et al., 2000; Canet, et al., 2015; Gupta, et al., 2011; Kor, et al., 2014). This measure addresses the prevalence of PRF and the variance between hospitals in the incidence of PRF. PRF is a serious complication that can increase the risk of morbidity and mortality, with in-hospital mortality resulting from PRF estimated at 25% to 40% (Arozullah et al., 2000; Canet, et al., 2014). Surgical procedures complicated by PRF have 3.74 times higher adjusted odds of death than those not complicated by respiratory failure, 1.47 times higher odds of 90-day readmission, and 1.86 times higher odds of an outpatient visit with one of 44 postoperative conditions (e.g., bacterial infection, fluid and electrolyte disorder, abdominal hernia) within 90 days of hospital discharge (Miller, et al., 2001; Romano, et al., 2009). PRF is additionally associated with prolonged mechanical ventilation and the need for rehabilitation or skilled nursing facility placement upon discharge (Thompson, et al., 2018).

The incidence of PRF varies by hospital, with higher reported rates of PRF in nonteaching hospitals than teaching hospitals (Rahman, et al., 2013). Additionally, one study found that the odds of developing PRF increased by 6% for each level increase in hospital size from small to large (Rahman, et al., 2013). This suggests that there remains room for improvement in hospitals reporting higher rates of PRF.

The most widely used current measures of PRF are based on either claims data (CMS PSI 11) or proprietary registry data (NSQIP of the ACS). The proposed eCQM is closely modeled after the NSQIP measure of PRF, which has been widely adopted across American hospitals, and is intended to complement and eventually supplant CMS PSI 11, which is a component of the CMS PIS 90 Patient Safety and Adverse Events Composite.

With a systematic EHR-based patient safety measure in place, hospitals can more reliably assess harm reduction efforts and modify their efforts in near real-time. In this way, greater achievements in reducing postoperative respiratory failure and enhancing hospital performance on patient safety outcomes can be expected.

Performance Results from Beta Testing:

Risk-adjusted rates showed substantial variation in performance scores across the 12 test hospitals from 0.0 to 16.79 postoperative respiratory failures per 1,000 hospital encounters, with one facility having a risk-adjusted rate significantly below the average (2.54 per 1,000 patients; 95% CI 1.43, 3.65).

Performance scores were as follows:

- Minimum: 0.00
- Median: 2.70
- Mean: 3.67
- Maximum: 16.79

See **Table 1 and Exhibit 2 in the logic model attachment** for a distribution of performance scores across sites.

References:

1. Arozullah, A. M., Daley, J., Henderson, W. G., & Khuri, S. F. (2000). Multifactorial risk index for predicting postoperative respiratory failure in men after major noncardiac surgery. The National Veterans Administration Surgical Quality Improvement Program. *Annals of surgery*, 232(2), 242-253.
2. Canet, J., & Gallart, L. (2014). Postoperative respiratory failure: pathogenesis, prediction, and prevention. *Current opinion in critical care*, 20(1), 56-62.
3. Canet, J., Sabaté, S., Mazo, V., Gallart, L., de Abreu, M. G., Belda, J., Langeron, O., Hoefft, A., Pelosi, P., & PERISCOPE group (2015). Development and validation of a score to predict postoperative respiratory failure in a multicentre European cohort: A prospective, observational study. *European journal of anaesthesiology*, 32(7), 458-470.
4. Gupta, H., Gupta, P. K., Fang, X., Miller, W. J., Cemaj, S., Forse, R. A., & Morrow, L. E. (2011). Development and validation of a risk calculator predicting postoperative respiratory failure. *Chest*, 140(5), 1207-1215.
5. Kor, D. J., Lingineni, R. K., Gajic, O., Park, P. K., Blum, J. M., Hou, P. C., Hoth, J. J., Anderson, H. L., 3rd, Bajwa, E. K., Bartz, R. R., Adesanya, A., Festic, E., Gong, M. N., Carter, R. E., & Talmor, D. S. (2014). Predicting risk of postoperative lung injury in high-risk surgical patients: a multicenter cohort study. *Anesthesiology*, 120(5), 1168-1181.
6. Miller, M. R., Elixhauser, A., Zhan, C., & Meyer, G. S. (2001). Patient Safety Indicators: using

administrative data to identify potential patient safety concerns. *Health services research*, 36(6 Pt 2), 110-132.

7. Rahman, M., Neal, D., Fargen, K. M., & Hoh, B. L. (2013). Establishing standard performance measures for adult brain tumor patients: a Nationwide Inpatient Sample database study. *Neuro-oncology*, 15(11), 1580-1588.

8. Romano, P. S., Mull, H. J., Rivard, P. E., Zhao, S., Henderson, W. G., Loveland, S., Tsilimingras, D., Christiansen, C. L., & Rosen, A. K. (2009). Validity of selected AHRQ patient safety indicators based on VA National Surgical Quality Improvement Program data. *Health services research*, 44(1), 182-204.

9. Thompson, S. L., & Lisco, S. J. (2018). Postoperative Respiratory Failure. *International anesthesiology clinics*, 56(1), 147-164.

2.5 Health Care Quality Landscape

Currently there are two related but non-competing quality measures for PRF: Postoperative Respiratory Failure Rate Patient Safety Indicator 11 (endorsement removed, NQF 0533; steward AHRQ) and Risk-Adjusted Postoperative Prolonged Intubation (Ventilation) (NQF #0129; steward: The Society of Thoracic Surgeons). This eCQM focuses on a slightly different population than the PSI 11 measure (Postoperative Respiratory Failure Rate Patient Safety Indicator 11 (endorsement removed, NQF 0533, Steward: AHRQ). The AHRQ PSI 11 measure is constructed with claims data to measure postoperative respiratory failure.

The Risk-Adjusted Postoperative Prolonged Intubation measure focuses on post-operative respiratory failure but in a narrow group of patients undergoing isolated CABG and uses registry data. This new measure will be an eCQM measure and will focus on a broader population than the CABG measure and would fill a gap in measurement for the all-payer population.

This eCQM also incorporates features of the manually abstracted measures “Unplanned Intubation” and “On Ventilator >48 Hours” of the ACS’ National Surgical Quality Improvement Program (and similar measures from the Society of Thoracic Surgeons’ Adult Cardiac Surgery Registry) (Risk-Adjusted Postoperative Prolonged Intubation (Ventilation) (NQF #0129). Steward: The Society of Thoracic Surgeons).

Adoption of this eCQM has the potential to improve the quality of care for surgical patients and, therefore, increase patient safety, which is a priority area identified by the National Quality Strategy (Rosen et al., 2013). Additionally, with a systematic EHR-based patient safety measure in place, hospitals can more reliably assess harm reduction efforts and modify their efforts in near real-time. In this way, greater achievements in reducing postoperative respiratory failure and enhancing hospital performance on patient safety outcomes can be expected.

Reference:

Rosen, A. K., Loveland, S., Shin, M., Schwartz, M., Hanchate, A., Chen, Q., Kaafarani, H. M., & Borzecki, A. (2013). Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Medical care*, 51(1), 37-44.

2.6 Meaningfulness to Target Population

A Technical Expert Panel (TEP) meeting to discuss the PRF measure specification was held in August of 2022 and a follow-up meeting was held in September 2023 to discuss testing results. TEP members consist of clinicians and other stakeholders, as well as three patient and caregiver representatives. At the meetings, we polled the group on measure importance. All patient/caregiver representatives agreed that the measure focuses attention on an outcome that holds the potential for substantial impact on the health status and health outcomes of individual patients as well as improving the health status of communities and populations. One respondent noted that this is an important measure that can contribute to organizational learning around the management of patients on mechanical ventilation.

3.1 Contributions Towards Closing Care Gaps

Disparities in the incidence of PRF across hospitals suggest that there is an important opportunity to reduce the occurrence of these events. One report from the Leapfrog group analyzed hospital discharge data from 15 states through the State Inpatient Databases processed by the AHRQ Healthcare Cost and Utilization Project and calculated the observed and adjusted rates for the claims-based PSI 11 measure (Postoperative Respiratory failure) across all hospitals, as well as by hospital Leapfrog safety grade (A through F) hospitals by race (Gangopadhyaya et al., 2023). The authors found non-Hispanic black and Hispanic patients had higher rates of PRF in comparison to white patients. All patients (White, Black, and Hispanic), on average, had lower rates of postoperative respiratory failure if they could access A-graded hospitals rather than C/D/F-graded hospitals. When adjusting for patient-level characteristics, Black patients experience rates of PRF that are 1.1 per 1,000 at-risk discharges higher (i.e., 17 percent higher relative to overall averages).

The Leapfrog study also analyzed disparities in PSI 11 rates between public and privately insured

patients. The study found differences in postoperative respiratory failure rates between Medicare-insured patients and privately insured patients, and rates increase as Hospital Safety Grade falls. Across all hospitals, the postoperative respiratory failure rate is 1.9 and 2.2 per 1,000 at-risk discharges higher for Medicare- and Medicaid-insured patients, respectively, relative to privately insured patients (Gangopadhaya et al., 2023).

A second study by Shen et al (2016) focused on racial and payer status disparities for PRF patients. The authors did not find differences among white, African American, and Hispanic racial groups, but did find that Medicaid patients were more likely to incur PRF than their privately insured counterparts (OR 1.24; CI 1.00, 1.53).

Another study by Burton and Gabriel (2023) focused on the primary endpoint of unintended endotracheal intubation or placement of other breathing device with ventilator support in patients unable to maintain airway patency within 30 days of carotid endarterectomy and found intubated patients were 2.2 times more likely to be Black/African American compared to White American (9.8% versus 4.4%, $p < 0.001$) (Burton and Gabriel., 2023). Based on the logistic regression analysis, the odds of short-term unanticipated intubation were increased by 77% for Black/African Americans compared to Whites (OR: 1.77, 95% CI: 1.11-2.68, $p=0.010$).

As part of our review of PSI data, we obtained information on payer status (up to two payers for each patient) to capture patients who are insured by Medicaid, dual eligible (Medicare and Medicaid), or uninsured. We plan to report findings from this analysis to address stakeholder concerns that this measure will help identify disparities associated with payer status. In general, there was not significant evidence of social disparities for PSI 11. Additionally, as part of the 2020 Comprehensive Reevaluation for PSI 90 (of which PSI 11 is a component) to maintain National Quality Forum (NQF) endorsement, AIR conducted an analysis of postoperative respiratory failure rate disparities (see **Table 13 in the logic model attachment**).

Using data from 12 hospitals we conducted a social disparities analysis and found:

- Hispanic patients have similar risk of PRF (OR=0.96; 95% CI, 0.42-2.20) as non-Hispanic patients, after adjusting for age and other factors in the risk-adjustment model.
- Black patients (OR=1.45; 95% CI, 0.77-2.75) and patients of "other" race (OR=0.92; 95% CI, 0.47-1.78) have similar risk of PRF as White patients, after adjusting for age and other factors in the risk-adjustment model.
- Risk of fall with injury is unrelated to Medicaid or uninsured status (OR=1.24; 95% CI, 0.72-2.12), or dual eligibility among Medicare beneficiaries, after adjusting for age and other factors in the risk-adjustment model
- Analyses of observed, expected, and risk-adjusted rates in all of the above patient cohorts confirm that the comorbidities and physiologic factors in the risk-adjustment model account

for some increased risk of PRF among Black patients (average expected rate 0.330% versus 0.296%), and that any residual bias is not statistically significant.

References:

1. Burton, B. N., & Gabriel, R. A. (2019). Racial disparities in postoperative respiratory failure after carotid endarterectomy. *Journal of clinical anesthesia*, 57, 139-140.
2. Gangopadhyaya, A., Pugazhendhi, A. Austin, M., Campione, A., and Danforth, M. (2023). Racial, Ethnic, and Payer Disparities in Adverse Safety Events: Are there Differences across Leapfrog Hospital Safety Grades? A report from the Leapfrog Group.
3. Shen, J. J., Cochran, C. R., Mazurenko, O., Moseley, C. B., Shan, G., Mukalian, R., & Neishi, S. (2016). Racial and Insurance Status Disparities in Patient Safety Indicators among Hospitalized Patients. *Ethnicity & disease*, 26(3), 443-452.

4.1 Feasibility Assessment

Thirteen hospitals across Cerner, Meditech, and Epic EHRs participated in the evaluation of feasibility. All hospital sites confirmed that the data elements used in the measure are captured within the EHR in a structured and codified manner either using nationally accepted terminology standards or local system codes that could be easily mapped. However, one Meditech hospital did not always use their structured fields to capture mechanical ventilation. For this reason, the site opted to not proceed with reliability and validity phases of testing.

While mechanical ventilation was captured in structured fields at all sites, documentation was not standardized. For example, some information was found in respiratory free text notes and start/end times were not discrete. Recognizing mechanical ventilation may have variability, we also evaluated intubation and extubation documentation for consideration in the measure specification. Though these two elements were more frequently captured, there are opportunities to expand electronic capture. For example, two of 13 sites documented rapid response interventions (including intubation) on paper and scanned into the EHR and others documented intubation/extubation in anesthesia free text notes for certain procedural areas (e.g., gastrointestinal lab, cardiac cath lab).

Please see **Table 2 in logic model attachment** for combined feasibility scores for data availability, data accuracy, data standards, and workflow across all 13 hospitals.

4.2 Attach Feasibility Scorecard

[PRF_COMBINED_Feasibility_Scorecard_EXTERNAL_11 01 2023.xlsx](#)

4.3 Feasibility Informed Final Measure

Due to variable documentation for mechanical ventilation (as described above), the measure also accommodates the use of intubation and extubation outside of a procedural area to trigger a postoperative respiratory event.

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 the use of this eCQM. Value sets are housed in the Value Set Authority Center (VSAC), which is provided by the National Library of Medicine (NLM), in coordination with the Office of the National Coordinator for Health Information Technology and the Centers for Medicare & Medicaid Services.

5.1.1 Data Used for Testing

We recruited 5 health systems consisting of 13 individual hospital sites. One hospital in the Southeast region only participated in alpha (feasibility) testing. We collected data for calendar year 2022 (January 1, 2022 - December 31, 2022) from 12 hospitals.

5.1.2 Differences in Data

Measure score level reliability testing used data from the full denominator population in Hospitals 1-12.

Measure data element level validity testing, on the other hand, were based on subsamples drawn from the measure initial population using the approach of random sampling without replacement. These subsamples served as the foundation upon which clinical abstractors compared data exported from the EHR (eData) to data manually abstracted from patients' medical charts (mData, or "gold standard"). This process is commonly known as the parallel-form comparison. When drawing the subsamples, we held constant the distribution of patient characteristics exhibited in the initial population to the extent possible (e.g., % of male, % of white, % of black, etc. in the abstraction sample are comparable to those in the initial population to the extent possible).

5.1.3 Characteristics of Measured Entities

Hospital test site characteristics are shown in **Table 3 in the logic model attachment**.

- **Vendor and location:** One hospital used Cerner as their EHR and another used Meditech as their EHR, both are headquartered in the Southeastern region of the United States. Eleven hospital used Epic as their EHR and are headquartered in various regions

(Southeast, Northeast, and West).

- **Bed size:** Four hospitals had between 100-199 beds, five hospitals had between 200-499 beds, and four hospitals had >499 beds.
- **Teaching Status:** Of the 13 hospitals, two were non-teaching hospitals, five were major teaching hospitals and six were community teaching hospitals. Teaching intensity is often measured by the ratio of interns and residents to beds. In this report, major teaching hospitals are those with an intern- and resident-to-bed ratio (IRB) of 0.25 (one resident for every four beds) or above and at least 50 beds, while community teaching hospitals include hospitals with an IRB of less than 0.25 or teaching hospitals with fewer than 50 beds).

5.1.4 Characteristics of Units of the Eligible Population

We collected data for calendar year 2022 (1/1/2022 and 12/31/2022) from 12 test sites. **Tables 4 and 5 in the logic model** attachment provide information on measure denominator population including age, sex, race, ethnicity, primary payer, and diagnoses. Measure denominator encounters ranged from a low of 73 to a high of 10,909 across test sites.

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 p. 12 for the methodology**.

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).

- Minimum: 0.152
- 25th percentile: 0.660
- Median: 0.732
- 75th percentile: 0.880
- Maximum: 0.964

5.2.4 Interpretation of Reliability Results

HH-PRF demonstrates high signal-to-noise reliability at most test facilities. ICC estimates ranged from 0.152 to 0.964 across test sites, with a mean and median equal to 0.71 and 0.73, respectively. ICCs at 10 of the 12 hospitals were at least 0.6 with 2 hospitals having lower values (0.152 and 0.441) due to very small numerators and denominators (i.e., 73 and 322 in the denominators, respectively). Decile analysis was not possible with only 12 facilities reporting complete data. Overall, testing results showed that HH-PRF, as currently specified, can distinguish the true performance in hospital postoperative respiratory failure rates from one hospital to another.

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	0.733	0.152											0.964
Mean Performance Score	12	1											1
N of Entities	30387	73											10909

5.3.1 Level(s) of Validity Testing Conducted

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

5.3.3 Method(s) of Validity Testing

To empirically assess data element validity, we compared data exported from the EHR (eData) to

data manually abstracted from patients' medical charts (mData) for a subsample of measure initial population. We then quantified the validity by calculating four statistics that tell us if the measure is subject to false positives and false negatives:

- **Positive Predictive Value (PPV)**—describes the probability that a patient who experienced the harm during hospitalization, per the EHR, is confirmed as a positive case per the clinical abstractor.
- **Sensitivity**— describes the probability that an encounter where the patient experienced the harm per the mData was correctly classified as having the same in the eData.
- **Negative Predictive Value (NPV)**—describes the probability that a patient who did not experience the harm per the eData is confirmed as a negative case with mData (either because the encounter is excluded from the denominator or numerator negative).
- **Specificity**— describes the probability that a patient who did not experience a harm per clinical abstraction was correctly classified as not experiencing the harm by the eData.

This process of data comparison is frequently known as the parallel-form comparison. As formulas are not allowed in the online form, **see logic model attachment p.13 for methodology.**

To assess measure score validity, we used face validity. Specifically, we reviewed the measure specification and results with members from our Hospital Harm Technical Expert Panel (TEP). We collected feedback on the precision of the measure specifications, importance of the measure outcome, and whether the performance scores can be used to distinguish good from poor hospital-level quality.

To evaluate the empirical impact of each exclusion criterion:

1. Using the full denominator data, we removed exclusion criteria one at a time from the measure logic and calculated the marginal and relative increase in the number of numerator and denominator encounters as a result.
2. Using the abstraction data, we compared each excluded sample case to the electronic information stored in the patient's medical record to assess whether the automated exclusion truly met the clinical criteria for exclusion.

5.3.4 Validity Testing Results

See **tables 6 and 7 in logic model** attachment for exclusion testing results. Exclusions occur for 0.05% - 4.04% of all initial population cases. Additionally, the exclusions affect both the number of denominator and numerator cases, affecting the measure score. In particular, exclusion 8 (neuromuscular disorder or degenerative neurological disorder) causes an 88.42% decrease in numerator cases, and exclusion 9 (high-risk airway procedures necessitating prophylactic endotracheal tube retention) causes a 27% decrease in numerator cases and a 4.6% decrease in denominator cases. This demonstrates that the exclusions occur frequently enough to justify their

use in the measure.

See **tables 8-12 in the logic model attachment** for PPV, sensitivity, NPV, and specificity values across sites

Face validity results are as follows:

- 15 of 15 TEP members (100%) voted “yes” that the measured outcome (rate of in-hospital postoperative respiratory failure) was important to measure and can improve care for patients.
- 15 of 15 TEP members (100%) voted “yes” that the measure specifications were precise and that it appears to measure what it is supposed to (i.e., face validity).
- 12 of 15 TEP members (80%) voted “yes” that the measure's performance scores provide an accurate reflection of hospital-level quality, and scores resulting from the measure Hospital Harm: Postoperative Respiratory Failure (PRF), as specified, can be used to distinguish good from poor hospital level quality related to hospital-acquired PRF. TEP members who voted “no” and other non-voting attendees felt it was premature to say “yes” without data from a more diverse group of hospitals (e.g., more non-teaching hospitals, other EHR vendors) in order to extrapolate results for generalizability. We explained the difficulty with recruiting hospital test sites and indicated that the teaching hospitals included in our pilot test sites include both major teaching and community teaching hospitals (sometimes called “minor teaching” because they do not sponsor multiple residency programs).

5.3.5 Interpretation of Validity Results

All exclusions occur frequently enough to justify their use in the measure. Exclusions are present for between 0.05% (Exclusions 2 and 3) and 4.04% (Exclusion 8) of patients. Given the threats to measure validity without these exclusions, these exclusions should be retained in the measure.

Testing results indicate strong concordance and inter-rater agreement between data exported from the EHR and data in the patient chart. For the measure numerator, PPV denotes the probability that an EHR-reported postoperative respiratory failure is valid based on the clinical review of patients’ medical records. Numerator PPV across all test sites was 89.6%. The primary reason for discordance was an isolated issue at one test site where respiratory therapy documented intubation erroneously. This issue can be addressed during implementation with improved documentation practices. For measure denominator exclusions, PPV denotes the probability that cases excluded from the measure per the EHR truly met the clinical rationale for exclusion. Denominator exclusion PPV across all test sites was 99.5%.

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)

5.4.1 Methods Used to Address Risk Factors

Statistical risk adjustment model with risk factors

5.4.2 Conceptual Model Rationale

As Canet and colleagues have described, “the pathogenesis of PRF depends on factors related to patient status as well as anaesthetic and surgical procedure” (Canet, 2015). The conceptual model shows how these patient risk factors, intraoperative factors, and perioperative management are thought to interact in contributing to the development of PRF.

Patient specific characteristics that increase the risk of PRF include age, gender, BMI, smoking status, comorbidities such as COPD, ASA class, preoperative vital signs (systolic blood pressure), laboratory values (arterial blood pH, pCO₂, sodium, hemoglobin, hematocrit), and complications present on admission. Some of these prior studies are cited in detail below.

Based on data from the American College of Surgeons National Surgical Quality Improvement Program on all elderly vascular and general surgery patients undergoing operations from 2005 to 2008 (Nafiu, 2011), univariate predictors of unplanned postoperative intubation (UPI) were older age, chronic obstructive pulmonary disease, low pre-operative functional status as well as emergency operation.

Svensson and colleagues (1991) analyzed data from June 1960 to September 1990 on 1414 patients who underwent repair of thoracoabdominal aortic aneurysms. The independent predictors of respiratory failure were chronic pulmonary disease, smoking history, cardiac and renal complications. In patients with chronic pulmonary disease, the only independent predictor was FEF₂₅ ($p = 0.030$).

In a cohort study of 44 VA medical centers (Arozullah, 2000), PRF developed in 2,746 patients (3.4%). The respiratory failure risk index was developed from a simplified logistic regression model and included abdominal aortic aneurysm repair, thoracic surgery, neurosurgery, upper abdominal surgery, peripheral vascular surgery, neck surgery, emergency surgery, albumin level less than 30 g/L, blood urea nitrogen level more than 30 mg/dL, dependent functional status, chronic obstructive pulmonary disease, and age.

Canet and colleagues (2015) reported a prospective observational study of a multicenter cohort and described a predictive score for PRF that includes seven independent risk factors: low preoperative SpO₂; at least one preoperative respiratory symptom; preoperative chronic liver

disease; history of congestive heart failure; open intrathoracic or upper abdominal surgery; surgical procedure lasting at least 2 h; and emergency surgery.

Ramachandran and colleagues (2011) analyzed data from 222,094 adult patients who underwent nonemergent, noncardiac surgery in the American College of Surgeons-National Surgical Quality Improvement Program database. Independent predictors of unanticipated early postoperative intubation included current ethanol use, current smoking, dyspnea, chronic obstructive pulmonary disease, diabetes mellitus needing insulin, active heart failure, hypertension requiring medication, abnormal liver function, cancer, prolonged hospitalization, recent weight loss, body mass index less than 18.5 or ≥ 40 kg/m, medium-risk surgery, high-risk surgery, very-high-risk surgery, and sepsis.

Johnson and colleagues (2007) analyzed data from 14 academic and 128 Veterans Affairs Medical Centers from October 2001 through September 2004, and developed a predictive model for PRF using logistic regression. Independent risk factors for PRF included Current Procedural Terminology group, American Society of Anesthesiologists classification, emergency operations, complex operations (work relative value units), preoperative sepsis, and elevated creatinine. Older patients, male patients, smokers, and those with a history of heart failure or COPD were also predisposed. The model's discrimination (c-statistic) was excellent, with no decrement from development (0.856) to validation (0.863) samples.

Burton and colleagues (2018) used data from the Nationwide Inpatient Sample from 2010 to 2014 to identify adult patients who underwent sinus surgery. In this population, the rate of PRF was 3.35% and independent risk factors included pneumonia, bleeding disorder, alcohol dependence, nutritional deficiency, heart failure, paranasal fungal infections, and chronic kidney disease.

Association with hospital and health system characteristics

Several studies have examined the association between postoperative respiratory failure and hospital or health system characteristics. In a multivariable analysis of Nationwide Inpatient Sample (NIS) data from the Healthcare Cost and Utilization Project (HCUP), Rahman and colleagues (2013) found that postoperative respiratory failure was less likely in patients admitted to nonteaching hospitals than those admitted to teaching hospitals (OR 0.89, 95% CI 0.85 to 0.93). The odds of developing postoperative respiratory failure increased by 6% for each level increase in hospital size from small to large (OR 1.06, 95% CI 1.03 to 1.09). Using data from 116 VA hospitals and NIS data from 992 community hospitals, Rivard and colleagues (2010) reported lower risk-adjusted rates of PRF in VA hospitals (3.86 per 1,000, 95% CI 2.83 to 4.88) than in the NIS (4.87 per 1,000, 95% CI 3.92 to 5.81).

Mediating Factors

Several care processes and intermediate factors (or mediators) may contribute to the occurrence of PRF. Some of these factors are within the hospital's/surgeon's control, while others may reflect patient's specific needs, and are therefore not considered as risk factors. These factors include procedure related risk factors such as surgical site, anesthesia type, fluid management, and duration of surgery (which reflects both the complexity of the operation and the skill of the surgical team).

Analyzing data on 50,367 patient admissions for common adult surgical procedures using an anesthesia information system between 2004 and 2009, Blum et al. (2013) identified intraoperative risk factors associated with respiratory failure among patients with similar preoperative risk: ventilator drive pressure (OR=1.17), fraction inspired oxygen (OR=1.02), erythrocyte transfusion (OR=5.36), and crystalloid administration in liters (OR=1.37). The number of different anesthetics administered during the admission was associated with higher risk of ARDS (OR=1.37). Fair evidence also suggests that short-acting neuromuscular blocking agents result in lower rates of residual neuromuscular blockade and may reduce risk for pulmonary complications (Kor, 2014).

In a multivariable analysis of the National Surgical Quality Improvement Program (NSQIP) database of adult inpatients who underwent neurosurgery under general anesthesia (2005-2010), Shalev and co-authors found that operative time exceeding 3 hours was associated with increased risk of reintubation (OR 2.9; 95%CI 1.8-4.8). In a retrospective time-matched cohort study, Attaallah (2019) found that operative-specific risk factors including ASA status, elective case type, and surgical duration were significantly associated with postoperative respiratory failure.

Lukannek and colleagues (2019) analyzed data from a registry of adult patients undergoing non-cardiac surgery between 2005 and 2017 at two independent healthcare networks. Intraoperative predictors of early postoperative tracheal re-intubation included early post-tracheal intubation desaturation; prolonged duration of surgery; high fraction of inspired oxygen; high vasopressor dose; blood transfusion; the absence of volatile anesthetic use; and the absence of lung-protective ventilation.

Social risk factors

Social factors or social determinants of health, SDOH, based on as SDOH CDC domains, have been studied for surgical patients by the American College of Surgeons (ACS) and others. For example, the Strong for Surgery initiative uses checklists to screen patients for risk factors that "can lead to surgical complications, and to provide appropriate interventions to ensure better surgical outcomes." Strong for Surgery targets several topics that have been shown to be associated with surgical outcomes such as nutrition, smoking, and glycemic control, and

encourages surgical teams to mitigate associated risks through preoperative interventions. The residual impact of these social factors is captured through measured patient characteristics such as smoking, ASA classification, weight loss, obesity, and laboratory test results such as serum albumin. Some social risk factors, such as social network characteristics, access to transportation, etc., are likely to have effects mediated through hospital choice. For these reasons, there is little conceptual rationale for adjusting for social risk factors in the risk-adjustment model for PRF.

References:

1. Ramachandran, S. K., Nafiu, O. O., Ghaferi, A., Tremper, K. K., Shanks, A., & Kheterpal, S. (2011). Independent predictors and outcomes of unanticipated early postoperative tracheal intubation after nonemergent, noncardiac surgery. *Anesthesiology*, *115*(1), 44-53. <https://doi.org/10.1097/ALN.0b013e31821cf6de>
2. Arozullah, A. M., Daley, J., Henderson, W. G., & Khuri, S. F. (2000). Multifactorial risk index for predicting postoperative respiratory failure in men after major noncardiac surgery. The National Veterans Administration Surgical Quality Improvement Program. *Annals of surgery*, *232*(2), 242-253. <https://doi.org/10.1097/00000658-200008000-00015>
3. Canet, Jaume; Sabaté, Sergi; Mazo, Valentín; Gallart, Lluís; de Abreu, Marcelo Gama; Belda, Javier; Langeron, Olivier; Hoefft, Andreas; Pelosi, Paolo For the PERISCOPE group. Development and validation of a score to predict postoperative respiratory failure in a multicentre European cohort: A prospective, observational study. *European Journal of Anaesthesiology* *32*(7):p 458-470, July 2015. | DOI: 10.1097/EJA.0000000000000223
4. Johnson, R. G., Arozullah, A. M., Neumayer, L., Henderson, W. G., Hosokawa, P., & Khuri, S. F. (2007). Multivariable predictors of postoperative respiratory failure after general and vascular surgery: results from the patient safety in surgery study. *Journal of the American College of Surgeons*, *204*(6), 1188-1198. <https://doi.org/10.1016/j.jamcollsurg.2007.02.070>
5. Lukannek, C., Shaefi, S., Platzbecker, K., Raub, D., Santer, P., Nabel, S., Lecamwasam, H. S., Houle, T. T., & Eikermann, M. (2019). The development and validation of the Score for the Prediction of Postoperative Respiratory Complications (SPORC-2) to predict the requirement for early postoperative tracheal re-intubation: a hospital registry study. *Anaesthesia*, *74*(9), 1165-1174. <https://doi.org/10.1111/anae.14742>
6. Amy Young, Satya Krishna Ramachandran; Clinical Prediction of Postoperative Respiratory Failure. *Anesthesiology* 2013; *118*:1247-1249
doi: <https://doi.org/10.1097/ALN.0b013e31829303c7>
7. Nafiu, O. O., Ramachandran, S. K., Ackwerh, R., Tremper, K. K., Campbell, D. A., Jr, & Stanley, J. C. (2011). Factors associated with and consequences of unplanned post-operative intubation in elderly vascular and general surgery patients. *European journal of anaesthesiology*, *28*(3), 220-224. <https://doi.org/10.1097/EJA.0b013e328342659c>
8. Svensson, L. G., Hess, K. R., Coselli, J. S., Safi, H. J., & Crawford, E. S. (1991). A prospective study of respiratory failure after high-risk surgery on the thoracoabdominal aorta. *Journal of vascular surgery*, *14*(3), 271-282.
9. Attaallah, A. F., Vallejo, M. C., Elzamzamy, O. M., Mueller, M. G., & Eller, W. S. (2019). Perioperative risk factors for postoperative respiratory failure. *Journal of perioperative practice*, *29*(3), 49-53. <https://doi.org/10.1177/1750458918788978>
10. Blum JM, Stentz MJ, Dechert R, et al. Preoperative and intraoperative predictors of

- postoperative acute respiratory distress syndrome in a general surgical population. *Anesthesiology*. 2013;118(1):19-29.
11. Brueckmann B, Villa-Urbe JL, Bateman BT, et al. Development and validation of a score for prediction of postoperative respiratory complications. *Anesthesiology*. 2013;118(6):1276-1285.
 12. Canet J, Sabaté S, Mazo V, et al. Development and validation of a score to predict postoperative respiratory failure in a multicentre European cohort: A prospective, observational study. *Eur J Anaesthesiol*. 2015;32(7):458-470.
 13. Gupta H, Gupta PK, Fang X, et al. Development and validation of a risk calculator predicting postoperative respiratory failure. *Chest*. 2011;140(5):1207-1215.
 14. Hua M, Brady JE, Li G. A scoring system to predict unplanned intubation in patients having undergone major surgical procedures. *Anesthesia and analgesia*. 2012;115(1):88-94.
 15. Johnson AP, Altmark RE, Weinstein MS, Pitt HA, Yeo CJ, Cowan SW. Predicting the Risk of Postoperative Respiratory Failure in Elective Abdominal and Vascular Operations Using the National Surgical Quality Improvement Program (NSQIP) Participant Use Data File. *Annals of surgery*. 2017;266(6).
 16. Kor DJ, Lingineni RK, Gajic O, et al. Predicting risk of postoperative lung injury in high-risk surgical patients: a multicenter cohort study. *Anesthesiology*. 2014;120(5):1168-1181.
 17. Kor DJ, Warner DO, Alsara A, et al. Derivation and diagnostic accuracy of the surgical lung injury prediction model. *Anesthesiology*. 2011;115(1):117-128.
 18. Canet J, Gallart L. Postoperative respiratory failure: Pathogenesis, prediction, and prevention. *Current Opinion in Critical Care*. 2014;20(1):56-62.
 19. Rahman M, Neal D, Fargen KM, Hoh BL. Establishing standard performance measures for adult brain tumor patients: a Nationwide Inpatient Sample database study. *Neuro Oncol*. 2013;15(11):1580-1588.
 20. Rivard, P. E., Elixhauser, A., Christiansen, C. L., Shibe Zhao, & Rosen, A. K. (2010). Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Medical care research and review : MCRR*, 67(3), 321-341. <https://doi.org/10.1177/1077558709347378>
 21. Rosen, A. K., Singer, S., Shibe Zhao, Shokeen, P., Meterko, M., & Gaba, D. (2010). Hospital safety climate and safety outcomes: is there a relationship in the VA?. *Medical care research and review : MCRR*, 67(5), 590-608. <https://doi.org/10.1177/1077558709356703>
 22. Strong for Surgery. American College of Surgeons. Available at: www.facs.org/quality-programs/strong-for-surgery. Accessed March 15, 2021
 23. Burton BN, Gilani S, Swisher MW, Urman RD, Schmidt UH, Gabriel RA. Factors Predictive of Postoperative Acute Respiratory Failure Following Inpatient Sinus Surgery. *Annals of Otolaryngology, Rhinology & Laryngology*. 2018;127(7):429-438. doi:10.1177/0003489418775129

5.4.2a Attach Conceptual Model

[Graphic for PRF RA Conceptual Model 11 01 2023.zip](#)

5.4.3 Variable Distribution Across Measured Entities

Tables 4 and 5 in the logic model attachment show substantial variation in the distribution of risk variables across the 12 measured entities. For example, mean age varied from 19.9 years at Site 7 to 70.4 years at Site 1. The percentage of Black patients varied from 5.4% at Site 7 to 40.6% at Site 2. The percentage of Hispanic patients varied from 0.3% at Site 1 to 32.9% at Site 7.

The percentage of Medicaid-enrolled patients varied from 5.1% at Site 2 to 45.2% at Site 7. Many diagnoses present on admission also demonstrated substantial variation across sites; for example:

- Deficiency anemias varied from 1.4% at Site 7 to 18% at Site 1
- Diabetes with chronic complications from 0.0% at Site 7 to 17.1% at Site 1
- Congestive heart failure varied from 1.3% at Site 5 to 21.7% at Site 1
- Peripheral vascular disease varied from 0.9% at Site 5 to 18.0% at Site 1

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

The final risk-adjustment model was estimated using multivariable probit regression to optimize calibration, after testing both logistic and Poisson 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 lab values, which were categorized and then dichotomized for analytic purposes, and age, which was tested in both piecewise linear and categorical forms.

Data sources included:

- ICD-10-CM diagnosis codes for comorbidities present on admission, including acquired immune deficiency syndrome (AIDS), alcohol abuse, deficiency anemia, autoimmune conditions, chronic blood loss anemia, leukemia, lymphoma, metastatic cancer, solid tumor without metastasis, cerebrovascular disease, coagulopathy, dementia, depression, diabetes (with and without chronic complications, drug abuse, congestive heart failure, hypertension (complicated and uncomplicated), liver disease (mild and moderate to severe), chronic pulmonary disease, neurological disorders, seizures and epilepsy, obesity, paralysis, peripheral vascular disease, psychoses, pulmonary circulation disease, renal failure (moderate and severe), hypothyroidism, other thyroid disorders, peptic ulcer with bleeding, valvular disease, and weight loss.
- Anesthesia, mechanical ventilation, intubation and extubation record for surgery;
- EHR lab values for white blood cells (leukocytes), albumin, bilirubin, BUN, creatinine, hematocrit, temperature, heart rate, pH arterial blood gas (ABG), partial pressure of oxygen in the arterial blood (PaO₂), and sodium.
- EHR demographic fields for age, sex, race, ethnicity, and primary payer.

After feature selection with 100-fold cross-validation and testing on the hold-out test set, the only retained risk factors were weight loss POA, deficiency anemias POA, heart failure POA, diabetes with chronic complications POA, moderate to severe liver disease POA, peripheral vascular disease POA, pulmonary circulation disease POA, valvular disease POA, ASA categories 3 through 5, and lab values for oxygen (partial pressure), leukocytes, albumin, BUN, bilirubin, and pH of arterial blood. We used APACHE II or APACHE III categorizations of these laboratory values, as appropriate, and aggregated categories to achieve the optimal separation of low-risk and high-risk

patients. In accord with APACHE categorization methods, missing values were assigned to the “normal” or reference category for each lab test. We tested models forcing in age (in both piecewise linear and categorical forms) and sex but found that these effects were neither statistically nor clinically significant. Including age and sex led to no meaningful improvement in any metric of model performance (e.g., AUC, Brier score, AIC/BIC).

Guided by the conceptual model, we developed the baseline risk adjustment model for PRF 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 and all laboratory tests, 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 PRF was linearly related to age from about 50 to 90 years of age.
3. Fit one model using the least absolute shrinkage and selection operator (LASSO) on the training set using 100-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 15 or 16 features, we did not use other penalized regression approaches (e.g., Elastic Net).
5. The final risk-adjustment model was a multivariable probit regression model, estimated on the entire dataset using the set of features selected by Lasso through 100-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.
2. H. Zou and T. Hastie, “Regularization and variable selection via the elastic net,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301-320, 2005.

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

[PRF RISK MODEL 11 01 2023.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 postoperative respiratory failure had a higher expected value than a randomly selected patient who did not experience that event. The AUC was 0.826 in the holdout test set (based on least absolute shrinkage and selection operator or LASSO regression) and 0.912 for the final probit 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 0.098 in the holdout test set (based on Lasso), indicating poor prediction at the individual patient level but good performance relative to a random classifier with AUPRC=0.0030.
- 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 78% of events occurred in the highest-risk decile, and nearly 88% occurred in the highest-risk quintile, the decile analysis is statistically unstable.
- A preferred approach in this situation is to estimate calibration belts suggested by Nattino et al. (2017). Calibration belts are an advance over the conventional Hosmer-Lemeshow plot, as the latter has the limitation of undue sensitivity to the choice of bins and extreme fluctuations in the observed-to-expected ratios in bins with few harm events. The null hypothesis of perfect calibration is barely rejected at the $p < 0.05$ level (i.e., $p = 0.049$), but the 95% confidence boundaries never cross the bisector.

References:

Nattino, G., Lemeshow, S., Phillips, G., Finazzi, S., & Bertolini, G. (2017). Assessing the calibration of dichotomous outcome models with the calibration belt. *The Stata Journal*, 17(4), 1003-1014

5.4.5a Attach Calibration and Discrimination Testing Results

[PRF CALIBRATION AND DISCRIMINATION TESTING 11 01 2023.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, Payment Program

6.2.1 Actions of Measured Entities to Improve Performance

Some cases of postoperative respiratory failure (PRF) are potentially preventable with optimal care (see clinical practice guidelines in **tables 14-21 of the logic model attachment**). Factors that might contribute include careful management of intra- and perioperative ventilator use and fluids, reducing surgical duration, using regional anesthesia, preventing wound infection, and optimizing pain control (Stocking et al, 2022; Encinosa et al, 2008; Zrelak , 2012). The proposed measure would enable hospitals to track and trend PRF rates to assess harm reduction efforts and modify their quality improvement efforts more reliably. The measure would also help to identify hospitals that have persistently high PRF rates. We collected feedback from 5 measured entities (hospital systems) on measure usability. All 4 measured entities (100%) agreed that the information produced by the performance measure is easy to understand and useful for decision making. Additionally, we polled 3 patients/family caregivers and all agreed that the measure outcome is important to know and can help improve care for patients.

References:

1. Stocking, J. C., Drake, C., Aldrich, J. M., Ong, M. K., Amin, A., Marmor, R. A., Godat, L., Cannesson, M., Gropper, M. A., Romano, P. S., Sandrock, C., Bime, C., Abraham, I., & Utter, G. H. (2022). Outcomes and risk factors for delayed-onset postoperative respiratory failure: a multi-center case-control study by the University of California Critical Care Research Collaborative (UC3RC). *BMC anesthesiology*, 22(1), 146.
2. Encinosa, W. E., & Hellinger, F. J. (2008). The impact of medical errors on ninety-day costs and outcomes: an examination of surgical patients. *Health services research*, 43(6), 2067-2085.
3. Zrelak, P. A., Utter, G. H., Sadeghi, B., Cuny, J., Baron, R., & Romano, P. S. (2012). Using the Agency for Healthcare Research and Quality patient safety indicators for targeting nursing quality improvement. *Journal of nursing care quality*, 27(2), 99-108.

Developer POC email

amichie@air.org

Measure Developer POC

Anna Michie
American Institutes for Research (AIR)
1400 Crystal Drive
10th Floor
Arlington, VA 22202

United States

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

Measured entity (reliability and validity) methodology and results (if available), Person or encounter-level (reliability and validity) methodology and results (if available)

Responder for Survey

Patient

Specify number of risk factors

16 (deficiency anemias, congestive heart failure, diabetes with chronic complications, moderate to severe liver disease, peripheral vascular disease, pulmonary circulation disease, valvular disease, weight loss, ASA category 3, ASA category 4 or 5, partial pressure oxygen, leukocyte count, albumin, BUN, bilirubin, and pH of arterial blood)

The measure developer is different from the measure steward

Yes

Steward Address

Donta Henson
CMS
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