NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Purple text represents the responses from measure developers.

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 0531

Corresponding Measures:

De.2. Measure Title: Patient Safety Indicator (PSI) 90: Patient Safety and Adverse Events Composite

Co.1.1. Measure Steward: Centers for Medicare and Medicaid Services

De.3. Brief Description of Measure: The PSI 90 composite measure summarizes patient safety across multiple indicators for the CMS Medicare fee-for-service population.

1b.1. Developer Rationale: Not applicable (composite measure)

S.4. Numerator Statement: PSI 03: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for pressure ulcer stage III or IV (or unstageable).

PSI 06: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for iatrogenic pneumothorax.

PSI 08: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for hip fracture.

PSI 09: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with: any secondary ICD-10-CM diagnosis codes for perioperative hemorrhage or hematoma and any-listed ICD-10-CM procedure codes for treatment of hemorrhage or hematoma (Note: The ICD-10-CM specification is limited to postoperative hemorrhage or hematoma).

PSI 10: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for acute renal failure and any-listed ICD-10-CM procedure codes for dialysis.

PSI 11: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with either: any secondary ICD-10-CM diagnosis code for acute respiratory failure; or any-listed ICD-10-CM procedure codes for a mechanical ventilation for 96 consecutive hours or more that occurs zero or more days after the first major operating room procedure code (based on days from admission to procedure); or any-listed ICD-

10-CM procedure codes for a mechanical ventilation for less than 96 consecutive hours (or undetermined) that occurs two or more days after the first major operating room procedure code (based on days from admission to procedure); or any-listed ICD-10-CM procedure codes for a reintubation that occurs one or more days after the first major operating room procedure code (based on days from admission to procedure).

PSI 12: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with a secondary ICD-10-CM diagnosis code for proximal deep vein thrombosis or a secondary ICD-10-CM diagnosis code for pulmonary embolism.

PSI 13: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for sepsis.

PSI 14: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any-listed ICD-10-PCS procedure codes for repair of the abdominal wall and any-listed ICD-10-CM diagnosis code for disruption of internal surgical wound

PSI 15: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for accidental puncture or laceration during a procedure and second abdominopelvic operation >=1 day after an index abdominopelvic operation.

S.6. Denominator Statement: PSI 03: Surgical or medical discharges, for patients ages 18 years and older. Surgical and medical discharges are defined by specific MS-DRG codes.

PSI 06: Surgical and medical discharges, for patients ages 18 years and older. Surgical and medical discharges are defined by specific MS-DRG codes.

PSI 08: Discharges, for patients ages 18 years and older, in a medical DRG or in a surgical DRG, with any listed ICD-10-PCS procedure codes for an operating room procedure.

PSI 09: Surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Surgical discharges are defined by specific MS-DRG codes.

PSI 10: Elective surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Elective surgical discharges are defined by specific MS-DRG codes with admission type recorded as elective.

PSI 11: Elective surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Elective surgical discharges are defined by specific MS-DRG codes with admission type recorded as elective.

PSI 12: Surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Surgical discharges are defined by specific MS-DRG codes.

PSI 13: Elective surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Elective surgical discharges are defined by specific MS-DRG codes with admission type recorded as elective.

PSI 14: Discharges, for patients ages 18 years and older, with any-listed ICD-10-CM procedure codes for abdominopelvic surgery, open approach, or with any-listed ICD-10-PCS procedure codes for abdominopelvic surgery, other than open approach.

PSI 15: Surgical and medical discharges, for patients ages 18 years and older, with any ICD-10-PCS procedure code for an abdominopelvic procedure

S.8. Denominator Exclusions: PSI 03:

- Length of stay of less than 3 days

- Principal ICD-10-CM diagnosis code for pressure ulcer stage III or IV (or unstageable)
- All secondary ICD-10-CM diagnosis codes for pressure ulcer III or IV (or unstageable) present on admission. If more than one diagnosis of pressure ulcer is present, all diagnoses must be present on admission for the discharge to be excluded
- Any listed ICD-10-CM diagnosis code for severe burns (>20% body surface area)
- Any listed ICD-10-CM diagnosis code for exfoliative disorders of the skin (>20% body surface area)
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 06:

- Principal ICD-10-CM diagnosis code for iatrogenic pneumothorax
- Any secondary ICD-10-CM diagnosis code for iatrogenic pneumothorax present on admission, among patients qualifying for the numerator
- Any listed ICD-10-CM diagnosis codes for specified chest trauma (rib fractures, traumatic pneumothorax and related chest wall injuries)
- Any listed ICD-10-CM diagnosis codes for pleural effusion
- Any listed ICD-10-PCS procedure codes for thoracic surgery
- Any listed ICD-10-CM procedure codes for cardiac procedure;
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 08:

- Principal ICD-10-CM diagnosis code for hip fracture
- Any secondary ICD-10-CM diagnosis code for hip fracture present on admission, among patients otherwise qualifying for the numerator
- Principal ICD-10-CM diagnosis code for seizure
- Principal ICD-10-CM diagnosis code for syncope
- Principal ICD-10-CM diagnosis code for stroke and occlusion of arteries
- Principal ICD-10-CM diagnosis code for coma
- Principal ICD-10-CM diagnosis code for cardiac arrest
- Principal ICD-10-CM diagnosis code for poisoning
- Principal ICD-10-CM diagnosis code for trauma
- Principal ICD-10-CM diagnosis code for delirium and other psychoses
- Principal ICD-10-CM diagnosis code for anoxic brain injury
- Any listed ICD-10-CM diagnosis codes for metastatic cancer
- Any listed ICD-10-CM diagnosis codes for lymphoid malignancy
- Any listed ICD-10-CM diagnosis codes for bone malignancy
- MDC 14 (pregnancy, childbirth, and puerperium)

- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 09:

- Principal ICD-10-CMS diagnosis code for perioperative hemorrhage or postoperative hematoma
- Any secondary ICD-10-CM diagnosis present on admission for perioperative hemorrhage or postoperative hematoma, among discharges that otherwise qualify for the numerator
- The only operating room procedure is for treatment of perioperative hemorrhage, or hematoma and with any secondary ICD-10-CM diagnosis codes for perioperative hemorrhage or hematoma
- Treatment of postoperative hemorrhage or hematoma occurs one day or more before the first operating room procedure, and with any secondary ICD-10-CM diagnosis codes for postoperative hemorrhage or hematoma
- With any listed ICD-10-CM diagnosis codes for coagulation disorders
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 10:

- Principal ICD-10-CM diagnosis code for acute renal failure, cardiac arrest, cardiac dysrhythmia, shock or chronic kidney failure
- Any secondary ICD-10-CM diagnosis code for acute kidney failure, cardiac arrest, cardiac dysrhythmia, shock or chronic kidney failure, present on admission, among patients otherwise qualifying for the numerator
- Any dialysis procedure that occurs before or on the same day as the first operating room procedure
- Any dialysis access procedure occurring before or on the same day as the first operating room procedure
- Principal ICD-10-CM (or secondary diagnosis present on admission) for urinary tract obstruction
- Any ICD-10-CM diagnosis code present on admission for solitary kidney disease and any ICD-10-PCS procedure code for partial nephrectomy
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 11:

- Principal ICD-10-CM diagnosis code for acute respiratory failure
- Any secondary ICD-10-CM diagnosis code for respiratory failure present on admission, among patients otherwise qualifying for the numerator
- Only operating room procedure is tracheostomy
- Procedure for tracheostomy occurs before the first operating room procedure
- Any listed ICD-10-CM diagnosis codes for neuromuscular disorder
- Any listed ICD-10-PCS procedure codes for laryngeal or pharyngeal, nose, mouth pharynx or facial surgery

- Any listed ICD-10-CM procedure codes for esophageal resection
- Any listed ICD-10-CM procedure codes for lung cancer
- Any listed ICD-10-CM diagnosis codes for degenerative neurological disorder
- Any listed ICD-10-CM procedure codes for lung transplant
- MDC 4 (diseases/disorders of respiratory system);
- MDC 5 (diseases/disorders of circulatory system);
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 12:

- Principal ICD-10-CM diagnosis code for proximal deep vein thrombosis (DVT) or pulmonary embolism (PE),
- Any secondary ICD-10-CM diagnosis code for DVT or PE present on admission, among patients otherwise qualifying for the numerator
- Procedure for interruption of vena cava occurs before or on the same day as the first operating room procedure
- Only operating room procedure was interruption of vena cava
- Any listed ICD-10-CM diagnosis code for acute brain or spinal injury present on admission
- Any listed ICD-10-PCS procedure code for extracorporeal membrane oxygenation (ECMO)
- Procedure for pulmonary arterial thrombectomy occurs before or on the same day as the first operating room procedure
- Only operating room procedure was for pulmonary arterial thrombectomy
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 13:

- Principal ICD-10-CM diagnosis code for sepsis or infection
- Any secondary ICD-10-CM diagnosis code for sepsis or infection present on admission, among patients otherwise qualifying for the numerator
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 14:

- Procedure for abdominal wall reclosure occurs on or before the day of the first open abdominopelvic surgery procedure, if any, and the day of the first laparoscopic abdominopelvic surgery procedure, if any
- Any listed ICD-10-CM diagnosis codes or any-listed ICD-10-PCS procedure codes for immunocompromised state

- Principal ICD-10-CM diagnosis code for disruption of internal operation wound
- Any secondary ICD-10-CM diagnosis code for disruption of internal operation wound present on admission
- Length of stay less than two (2) days-MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 15:

- Principal ICD-10-CM diagnosis code for accidental puncture or lacerations during a procedure
- Any secondary ICD-10-CM diagnosis code for accidental puncture or laceration during a procedure, among patients otherwise qualifying for the numerator
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

De.1. Measure Type: Composite

S.17. Data Source: Claims

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Jun 19, 2009 **Most Recent Endorsement Date:** Dec 10, 2015

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not applicable.

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

- 1a. Evidence

1a. Evidence. The evidence requirements for a health outcome measure include providing empirical data that demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service; if these data not available, data demonstrating wide variation in performance, assuming the data are from a robust number of providers and results are not subject to systematic bias. For measures derived from patient report, evidence also should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

Evidence Summary

- The developer provided detailed literature reviews of the evidence for each the outcome measures the 531's components measures with information showing that one or more healthcare actions can be performed to reduce the incidence of each measure.

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- Additional detailed literature review is provided, updated from the prior one to include new literature.

Question for the Committee:

- *Is there at least one thing that the provider can do to achieve a change in the measure results for each of the component measure?*

Questions for the Committee:

- *Does the Committee agree the evidence basis for the measure has not changed and there is no need for repeat discussion and vote on Evidence?*

Guidance from the Evidence Algorithm

Box 1 -> Box 2 -> PASS

Preliminary rating for evidence: ☒ Pass ☐ No Pass

- 1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

= Medicare Fee-for-Service Reference Population Rate and Distribution of Hospital Performance on PSI 90 (Patient Safety Composite)

Year	N	Mean	SD	min	p10	p25	Med	p75	p90	max
2016-2017	3212	0.995	0.174	0.567	0.842	0.906	0.970	1.036	1.181	5.326
2017-2018	3212	0.994	0.166	0.530	0.845	0.907	0.970	1.029	1.178	3.791
2018-2019	3212	0.996	0.161	0.629	0.849	0.913	0.971	1.032	1.174	2.588

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

Distribution of Hospital Performance on PSI 90 Specific Indicators

	2017		2018		2019	
Component	N	Mean	N	Mean	N	Mean
PSI 03	3207	0.593	3208	0.618	3208	0.580
PSI 06	3211	0.257	3211	0.243	3210	0.221
PSI 08	3209	0.108	3209	0.106	3207	0.100
PSI 09	3086	2.408	3083	2.532	3075	2.481

PSI 10	2970	1.323	2981	1.368	2960	1.337
PSI 11	2961	6.801	2973	5.755	2950	4.876
PSI 12	3086	3.784	3083	3.670	3077	3.572
PSI 13	2964	4.702	2966	4.743	2943	4.491
PSI 14	3036	0.977	3035	0.865	3021	0.879
PSI 15	3112	1.273	3104	1.226	3086	1.180

Disparities

- Disparities exist for the PSI 90 component measures, but the developer reports no consistent pattern across these components. This finding is not surprising as the PSI 90 component measures focus on hospital-acquired complications of care.
- For example, men have at least 20% higher observed rates than women for PSI 03 (Pressure UIcer), PSI 09 (Postoperative Hemorrhage or Hematoma), PSI 10 (Postoperative Acute Kidney Injury Requiring Dialysis), PSI 11 (Postoperative Respiratory Failure), PSI 13 (Postoperative Sepsis), and PSI 14 (Postoperative Wound Dehiscence). Men have at least 20% lower observed rates than women for PSI 06 (Iatrogenic Pneumothorax), PSI 08 (In-hospital Fall with Hip Fracture), and PSI 15 (Unrecognized Accidental Puncture or Laceration).
- All of the PSI risk-adjustment models include sex, age groups, and sex-age interactions. Therefore, the observed disparities across age and sex categories greatly diminish or disappear after risk-adjustment, as intended.
- Across racial-ethnic categories, the Medicare FFS data show at least 25% higher adjusted rates among Black patients, relative to White patients, for only three PSIs: PSI 03 (Pressure Injury), PSI 12 (Perioperative Deep Vein Thrombosis or Pulmonary Embolism), and PSI 15 (Unrecognized Accidental Puncture or Laceration).
- Comparing Hispanic patients with White patients, Hispanics had at least 20% higher adjusted PSI rates only for PSI 14 (Postoperative Wound Dehiscence) and PSI 15 (Unrecognized Accidental Puncture or Laceration). For all other PSI 90 component measures, Black and Hispanic patients had lower or similar adjusted rates, when compared with White patients.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?
- Is the committee satisfied with the developer's description of the disparities data?

Preliminary rating for opportunity for improvement: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

RATIONALE:

1c. Composite – Quality Construct and Rationale

Maintenance measures – same emphasis on quality construct and rationale as for new measures.

1c. Composite Quality Construct and Rationale. The quality construct and rationale should be explicitly articulated and logical; a description of how the aggregation and weighting of the components is consistent with the quality construct and rationale also should be explicitly articulated and logical.

- CMS Patient Safety Indicator (PSI) 90, also known as the Patient Safety and Adverse Events Composite, combines information from 10 common patient safety events that may occur in hospitalized patients.
- It was created to provide a simple and transparent single metric that can be used to better understand, communicate, and track patient safety in U.S. hospitals. The underlying concept, as described by the Institute of Medicine, is that safety is "freedom from accidental injury"(1) and that safe care "involves making evidence-based clinical decisions to maximize the health outcomes of an individual and to minimize the potential for harm"(2).
- This concept is closely linked to CMS's priority to implement quality initiative assuring quality healthcare for Medicare Beneficiaries using tools to achieve effective, safe, efficient, patient-centered, equitable, and timely care.
- Section 3008 of the Affordable Care Act established the Hospital-Acquired Condition (HAC) Reduction Program to encourage hospitals to reduce HACs. Beginning with Fiscal Year (FY) 2015 discharges (i.e. October 1, 2014), the HAC Reduction Program requires the Secretary of Health and Human Services (HHS) to adjust payments. As set forth in the Affordable Care Act, CMS may reduce payments for the worst-performing 25 percent of hospitals by up to one percent, and publicly report hospitals' measure scores, domain scores, and HAC Reduction Program data.

- CMS PSI 90 combines the standardized morbidity ratios (observed/expected ratio) from 10 component indicators: PSI 03 Pressure Ulcer, PSI 06 Iatrogenic Pneumothorax, PSI 08 In-Hospital Fall with Hip Fracture, PSI 09 Postoperative Hemorrhage or Hematoma, PSI 10 Postoperative Acute Kidney Injury Requiring Dialysis, PSI 11 Postoperative Respiratory Failure, PSI 12 Perioperative Pulmonary Embolism or Deep Vein Thrombosis, PSI 13 Postoperative Sepsis Rate, PSI 14 Postoperative Wound Dehiscence, and PSI 15 Unrecognized Abdominopelvic Accidental Puncture or Laceration Rate.

- From a conceptual perspective, the CMS PSI 90 composite should reflect the likelihood of harm associated with a wide range of potentially preventable adverse events.
- Within this conceptualization, each PSI is an individual predictor of an important and relevant aspect of harm. Thus, the likelihood of harm is expressed as the probability of a potentially preventable adverse event (such as postoperative sepsis) times the average net severity of harm associated with that event.
- In this conceptualization, CMS PSI 90 is modeled as a heterogeneous, formative index, meaning that the composite is "formed from" a set of measured components, each of which reflects different but overlapping aspects of care.
- The use of administrative data provides an inexpensive and fairly comprehensive approach to measuring a wide range of elements of the construct of harm while each component indicator contributes a unique aspect of harm.

- The final weight for each component measure is the product of harm weights and volume weights (numerator weights). Harm weights are calculated by multiplying empirical estimates of probability of excess harms associated with the patient safety event by the corresponding utility weights linked to each of the harms (1-disutility).

- Disutility is the measure of the severity of the adverse events associated with each of the harms (i.e.., outcome severity, or at least preferred states from the patient perspective). The harm weights are calculated using linked claims data for two years of Medicare Fee for Service beneficiaries.
- Volume weights, the second part of the final weight, are calculated based on the number of safety-related events for the component indicators in the fee-for-service reference population.

Questions for the Committee:

- Are the quality construct and a rationale for the composite explicitly stated and logical?
- Is the method for aggregation and weighting of the components explicitly stated and logical?

Preliminary rating for composite quality construct and rationale:
☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure.

- Pass
- Agree that the evidence basis has not changed
- Strong evidence from literature to support measure focus, and good review on relationships between these events and poor patient outcomes.
- Evidence provided for each of the outcomes measures
- Unaware of any new evidence. Outcome measures seem important, although limited in scope. The evidence does not seem to require additional discussion.
- Evidence exists and is applicable to the measures. Yes, at least one thing can be done to improve in each area.
- Agree that evidence is Pass. Abundant evidence presented for each sub component that supports relationship between the measured outcome and at least one healthcare action.

1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?

- High rating, no concerns.
- Yes, and there is a continued performance gap

- Yes, meaningful differences in hospitals were observed. Disparities data by race presented to suggest substantial differences in quality for Black and Hispanic population (worse outcomes) for many measures.
- High gap
- Performance gap seems sufficiently wide to warrant the composite measure. Some disparities were apparent in specific components.
- Gaps have been demonstrated to exist and satisfactory disparities data was supplied
- Presented data support presence of disparities in sub measures and significant variation in both submeasures and in the composite.

1c. Composite Performance Measure - Quality Construct (if applicable): Are the following stated and logical: overall quality construct, component performance measures, and their relationships; rationale and distinctive and additive value; and aggregation and weighting rules? Weighting rules appear explicitly stated.

- Yes
- Yes, the rationale is clear for framing as an overall patient safety metric.

- Agree with moderate prelim rating
- Leave to the statisticians. I have concerns that some important measures may be overlooked because of exclusions such as syncope, delirium, and iatrogenic stroke.
- Quality construct is acceptable
- The components are weighted by volume and by importance. The latter is assigned by empirically estimating excess harm and then assigning utility based on clinician input and literature to each harm. It is a complicated measure but this approach is justified. In terms of selection of the 10 subcomponents, possibly because this is a maintenance review there's less information (i.e. why these 10 and not others like iatrogenic infections, possibly because other measures already capture those?). It's not that I don't think these 10 outcomes are reasonable, just would have liked information on how they were selected in the first place. No other concerns.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

2c. For composite measures: empirical analysis support composite approach

- Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period

and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

- Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

Complex measure evaluated by Scientific Methods Panel? ☒ **Yes** ☐ **No**

SMP Ratings:

- **Reliability:** H-2; M-5; L-0; I-1
- **Validity:** H-2; M-4; L-1; I-1
- **Composite:** H-2; M-4; L-1; I-1

Evaluators: Scientific Methods Panel

Methods Panel Review (Combined)

Methods Panel Evaluation Summary:

This measure was reviewed by the Scientific Methods Panel and discussed on the call. A summary of the measure and the Panel discussion is provided below.

Specifications:

Appendix A. Several changes have been made to the definitions of each of the component measures over time to improve them.

Reliability

- Component reliability was reported using signal-to-noise ratios:
- Weighted Mean Signal-to-Noise Reliability for PSI 90 Component Measures across Hospitals in Medicare Fee-for-Service Data, 2016-2018 (Comparing Current CMS v10.0 with Previous CMS v9.0)

PSIs	CMS v10.0 signal-to-noise reliability (N)	CMS v10.0 signal-to-noise reliability (Weighted mean)	CMS v9.0 signal-to-noise reliability (Weighted mean)
PSI 03	3,294	0.777	0.784
PSI 06	3,305	0.400	0.388
PSI 08	3,303	0.152	0.208
PSI 09	3,130	0.469	0.485

PSIs	CMS v10.0 signal-to-noise reliability (N)	CMS v10.0 signal-to-noise reliability (Weighted mean)	CMS v9.0 signal-to-noise reliability (Weighted mean)
PSI 10	3,008	0.489	0.509
PSI 11	2,998	0.652	0.654
PSI 12	3,131	0.610	0.609
PSI 13	2,994	0.554	0.567
PSI 14	3,060	0.167	0.261
PSI 15	3,152	0.443	0.443

- Split-Sample Reliability Testing was Conducted to Assess the Composite as well as Test-Retest Reliability
- Split Sample PSI 90 Reliability at Hospital Level in Medicare Fee-for-Service Data, 2016-2018

Reliability Assessment	24 months of data	36 months of data
Hospitals meeting 3 case minimum	3,305	3,305
Median Intracluster Correlation Coefficient (ICC)	0.74	0.81
% Hospitals meeting ICC≥0.6	67%	76%
% Hospitals meeting ICC≥0.4	83%	89%

- Test-Retest PSI 90 Consistency at Hospital Level in Medicare Fee-for-Service Data, 2016-2019

Reliability Assessment	24 months of data	36 months of data
Hospitals meeting 3 case minimum	3,305	3,305
Median Intracluster Correlation Coefficient (ICC)	0.61	0.70
% Hospitals meeting ICC≥0.6	51%	62%
% Hospitals meeting ICC≥0.4	72%	81%

Validity

- Validity testing was conducted at three levels: face, component and composite-level using convergent validity
    - Component Validity:
    - Predictive Validity of PSI 90 Components at the Patient Level, Showing the Average Marginal Effect of Each PSI Event on Subsequent Adverse Outcomes (after Adjusting for Confounding Factors through Inverse Probability Propensity Weighting)

Adverse Outcome (absolute diff in days or %)	PSI 03	PSI 06	PSI 08	PSI 09	PSI 10	PSI 11	PSI 12	PSI 13	PSI 14	PSI 15
Hospital Length of Stay (days)	9.3	4.6	4.5	5.1	11.4	7.1	8.0	12.0	12.2	14.2
30-day readmission	5.0%	0.0%	9.1%	4.7%	6.3%	5.2%	5.8%	4.8%	8.4%	7.3%
Death (30*/180 days)	27.0 %	13.0 %*	7.3%*	4.5%*	32.7%	18.6%	13.4%	28.6%	10.8%	10.1% *
Long-term SNF admission	9.3%	0.0%	25.3%	3.1%	2.9%	6.3%	5.0%	6.5%	10.2%	5.4%
SNF Length of Stay (days)	8.8	0	18.6	2.5	1.8	4.7	3.9	4.9	8.2	4.8
Late complication**	7.4%	N/A	N/A	10.7%	N/A	N/A	N/A	N/A	1.7%	9.4%
Late operation***	4.9%	N/A	0.12%	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Late incisional hernia	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1.4%	N/A
Tracheostomy	N/A	N/A	N/A	N/A	N/A	14.0%	N/A	N/A	N/A	N/A
DVT/PE/bleed (ED visit)	N/A	N/A	N/A	N/A	N/A	N/A	47.6%	N/A	N/A	N/A
Long-term dialysis	N/A	N/A	N/A	N/A	4.3%	N/A	N/A	N/A	N/A	N/A

- Convergent Validity Between PSI 90 and Infection-Related Outcome Measures by Spearman Rank Correlation, Using Different Data Periods

Hospital Compare Measures	Hospitals	PSI 90 – CMS v10, 2016-2018	PSI 90 – CMS v10, 2017-2019
Hip/knee complication rate	2,387	0.149***	0.136***
Central line-associated bloodstream infection (CLABSI)	2,273	0.042*	0.040
Catheter-associated urinary tract infection	2,536	0.047*	0.060**
Clostridium difficile (C. diff) infection	2,946	0.054**	0.060**
Surgical-site infection (SSI) following abdominal hysterectomy/colon procedure	2,425	0.108***	0.104***
Methicillin-resistant Staphylococcus aureus (MRSA) bacteremia	2,131	0.058*	0.072**
Total healthcare-acquired condition (HAC) score	3,188	0.420***	0.347***

- Convergent Validity Between PSI 90 and 30-day Readmission Measures by Spearman Rank Correlation, Using Different Data Periods

Hospital Compare Measures	Hospitals	PSI 90 – CMS v10, 2016-2018	PSI 90 – CMS v10, 2017-2019
30-day readmission: Acute Myocardial Infarction (AMI)	2,061	0.024	0.016
30-day readmission: Coronary Artery Bypass Graft (CABG)	994	0.058	0.054
30-day readmission: Chronic Obstructive Pulmonary Disease (COPD)	2,836	0.037*	0.043*
30-day readmission: Heart Failure	2,856	0.059**	0.045*
30-day readmission: Hip and Knee	2,460	0.084***	0.096***
30-day readmission: Pneumonia	2,927	0.065***	0.070***
30-day readmission: Hospital-wide	3,140	0.138***	0.145***

- Convergent Validity Between PSI 90 and Leapfrog Survey Safe Practice Scores, Based on Mean PSI 90 Values by Category of Response

Safe Practice Score: Overall Performance	Hospitals (%)	Mean Score (SD)	Relative risk (compared with "fully meets standard")
Fully meets standard	1,493 (49.2%)	1.002 (0.215)	1
Substantial progress	183 (6.0%)	0.983 (0.179)	0.981
Some progress	45 (1.5%)	1.009 (0.193)	1.007
Willing to report	38 (1.3%)	1.033 (0.192)	1.031
Declined to respond	1,274 (42.0%)	0.986 (0.174)	0.984

- Convergent Validity Between PSI 90 and Leapfrog Survey Safe Practice Scores by Spearman Rank Correlation, Excluding Hospitals that Declined to Respond

Performance on Safe Practice Measures	Hospitals	Mean Score (SD)	PSI 90 – CMS v10, 2017-2019
Culture of safety leadership structures and systems (out of 120 points)	1,759	116.92 (8.46)	0.034
Culture measurement, feedback and interventions (out of 120 points)	1,759	116.47 (12.95)	-0.020
Risks and hazards (out of 100 points)	1,759	97.25 (9.63)	-0.017
Nursing workforce (out of 100 points)	1,759	97.60 (9.09)	-0.021
Hand hygiene (out of 60 points)	1,759	57.22 (7.93)	-0.017

- Known Groups Validity for PSI 90

Known Groups Category	Hospitals (%)	Mean	SD
Hospital Teaching	--	--	--

Known Groups Category	Hospitals (%)	Mean	SD
Resident FTE/bed ratio = 0	2,564 (79.3%)	0.978	0.168
Resident FTE/bed ratio (0 - 0.25)	471 (14.6%)	1.013	0.194
Resident FTE/bed ratio ≥0.25	196 (6.1%)	1.125	0.372
Hospital Nursing*	--	--	--
Nurse FTE/bed ratio <1.0	950 (29.4%)	0.985	0.154
Nurse FTE/bed ratio (1.0 – 2.0)	1,700 (52.6%)	0.995	0.188
Nurse FTE/bed ratio >2.0	581 (18.0%)	0.992	0.258
Nursing Skill Mix**	--	--	--
Low (<0.85)	765 (23.7%)	0.995	0.175
Medium (0.85-0.975)	1,359 (42.0%)	1.002	0.218
High (>0.975)	1,107 (34.3%)	0.976	0.172

- - cell intentionally left blank

- Face validity
- On July 20, 2020, TEP members voted 12-1 in favor of continued use of PSI 90, subject to reassessment as additional validation data and measures become available.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The Scientific Methods Panel is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The Scientific Methods Panel is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

Questions for the Committee regarding composite construction:

- Do you have any concerns regarding the composite construction approach (e.g., do the component measures fit the quality construct and add value to the overall composite? Are the aggregation and weighting rules consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible?)?
- The Scientific Methods Panel is satisfied with the composite construction. Does the Committee think there is a need to discuss and/or vote on the composite construction approach?

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Preliminary rating for composite construction: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?

- Appropriate.
- No concerns
- A lot of concern about low volume hospitals as the reliability appears very poor in those cases and it is unclear how many would be excluded. Lots of variation between components which may affect some facilities differently.
- Agree with prelim moderate rating
- none
- Specifications are adequate, no discussion or vote needed
- I do not have concerns.

2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?

- No concerns
- No concerns
- For low volume hospitals
- Agree with prelim moderate rating
- no
- No
- Seems reliability for the composite measure is acceptable but much more variable for individual sub measures.

2b1. Validity -Testing: Do you have any concerns with the testing results?

- No concerns
- No concerns
- No, face and construct validity are strong
- Agree with prelim moderate rating
- no
- No
- No concerns.

2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment) 2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately

excluded from the measure? 2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?

- No
- None
- Risk adjustment was sound. Social risk factor analysis showed some substantial variation by race, and it wasn't clear if this was structural in terms of higher complexity for minority populations not captured by the risk adjustment or poorer quality of care within a hospital. The latter would be concerning, and would seem that social factors should have been more deeply explored because of the direct links between hands-on patient care, education, and some of these patient outcomes.
- Agree with prelim moderate rating
- Unable to evaluate ' Inverse Probability Propensity Weighting.'
- No comment
- Risk adjustment is appropriate. Exclusion is appropriate.
- 2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data) 2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?
- No concerns
- No concerns
- No
- Agree with prelim moderate rating
- Seems adequate, although one wonders about the value of the composite measure in light of more specific details for each component measure.
- No
- No concerns.
- 2c. Composite Performance Measure - Composite Analysis (if applicable): Do analyses demonstrate the component measures fit the quality construct and add value? Do analyses demonstrate the aggregation and weighting rules fit the quality construct and rationale?
- Adequate- review of older documents better explicate aggregation.
- Yes
- Weighting was really unclear, and how that might impact some subsets of hospitals is unclear
- Agree with prelim moderate rating
- Component measures seem to fit; aggregated seems difficult to justify.
- Yes

- No concerns but would have liked more information on how the excess harm and disutility scores are estimated.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- All data elements are in defined fields in electronic claims

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Are the required data elements available in electronic form, e.g., EHR or other electronic sources?
- Is the data collection strategy ready to be put into operational use?

Preliminary rating for feasibility: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?

- No concerns
- No concerns
- Very feasible
- High feasibility
- none
- Measure is feasible
- No concerns.

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

- 4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after

initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No** ☐ **UNCLEAR**

Accountability program details

Public Reporting

CMS Medicare Hospital Compare Program

https://www.medicare.gov/HospitalCompare/Data/Serious-Complications.html

CMS Medicare Hospital Compare Program

https://www.medicare.gov/HospitalCompare/Data/Serious-Complications.html

Payment Program

https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/HAC-Reduction-Program

CMS Hospital-Acquired Condition (HAC) Reduction Program (HACRP)

CMS Hospital Value-Based Purchasing Program (HVBP)

https://www.qualitynet.org/inpatient/hvbp/measures

Regulatory and Accreditation Programs

Statewide Quality Advisory Committee (Massachusetts)

http://chiamass.gov/sqms/

4a.2. Feedback on the measure by those being measured or others.

- Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- Feedback is obtained from users through a variety of channels, particularly through the technical assistance mechanisms that are used to support users.
- In addition, CMS incorporates input on PSI implementation from technical expert panels convened to support PSI development and maintenance, stakeholder committees such as the NQF standing committees, and peer-reviewed or other research publications.

Additional Feedback:

- The CMS PSIs are updated annually, including updating indicator technical specifications in accordance with the latest coding guidance; suggestions from users and other stakeholders obtained through Technical Assistance, committees, or workgroups; and the latest clinical and scientific research.
- CMS regularly reviews these sources, identifies possible indicator updates, and prioritizes updates for each indicator and software update based on expected impact on users.

***Questions for the Committee**:*

- How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?
- How has the measure been vetted in real-world settings by those being measured or others?

Preliminary rating for Use: ☒ Pass ☐ No Pass

- 4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- Over this three-year period from July 2016 through June 2018, based on national Medicare fee-for-service claims data as described in the Testing attachment, PSI 90 has shown minimal change in mean and median values.
- However, the 75th, 90th, 95th, and 100th percentile values have decreased, suggesting that the hospitals with the highest PSI event rates have been able to reduce their rates.
- Data from before October 1, 2015 cannot be compared with later data due to the code set conversion from ICD-9-CM to ICD-10-CM/PCS.
- However, these results for PSI 90 do not tell the full story, because each component indicator is separately risk-adjusted and reliability-adjusted at the hospital level before it is put into PSI 90. The observed rates of the component indicators are also shown in 1b above and eight of the ten components demonstrate consistent improvement over time between the 7/1/2016-6/30/2017 year and the 7/1/2018-6/30/2019 year.
- Specifically, overall national observed rates of PSI 03, 06, 08, 09, 10, 11, 12, 13, 14, and 15 have decreased by 2.2%, 14.0%, 7.6%, -3.0%, -1.0%, 28.3%, 5.6%, 4.5%, 10.0%, and 7.3%, respectively.
- For all components except PSI 14, the overall national observed rate in 2018-19 was lower than the corresponding rate in 2017-18.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

1) Several large hospitals, such as New York University Langone Medical Center and the University of Washington Medical Center have established "prebilling review processes" with "prompt review of documentation and coding to confirm accuracy [of potential PSI diagnoses] and to identify opportunities to improve care quality and safety."

The AHRQ QI Toolkit offers specific guidance to hospitals and quality improvement leaders about "how to establish an effective coding communication and review process." The implication of these efforts is that some of the observed decrease in the incidence of this event over the last decade may be due to more accurate clinical documentation and coding, rather than to true improvements in patient outcomes and quality of care. Therefore, users should be cautious about interpreting recently observed changes in the

incidence of component events. There is no evidence that more accurate clinical documentation and coding have had any negative consequences for individuals or populations. Any harm from increasing providers' attention to documentation is likely to be counterbalanced by the benefits of more accurate data and more careful reflection on adverse events. In addition, these efforts appear to lead to "one-time corrections" in PSI rates, as hospitals implement processes to prevent overreporting, but do not affect the prior or subsequent trend lines. For example, both the University of Washington Medical Center and Cedars Sinai Medical Center (CSMC) reported that concurrent review of clinical documentation was only the first step toward improving PSI performance. CSMC noted that "task forces that include staff from many different departments and disciplines are assigned to carry out a "leave-no-stone-unturned" search for opportunities to prevent harm across the board… all ideas are important..."

Some users have raised a specific concern about unintended consequences of PSI 12, Perioperative Pulmonary Embolism or Deep Vein Thrombosis Rate. Specifically, the concern is that higher rates are a result of "increased vigilance in detection" at some hospitals. To address these concerns, CMS has made two important changes to PSI 12 to make it less sensitive to overdiagnosis bias: (1) PSI 12 now captures only proximal (groin/thigh), not distal (calf) vein thromboses; and (2) PSI 12 no longer captures solitary subsegmental pulmonary emboli. With these changes, CMS is now seeing a decreasing temporal trend in PSI 12 rates (down 10.2% from 7/1/2016-6/30/2017 to 7/1/2017-6/30/2018) and no change in case fatality over time. These results provide reassurance that the current specification of PSI 12 is not sensitive to overdiagnosis bias, because it focuses on clinically important events that are consistently diagnosed and treated across all hospitals.

Potential harms None

Additional Feedback: None

***Questions for the Committee**:*

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- Do the benefits of the measure outweigh any potential unintended consequences?

Preliminary rating for Usability and use: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

- **Committee Pre-evaluation Comments:**
  Criteria 4: Usability and Use

- **4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided? 4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?**

- High feasibility and transparency. Pass
- Drive patients to safer care delivery systems
- Public reporting of data is ongoing, and currently being used in an accountability program.
- Agree with pass prelim rating

- ok
- This measure makes an easy snapshot of overall care but identifying salient issues requires disaggregating the measures.
- Currently being used and publicly reported on hospitals. Rating Pass.

- **4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations? 4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.**

- The benefits of the measure outweigh risks.
- Appreciate the CMS changes related to over diagnosis bias
- Benefits outweigh harms, and no unintended consequences seem to have occurred with implementation though racial differences in hospitals may warrant future study and consideration for adjustment
- Benefits > harms
- Benefits seem to far outweigh risks, although this is a question best addressed by clinicians.
- Still need to separate measures to make them actionable.
- It's not entirely clear to me why the individual measures have improved but the overall composite has not, is this is a sign that perhaps the overall composite measure is too broad to capture information that's important to inform patient safety?

Criterion 5: Related and Competing Measures

Related or competing measures
None
Harmonization
None

- **Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures**

- **5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?**

- N/A
- None
- No
- None listed
- The composite measure seems comprehensive.
- No comment
- None.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 01/15/2021

- Comment by: **American Medical Association**

The American Medical Association (AMA) appreciates the opportunity to comment on #531, Patient Safety Indicator (PSI) 90: Patient Safety and Adverse Events Composite. We are disappointed to see that only 67% of all hospitals were able to achieve an intraclass correlation coefficients (ICC) of =0.6 in the split sample testing and only 51% in the test-retest using 24 months of data. We believe that measures must require higher case minimums to allow the overwhelming majority of hospitals to achieve an ICC of 0.6 or higher.

In addition, the AMA is extremely concerned to see that the measure developer used the recommendation to not include social risk factors in the risk adjustment models for measures that are publicly reported as outlined in the recent report to Congress by Assistant Secretary for Planning and Evaluation (ASPE) on Social Risk Factors and Performance in Medicare's Value-based Purchasing program (ASPE, 2020). We believe that while the current testing may not have produced results that would indicate incorporation of the two social risk factors included in testing, this measure is currently used both for public reporting and value-based purchasing. A primary limitation of the ASPE report was that none of the recommendations adequately addressed whether it was or was not appropriate to adjust for social risk factors in the same measure used for more than one accountability purpose, which is the case for here. This discrepancy along with the fact that the additional analysis using the American Community Survey is not yet released must be addressed prior to any measure developer relying on the recommendations within this report.

We request that the Standing Committee evaluate whether the measure meets the scientific acceptability criteria.

Reference:

Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health & Human Services. Second Report to Congress on Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program. 2020. https://aspe.hhs.gov/social-risk-factors-and-medicares-value-based-purchasing-programs

Comment by: Federation of American Hospitals

The Federation of American Hospitals (FAH) appreciates the opportunity to comment on Measure #531, Patient Safety Indicator (PSI) 90: Patient Safety and Adverse Events Composite. FAH is concerned that the majority of hospitals (67% in the split sample and 51% in the test-retest) were unable to achieve an intraclass correlation coefficients (ICC) of equal to or greater than 0.6. We believe that the developer must increase the minimum sample size to a higher number to ensure that at least 90% of the hospitals achieve an ICC of 0.6 or higher.

In addition, the FAH is very concerned to see that the measure developer's rationale to not include social risk factors in the risk adjustment model was in part based on the recommendations from the report to Congress by Assistant Secretary for Planning and Evaluation (ASPE) on Social Risk Factors and Performance in Medicare's Value-based Purchasing program released in March of last year (ASPE, 2020). A fundament

flaw within the ASPE report was the lack of any recommendation addressing how a single measure with multiple accountability uses should address inclusion of social risk factors as is the case with this measure, which is both publicly reported and included in the Hospital Value-Based Purchasing program. Regardless of whether the testing of social risk factors produced results that were sufficiently significant, the FAH believes that no developer should rely on the recommendations of this report until the question of how to handle multiple uses is addressed along with the additional analysis using the American Community Survey.

As a result, the FAH requests that the Standing Committee carefully consider whether the measure as specified meets the scientific acceptability criteria.

Reference:

Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health & Human Services. Second Report to Congress on Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program. 2020. https://aspe.hhs.gov/social-risk-factors-and-medicares-value-based-purchasing-programs

- **Of the 1 NQF member who have submitted a support/non-support choice:**
  - **0 support the measure**
  - **1 does not support the measure**

- Combined Methods Panel Scientific Acceptability Evaluation

Evaluating Scientific Acceptability: Instructions

Scientific Acceptability: Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. Measures must be judged to meet the subcriteria for both reliability and validity to pass this criterion.

Instructions for filling out this form:

- Please complete this form for each measure you are evaluating. Relevant measure documents are at the bottom of the site.
- If you are unable to check a box, please highlight or shade the box for your response.
- You must answer the "overall rating" item for both Reliability and Validity. Also, be sure to answer the composite measure question at the end of the form **if your measure is a composite.**
- For several questions, we have noted which sections of the submission documents you should ***REFERENCE*** and provided ***INSTRUCTON BOXES*** in comment bubbles to help you answer them.
- ***Please refer to the 2019 (pages 17-27) and the 2-page when evaluating your measures***.
- ***Please base your evaluations solely on the submission materials provided by developers.*** NQF strongly discourages the use of outside articles or other resources, even if they are cited in the submission materials. If you require further information or clarification to conduct your evaluation, please communicate with NQF staff as soon as possible (). Is it possible that we can obtain the needed information, but only if requested in a timely manner.
- **Remember** that testing at either the data element level **OR** the measure score level is accepted for some types of measures, but not all (e.g., instrument-based measures, composite measures), and

therefore, the embedded rating instructions may not be appropriate for all measures. Please review the box below to guide your rating.

- If a measure you are evaluating includes multiple measures (e.g., the Hopsital CAHPS measure submsission acutally includes 11 performance measures), all included measures must be rated. You may decide that one rating applies to all included measures, or you may need to provide separate ratings (e.g., if results are substantially better for one measure than for another).

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 0531

Measure Title: Patient Safety Indicator (PSI) 90: Patient Safety and Adverse Events Composite

Type of measure:

☐ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☒ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☒ **Composite**

Data Source:

☒ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☒ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**

   Panel Member #1: No concerns.

   Panel Member #2: None.

   Panel Member #3: none

   Panel Member #4: No concerns.

   Panel Member #5: Specifications are very clear.

   A minor point - S.22 states eleven components. It may be useful to clarify that 14 has two components: A & B.

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**

4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure ☒ Yes ☐ No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☒ **Yes** ☐ **No**

   Panel Member #1: N/A
6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2
   Panel Member #1: Developer used a split sample framework to compute an ICC which is appropriate for this composite measure.

   Panel Member #2: Split-sample and test-retest per hospital for two years of data.

   Panel Member #3: Split-sample reliability testing:

   - Component level: ICC ranged between 0.010 and 0.77.
   - Composite level: overall ICC was 0.61 in the Medicare 2016-2019 FFS data

   Panel Member #4: Appropriate..
   For component PSIs, they calculated measure reliability using estimated intraclass correlation coefficients (ICC), a type of signal-to-noise analysis.

   For the composite score, they used split-half and test-retest approaches to estimate the reliability.

   Panel Member #5: Reliability was tested at both the component level and the composite level.

   Methods included signal to noise analyses (components) and split-half and test-retest approaches (composite).

   Panel Member #6: Signal to noise for the individual components but split sample and test-retest for the component

   Panel Member #7: "The signal-to-noise reliability approach does not apply to PSI 90 as a composite measure, because PSI 90 is a weighted average of risk-adjusted, reliability-adjusted (smoothed) component measures. In other words, each hospital's own signal-to-noise reliability is used as a shrinkage parameter to determine how far to shrink that hospital's estimate toward the national reference mean of 1.0. Through this process, noise variance is essentially removed. Therefore, we apply split-half and test-retest approaches to estimate the reliability of smoothed measures such as PSI 90."

   Panel Member #8: Split-sample and test-retest
7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   Panel Member #1: PSI 90 is a reliable hospital-level measure, with a median ICC of 0.61 using 24 months of data and ICC of 0.70 using 36 months using the most current data. These findings indicate the need to use the longest possible data stream to bolster reliability.

   Panel Member #2: Wide variation in the signal to noise statistics for the individual components of the composite – no discussion of the impact of that result on the reliability of the composite.

   Panel Member #4: Components: Signal-to-noise reliability varies across the PSI 90 component measures, with more frequent events (i.e., PSI 03, PSI 11, PSI 12, PSI 13) having higher signal-to-noise reliability (weighted mean >0.5) than rare events (i.e., PSI 08, PSI 14, with weighted mean <0.3),

highlighting the importance of PSI 90 as a composite measure that draws statistical strength from all of its component measures.

Composite: the measure demonstrates moderate-to-high score reliability at the hospital level, with an overall split half (intracluster correlation coefficient, ICC) reliability estimate of 0.74 based on 24 months of Medicare FFS claims data, and 67% of facilities exceeding ICC=0.6.

Panel Member #5: Results suggest acceptable levels of reliability only for the composite measure, as also noted by the developers.

Panel Member #6: Although data for the composite measure seemed acceptable, split sample 0.74 and test retest 0.61, the testing for individual components varied considerably from 0.026 to 0.668 using current CMSv10; the meaning of this discrepancy is not clear but concerning

Panel Member #7: "As shown in Table 6, PSI 90 demonstrates moderate-to-high score reliability at the hospital level, with an overall split half (intracluster correlation coefficient, ICC) reliability estimate of 0.74 based on 24 months of Medicare FFS claims data, and 67% of facilities exceeding ICC=0.6. If even higher reliability were desired, the data period could be increased to 36 months, with split half reliability of 0.81 and 76% of facilities exceeding ICC=0.6 (Table 6). As the reliability distribution in Figure 1 shows, only 2-3% of hospitals have very low reliability (ICC<0.05). CMS anticipates excluding most of these low-reliability hospitals from public reporting using a minimum volume threshold (e.g., 25 denominator records) and a missing data threshold…"

Panel Member #8: "As shown in Table 6, PSI 90 demonstrates moderate-to-high score reliability at the hospital level, with an overall split half (intracluster correlation coefficient, ICC) reliability estimate of 0.74 based on 24 months of Medicare FFS claims data, and 67% of facilities exceeding ICC=0.6. If even higher reliability were desired, the data period could be increased to 36 months, with split half reliability of 0.81 and 76% of facilities exceeding ICC=0.6 (Table 6). As the reliability distribution in Figure 1 shows, only 2-3% of hospitals have very low reliability (ICC<0.05). CMS anticipates excluding most of these low-reliability hospitals from public reporting using a minimum volume threshold (e.g., 25 denominator records) and a missing data threshold as described further below"

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

   ☒ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and <u>all</u> testing results):

    ☒ **High** (NOTE: Can be HIGH **only if** score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has **not** been conducted)

☐ **Low** (NOTE: Should rate **LOW** if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

☒ **Insufficient** (NOTE: Should rate ***INSUFFICIENT*** if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

Panel Member #1: See comments for item #7.

Panel Member #2: Moderate reliability statistics – 20% of hospitals have ICC below 0.4. There is discussion of applying a low volume threshold, but the impact on the reliability is not formally assessed

Panel Member #3: Composite level: overall ICC was 0.61 in the Medicare 2016-2019 FFS data

Panel Member #4: Used appropriate testing methods, looked at both reliability of components and composite. Composite demonstrated moderate-to-high score reliability.

Panel Member #5: The moderate rating is derived from the moderate level of reliability for both the split-half ICC (0.74, 24-month period), and test-retest reliability (0.75). Having said that, the selection between moderate and high is not clear. Since higher reliability would be desired, I selected moderate, but I think a high rating would also be appropriate. As there are no consequences to this differentiation, so I am comfortable selecting the moderate rating. Also, the moderate rating was selected mainly because the measure was reliably only at the composite level. It would have been better to see high reliability for all components.

Panel Member #6: Given the disparity in reliability across different components, even if overall model appears to function, the impact on any given institution with a higher proportion of those components with low reliability because unstable and potentially misrepresentative of overall quality

Panel Member #7: Overall reliability statistics are OK, however, low volume sites remain an issue and individual PSIs 8 and 14 are problematic.

Panel Member #8: Moderate to food overall reliability statistics, except for low volume sites.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

Submission document: Testing attachment, section 2b2.

Panel Member #1: None.

Panel Member #2: None.

Panel Member #3: none

Panel Member #4: None.

Panel Member #5: Number of exclusions listed, and the data presented are somewhat overwhelming. However, I have no concerns. Exclusions were supported by expert panels convened by AHRQ and/or CMS.

Panel Member #6: Measure exclusions seem reasonable and well-documented

Panel Member #7: None.

Panel Member #8: None.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

Submission document: Testing attachment, section 2b4.

Panel Member #1: None.

Panel Member #2: None.

Panel Member #3: none

Panel Member #4: There is reasonable variation in scores. 10% of hospitals are outside of "as expected".

Panel Member #5: No concerns. About 10% of hospitals were either high or low performers, which still leaves room for improvement for the low performers (6%).

However, this measure has the potential of being topped out, which calls for a consideration for future revisions.

Panel Member #6: If you accept the validity of the measure, there is a reasonable distribution of performance that does distinguish high and low "performers"

Panel Member #7: None.

Panel Member #8: None.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**

Submission document: Testing attachment, section 2b5.

Panel Member #1: N/A

Panel Member #2: None.

Panel Member #3: none

Panel Member #4: Not applicable.

Panel Member #6: Not applicable.

Panel Member #7: None.

Panel Member #8: None.

15. **Please describe any concerns you have regarding missing data.**

Submission document: Testing attachment, section 2b6.

Panel Member #1: None.

Panel Member #2: None.

Panel Member #3: none

Panel Member #4: No strong concerns. A limited number of hospitals may be potentially penalized by the imputation approach.

Panel Member #5: No major concerns, as missing data were negligible at the component measure level, and apparent for at least one missing component in only 10.9 % of hospitals. CMS's proposed threshold for exclusion from public reporting of hospitals with 4 or more missing PSI90 components is encouraged, as these hospitals account for 6.3% of hospitals, i.e., 58% (6.3/10.9) of hospitals with at least one missing component, further mitigating potential for bias due to missing data.

Panel Member #6: Level of missing data seems low in general, although there is a small subset of hospitals that have considerable missing data—developers raise the possibility of excluding hospitals with a threshold of missing data—this appears to be a wise idea

Panel Member #7: None.

Panel Member #8: None.

16. **Risk Adjustment**

16a. **Risk-adjustment method** ☐ **None** ☒ **Statistical model** ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☐ Yes ☒ No ☐ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☒ Yes ☐ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☒ Yes ☒ No

Panel Member #5: Developers stated that "Because the PSIs focus on adverse events occurring within acute care hospitals, often after a major operating room procedure, social risk factors are not included in the conceptual approach.". There is no explanation of this rationale, and why it holds differently for social risk factors compared to other clinical factors.

Developers adopted the present ASPE Report to Congress on Social Risk Factors (2020) which recommended not to adjust composite scores for social risk factors.

16d. **Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☒ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☒ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☒ Yes ☐ No

16d.5. Appropriate risk-adjustment strategy included in the measure? ☒ Yes ☒ No

16e**. Assess the risk-adjustment approach**

Panel Member #1: The risk adjustment model discrimination is good.

Panel Member #2: Model calibration statistics are presented for the components. Hosmer-lemenshow plots are provided – many are very noisy. No statistical testing provided.

Panel Member #3: Model discrimination (C statistic) is acceptable for the component outcomes. Model calibration (assessed using OE ratios for risk deciles) is acceptable across most deciles of risk (with the exception of lowest-risk deciles which have few events).

Panel Member #4: Appropriate approach; high c-statistics

Panel Member #5: No concerns.

Panel Member #6:

1) Model development describes elimination of elements with variable inflation factor>1000—usually 5 or 10 is considered a reasonable threshold; hat said, discrimination of final model appears strong—developed in CMS and tested in HCUP
2) Rationale for exclusion of social risk factors is not logical. Example given is for central line infection, which is not a metric in the composite. However, the possibility for renal failure, respiratory failure, etc. to be related to variation is underlying care that are not adequately reflected by the diagnosis codes (e.g. someone who has comorbidities that are well managed vs. one who has the same comorbidities that are not well managed due to lack of access to

comparable care) is not considered in the modeling. In other words, underlying conditions, the severity of which are not adequately captured by the diagnosis codes, can impact postoperative and in-hospital complications despite comparable levels of in-hospital care. This factor is not accounted for in the model and therefore this model may penalize hospitals who care for such at-risk patients, potentially impacting needed resources as well as providing incentive not to care for such patients.

Panel Member #7: Discrimination is good, however, some H-L plots are less than ideal. PSI-90 is a behemoth that addressing risk adjustment for each composite may require a separate team if such is truly desired.

For cost/resource use measures ONLY:

17. **Are the specifications in alignment with the stated measure intent?**

    ☐ **Yes** ☐ **Somewhat** ☐ **No (If "Somewhat" or "No", please explain)**

18. **Describe any concerns of threats to validity related to attribution, the costing approach, carve outs, or truncation (approach to outliers):**

VALIDITY: TESTING

19. **Validity testing level:** ☒ **Measure score** ☐ **Data element** ☐ **Both**

20. **Method of establishing validity of the measure score:**

    ☒ **Face validity**

    ☒ **Empirical validity testing of the measure score**

    ☐ **N/A (score-level testing not conducted)**

21. **Assess the method(s) for establishing validity**

    Submission document: Testing attachment, section 2b2.2

    Panel Member #1: The developer used multiple approaches to evaluate validity including predictive validity for the component measures, for a range of related outcomes, and demonstrated construct validity for the composite measure using a known-group validity framework. These approaches are sound.

    Panel Member #2: Construct validity against other metrics shows good validity.

    Panel Member #3: Assessed empiric validity on multiple levels, including predictive validity (which they define as the average marginal effect of each PSI event on subsequent adverse outcomes), convergent validity between PSI190 and (i) infection outcomes, (ii) 30-day readmissions, (iii) Leapfrog survey safe practice scores, and structural factors (e.g. nurse staffing ratios). The findings of these analyses suggest that this measure is valid.

    Panel Member #4: Compared PSI results with related measures of patient safety and outcomes at the hospital level, publicly available on https://data.Medicare.gov. Compared hospital-level PSI rates with rates of complications for hip/knee replacement patients, risk-standardized 30-day readmission rates (e.g., hospital-wide unplanned all-cause readmissions) and health care-associated infection measures from the National Healthcare Safety Network.

    Also looked at correlation with Leapfrog's Safe Practices and structural measures (resident-to-bed ratio, nurse-to-bed ratio, and nurse skill mix)

    Panel Member #5: Face validity and empirical validity were assessed at both the component and the composite levels.

Convergent validity was tested against a set of related measures, with moderate correlations expected.

Known groups construct validity was also assesses at the hospital level, expecting that the composite scores would be able to discriminate between groups of hospitals that have different levels of three related measures (Hospital resident-to-bed ratio, Hospital nurse-to-bed ratio, and Hospital nurse skill mix) in expected ways.

Panel Member #6: Demonstration of impact of each factor on subsequent adverse events as well as correlation with other external metrics of hospital quality.

Panel Member #7: Adequate.

Panel Member #8: Reasonable.

22. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

Panel Member #1: The composite results were the most compelling evaluations of validity so I focus on those tests. The validity testing showed weak correlations between the measure score and hospital compare infection-related outcomes, with the exception for the HAC score which had a strong correlation. Correlations with readmission measures were even weaker ($r < 0.10$ with the exception of 30-day hospital-wide readmissions which was .145). This is modest evidence of validity.

Panel Member #2: Good correlation with other gold standard metrics**.**

Panel Member #3: Assessed empiric validity on multiple levels, including predictive validity (which they define as the average marginal effect of each PSI event on subsequent adverse outcomes), convergent validity between PSI190 and (i) infection outcomes, (ii) 30-day readmissions, (iii) Leapfrog survey safe practice scores, and structural factors (e.g. nurse staffing ratios). The findings of these analyses suggest that this measure is valid.

Panel Member #4: Mixed results, but as the measure developer noted, patient safety is a complex construct, so one wouldn't expect high correlations.

Panel Member #5: Face validity was established in 2014 and reassessed in 2019, with an overall strong support for the composite measure and its components. Overall, validity was supported for some assessments, sufficiently supporting the overall validity of measure 0531.

Panel Member #6:

1) As regards the modeling of impact of the various complications on other subsequent adverse events such as hospital mortality or length of stay, etc., results are presented in table form after description of the statistical approach without any demonstration of the specific models developed, nor the results of those models (Table 9) For example PSI03 has a 27% incremental increase in mortality—is this based on an odds ratio of 1.27 after appropriate logistic regression with inverse probability of treatment weighting? What was the model? What were the confidence intervals and p values? May be very well done but we have no idea from the data presented.
2) Tables 10 and 11 demonstrated very low apparent correlation with other measures which would be expected to correlate closely; not clear if the numbers presented for results of Spearman rank test represent the r value, but if they do the numbers, many of which are less than 0.1, suggest that the measure has a very low ability to account for the variability seen in the external measures evaluated

Panel Member #7: Bad things are correlated with bad things.

Panel Member #8: Reported relationship are as expected.

23. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

    Submission document: Testing attachment, section 2b1.

    ☒ **Yes**

    ☒ **No**

    ☐ **Not applicable** (score-level testing was not performed)

    Panel Member #6: There are two major challenges to the validity of this measure that are not addressed by the developers:

    1) Validation that the diagnosis code being used actually represents the clinical entity of concern. For example, pulmonary physicians routinely consulted to manage the ventilator in postoperative patients with automatically use the diagnosis of respiratory failure in order to validate their consult, even though the patient is a routine postoperative patient who expected to be on the ventilator for a limited amount of time.
    2) More problematic is the presumption that all of these measures are necessarily preventable. The specific intention of the metric is to measure preventable patient harm. Successful processes that are well documented to prevent renal failure or pulmonary failure are limited at best. Central line infection is one thing, but sepsis in an otherwise sick patient may or may not be preventable. Without clear evidence that there are validated processes that prevent the complication, the rationale for inclusion is incomplete.

24. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**

    NOTE that data element validation from the literature is acceptable.

    Submission document: *Testing attachment, section 2b1.*

    ☒ **Yes**

    ☐ **No**

    ☒ **Not applicable** (data element testing was not performed)

    Panel Member #5: considering components as data elements.

25. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

    ☒ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

    ☒ **Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

    ☒ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

26. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

Panel Member #1: Convergent validity measured via correlations with infection related outcomes and readmissions for the composite score was weak and there were also very tenuous relationships between PSI scores and evidence-based safe practices.

Panel Member #3: Assessed empiric validity on multiple levels, including predictive validity (which they define as the average marginal effect of each PSI event on subsequent adverse outcomes), convergent validity between PSI190 and (i) infection outcomes, (ii) 30-day readmissions, (iii) Leapfrog survey safe practice scores, and structural factors (e.g. nurse staffing ratios). The findings of these analyses suggest that this measure is valid.

Panel Member #4: Measure developer did comprehensive comparisons to other measures that intend to capture patient safety (process, structure, and outcome measures). Mixed results, for sure, but they did not see any consistent pattern that would indicate their measure is not capturing patient safety.

Panel Member #6: Lack of documentation of modeling for impact of these complications, lack of adequate rationale for non-inclusion of potentially available social risk factors, and lack of validation that each of these complications is preventable considerably weaken the validity of the metric.

FOR COMPOSITE MEASURES ONLY: Empirical analyses to support composite construction

27. **What is the level of certainty or confidence that the empirical analysis demonstrates that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct?**

    ☒ **High**

    ☒ **Moderate**

    ☒ **Low**

    ☒ **Insufficient**

28. **Briefly explain rationale for rating of EMPIRICAL ANALYSES TO SUPPORT COMPOSITE CONSTRUCTION**

    Panel Member #1: The weighting system was not clearly explained although the item spearman rank correlations seem reasonable.

    Panel Member #2: The reliability and validity of some of the component measures is low – they tend to be weighted lower in the composite calculation, but unclear why/if they should be eliminated altogether.

    Panel Member #3: Analysis shows that item-total correlations are higher than component weights.

    Panel Member #4: Correlation coefficients between the component measures are relatively modest, indicating that each component may be contributing something unique to the overall composite. Have taken great care in assigning relative weights, including incorporation of excess harm and disutility.

    Panel Member #5: See comments above under validity

    Panel Member #6: There appears to be a complex weighting system based on the contribution of the metric to the overall score—which itself is circular logic, since it is the weighting system that determines the weight in the overall score. Origin of the "harm weights" (??based on the harm models inadequately described above??) and the disutility scores (expert opinion??) is unclear.

    Panel Member #7: "One tenet of this composite is that each component measure is correlated with an aspect of each hospital's underlying quality of care. Therefore, we expect to observe positive hospital-level correlations among the individual measures within the composite." (And they were positive.)

"For each component indicator in the modified version of PSI 90 composite, two sets of values need to be computed or estimated. The first is the excess risk of each harm outcome (risk difference) that may occur in association with the component PSI event. These harm risks are multiplied by harm-specific disutility scores, which reflect the relative valuation of various outcome states by patients and clinicians, and then summed across all of the harms relevant to a component PSI, to obtain the summed harm weight for each PSI. Next numerator weights are calculated from the volume (count) of each PSI component event in the CMS FFS reference population. Finally, the volume weight for each PSI is multiplied by its summed harm weight, and the resulting product is rescaled across all 10 components so that the sum of the final weights is 1."

ADDITIONAL RECOMMENDATIONS

29. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below**

Developer Submission

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

PSI_90_NQF_Evidence_Attachment_Master-637395159640759611.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

Yes

- 1a. Evidence (subcriterion 1a)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite Measure Number)

Measure Title: PSI 03 Pressure Ulcer Rate (Component Measure)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

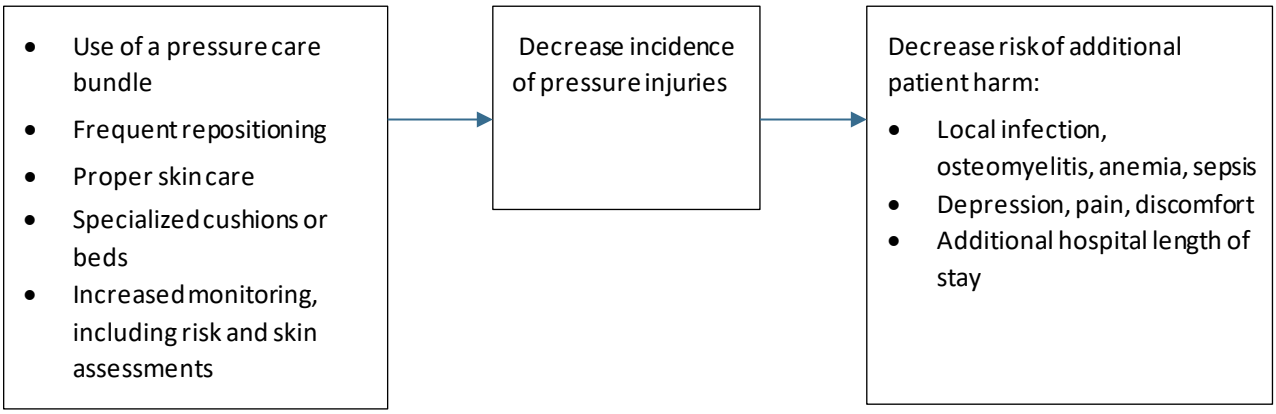
☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite: f

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.



Pressure injuries, also called pressure ulcers, are serious events and one of the most common patient harms. Pressure ulcers can be prevented by addressing modifiable risk factors such as friction, humidity, temperature, continence, medication, shearing forces, unrelieved pressure, and poor nutrition. One approach that has been very successful in decreasing hospital associated pressure ulcer includes the use of a pressure ulcer care bundle based on best practices and tailored to individual hospital settings. Critical components of the pressure ulcer care bundle include: use of a comprehensive skin assessment and standardized pressure ulcer risk assessment along with structured interventions to address areas of risk (such as the use of mechanical loading and support surfaces).

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the *systematic review of the body of evidence* that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review

Evidence

Source of Systematic Review:

- Title
- Author
- Date
- Citation, including page number
- URL

Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.

Grade assigned to the evidence associated with the recommendation with the definition of the grade

Provide all other grades and definitions from the evidence grading system

Grade assigned to the recommendation

with definition of the grade

Provide all other grades and definitions from the recommendation grading system

Body of evidence:

Quantity – how many studies?

Quality – what type of studies?

Estimates of benefit and consistency across studies

What harms were identified?

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with Process of Care

To prevent hospital-acquired pressure ulcers, evidence-based guidelines[1-4] and systematic reviews[5] recommend that skin assessments be performed at admission and daily during the inpatient stay, with particular attention to bony prominences and skin adjacent to external/medical devices. These assessments should include complete documentation of all skin lesions and pressure ulcers along with staging (including location, tissue type, shape, size, presence of sinus tracts/tunneling, undermining, exudate amount and type, presence/absence of infection, and wound edges). Documentation in the medical record should include skin temperature, skin color, skin texture/turgor, skin integrity, and moisture status. In addition, evidence-based guidelines[2,6] and systematic reviews[5] recommend that nutritional assessments be performed at entry to new health care settings and whenever patient status changes.

Evidence-based guidelines and systematic reviews also recommend that at-risk patients are placed on a pressure-reducing surface rather than a standard hospital mattress.[1-3,5,7] In a systematic review of 120 studies reporting on pressure ulcer risk assessment and prevention, Chou and colleagues examined the effectiveness of various interventions for reducing pressure ulcers for hospitalized patients.[8] They concluded that fair-quality randomized trials consistently found that in higher risk patients, advanced static support surfaces were associated with lower risk of pressure ulcers compared to standard mattresses (relative risk range, 0.20 to 0.60), with no clear differences among different advanced static support surfaces. Evidence on the comparative effectiveness of more advanced dynamic support surfaces was limited; some trials showed no significant differences between dynamic and static support surfaces. In lower risk populations of patients undergoing surgery, two trials found use of a foam overlay associated with a higher risk of pressure ulcers compared with a standard operating room mattress. Evidence on the effectiveness of other preventive interventions (e.g., nutritional supplementation; repositioning; pads and dressings; lotions, creams, and cleansers; corticotrophin injections; polarized light therapy; and

intraoperative warming therapy) was sparse and insufficient to reach reliable conclusions. In a systematic review of 25 studies reporting on pressure ulcer prevention strategies in the ICU, meta-analyses found a statistically significant effect of silicone foam dressing in reducing sacral HAPUs in critically ill patients (effect size, 0.12; p<0.00001). Evidence on the effectiveness of other strategies (nutrition, skin care, position/repositioning, support surfaces) was limited and insufficient to reach reliable conclusions.[9] A 2020 Cochrane review of eight trials concluded that there is an absence of high-quality evidence to evaluating the effectiveness of repositioning frequency and positioning for pressure ulcer prevention.[10] Another 2020 systematic review including both trial and observational studies assessing the effects of different repositioning regimens concluded that there is low-certainty evidence that more frequent repositioning (every 2-3 hours versus 4-6 hours; OR, 0.75; 95% CI, 0.61-0.90, p=0.03) and use of a turning team (OR, 0.49; 95% CI, 0.27-0.86, p=0.01) can reduce pressure ulcer incidence in at-risk adult patients.[11] A 2018 Cochrane review including six trials comparing silicone dressings with no dressings found low-certainty evidence that silicone dressings reduce HAPUs (RR, 0.25; 90% CI, 0.16-0.41).[12] An observational cohort study conducted in 38 acute care hospitals between 2010 and 2015 found that adoption of a prophylactic foam sacral dressing as part of a HAPU prevention protocol resulted in reduced HAPU rates; the average hospital experienced one fewer HAPU per quarter following implementation of the dressing.[13] Elsabrout et al. found that a hospital-wide mattress switch-out program resulted in a 66.6% decrease in Stage III and IV HAPUs and a cost savings of $714,724.[14]

Association with hospital and health system characteristics

Two studies have showed that low-volume hospitals have higher rates of pressure ulcers than higher volume hospitals. In an analysis of Diagnosis Procedure Combination/per-diem payment system (DPC/PDPS) data from 1,383,872 patients discharged from 188 hospitals in Japan (2008-2010), Kitazawa and colleagues found that low-volume hospitals (< 33rd percentile by volume) had higher rates (8.0 per 1,000 discharges; 95% CI 5.1 to 10.09) of pressure ulcers (PSI 03, version 4.2) than mid-volume (4.5 per 1,000 discharges; 95% CI 3.5 to 5.5) and high volume hospitals (3.8 per 1,000 discharges; 95% CI 3.0 to 4.6) ($p < 0.05$).[15] Likewise, in an analysis of Medicare claims data for patients undergoing any of six types of cancer resection in 2005-2009, Short et al. found that the pressure ulcer rate was higher (0.78%) at low-volume hospitals than at high volume hospitals (0.59%), but not different between teaching and non-teaching hospitals.[16] Seemingly at odds with these findings, a cross-sectional study by Choi et al. using 2009 unit-level data from the National Database of Nursing Quality Indicators (NDNQI) linked with the NDNQI RN Survey found the odds of hospital-acquired pressure ulcer (HAPU) were higher (OR 1.27; $p < .05$) in hospitals with 300 or more beds compared to hospitals with <300 beds.[17] However, these associations were not consistent across unit types.

The impact of Magnet-hospital designation was assessed by three different studies. In an analysis of quarterly unit-level data from the NDNQI (2008-2010), including 10,935 unit-quarter observations (2,294 adult units in 465 hospitals from 47 U.S. states), Park et al found that hospital magnet status was significantly associated with lower unit-acquired pressure ulcer (UAPU) rates (OR 0.84; p =0.049). In a cross-sectional study using 2009 unit-level data from the NDNQI linked with the NDNQI RN Survey, Choi et

al found[17] the odds of HAPU occurrence were lower in Magnet hospitals (OR 0.81; p < .05) than in non-Magnet hospitals. However, these associations were not consistent across unit types. Magnet status was significantly associated with the UAPU rate in step-down (OR 0.76; p < .05) and medical units (OR 0.64; p < .001), but not in critical care units, rehabilitation units, surgical units and combined medical-surgical units. Mills & Gillespie, however, found conflicting results using five years (2001-2005) of data from the Health Care Cost and Utilization Project (HCUP) Nationwide Inpatient Sample (NIS), merged with annual surveys conducted by the American Hospital Association. They found no differences in HAPU rates between 80 Magnet hospitals and 80 non-Magnet hospitals (p > .05).[18]

Rosen and colleagues used PSI 03 (version 3.1a) to explore associations between safety climate, as measured through 9,309 responses to the Patient Safety Climate in Healthcare Organizations survey, and hospital safety performance in 2005-2006. Among the 30 Veterans Health Administration (VA) hospitals that participated in the survey, "greater fear of blame and punishment for making mistakes" was significantly associated with higher rates of pressure ulcer (p < .05), stronger endorsement of "overall emphasis on safety" was significantly associated with lower rates of pressure ulcer (p < .05), and greater endorsement of "unit safety norms" was associated with lower rates of pressure ulcer (p < 0.1), in models adjusting for hospital teaching status, metropolitan area, and nurse-staffing ratio.[19] A different study by Rosen et al. (2006) used VA Patient Treatment File (PTF) data to examine risk-adjusted PSIs (version 2.1) for acute care VA hospitals for fiscal years 2001 to 2004.[20] The PSIs were assessed to characterize adverse patient safety events, determine trends in PSIs over time, and evaluate potential predictors of hospital safety. This study did not find PSI 03 to be significantly associated with any single hospital characteristic (e.g. bed size, teaching status, location).

Since 1991, the VA system has used risk-adjusted pressure ulcer rates as a system-wide outcome to monitor hospital performance and also supported research to explore best practices targeting pressure ulcer prevention in the inpatient setting. To assess the impact of this work, Chen et al. compared PSI 03 rates among 266,203 veteran dual users who were hospitalized in both VA facilities and private sector facilities between 2002 and 2007 using AHRQ PSI software (version 3.1a).[21] Rates of pressure ulcer among these dual users were significantly lower in VA hospitals than in private sector hospitals: 25.9 versus 44.4 per 1,000 population. After adjustment for age, sex and 27 comorbidities, the pressure ulcer rate among VA hospitalizations was 20.4 per 1,000 discharges (95% CI 19.9 to 21.0) while among Medicare hospitalizations, the rate was 27.8 per 1,000 discharges (95% CI 27.3 to 28.3). Among veteran dual users, the odds of experiencing a PSI 03 event was 35% lower in the VA versus in the private sector (OR 0.65, 95% CI 0.63 to 0.68).

Associations with nursing staff characteristics

Multiple studies have examined the association between pressure ulcers and hospital nurse staffing characteristics, including registered nurse (RNs) turnover, hours, and education level. In a longitudinal study of 23 nursing units in two hospitals from October 2009 through December 2011, Warashawsky et al found that patients on units with nurse manager turnover (OR 3.16; 95% CI 1.49 to 6.70) were more likely to develop pressure ulcers than patients on other units.[22] In an analysis of quarterly unit-level data from

the NDNQI (2008-2010), including 10,935 unit-quarter observations (2,294 adult units in 465 hospitals from 47 U.S. states), Park et al found a significant lagged effect of RN turnover on HAPU rates, but not a concurrent effect.[23] For every 10 percentage-point increase in RN turnover in a quarter, the odds of a patient having a pressure ulcer increased by 4% in the next quarter ($p = 0.038$). Similarly, in a cross-sectional study using 2009 unit-level data from the NDNQI linked with the NDNQI RN Survey of 77,826 nurse respondents on 3,329 units at 561 different hospitals, Choi et al found that longer RN tenure on the current unit was related to lower HAPU rates among older adult patients (OR 0.97; $p < .05$), although when examined by unit type, this relationship only remained significant for step-down units (OR 0.97; $p < .05$).[17]

Choi et al also found[17] that UAPU rates were associated with both RN job satisfaction and hours per patient day. RN workgroup job satisfaction was significantly and inversely associated (OR 0.98; $p < .001$) with the odds of UAPU after adjusting for other unit (nurse staffing, RN education level, and RN unit tenure) and hospital (bed size, teaching status, and Magnet status) characteristics. However, the association between RN job satisfaction and UAPU rates varied by unit type: Higher job satisfaction among RN workgroups was significantly associated with lower UAPU rates among older adults on critical care (OR 0.97; $p < .001$), medical (OR 0.98; $p < .05$), and rehabilitation units (OR 0.97; $p < .05$), but no significant relationship was found in step-down, surgical, and medical-surgical units. In the model with all sample units, RN hours per patient day and licensed practical nurse (LPN) hours per patient day were also significantly related to UAPU rates, but in the opposite direction of what was expected: the odds of UAPU occurrence increased for each additional RN hour per patient day (OR 1.05; $p < .05$) and additional LPN hour per patient day (OR 1.14; $p < .05$). In the unit-specific model, this relationship between increased hour per patient day and increased HAPU rates only remained significant in step-down units for additional RN hours per patient day (OR 1.07; $p < .05$). Park et al., however, found the opposite: more RN hours per patient day were associated with lower UAPU rates, controlling for other variables (unit type, non-RN staffing, hospital magnet status, hospital size, case mix index), (OR 0.952; $p = 0.11$). The significant effect of RN staffing on UAPUs remained without including the RN turnover variables as predictors (OR 0.950; $p = .009$).[23]

In a cross-sectional study using 2005 data from 21 University Health System Consortium hospitals, Blegen et al. found that hospitals with higher percentages of RNs with baccalaureate or higher degrees had lower rates of hospital-acquired pressure ulcers (Pearson $r = 0.500$; $p < 0.05$), as measured by PSI 03 (version 3.1).[24] The effect of nursing education was stronger than the effect of nurse staffing (as measured by hours of direct patient care by RN, licensed practical nurse [LPN] or nursing assistant [NA] per patient day). Results were similar when using a regression model to adjust for nurse staffing, Medicare case mix index (CMI), and Hospital Technology and Safety Net status; there was a trend towards lower pressure ulcer rates in general hospital units and in intensive care units when the proportion of baccalaureate -prepared RNs was higher ($p < 0.10$). Using NDNQI data, Boyle et al found that hospitals employing certified wound, ostomy, and continence (WOC) specialty nurses had lower HAPU rates, as well as better pressure injury risk assessment and prevention practices. The study found that the prevalence of stage III and IV pressure ulcers at hospitals employing specialty certified nurses was much lower (0.27%) than at hospitals that did not employ specialty nurses (0.51%).[25] Another study using longitudinal data from NDNQI found that HAPU rates decreased after nurse practitioners took on the role of wound care consultants (OR, 0.20; 95% CI, 0.15-0.27).[26]

Finally, Aydin et al. used data from a convenience sample of 789 medical-surgical units from 215 hospitals from CALNOC, a nursing-sensitive benchmarking registry, to model the predictive power of nursing staff characteristics on HAPU prevalence.[27] The percent of patients with HAPUs decreased as total nursing hours per patient day (HPPD) increased, the years of RN experience increased, and percent of hours provided by contract staff decreased. Thus, at 5 HPPD, with 6 years of mean experience and 10 percent of hours provided by contact staff, HAPUs affected 5.2% of patients, versus only 2.2% when RNs had 16 years of experience and no contract staff hours were used.

Association with other outcomes

Pressure ulcers commonly lead to further patient harm, including local infection, osteomyelitis, anemia, and sepsis, cellulitis, pyoderma, bacteremia, septic arthritis, necrotizing fasciitis, and gas gangrene/gangrene, and or flap failure;[28-31] these complications often require intensive care or surgical procedures including wound debridement and skin graft or flap.[32] Pressure ulcers can also lead to significant depression, pain, and discomfort to patients.[28-30]

Multiple studies in adult populations have found that the occurrence of pressure ulcers is associated with longer length of stay in the hospital and greater costs. Brito et al conducted a multicenter, cross-sectional study of 473 adults admitted to hospitals in different geographic regions of Brazil (2009-2011).[33] In multivariable analyses, the presence of pressure ulcers was directly associated with length of stay of more than 8 days (OR 3.85; 95% CI 1.53 to 9.73). Using data on Medicare fee-for-service patient discharges (n=51,842) in 50 U.S. states over a two-year period (2006 –2007), Moore found that patients with a HAPU had a statistically significant longer length of stay than those without a HAPU (11.6±10.1 days vs 4.9±5.2 days, p<0.001).[34] In a cross-sectional study using the 2008 NIS, Lee et al analyzed 10,660 hospitalizations with head and neck cancers who underwent radical neck dissections.[35] Using multivariable linear regression analysis (controlling for patient age, sex, race, comorbidities, insurance, type of radical neck dissection, hospital region and hospital teaching status), the authors found that patients who experienced pressure ulcers stayed 5.6 extra days in the hospital ($p < 0.0001$) and incurred $49,153 in extra hospitalization charges ($p = 0.003$), compared with patients who did not. In an analysis of Medicare claims data for patients undergoing any of six types of cancer resection in 2005-2009, Short et al found that after adjusting for patient (age, sex, race, income), hospital (hospital volume, surgeon volume, surgeon specialty designation, hospital resources, patient characteristics) and tumor factors (tumor stage, site), costs increased by 28% to 60% after a pressure ulcer ($p < 0.001$).[16] In an analysis of 3,466,596 inpatient visits in the year 2009 from the Premier Hospital Database, Mallow et al found that the prevalence of in-hospital pressure ulcer in this general population sample was 18.3 per 1000 visits, the median cost associated with pressure ulcers was $1,017, and the total cost of pressure ulcers was $478,501,000 for 2009.[36] Zhan and Miller used AHRQ PSI software on 7.45 million discharges in the HCUP NIS (2000) and found that patients who experienced a PSI 03 event had a higher mean (SD) unadjusted length of stay (160.32 [0.09] vs. 90.79 [0.006]), charges ($45,987 [375] vs. $28,100 [29]), and mortality (130.85 [0.17] vs. 40.01 [0.01]) than patients who did not.[37] However, statistical differences of these comparisons were not reported. Ramanathan et al. retroactively examined data on surgical patients hospitalized between 2011 and 2012 at

an academic medical center and found that hospitalizations that included pressure ulcers (PSI 03, version 3.1) were associated with a 48.0 day mean hospital LOS, 80% included an intensive care unit stay, and 33.3% died in hospital.[31] 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 03 event were associated with an additional 2.26 hospital days compared to patients without a PSI 03 event ($p<0.001$).[38] Using HCUP NIS data for 2008 to 2012, Bauer et al. identified statistically significant differences in median length of stay between patients with and without pressure ulcers (7 days versus 3 days) and significant differences in total cost ($36,500 versus $17,000).[39] Bath et al. used Medicare data (MedPAR) from 2009 to 2012 and found that the likelihood of 30-day readmission among patients undergoing abdominal aortic aneurysm repair was greater among patients with a pressure ulcer event (OR=2.88, $p<0.001$).[40]

Population group disparities

Table 3 presents population group disparities for component measure PSI 03 Pressure Ulcer Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 3. PSI 03 Pressure Ulcer Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	66766	0.509	0.461
White	5399715	0.564	0.572
Black	879228	0.954	0.781
Other	86037	0.628	0.480
Asian	100197	0.659	0.498
Hispanic	155142	0.535	0.458
North American Native	49801	1.024	0.803

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Gender	*	*	*
Female	3642129	0.512	0.614
Male	3094757	0.745	0.588
Age	*	*	*
<50	383026	0.509	0.587
50-54	209771	0.682	0.636
55-59	319073	0.787	0.647
60-64	402353	0.833	0.635
65-69	1070923	0.660	0.601
70-74	1078841	0.608	0.579
75-79	1005215	0.652	0.584
80-84	899429	0.609	0.599
85-89	759021	0.506	0.578
90 plus	609234	0.486	0.623

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review presented below constitute the most recent

update, conducted in August 2020. Search terms included relevant MeSH terms (pressure ulcer) and keywords (pressure ulcer, decubitus ulcer, pressure sore, skin ulcer). Studies focused on long-term care settings or obstetric care were excluded. We combined this clinical search string with MeSH terms (patient admission) and keywords (hospitals, patient admission, inpatient, patient safety, quality, indicator, epidemiologic, statistic, AHRQ, prevalence, incidence, or utilization) to identify studies examining inpatient care and quality measurement. Search was limited to English publications. We also tested more inclusive search strings. To provide the most up-to-date evidence, we summarize below the most recent evidence.

1a.4.3. Provide the citation(s) for the evidence.

1. Black JM, Cuddigan JE, Walko MA, Didier LA, Lander MJ, Kelpe MR. Medical device related pressure ulcers in hospitalized patients. *Int Wound J.* 2010;7(5):358-365.
2. WOCN 2016 Guideline for Prevention and Management of Pressure Injuries (Ulcers): An Executive Summary. *J Wound Ostomy Continence Nurs.* 2017;44(3):241-246.
3. Registered Nurses' Association of Ontario. *Risk assessment and prevention of pressure ulcers (Revised).* Toronto, Canada: Registered Nurses' Association of Ontario;2005.
4. Institute for Clinical Systems Improvement (ICSI). Health Care Protocol: Pressure Ulcer Prevention and Treatment Protocol. http://lnx.mednemo.it/wp-content/uploads/2010/11/febb7157.pdf. Published January 2012. Accessed August 17, 2020.
5. Reddy M, Gill SS, Rochon PA. Preventing pressure ulcers: a systematic review. *Jama.* 2006;296(8):974-984.
6. Posthauer ME, Banks M, Dorner B, Schols JM. The role of nutrition for pressure ulcer management: national pressure ulcer advisory panel, European pressure ulcer advisory panel, and pan pacific pressure injury alliance white paper. *Adv Skin Wound Care.* 2015;28(4):175-188; quiz 189-190.
7. Chou R, Dana T, Bougatsos C, et al. Pressure ulcer risk assessment and prevention: a systematic comparative effectiveness review. *Ann Intern Med.* 2013;159(1):28-38.
8. Chou R, Dana T, Bougatsos C, et al. AHRQ Comparative Effectiveness Reviews. In: *Pressure Ulcer Risk Assessment and Prevention: Comparative Effectiveness.* Rockville (MD): Agency for Healthcare Research and Quality (US); 2013.
9. Tayyib N, Coyer F. Effectiveness of Pressure Ulcer Prevention Strategies for Adult Patients in Intensive Care Units: A Systematic Review. *Worldviews Evid Based Nurs.* 2016;13(6):432-444.
10. Gillespie BM, Walker RM, Latimer SL, et al. Repositioning for pressure injury prevention in adults. *Cochrane Database Syst Rev.* 2020;6(6):Cd009958.
11. Avsar P, Moore Z, Patton D, O'Connor T, Budri AM, Nuget L. Repositioning for preventing pressure ulcers: a systematic review and meta-analysis. *J Wound Care.* 2020;29(9):496-508.
12. Moore ZE, Webster J. Dressings and topical agents for preventing pressure ulcers. *Cochrane Database Syst Rev.* 2018;12(12):Cd009362.

13. Padula WV. Effectiveness and Value of Prophylactic 5-Layer Foam Sacral Dressings to Prevent Hospital-Acquired Pressure Injuries in Acute Care Hospitals: An Observational Cohort Study. *J Wound Ostomy Continence Nurs.* 2017;44(5):413-419.
14. Elsabrout K, Orbacz E, McMahon LA, Apold S. Large-Scale Hospital Mattress Switch-Out Leads to Reduction Hospital-Acquired Pressure Ulcers: Operationalization of a Multidisciplinary Task Force. *Worldviews Evid Based Nurs.* 2018;15(3):161-169.
15. Kitazawa T, Matsumoto K, Fujita S, et al. Perioperative patient safety indicators and hospital surgical volumes. *BMC research notes.* 2014;7:117.
16. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer.* 2014;120(7):1035-1041.
17. Choi J, Bergquist-Beringer S, Staggs VS. Linking RN workgroup job satisfaction to pressure ulcers among older adults on acute care hospital units. *Res Nurs Health.* 2013;36(2):181-190.
18. Mills AC, Gillespie KN. Effect of Magnet hospital recognition on 2 patient outcomes. *J Nurs Care Qual.* 2013;28(1):17-23.
19. Rosen AK, Singer S, Zhao S, Shokeen P, Meterko M, Gaba D. Hospital Safety Climate and Safety Outcomes: Is There a Relationship in the VA? *Medical Care Research and Review.* 2010;67(590).
20. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Med Care.* 2006;44(9):850-861.
21. Chen Q, Hanchate A, Shwartz M, et al. Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicator Rates Among Veteran Dual Users. *Am J Med Qual.* 2013;29(4):335-343.
22. Warshawsky N, Rayens MK, Stefaniak K, Rahman R. The effect of nurse manager turnover on patient fall and pressure ulcer rates. *J Nurs Manag.* 2013;21(5):725-732.
23. Park SH, Boyle DK, Bergquist-Beringer S, Staggs VS, Dunton NE. Concurrent and lagged effects of registered nurse turnover and staffing on unit-acquired pressure ulcers. *Health Serv Res.* 2014;49(4):1205-1225.
24. Blegen MA, Goode CJ, Park SH, Vaughn T, Spetz J. Baccalaureate education in nursing and patient outcomes. *J Nurs Adm.* 2013;43(2):89-94.
25. Boyle DK, Bergquist-Beringer S, Cramer E. Relationship of Wound, Ostomy, and Continence Certified Nurses and Healthcare-Acquired Conditions in Acute Care Hospitals. *J Wound Ostomy Continence Nurs.* 2017;44(3):283-292.
26. Irvin C, Sedlak E, Walton C, Collier S, Bernhofer EI. Hospital-acquired pressure injuries: The significance of the advanced practice registered nurse's role in a community hospital. *J Am Assoc Nurse Pract.* 2017;29(4):203-208.
27. Aydin C, Donaldson N, Stotts NA, Fridman M, Brown DS. Modeling hospital-acquired pressure ulcer prevalence on medical-surgical units: nurse workload, expertise, and clinical processes of care. *Health Serv Res.* 2015;50(2):351-373.
28. Brem H, Maggi J, Nierman D, et al. High cost of stage IV pressure ulcers. *Am J Surg.* 2010;200(4):473-477.

29. National Pressure Ulcer Advisory Panel (NPUAP). NPUAP Pressure Injury Stages. https://npiap.com/page/PressureInjuryStages. Published 2016. Accessed August 17, 2020.
30. Gunningberg L, Donaldson N, Aydin C, Idvall E. Exploring variation in pressure ulcer prevalence in Sweden and the USA: benchmarking in action. *J Eval Clin Pract.* 2012;18(4):904-910.
31. Ramanathan R, Leavell P, Wolfe LG, Duane TM. Agency for Healthcare Research and Quality patient safety indicators and mortality in surgical patients. *Am Surg.* 2014;80(8):801-804.
32. Berry M, Kiel D. Falls: Prevention in nursing care facilities and the hospital setting. In: Post T, ed. *UpToDate.* Waltham, MA2020.
33. Brito PA, de Vasconcelos Generoso S, Correia MI. Prevalence of pressure ulcers in hospitals in Brazil and association with nutritional status--a multicenter, cross-sectional study. *Nutrition (Burbank, Los Angeles County, Calif).* 2013;29(4):646-649.
34. Moore Z. US Medicare data show incidence of hospital-acquired pressure ulcers is 4.5%, and they are associated with longer hospital stay and higher risk of death. *Evidence-based nursing.* 2013;16(4):118-119.
35. Lee MK, Dodson TB, Karimbux NY, Nalliah RP, Allareddy V. Effect of occurrence of infection-related never events on length of stay and hospital charges in patients undergoing radical neck dissection for head and neck cancer. *Oral surgery, oral medicine, oral pathology and oral radiology.* 2013;116(2):147-158.
36. Mallow PJ, Pandya B, Horblyuk R, Kaplan HS. Prevalence and cost of hospital medical errors in the general and elderly United States populations. *Journal of medical economics.* 2013;16(12):1367-1378.
37. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality & safety in health care.* 2003;12 Suppl 2:ii58-63.
38. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *Am J Med Qual.* 2017;32(6):583-590.
39. Bauer K, Rock K, Nazzal M, Jones O, Qu W. Pressure Ulcers in the United States' Inpatient Population From 2008 to 2012: Results of a Retrospective Nationwide Study. *Ostomy Wound Manage.* 2016;62(11):30-38.
40. Bath J, Dombrovskiy VY, Vogel TR. Impact of Patient Safety Indicators on readmission after abdominal aortic surgery. *J Vasc Nurs.* 2018;36(4):189-195.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite Measure)

Measure Title: PSI 06 Iatrogenic Pneumothorax Rate (Component Measure)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome**: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process:** [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency:** [6] evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

PSI 06 targets iatrogenic pneumothorax/pneumothoraces caused by diagnostic and therapeutic interventional-based procedures in the hospitalized patient. The most common procedures associated with iatrogenic pneumothoraces (as targeted by this indicator) include central line placement, pacemaker placement or manipulation, barotrauma from positive pressure ventilation, feeding tube placement, and other procedures close to the thoracic cavity. Operator technical skill and experience has been shown to be inversely related to the rate of iatrogenic pneumothorax. Use of ultrasound guidance during central venous catheter placement and judicious site selection (such as use of the internal jugular vein) are associated with lower indicator rates.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the *systematic review of the body of evidence* that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review

Evidence

Source of Systematic Review:

- Title
- Author
- Date
- Citation, including page number
- URL

*

Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.

*

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade

*

Provide all other grades and definitions from the evidence grading system

*

Grade assigned to the **recommendation**
with definition of the grade

*

Provide all other grades and definitions from the recommendation grading system

*

Body of evidence:

- Quantity – how many studies?
- Quality – what type of studies?

*

Estimates of benefit and consistency across studies

*

What harms were identified?

*

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?

*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with process of care

Two studies have shown that implementation of guidelines (American College of Surgeons and UK National Institute for Clinical Excellence)[1,2] is associated with a substantial reduction in the incidence of procedure-associated iatrogenic pneumothorax (i.e., 0 of 1,978 procedures in Cavanna's series, and 0 of 169 procedures in Wigmore's series).[3,4]

Several studies have focused on the potential to prevent iatrogenic pneumothorax. Research findings summarized in a narrative review by Wrightson et al (2010) recommend that the use of a lateral approach (versus posterior approach) to thoracentesis and use of blunt dissection (versus trocar use) for chest tube insertion can reduce the risk of pneumothorax. A meta-analysis of 6 randomized trials with 579 participants showed that the risk of any procedural complication, including pneumothorax, is reduced when internal jugular (IJ) venous catheters are inserted with real-time ultrasound guidance (relative risk [RR] = 0.43; 95% confidence interval [CI] = 0.22-0.87).[5] A subsequent randomized trial that involved 450 critically ill adults who underwent real-time ultrasound-guided cannulation of the IJ vein and 450 comparison patients for whom the landmark technique was used confirmed that ultrasound reduces the risk of pneumothorax (i.e., from 2.4% to 0%, P < .001) and other complications.[6] More recently, a meta-analysis of 24 studies (of which only two were randomized trials) reported pneumothorax rates following 6,605 unique thoracentesis procedures.[7] Of the 6 comparative studies that reported pneumothorax rates with and without ultrasonography guidance, ultrasonography-guided thoracentesis was associated with a significantly lower risk of pneumothorax than unguided thoracentesis (OR 0.3, 95% CI 0.2-0.7). Among these studies, two randomized controlled trials found a similar effect size, but the difference was not significant (OR 0.3, 95% CI 0.0-2.8). A more recent retrospective study of 394 ICU patients at a single tertiary referral center found that the use of real-time ultrasound guidance was associated with a lower rate of iatrogenic pneumothorax compared to ultrasound-marked procedures (0.63% vs. 4.43%; p=0.02).[8]

Buckley et al. measured the rate of iatrogenic pneumothorax to evaluate quality improvement efforts based on a Plan-Do-Study-Act (PDSA) methodology to improve clinical outcomes at a single institution.[9] Beginning in 2005, the PDSA intervention consisted of providing quality improvement education to residents and fellows in the medical intensive care unit (MICU) and providing training on central venous catheter insertion techniques known to reduce iatrogenic pneumothorax rates. Iatrogenic pneumothorax rates decreased from 0.31% at the beginning of the intervention to 0.17% approximately 3 years after the intervention was first implemented (chi-square with Yates correction, p < 0.001). Beginning in 2007, other improvements and areas of evaluation included expanding ultrasound catheter insertion guidance to fellows and residents and advocating for the use of peripheral rather than central catheters.

Association with hospital and health system characteristics

Several studies have explored the association between health system characteristics and the prevalence of iatrogenic pneumothorax. For example, in 2010, Rivard and colleagues compared the relationship between Agency for Healthcare Research and Quality (AHRQ) Patient Safety Indicators (PSIs) rates, including PSI6 (version 2.1), and various hospital characteristics in VA vs. community non-Federal hospitals.[10] Using VA and Nationwide Inpatient Sample (NIS) data from 2003 through 2004 (n=116 VA hospitals, n=992 community non-Federal hospitals from NIS), they found that the risk-adjusted rate iatrogenic pneumothorax was 1.34 per 1000 (95% CI 1.14 to 1.53) in VA hospitals and 0.78 per 1000 (95% CI 0.72 to 0.83) in non-VA hospitals (from the NIS dataset). In both VA and non-VA (NIS) hospitals, rates of PSI 06 were significantly higher in major teaching hospitals than in nonteaching hospitals [(VA OR 2.51, 95% CI 1.30 to 4.86) (NIS OR 1.59, 95% CI 1.33 to 1.91)]. Rates of this indicator were significantly associated with nurse staffing hours in VA hospitals only (OR 1.03, 95% CI 1.00 to 1.07).

Using fiscal year 2004 data from the Veterans Health Administration (VA) and calendar year 2003 Healthcare Cost and Utilization Project (HCUP) National Inpatient Sample (NIS), Rivard and co-authors (2008) used the AHRQ PSIs (version 2.1, rev. 3a) to compare the rates of adverse patient safety events between federal (VA) and nonfederal (NIS) hospitals and between teaching and nonteaching hospitals.[11] They found that risk-adjusted rate of PSI 06 overall was higher in federal hospitals than nonfederal hospitals [(VA 1.3 cases per 1,000 discharges) (NIS 0.8 cases per 1,000 discharges)], though the results were not significant. These findings were consistent across major teaching hospitals [(VA 1.4 cases per 1,000 discharges, 99% CI 1.1 to 1.7) (NIS 1.2 cases per 1,000 discharges, 99% CI 1.0 to 1.4)], minor teaching hospitals [(VA 1.3 cases per 1,000 discharges, 99% CI 0.8 to 1.9) (NIS 0.8 cases per 1,000 discharges, 99% CI 0.7 to 1.0)], and nonteaching hospitals [(VA 0.7 cases per 1,000 discharges, 99% CI 0.2 to 1.1) (NIS 0.6 cases per 1,000 discharges, 99% CI 0.6 to 0.7)]. Among both federal and nonfederal teaching and non-teaching hospitals, major teaching hospitals had higher risk-adjusted PSI 06 rates than nonteaching hospitals; however, the difference was only significant in nonfederal hospitals. The rate of PSI6 was significantly greater in nonfederal major teaching hospitals compared to nonfederal nonteaching hospitals (OR 1.45, 99% CI 1.13 to 1.85, $p<0.01$). The rate of PSI6 was greater in federal major teaching hospitals than in federal nonteaching hospitals, although this relationship was not significant (OR 1.63, 99% CI 0.59 to 4.51).

Chen et al. analyzed rates of PSI 06 (version 3.1a) among veteran dual users (i.e., those with hospitalizations in both the Veterans Health Administration [VA] and the private sector through Medicare fee-for-service coverage) during 2002 to 2007 and found the risk-adjusted rate of PSI 06 in the VA (0.8; 95% CI 0.7 to 0.9) to be significantly higher than in the private sector (0.5; 95% CI 0.5 to 0.6), however when risk-adjusted PSI rates for Medicare hospitalizations were recalculated using VA expected rates, risk-adjusted rates for PSI 06 were no longer significantly different across the two settings.[12] This study found no significant differences in the risk-adjusted odds among dual users of developing PSI 06 between those hospitalized in the VA and those hospitalized in the private sector.

One study by Li et al. compared rates of several PSIs (version 3.0) in Iowa hospitals between 1997 to 2004.[13] The authors examined the difference in PSI rates between critical access hospitals (CAHs) and Rural Prospective Payment System (PPS) hospitals and found that CAHs had significantly better performance than rural PPS hospitals for PSI 06 in 2001, 2003, and 2004 ($p<0.05$). In 2001, the PSI 06 rate was 0.07 cases per 1,000 discharges for CAHs and 0.21 cases per 1,000 discharges for PPS hospitals ($p<0.05$). In 2003, the PSI 06 rate was 0.26 cases per 1,000 discharges for CAHs and 0.46 cases per 1,000 discharges for PPS hospitals ($p<0.05$). In 2004 the PSI 06 rate was 0.14 cases per 1,000 discharges for CAHs and 0.29 cases per 1,000 discharges for PPS hospitals ($p<0.05$). Further analyses found that the odds of poor performance on PSI 06 were significantly lower among CAHs compared to PPS hospitals (OR 0.29, 95% CI 0.15 to 0.56). To examine the effect of CAH conversion on patient safety, Li et al. also compared PSIs within-hospitals before and after conversion from a PPS hospital to a CAH. Conversion from a PPS hospital to a CAH was associated with non-significant improved risk-adjusted rates of iatrogenic pneumothorax. The PSI rate of PPS hospitals decreased by an average of 0.090 cases per 1,000 discharges when they converted to CAHs ($p=0.34$). Of the 66 hospitals that converted, 18 had better performance on PSI 06 after conversion compared to 8 that had worse performance. In adjusted analyses controlling for comorbidities, selection bias and history bias, the odds ratios of poor performance in CAH hospitals compared with rural PPS hospitals was 0.30 (95% CI 0.14 to 0.64).

In an analysis of patient-level Medicare claims data for patients undergoing any of 6 cancer resections for the years 2005-2009, Short et al. found that the iatrogenic pneumothorax rate was lower at high procedure volume hospitals than at low-volume hospitals (0.67% vs 0.76%), at rural hospitals than urban hospitals (0.27% vs 0.69%), and at non-teaching hospitals than teaching hospitals (0.65% vs 0.71%) (statistical values not provided).[14]

In another study, Rosen et al. (2010) used PSI 06 (version 3.1a), to explore the potential relationship between safety climate, as measured through more than 4500 responses to the Patient Safety Climate in Healthcare Organizations survey, and hospital safety performance, and found that among the 30 Veteran's Health Administration hospitals that participated in the survey, the rate of iatrogenic pneumothorax was only marginally associated with the overall 11 dimensions of patient safety culture included in the analysis ($p>0.05$).[15] Analyses were adjusted for major teaching status, metropolitan area and nurse-staffing ratio. The relationship between the indicator rate and patient safety culture dimensions remained non-significant when senior managers and frontline staff were analyzed separately. An additional study by Rosen et al. (2006) used VA Patient Treatment File (PTF) data, to examine risk-adjusted PSIs (version 2.1, revision 2) for acute care VA hospitals for fiscal years 2001 to 2004.[16] The PSIs were assessed to characterize adverse patient safety events, determine trends in PSIs over time, and evaluate potential predictors of hospital safety. The only hospital characteristic (e.g. bed size, teaching status) they found to be associated with PSI 06 rates was a measure of hospital leadership, a component of a quality improvement score given to hospitals.

A study by Anhang Price et al compared patient safety events in VA versus matched non-VA hospitals and did not identify a significant difference in PSI 06 rates ($p=0.177$).[17] Using NIS data from 2000 to 2012, John et al found that the incidence of iatrogenic pneumothorax was higher in teaching hospitals compared to non-teaching hospitals.[18]

Association with other outcomes

Numerous studies have examined the relationship between **iatrogenic pneumothorax** and outcomes including length of stay in the hospital, costs, mortality, and readmissions. Rosen et al. (2013), examined whether PSI events, experienced within index hospitalizations, increased the likelihood of readmission within Veterans Health Administration (VA) hospitals.[19] They found that iatrogenic pneumothorax resulted in significantly higher rates of all-cause readmissions (18.0%) compared to those hospitalizations without an event (14.3%; p<0.0001). In a multivariate analysis using AHRQ comorbidity software (version 3.5) - controlling for age, sex, comorbidities, and other PSI events - hospitalizations with a PSI 06 event were 22% more likely to result in subsequent readmissions (OR 1.22; 95% CI 1.03 to 1.45).

In an analysis of patient-level Medicare claims data for patients undergoing any of 6 cancer resections for the years 2005-2009, Short et al. found that after adjusting for patient (age, sex, race, income), hospital (hospital volume, surgeon volume, surgeon specialty designation, hospital resources, patient characteristics) and tumor factors (tumor stage, site), costs increased significantly with iatrogenic pneumothorax among 3 of the 6 types of cancer resection patients (p<0.01).[14]

Based on an analysis of the 501,908 hospitalizations involving a brain tumor in the NIS between 2002 and 2010, Rahman et al. found that patients with iatrogenic pneumothorax had significantly longer LOS (p < 0.0001) than patients without iatrogenic pneumothorax.[20] Finally, Ramanathan et al. retroactively examined data on surgical patients hospitalized between 2011 and 2012 at an academic medical center and found that hospitalizations that included iatrogenic pneumothorax (PSI 06, version 3.1) were associated with a 13.0 day mean hospital LOS, 0% included an intensive care unit stay, and 0% died in hospital.[21]

Zhan and Miller used AHRQ PSI software on 7.45 million discharges in the HCUP Nationwide Inpatient Sample (NIS, 2000) and found that compared to those that did not experienced a PSI 06 event, those that did had a higher mean (SD) unadjusted length of stay (130.78 [0.25] vs. 40.59 [0.003]), charges ($55,286 [1454] vs. $13,384 [11]), and percent mortality (160.11 [0.59] vs. 20.56 [0.006]).[22] However, statistical differences of these comparisons were not reported. The overall rate of iatrogenic pneumothorax was 0.67 per 1000 discharges at risk.

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 06 event were associated with an additional 1.41 hospital days compared to patients without an event (p=0.006) and an increased risk of 30-day unplanned readmissions (OR=3.30, p<0.001).[23]

Population group disparities

Table 4 presents population group disparities for PSI 06 Iatrogenic Pneumothorax Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 4. PSI 06 Iatrogenic Pneumothorax Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	94976	0.211	0.250
White	7138286	0.253	0.235
Black	1150073	0.202	0.240
Other	111270	0.270	0.252
Asian	127710	0.211	0.177
Hispanic	201697	0.159	0.188
North American Native	67236	0.134	0.153
Gender	*	*	*
Female	4782258	0.269	0.232
Male	4108990	0.211	0.237
Age	*	*	*
<50	533851	0.103	0.227
50-54	286839	0.129	0.235
55-59	427643	0.124	0.201
60-64	528012	0.155	0.233
65-69	1498806	0.221	0.246
70-74	1468061	0.217	0.229
75-79	1318805	0.291	0.243
80-84	1146640	0.328	0.227
85-89	940204	0.333	0.240
90 plus	742387	0.273	0.222

*cell intentionally left blank

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review results presented below constitute the most recent update, conducted in August 2020. Search terms included relevant MeSH terms (pneumothorax, iatrogenic disease) We combined this clinical search string with a MeSH term (postoperative complications) to identify complications following surgery. Search was limited to English publications. We also tested more inclusive search strings. To provide the most up-to-date evidence, we summarize below the most recent evidence.

1a.4.3. Provide the citation(s) for the evidence.

1. American College of Surgeons. Revised Statement on Recommendations for Use of Real-Time Ultrasound Guidance for Placement of Central Venous Catheters. http://www.facs.org/fellows_info/statements/st-60.html Published 2011. Accessed March 4, 2015.
2. National Institute for Clinical Excellence. *Guidance on the Use of Ultrasound Locating Devices for Placing Central Venous Catheters.* London, UK: National Institute for Clinical Excellence,;2002.
3. Cavanna L, Civardi G, Vallisa D, et al. Ultrasound-guided central venous catheterization in cancer patients improves the success rate of cannulation and reduces mechanical complications: a prospective observational study of 1,978 consecutive catheterizations. *World journal of surgical oncology.* 2010;8:91.
4. Wigmore TJ, Smythe JF, Hacking MB, Raobaikady R, MacCallum NS. Effect of the implementation of NICE guidelines for ultrasound guidance on the complication rates associated with central venous catheter placement in patients presenting for routine surgery in a tertiary referral centre. *British journal of anaesthesia.* 2007;99(5):662-665.
5. Hind D, Calvert N, McWilliams R, et al. Ultrasonic locating devices for central venous cannulation: meta-analysis. *Bmj.* 2003;327(7411):361.
6. Karakitsos D, Labropoulos N, De Groot E, et al. Real-time ultrasound-guided catheterisation of the internal jugular vein: a prospective comparison with the landmark technique in critical care patients. *Crit Care.* 2006;10(6):R162.

7. Gordon CE, Feller-Kopman D, Balk EM, Smetana GW. Pneumothorax following thoracentesis: a systematic review and meta-analysis. *Archives of internal medicine.* 2010;170(4):332-339.
8. Helgeson SA, Fritz AV, Tatari MM, Daniels CE, Diaz-Gomez JL. Reducing Iatrogenic Pneumothoraces: Using Real-Time Ultrasound Guidance for Pleural Procedures. *Crit Care Med.* 2019;47(7):903-909.
9. Buckley JD, Joyce B, Garcia AJ, Jordan J, Scher E. Linking residency training effectiveness to clinical outcomes: a quality improvement approach. *Jt Comm J Qual Patient Saf.* 2010;36(5):203-208.
10. Rivard PE, Elixhauser A, Christiansen CL, Shibei Z, Rosen AK. Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Medical care research and review : MCRR.* 2010;67(3):321-341.
11. Rivard PE, Christiansen CL, Zhao S, Elixhauser A, Rosen AK. Is There an Association Between Patient Safety Indicators and Hospital Teaching Status? In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 2: Culture and Redesign).* Rockville (MD)2008.
12. Chen Q, Hanchate A, Shwartz M, et al. Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicator Rates Among Veteran Dual Users. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2013;29(4):335-343.
13. Li P, Schneider JE, Ward MM. Effect of critical access hospital conversion on patient safety. *Health Serv Res.* 2007;42(6 Pt 1):2089-2108; discussion 2294-2323.
14. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer.* 2014;120(7):1035-1041.
15. Rosen AK, Singer S, Shibei Z, Shokeen P, Meterko M, Gaba D. Hospital safety climate and safety outcomes: is there a relationship in the VA? *Medical care research and review : MCRR.* 2010;67(5):590-608.
16. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Medical care.* 2006;44(9):850-861.
17. Anhang Price R, Sloss EM, Cefalu M, Farmer CM, Hussey PS. Comparing Quality of Care in Veterans Affairs and Non-Veterans Affairs Settings. *J Gen Intern Med.* 2018;33(10):1631-1638.
18. John J, Seifi A. Incidence of iatrogenic pneumothorax in the United States in teaching vs. non-teaching hospitals from 2000 to 2012. *J Crit Care.* 2016;34:66-68.
19. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Medical care.* 2013;51(1):37-44.
20. 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.
21. Ramanathan R, Leavell P, Wolfe LG, Duane TM. Agency for Healthcare Research and Quality patient safety indicators and mortality in surgical patients. *Am Surg.* 2014;80(8):801-804.
22. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality & safety in health care.* 2003;12 Suppl 2:ii58-63.

23. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2017;32(6):583-590.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite Measure)

Measure Title: PSI 08 In-Hospital Fall with Fracture Rate (Component Measure)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome:** [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome:** [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.

- **Intermediate clinical outcome:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process**: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency:** [6] evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

5. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

6. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

7. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

Falls leading to hip fracture among hospitalized patients can be prevented by identifying patients at high risk for falling and taking appropriate preventive actions. Interventions that have been shown to decrease the risk of falls in institutional settings include: use of adaptive equipment such as mobility aids; use of safety devices such as bed alarms, call lights and hip protectors; engaging the patient and family in safety; frequent toileting; attention to postoperative medication management (especially polypharmacy and use of select medications); and implementation of a standardized fall prevention protocol. Structural inventions at the hospital level include making the environment safer through use of handrails, no-slip bathing surfaces, improved lighting, and the provision of no-slip footwear.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	*
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions	*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	*
Provide all other grades and definitions from the evidence grading system	*
Grade assigned to the **recommendation** with definition of the grade	*
Provide all other grades and definitions from the recommendation grading system	*
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	*

Systematic Review	Evidence
Estimates of benefit and consistency across studies	*
What harms were identified?	*
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with processes of care

Inpatient hip fractures can be prevented by reducing falls during hospital stays, particularly among elderly patients. Accordingly, several studies have examined the effect of interventions aimed at either preventing in-hospital falls or decreasing the severity of injuries from falls. According to a meta-analysis by Cameron et al., multifactorial interventions reduce fall rates (rate ratio 0.69; 95% CI 0.49 to 0.96; 4 trials, 6478 participants) and the risk of falling (risk ratio 0.73; 95% CI 0.56 to 0.96; 3 trials, 4824 participants) in hospitals.[1] Several other recent studies have reported on interventions that significantly reduced the risk of falls during hospital stays.[2-6] These interventions included bed posters, patient education handouts, plans of care, fall risk alert cards with informational brochures, exercise programs, education programs, hip protectors, and pre-printed care plans for patients identified as at risk of falling. All intervention groups reported a significant reduction in the risk of falling in-hospital compared with the control groups. One non-blinded cluster randomized trial by Drahota et al. in eight hospitals in the United Kingdom between 2010 and 2011 estimated the impact of shock-absorbing flooring and found fewer injuries from falls in the intervention group (22.9%) than in the control group (42.4%) [injury incident rate ratio (IRR) 0.58; 95% CI 0.18 to 1.91].[7] In this study, there were no moderate to major injuries in the intervention group, compared with six in the control group [IRR 1.07; 95% CI 0.64 to 1.81].

Falls can also be prevented through careful attention to postoperative medication management and avoidance of polypharmacy. Two cross-sectional studies of older patients (one conducted in Taiwan and the other in the Netherlands) found that polypharmacy (daily use of ≥4 or 5 medications) is a significant risk factor for falling, and the risk increases with the number of medications used.[8,9]

Pierce et al. conducted univariable and multivariable analyses based on the medical records of patients in a 435-bed university hospital in New Mexico who fell in-hospital in 2010.[10] They found that 25% of falls were associated with injury and 4% were associated with serious injury. Furthermore, patients who reported hitting their head, patients with pre-fall confusion, and patients who received narcotics within 24 hours before falling were more likely to suffer injury than those who did not (OR 6.04, 2.00, and 5.12 respectively). Using multivariable analysis, they confirmed that receiving a narcotic prior to falling was the strongest clinical predictor of fall-related injury (OR 5.38; 95% CI 2.07 to 13.98, $p < 0.001$).

Other studies have examined the cost effectiveness of interventions aimed at preventing in-hospital falls and hip fractures. Stollenwerk et al. conducted a cost effectiveness analysis on the use of hip protectors for hospitalized patients in Germany.[11] They found that hip protectors could prevent 45.4% (95% CI 35.1% to 51.4%) of in-hospital hip fractures and save hospitals €52.2 ($72.60) per patient screened to be at risk of falling. Latimer et al. estimated the cost-effectiveness of a shock-absorbing floor intervention aimed at preventing serious injuries from falls among elderly patients in eight United Kingdom hospitals.[12] They found the shock-absorbing floor to be associated with an £843 cost saving per patient, but a quality-adjusted life year (QALY) loss of 0.006, yielding an incremental cost-effectiveness ratio of £134,903. A third study conducted in two Australian hospitals (n=1,206) by Haines et al. evaluated the cost-effectiveness of two different patient education models for the prevention of in-hospital falls.[3] One model included multimedia patient education materials, while the other combined these materials with a trained health professional follow-up. A control group received usual care and no patient education materials. There was no significant difference in fall rates between the control group and the group with only patient education materials. However, the patients who received multimedia educational materials as well as a healthcare provider follow-up had a significantly lower fall rate (8.72 vs. 4.01 falls per 1,000 patient days, adjusted hazard ratio = 0.43) and lower odds of falling (30 fallers and 280 non-fallers in control group vs. 20 fallers and 260 non-fallers in complete program, adjusted odds ratio = 0.51). If the percent of patients on a hospital ward who fall is 4% or greater, then the complete program of multimedia materials and professional follow-ups is cost-effective and likely to prevent falls and reduce future costs.

Association with hospital and health system characteristics

Several studies have examined the association between PSI 08 and hospital factors, such as staffing and assistance with falls. Staggs et al. conducted a cross-sectional analysis using data from the National Database of Nursing Quality Indicators (NDNQI) to compare assisted falls (falls for which a staff member was present to ease the patient's descent) and unassisted falls that occurred in-hospital.[13] Out of 166,883 falls (3.44 per 1,000 patient-days), 85.5% were unassisted, and unassisted falls had a higher odds of injury (adjusted odds ratio [aOR] 1.39; 95% CI 1.32 to 1.46) than assisted falls. Staggs and Dunton separately analyzed the rate of unassisted falls per inpatient hospital day in 1391 US hospitals in 2011 using data from the NDNQI. In medical-surgical units, each additional registered nurse (RN) hour per patient-day was weakly associated with a 2% (95% CI 0 to 3%) decrease in average fall rate.[14] In step-down and medical units, fall rates depended on the level of staffing: at low staffing levels, fall rates increased as staffing increased, but at moderate to high levels of staffing, the fall rate decreased as staffing increased. Higher levels of non-RN staffing were generally associated with higher fall rates.

Using fiscal year 2004 data from the Veterans Health Administration (VA) and calendar year 2003 data from the Healthcare Cost and Utilization Project (HCUP) Nationwide Inpatient Sample (NIS), Rivard and co-authors (2008)[6] found that the overall risk-adjusted rate of PSI 08 (version 2.1) was similar (0.3 per 1,000 discharges) between the VA and the NIS. Chen et al. analyzed rates of PSI 08 (version 3.1a) among veteran dual users (i.e., those with hospitalizations in both the Veterans Health Administration [VA] and private sector hospitals through Medicare) during 2002-2007 and found observed and risk-adjusted rates of PSI 08 in the VA (0.8 and 0.6 respectively; 95% CI 0.5 to 0.7) to be significantly lower than in the private sector (0.7 and 0.4 respectively; 95% CI 0.3 to 0.5).[15] However, they found no significant differences in the adjusted odds of developing PSI 08 between VA and private sector hospitalizations, among dual users (OR 1.20; 95% CI 0.80 to 1.81). Rosen et al. (2006) used VA Patient Treatment File (PTF) data to examine risk-adjusted PSI rates (version 2.1) across acute care VA hospitals for fiscal years 2001 to 2004;[16] the only hospital characteristic (e.g. bed size, teaching status) associated with PSI 08 rates was hospital location (i.e., metropolitan status) ($p < 0.01$). Finally, Rosen et al. (2010) used PSI 08 (version 3.1a) to explore associations between safety climate, as measured through more than 4,500 responses to the Patient Safety Climate in Healthcare Organizations survey, and hospital safety performance. Among the 30 VA hospitals that participated in the survey, the rate of postoperative hip fractures was not significantly associated with the 11 dimensions of patient safety culture included in the analysis.[17]

Analyzing NIS data on all hospitalizations between 2002 and 2010 involving coiling or clipping unruptured cerebral aneurysms, Fargen et al found hospital type (teaching vs nonteaching) and hospital bed size were not associated with PSI 08 incidence in this sample.[18] In an analysis of Medicare claims data for patients undergoing any of 6 cancer resections in 2005-2009, Short et al. found that the postoperative hip fracture rate was higher at high procedure volume hospitals than at low-volume hospitals (0.01% vs 0%), at urban hospitals than rural hospitals (0.01% vs 0%), and at non-teaching hospitals than teaching hospitals (0.02% vs 0.01%) (statistical values not provided).[19]

Association with other outcomes

Zapatero et al. analyzed clinical data (n=2,134,363) from the Basic Minimum Data Set (BMDS) which is part of the Spanish National Health Service. A total of 1127 (0.057%) patients were coded using the AHRQ PSI (version 4.3) for an in-hospital hip fracture.[20] Patients with an in-hospital hip fracture had a higher mortality rate (27.9% vs 9.4%, $p < 0.001$) and a longer mean length of stay (20.7 days vs 9.8 days, $p<0.001$) than those who did not experience a hip fracture. Costs were also higher for patients who experienced PSI 08 than for patients who did not (6927€ versus 3730€). Murray et al. studied 2003 data on Australian patients with hip fractures and found that several outcome measures were worse after hospital-acquired hip fractures than after hip fractures in the community.[21] These outcomes included higher in-hospital mortality (28% vs 9%, $p=0.03$), higher prevalence of discharge to nursing homes (33% vs 12%, $p=0.02$), lower prevalence of discharge back into the community (23% vs. 72%, $p <0.001$), lower prevalence of return to preadmission activities of daily living (ADL) (9% vs 56%, $p <0.001$), and higher median length of stay after fracture (46 versus 32 days, $p<0.01$). Based on an analysis of the 501,908 hospitalizations involving a brain tumor in the NIS between 2002 and 2010, Rahman et al found that patients with postoperative hip fractures had significantly longer stays than patients without postoperative hip fractures (7.6 vs 6.5 day mean length of stay, respectively; $p < 0.0001$).[22]

In an analysis of Medicare claims data for patients undergoing any of 6 cancer resections for the years 2005-2009, Short et al. found that after adjusting for patient (age, sex, race, income), hospital (hospital volume, surgeon volume, surgeon specialty designation, hospital resources, patient characteristics) and tumor factors (tumor stage, site), costs increased significantly with postoperative hip fracture among 2 of the 6 types of cancer resection patients ($p < 0.01$).[19] Zhan and Miller used AHRQ PSI software on 7.45 million discharges in the HCUP NIS (2000) and found that compared to those that did not experienced a PSI 08 event, those that did had a higher mean (SD) unadjusted length of stay (160.37 [0.58] vs. 50.39 [0.007]), charges $52,224 [1784] vs. $24,594 [35]), and percent mortality (90.93 [0.92] vs. 10.70 [0.01]).[23] Finally, Gray et al. retrospectively examined 57,000 inpatient discharges at six hospitals between July 2012 and June 2014 and found that hospitalizations with a PSI 08 event were associated with an additional 3.46 hospital days compared to patients without a PSI 03 event ($p<0.001$).[24]

Population group disparities

Table 5 presents population group disparities for PSI 08 In-Hospital Fall with Fracture Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 5. PSI 08 In-Hospital Fall with Fracture Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	84356	0.083	0.117
White	6266488	0.120	0.115
Black	1015824	0.040	0.050
Other	98333	0.102	0.109
Asian	112160	0.054	0.055
Hispanic	180702	0.066	0.083
North American Native	60503	0.066	0.080
Gender	*	*	*
Female	4168509	0.125	0.103
Male	3649857	0.085	0.114
Age	*	*	*
<50	459025	0.037	0.112
50-54	255344	0.086	0.149

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
55-59	383777	0.050	0.085
60-64	476589	0.076	0.097
65-69	1352972	0.070	0.105
70-74	1318953	0.090	0.113
75-79	1166006	0.121	0.097
80-84	989580	0.145	0.113
85-89	793387	0.187	0.116
90 plus	622733	0.148	0.095

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review results presented below constitute the most recent update, conducted in August 2020. Search terms included relevant MeSH terms (accidental falls, hip fracture) and keywords (inpatient falls). We combined this clinical search string with MeSH terms (patient admission) and keywords (hospitals, patient admission, inpatient, patient safety, or quality) to identify studies examining inpatient care. Search was limited to English publications. We also tested more inclusive search strings. To provide the most up-to-date evidence, we summarize below the most recent evidence.

1a.4.3. Provide the citation(s) for the evidence.

1. Cameron ID, Gillespie LD, Robertson MC, et al. Interventions for preventing falls in older people in care facilities and hospitals. *Cochrane Database Syst Rev.* 2012;12:CD005465.
2. Dykes PC, Carroll DL, Hurley A, et al. Fall prevention in acute care hospitals: a randomized trial. *Jama.* 2010;304(17):1912-1918.

3. Haines TP, Hill AM, Hill KD, et al. Cost effectiveness of patient education for the prevention of falls in hospital: economic evaluation from a randomized controlled trial. *BMC Med.* 2013;11:135.
4. Haines TP, Bennell KL, Osborne RH, Hill KD. Effectiveness of targeted falls prevention programme in subacute hospital setting: randomised controlled trial. *Bmj.* 2004;328(7441):676.
5. Healey F, Monro A, Cockram A, Adams V, Heseltine D. Using targeted risk factor reduction to prevent falls in older in-patients: a randomised controlled trial. *Age Ageing.* 2004;33(4):390-395.
6. Miake-Lye IM, Hempel S, Ganz DA, Shekelle PG. Inpatient fall prevention programs as a patient safety strategy: a systematic review. *Ann Intern Med.* 2013;158(5 Pt 2):390-396.
7. Drahota AK, Ward D, Udell JE, et al. Pilot cluster randomised controlled trial of flooring to reduce injuries from falls in wards for older people. *Age Ageing.* 2013;42(5):633-640.
8. Lai SW, Liao KF, Liao CC, Muo CH, Liu CS, Sung FC. Polypharmacy correlates with increased risk for hip fracture in the elderly: a population-based study. *Medicine (Baltimore).* 2010;89(5):295-299.
9. Ziere G, Dieleman JP, Hofman A, Pols HA, van der Cammen TJ, Stricker BH. Polypharmacy and falls in the middle age and elderly population. *Br J Clin Pharmacol.* 2006;61(2):218-223.
10. Pierce JR, Jr., Shirley M, Johnson EF, Kang H. Narcotic administration and fall-related injury in the hospital: implications for patient safety programs and providers. *The International journal of risk & safety in medicine.* 2013;25(4):229-234.
11. Stollenwerk B, Waldeyer R, Klein-Meding C, Müller D, Stock S. Cost effectiveness of external hip protectors in the hospital setting: a modeling study. *Nurs Econ.* 2014;32(2):89-98.
12. Latimer N, Dixon S, Drahota AK, Severs M. Cost--utility analysis of a shock-absorbing floor intervention to prevent injuries from falls in hospital wards for older people. *Age Ageing.* 2013;42(5):641-645.
13. Staggs VS, Mion LC, Shorr RI. Assisted and unassisted falls: different events, different outcomes, different implications for quality of hospital care. *Jt Comm J Qual Patient Saf.* 2014;40(8):358-364.
14. Staggs VS, Dunton N. Associations between rates of unassisted inpatient falls and levels of registered and non-registered nurse staffing. *Int J Qual Health Care.* 2014;26(1):87-92.
15. Chen Q, Hanchate A, Shwartz M, et al. Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicator Rates Among Veteran Dual Users. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2013;29(4):335-343.
16. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Medical care.* 2006;44(9):850-861.
17. Rosen AK, Singer S, Shibei Z, Shokeen P, Meterko M, Gaba D. Hospital safety climate and safety outcomes: is there a relationship in the VA? *Med Care Res Rev.* 2010;67(5):590-608.
18. Fargen KM, Rahman M, Neal D, Hoh BL. Prevalence of patient safety indicators and hospital-acquired conditions in those treated for unruptured cerebral aneurysms: establishing standard performance measures using the Nationwide Inpatient Sample database. *J Neurosurg.* 2013;119(4):966-973.
19. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer.* 2014;120(7):1035-1041.

20. Zapatero A, Barba R, Canora J, et al. Hip fracture in hospitalized medical patients. *BMC Musculoskelet Disord.* 2013;14:15.

21. Murray GR, Cameron ID, Cumming RG. The consequences of falls in acute and subacute hospitals in Australia that cause proximal femoral fractures. *Journal of the American Geriatrics Society.* 2007;55(4):577-582.

22. 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.

23. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality & safety in health care.* 2003;12 Suppl 2:ii58-63.

24. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2017;32(6):583-590.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite Measure)

Measure Title: PSI 09 Perioperative Hemorrhage and Hematoma Rate (Component Measure)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome**: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process**: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency**: [6] evidence not required for the resource use component.

- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

8. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

9. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

10. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

11. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

PSI 09 is intended to capture preventable and significant perioperative hemorrhage or hematoma events that are in excess of what is expected for the surgery type. The intent is to capture bleeding-related events that are severe or involve a delay in diagnosis or treatment requiring reoperation, as these events are

associated with a significant increase in risk to the patient. Such events are often associated with the technical skill and judgment of the surgeon, especially when the hemorrhage is not recognized during the initial procedure and requires reoperation on a subsequent day. Best practices to prevent perioperative hemorrhage and hematoma include taking steps to address and avoid technical errors such as inadequate ligation, cauterization, clipping, or stapling of blood vessels; failure to recognize transection of a minor vessel; or defects in vascular anastomoses. Additional patient management processes that can contribute to PSI 09 events include excessive anticoagulation; inadequate correction or reversal of coagulopathy; failure to replace clotting factors in cases involving large-volume blood loss; and intraoperative hypothermia.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	*
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	*
Provide all other grades and definitions from the evidence grading system	*
Grade assigned to the **recommendation** with definition of the grade	*
Provide all other grades and definitions from the recommendation grading system	*
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	*
Estimates of benefit and consistency across studies	*
What harms were identified?	*

Systematic Review	Evidence
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with processes of care

Several studies have identified variation in postoperative bleeding rates based on the operative approach chosen by the surgeon, including the use of transcervical arterial ligation for transoral robotic surgery,[1] robotic distal pancreatectomy,[2] and percutaneous approach for trans-femoral transcatheter aortic valve implantation.[3] A Cochrane review assessed the comparative effects of three anti-fibrinolytic drugs (aprotinin, tranexamic acid (TXA), and epsilon aminocaproic acid [EACA]) on blood loss during surgery, the need for red blood cell (RBC) transfusion, and other adverse events. The review concluded that aprotinin and EACA significantly reduced the need for reoperation due to bleeding, but that TXA did not.[4] Wiegmann et al analyzed claims data from 2007 through 2017 (22 million covered lives) found that patients preoperatively prescribed antithrombotic agents were 2.3 times more likely to develop postoperative bleeding complications ($p<0.0001$).[5]

Spertus et al. (2015) used percutaneous coronary intervention data from 9 US hospitals to compare the use of bleeding avoidance strategies and bleeding rates before and after implementation of a validated risk model to determine individual patient risk of bleeding [developed by the American College of Cardiology's National Cardiovascular Data Registry (NCDR) Catheterization PCI Registry]. They compared 7408 pre-intervention procedures with 3529 post-intervention procedures and found that the use of the risk stratification protocol was also associated with lower bleeding rates compared to non-interventional sites (1.0% v 1.7%; odds ratio 0.56, 0.40 to 0.78; 0.62, 0.44 to 0.87), after adjustment.[6]

A limited number of older studies evaluated the actual occurrence of process failures in association with PSI 09 events. In a case control study involving 1,025 Medicare discharges from acute-care hospitals in California and Connecticut in 1994, nurse-identified process of care failures were relatively frequent among major surgical cases with postprocedural hemorrhage or hematoma (29/44=66%), after excluding patients who had hemorrhage or hematoma at admission.[7] Specifically, "problems with technical care during a procedure were present in 12 of 17 surgical… cases of postprocedural hemorrhage or hematoma".[8] Physician reviewers identified potential quality problems in 37% of major surgery patients

with this event, versus 2% of unflagged controls.[8] However, cases flagged on this indicator and unflagged controls did not differ significantly on a composite of 17 generic process criteria, confirming previous findings in elderly Medicare beneficiaries from Massachusetts, Alabama, Iowa, and New York.[9]

Association with hospital and health system characteristics

Studies examining the impact of health system characteristics such as teaching status, safety climate, bed size, and nurse staffing hours on PSI 09 rates have been inconclusive.[10-13] Before mandatory present on admission (POA) reporting, rates were significantly higher at major teaching hospitals than at nonteaching hospitals in the Nationwide Inpatient Sample (OR 1.20 [95% CI 1.01 to 1.42]), but not in the Veterans Health Administration. Chen et al. analyzed rates of PSI 09 (version 3.1a) among veteran dual users (i.e., those with hospitalizations in both the VA and the private sector with Medicare coverage) during 2002 to 2007 and found the risk-adjusted rate of PSI 09 in the VA (3.3; 95% CI 3.0-3.6) to be significantly higher than in the private sector (2.1; 95% CI 1.9-2.4); dual users hospitalized in the VA had 1.73 times higher odds of PSI 09 than those hospitalized in the private sector (95% CI 1.48-2.03).[12] Rivard et al. (2010) examined over 4500 responses to the Patient Safety Climate in Healthcare Organizations survey **and found** that the PSI 09 rate was not significantly associated with any of the 11 dimensions of patient safety culture, adjusting for major teaching status, metropolitan area, and nurse-staffing ratio (p>0.10 for all comparisons).[10] A study using the national inpatient data from the Japanese Diagnosis Procedure Combination database reported postoperative bleeding and perforation in 331 (4.4%) and 13 patients (0.2%) who underwent colorectal endoscopic submucosal dissections (n=7567). "Multivariable logistic regression analysis showed that the very high hospital volume group had a significantly lower proportion of severe postoperative bleeding than the very low hospital volume group (OR = 0.48 [95 % CI, 0.27-0.83]; p = 0.009)".[14]

Association with other outcomes

PSI 09 events are associated with a number of important and significant patient harms such as increased postoperative infection, hypovolemic or hemorrhagic shock, reoperation, complications from blood transfusion (such as transfusion-related acute cardiac overload [TACO] and transfusion-related lung injury [TRALI]), mortality and resource use.[11,15-25]

Research has established associations between PSI 09 and other outcomes, including hospital readmissions, costs, length of stay, and mortality.[17-20,26] Cases from the 2000 Nationwide Inpatient Sample that were flagged by this PSI had 3.0% excess mortality, 3.9 days of excess hospitalization, and $21,431 in excess hospital charges, relative to carefully matched controls that were not flagged.[19] This finding was confirmed in the Veterans Health Administration system, where cases that were flagged by this PSI in 2001 had 5.1-8.0% excess mortality, 3.9-4.7 days of excess hospitalization, and $7,863-10,012 in excess hospital costs, relative to propensity-matched or multivariable regression-adjusted controls that were eligible but not flagged.[11] In another study based on State Inpatient Databases from seven states that permit linkage of serial hospitalizations, PSI 09 was associated with risk ratios of 1.03 (NS) for inpatient death, 1.18 (p<0.01) for readmission within three months, and 1.10 (NS) for readmission within one month, after adjusting for age, gender, payer, comorbidities, specific surgical DRGs, and APR-DRG severity levels.[21] Similarly, in a multivariable analysis of Veterans Health Administration data, hospitalizations with PSI 09 were 60% more likely to result in a readmission within 30 days than eligible hospitalizations without PSI 09 (18.8% versus 11.3%; OR=1.60, 95% CI 1.40 to 1.83), after adjusting for age, sex, comorbidities, and other PSI events

(Rosen et al., 2013).[17] Ramanathan et al. (2014) retroactively examined data on surgical patients hospitalized between 2011 and 2012 at a single academic medical center and found that hospitalizations with PSI 09 (version 3.1) were associated with a mean hospital LOS of 22.1 days, 64.5% included an intensive care unit stay, and 3.2% died in hospital.[22] 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 09 event were associated with an additional 2.23 hospital days compared to patients without a PSI 03 event ($p<0.001$).[27]

Several other studies have focused on narrower clinical cohorts. In an analysis of patient-level Medicare claims data for patients undergoing any of 6 cancer resections in 2005-2009, Short et al. found that after adjusting for patient factors (age, sex, race, income), hospital factors (hospital volume, surgeon volume, surgeon specialty designation, hospital resources, patient characteristics) and tumor factors (tumor stage, site), costs increased significantly in association with postoperative hemorrhage or hematoma for four of the six types of cancer resection patients ($p<0.001$).[18] Based on an analysis of the 501,908 hospitalizations involving a brain tumor in the NIS between 2002 and 2010, Rahman et al. (2013) found that patients with postoperative hemorrhage or hematoma had significantly longer length-of-stay (LOS) (13.1 days vs 6.5 days; $p < 0.0001$), on average, than patients without this complication.[20] In another NIS-based study limited to patients with breast cancer hospitalized for a mastectomy in 2011, Nwaogu et al. (2015) reported a 1.3 day increase in the mean length of stay ($P < 0.0001$), a $5495 increase in the mean cost per hospital stay ($P < 0.0001$), and a reoperation rate of 2.5% (42 of 201) associated with a bleeding complication (as defined by ICD-9-CM codes 998.11, 998.12, 39.98, and 86.04).[16] De la Garza-Ramos and colleagues (2016) estimated the incidence of in-hospital morbidity and mortality following surgery for malignant brain tumors using the NIS from 2002 to 2011; patients who had experienced a hemorrhage/hematoma complication (based on an expanded list of ICD-9-CM codes [998.1–998.13] compared to PSI 09) had 3.3 times higher odds of mortality (95% CI 1.6–6.6) than those who did not experience that surgical complication.[23] Finally, Ang and colleagues (2015) used 2013 data from the Florida Agency for Health Care Administration to evaluate trauma mortality using the AHRQ PSIs. Of the 939 PSI 09 events (version 4.5) in 50,596 trauma patients, there were 101 deaths. With an adjusted "failure to prevent" observed-to-expected ratio of 3.53, PSI 09 had the strongest influence on trauma mortality among the 10 PSIs reviewed.[28]

Population group disparities

Table 6 presents population group disparities for PSI 09 Perioperative Hemorrhage and Hematoma Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 6. PSI 09 Perioperative Hemorrhage and Hematoma Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	35227	2.498	2.642

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
White	1990623	2.396	2.579
Black	216349	3.480	2.921
Other	28434	3.622	3.352
Asian	26454	3.553	3.235
Hispanic	42594	2.676	2.315
North American Native	16747	2.986	2.917
Gender	*	*	*
Female	1240467	2.102	2.636
Male	1115961	3.015	2.633
Age	*	*	*
<50	99339	3.755	2.697
50-54	65309	3.491	2.854
55-59	102184	3.269	2.678
60-64	131922	2.903	2.582
65-69	554705	2.331	2.618
70-74	497712	2.429	2.608
75-79	384242	2.800	2.718
80-84	263996	2.379	2.540
85-89	165309	2.057	2.587
90 plus	91710	1.178	2.476

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review results presented below constitute the most recent update, conducted in August 2020. Search terms included relevant MeSH terms (hematoma, hemorrhage, hypovolemic shock, postoperative, perioperative, or surgical complications). We combined this clinical search string with MeSH terms (hospitals, patient admission, inpatient, indicator, epidemiol*, statistic, patient safety, AHRQ, prevalence, incidence, or utilization) to identify studies examining quality of inpatient care. The search was limited to English-language publications. For completeness we also tested more inclusive search strings. Below we have provided a summary of the most up-to-date evidence.

1a.4.3. Provide the citation(s) for the evidence.

1. Sharbel DD, Abkemeier M, Sullivan J, et al. Transcervical arterial ligation for prevention of postoperative hemorrhage in transoral oropharyngectomy: Systematic review and meta-analysis. *Head Neck.* 2020.
2. Lyu Y, Cheng Y, Wang B, Zhao S, Chen L. Comparison of 3 Minimally Invasive Methods Versus Open Distal Pancreatectomy: A Systematic Review and Network Meta-Analysis. *Surg Laparosc Endosc Percutan Tech.* 2020.
3. Abdelaziz HK, Megaly M, Debski M, et al. Meta-Analysis Comparing Percutaneous to Surgical Access in Trans-Femoral Transcatheter Aortic Valve Implantation. *Am J Cardiol.* 2020;125(8):1239-1248.
4. Henry DA, Carless PA, Moxey AJ, et al. Anti-fibrinolytic use for minimising perioperative allogeneic blood transfusion. *Cochrane Database Syst Rev.* 2011(3):CD001886.
5. Wiegmann AL, Khalid SI, Coogan AC, et al. Antithrombotic prescriptions for many general surgery patients significantly increases the likelihood of post-operative bleeding complications. *Am J Surg.* 2020;219(3):453-459.
6. Spertus JA, Decker C, Gialde E, et al. Precision medicine to improve use of bleeding avoidance strategies and reduce bleeding in patients undergoing percutaneous coronary intervention: prospective cohort study before and after implementation of personalized bleeding risks. *Bmj.* 2015;350:h1302.
7. Iezzoni LI, Davis RB, Palmer RH, et al. Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes. *Int J Qual Health Care.* 1999;11(2):107-118.
8. Weingart SN, Iezzoni LI, Davis RB, et al. Use of administrative data to find substandard care: validation of the complications screening program. *Med Care.* 2000;38(8):796-806.

9. Iezzoni LI, Lawthers A, Peterson L, McCarthy E, Palmer RH, Cahalane M. *Project to validate the Complications Screening Program: Health Care Financing Administration.* 1998. HCFA Contract 500-94-0055.

10. Rivard PE, Elixhauser A, Christiansen CL, Shibei Z, Rosen AK. Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Med Care Res Rev.* 2010;67(3):321-341.

11. Rivard PE, Christiansen CL, Zhao S, Elixhauser A, Rosen AK. Advances in Patient Safety: Is There an Association Between Patient Safety Indicators and Hospital Teaching Status? In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 2: Culture and Redesign).* Rockville (MD)2008.

12. Chen Q, Hanchate A, Shwartz M, et al. Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicator Rates Among Veteran Dual Users. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2013;29(4):335-343.

13. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Medical care.* 2006;44(9):850-861.

14. Odagiri H, Yasunaga H, Matsui H, Fushimi K, Iizuka T, Kaise M. Hospital volume and the occurrence of bleeding and perforation after colorectal endoscopic submucosal dissection: analysis of a national administrative database in Japan. *Dis Colon Rectum.* 2015;58(6):597-603.

15. In: *WHO Guidelines for Safe Surgery 2009: Safe Surgery Saves Lives.* Geneva2009.

16. Nwaogu IY, Bommarito K, Olsen MA, Margenthaler JA. Economic impact of bleeding complications after mastectomy. *J Surg Res.* 2015;199(1):77-83.

17. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Med Care.* 2013;51(1):37-44.

18. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer.* 2014;120(7):1035-1041.

19. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality & safety in health care.* 2003;12 Suppl 2:ii58-63.

20. 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.

21. Friedman B, Encinosa W, Jiang HJ, Mutter R. Do patient safety events increase readmissions? *Med Care.* 2009;47(5):583-590.

22. Ramanathan R, Leavell P, Wolfe LG, Duane TM. Agency for Healthcare Research and Quality patient safety indicators and mortality in surgical patients. *Am Surg.* 2014;80(8):801-804.

23. De la Garza-Ramos R, Kerezoudis P, Tamargo RJ, Brem H, Huang J, Bydon M. Surgical complications following malignant brain tumor surgery: An analysis of 2002-2011 data. *Clin Neurol Neurosurg.* 2016;140:6-10.

24. Mulholland MW, Doherty GM. *Complications in Surgery.* Philadelphia, PA: Lippincott, Williams & Wilkins; 2011.

25. Brunicardi F, Andersen D, Billiar T. *Schwartz's Principles of Surgery, 10e.* Columbus, OH: McGraw-Hill Education; 2015.

26. Ramanathan R, Leavell P, Stockslager G, Mays C, Harvey D, Duane TM. Validity of Agency for Healthcare Research and Quality Patient Safety Indicators at an academic medical center. *Am Surg.* 2013;79(6):578-582.

27. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2017;32(6):583-590.

28. 28. Ang D, McKenney M, Norwood S, et al. Benchmarking statewide trauma mortality using Agency for Healthcare Research and Quality's patient safety indicators. *J Surg Res.* 2015;198(1):34-40.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite Measure)

Measure Title: PSI 10 Postoperative Acute Kidney Injury Requiring Dialysis Rate (Component)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Outcome: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process:** [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency:** [6] evidence not required for the resource use component.

- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- **Process measures incorporating Appropriate Use Criteria:** See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

12. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

13. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

14. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

15. Measures of efficiency combine the concepts of resource use **and** quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

PSI 10 focuses on patients without severe chronic kidney disease at baseline who undergo an elective surgical procedure and then develop acute kidney failure (also referred to as acute kidney injury [AKI] or acute renal failure [ARF]) severe enough to require dialysis as a postoperative complication. It is thought that through better perioperative care, many of these events are preventable. Best practices to prevent postoperative kidney failure include identifying patients at risk (e.g. older age, hypovolemia, infection, etc.); avoiding nephrotoxic medications or using them with caution (e.g. ACE inhibitors, aminoglycosides,

NSAIDs, intravenous contrast, etc.); and using volume expansion, vasodilators, and inotropes as needed to avoid hypovolemia and hypotension.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	*
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	*
Provide all other grades and definitions from the evidence grading system	*
Grade assigned to the **recommendation** with definition of the grade	*
Provide all other grades and definitions from the recommendation grading system	*
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	*
Estimates of benefit and consistency across studies	*
What harms were identified?	*
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with process of care

Several studies have identified variation in AKI rates based on the operative approach chosen by the surgeon, including laparoscopic vs. open bariatric surgery (0.94% versus 3.87%; p < .01),[1] non-gastric vs. gastric bypass (0.82% versus 1.54%; p < .01)[1] and laparoscopic vs. open abdominal procedures (aRR 0.52 [95% CI 0.47 to 0.58]).[2] Among patients who underwent lung resection surgery in a tertiary care academic center (2006-2010; n=1129), postoperative AKI was associated with preoperative use of angiotensin II receptor blockers (OR 2.2, 95% CI: 1.1-4.4), intraoperative hydroxyethyl starch administration (OR 1.5, 95% CI: 1.1-2.1), and thoracoscopic (versus open) procedures (OR 0.37, 95% CI: 0.15-0.90). AKI in this study was also associated with increased rates of tracheal reintubation (12% vs 2%, P < 0.001) and postoperative mechanical ventilation (15% vs 3%, P < 0.001), suggesting that AKI may precede or occur concurrently with postoperative respiratory failure.[3] In a study of 119 cases flagged by PSI 10 (v3.1) from 28 acute care hospitals in the Veterans Health Administration, there were 73 true positives with AKI, of whom 37% died and 26% were discharged on dialysis. AKI was most commonly attributed to perioperative renal hypoperfusion (84% of true positives), followed by nephrotoxins (33%) including contrast (11%).[23] A recent systematic review including one randomized trial and four observational studies (n=10,468) concluded that preoperative aspirin (at any dose) is associated with reduced incidence of postoperative AKI (OR 0.68, 95% CI, 0.51-0.91, p=0.008).[4]

Studies examining the association between PSI 10 and patient or clinical characteristics found that strong predictors of AKI include male gender, hypertension, higher body mass index, ascites, preoperative sepsis, active congestive heart failure, renal insufficiency, peripheral vascular disease, increased age, Medicare payer, alcohol abuse, chronic lung disease, diabetes, smoking, schizophrenia, functional dependence, ventilator dependence, myocardial infarction, bleeding disorders, hematocrit, chronic steroid use, and cancer.[1-3,5-9]

Association with hospital and health system characteristics

Various hospital and health system characteristics have been shown to be associated with postoperative acute kidney injury (AKI). Using 2000-2010 NIS data, Spolverate et al found acute renal **f**ailure (ARF) occurred in 4.2% of patients who underwent liver resection for malignancy and was less common in high-volume hospitals than in low or intermediate volume hospitals (3.3% versus 4.7% and 5.0%, respectively).[10] In an analysis of Medicare claims data for patients undergoing any of six types of cancer resection (2005-2009), Short et al. found that the rate of **postoperative AKI** was lower at high procedure volume hospitals than at low-volume hospitals (0.06 vs 0.09), at rural hospitals than at urban hospitals (0 vs 0.09), and at non-teaching hospitals than at teaching hospitals (0.07 vs 0.10) (statistical values not provided).[11] Rivard et al. (2010) examined the relationship between PSI 10 (V2.1) and hospital characteristics in Veterans Health Administration (VHA) and nonfederal community hospitals, using data from the VA and the Healthcare Cost and Utilization Project (HCUP) Nationwide Inpatient Sample (NIS, 2003-2004). Rates of this indicator were not significantly associated with bed size, nurse staffing hours, or teaching status in either VA or NIS hospitals.[12] Using the NIS database (2006 to 2008), Masoomi et al analyzed **cl**inical data from morbidly obese patients who underwent bariatric surgery and also found no significant effect for teaching status of the hospital on the rate of ARF.[1]

Rosen et al. (2006) did not find risk-adjusted PSI 10 rates (version 2.1) to be significantly associated with any single hospital characteristic (e.g. bed size, teaching status, location) in the 2001-2004 VA Patient Treatment File (PTF).[13] In another study, Rosen et al. used the PSIs, including PSI 10 (version 3.1), to explore the potential relationship between safety climate, as measured through more than 4500 responses to the Patient Safety Climate in Healthcare Organizations survey, and hospital safety performance. They found that among the 30 VA hospitals that participated in the survey, the PSI 10 rate was not significantly associated with any of the 11 dimensions of patient safety culture ($p>0.10$ for all comparisons). The authors note that due to the small sample size, the relatively low rate of PSIs among VA hospitals, and narrow variation across hospitals in patient safety culture, statistical power to detect associations was limited for most PSIs, including PSI 10.[14]

Finally, a study by Hawkins et al, drawing on NSQIP data (2005-2010) for patients undergoing repairs of ruptured abdominal aortic aneurysms, found that after risk adjustment for factors including age, sex, and method of repair, the odds of renal insufficiency or failure (OR 0.54; 95% CI 0.31 to 0.95; $p = .034$) were significantly less for those operated on by vascular surgeons than for those who were operated on by general surgeons.[15]

Association with other outcomes

Cases from the 2000 Nationwide Inpatient Sample that were flagged by this PSI had 19.8% excess mortality, 8.9 days of excess hospitalization, and $54,818 in excess hospital charges, relative to carefully

matched controls that were not flagged (Zhan and Miller, 2003).[16] More recent studies have confirmed these findings, as summarized below, separately for each outcome.

AKI is associated with significantly increased mortality following a variety of surgical procedures. Ricciardi et al. analyzed NSQIP data (2005-2008) for patients undergoing colorectal surgery and found that patients who experienced postoperative kidney failure had 2.7 times higher odds of 30-day mortality than those who did not (95% CI 1.3 to 5.5).[17] Using the NIS database (2006-2008), Masoomi et al. studied morbidly obese patients who underwent bariatric surgery and found that patients with acute kidney failure had significantly greater in-hospital mortality than those without it (5.69% versus 0.04%, $p < .01$).[1] Similarly, Bensley et al.'s multivariable regression analysis of 450 NSQIP (2005-2010) patients who underwent open surgical repair of thoracoabdominal aortic aneurysms (TAAA) showed that postoperative kidney failure was a strong predictor of perioperative (30-day) mortality (OR=8.4; 95% CI 3.41 to 20.56).[18] Based on data on 15 intra-abdominal general surgery procedure categories (n = 457,656) in the NSQIP database (2005-2010), Kim et al. found that after adjusting for comorbidities and operative factors, perioperative AKI was associated with a 3.5-fold increase in the risk of 30-day mortality (aRR, 3.51, 95% CI 3.29 to 3.74).[2] Similarly, using NSQIP (2005-2006) data, Kheterpal et al. found that all-cause 30-day mortality among patients who developed AKI after general surgery was 42% vs 8.6% for matched cohorts who did not develop AKI (hazard ratio 7.5; 95% CI 5.2–10.8).[8] Corredor et al. conducted a systematic review and meta–analysis based on 9 observational studies with long-term follow-up of 35,021 cardiac surgery patients. Postoperative AKI was associated with a significantly increased risk of long-term mortality (HR 1.68, 95% CI, 1.45-1.95 based on 8 studies). Hobson et al. found that risk-adjusted 90-day mortality was 6.5% for patients with AKI compared to 4.4% for patients without AKI (risk-adjusted rate ratio 1.65, 95% CI 1.48 to 1.84) in a single-center cohort of 50,314 adult surgical patients undergoing major inpatient surgery.[19] In-hospital mortality was also significantly higher in the AKI group (4.2% vs. 2.1%, adjusted risk ratio 2.38, 95% CI 2.00-2.84). 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 10 event were associated with an additional 6.37 hospital days compared to hospitalizations without a PSI 10 event ($p<0.001$), as well as a significantly increased risk of in-hospital mortality (OR=168.91; $p<0.001$).[20]

In a study based on State Inpatient Databases from seven states that permit linkage of serial hospitalizations, PSI 10 was associated with risk ratios of 1.30 for readmission within three months, and 1.09 for readmission within one month, after adjusting for age, gender, payer, comorbidities, specific surgical DRGs, and APR-DRG severity levels).[21] Similarly, in a multivariable analysis of Veterans Health Administration data, hospitalizations with a PSI 10 event had 53% higher odds of resulting in a readmission within 30 days (OR 1.53; 95% CI 1.26 to 1.86), after adjusting for age, sex, comorbidities, and other PSI events (Rosen et al., 2013).[22] In a sample of over 2,000 coronary artery or valve surgery patients from seven hospitals (2008-2010), Brown et. al. found that patients without postoperative AKI had a 30-day readmission rate of 9.3% compared to 16.1%, 21.8%, and 28.6% among patients developing stage 1, 2, and 3 AKI, respectively ($p < 0.001$).[23] Adjusted odds ratios showed a similar progression of risk, with odds ratios of 1.81 (95% CI 1.35 to 2.44), 2.39 (95% CI 1.38 to 4.14), and 3.47 (95% CI 1.85 to 6.50) for patients developing stage 1, 2, and 3 AKI, compared to those without AKI. Among 501,908 hospitalizations with brain tumors in the 2002-2010 NIS, patients with a PSI 10 AKI event had significantly longer LOS (7.6 days vs 6.5 days; $p < 0.0001$) than patients without postoperative AKI.[24] Bath et al. used Medicare data (MedPAR) from 2009 to 2012 and found that the likelihood of 30-day readmission among patients

undergoing abdominal aortic aneurysm repair was greater among patients with postoperative AKI requiring dialysis (OR=1.88, p=0.0001).[25]

In a single-center cohort of 50,314 adult surgical patients, Hobson et al. used regression models to estimate the effect of postoperative AKI on hospital costs.[19] The risk-adjusted average cost of care was $42,600 for patients with any AKI compared with $26,700 for patients without AKI, but the average cost per patient rose to $62,600 for those in the failure (F) category. Short et al. found that after adjusting for patient (age, sex, race, income), hospital (hospital volume, surgeon volume, surgeon specialty designation, hospital resources, patient characteristics) and tumor factors (tumor stage, site), costs increased significantly with PSI 10 among patients with three of six types of cancer resection ($p < 0.001$).[11]

Population group disparities

Table 7 presents population group disparities for PSI 10 Postoperative AKI Requiring Dialysis Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 7: PSI 10 Postoperative AKI Requiring Dialysis Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	22980	1.480	1.678
White	1141162	1.332	1.339
Black	94263	1.772	1.422
Other	14517	0.896	0.730
Asian	11784	1.697	1.289
Hispanic	16533	1.391	1.224
North American Native	7416	1.753	1.581
Gender	*	*	*
Female	708991	0.990	1.351
Male	599664	1.814	1.337
Age	*	*	*
<50	40893	0.954	1.374
50-54	30371	0.823	1.086

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
55-59	49802	0.944	1.339
60-64	66518	1.443	1.414
65-69	362055	1.190	1.395
70-74	319420	1.406	1.386
75-79	230198	1.672	1.280
80-84	131561	1.703	1.338
85-89	59900	1.336	1.218
90 plus	17937	0.781	1.124

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review was conducted in August 2020. Search terms included relevant MeSH terms (Acute kidney injury or failure, or acute renal failure or insufficiency, or kidney tubular necrosis) and MeSH terms (hospitals, patient admission, inpatient, patient safety, quality, and perioperative, postoperative, or surgical complications, and indicator, epidemiol*, statistic, patient safety, AHRQ, prevalence, incidence, or utilization) to identify studies examining quality of inpatient care. The search was limited to English publications.

1a.4.3. Provide the citation(s) for the evidence.

1. Masoomi H, Reavis KM, Smith BR, Kim H, Stamos MJ, Nguyen NT. Risk factors for acute respiratory failure in bariatric surgery: data from the Nationwide Inpatient Sample, 2006-2008. *Surg Obes Relat Dis.* 2013;9(2):277-281.

2. Kim M, Brady JE, Li G. Variations in the risk of acute kidney injury across intraabdominal surgery procedures. *Anesth Analg.* 2014;119(5):1121-1132.
3. Ishikawa S, Griesdale DE, Lohser J. Acute kidney injury after lung resection surgery: incidence and perioperative risk factors. *Anesth Analg.* 2012;114(6):1256-1262.
4. Aboul-Hassan SS, Stankowski T, Marczak J, et al. The use of preoperative aspirin in cardiac surgery: A systematic review and meta-analysis. *J Card Surg.* 2017;32(12):758-774.
5. Liao CC, Shen WW, Chang CC, Chang H, Chen TL. Surgical adverse outcomes in patients with schizophrenia: a population-based study. *Ann Surg.* 2013;257(3):433-438.
6. Gupta H, Ramanan B, Gupta PK, et al. Impact of COPD on postoperative outcomes: results from a national database. *Chest.* 2013;143(6):1599-1606.
7. Ponce BA, Oladeji LO, Raley JA, Menendez ME. Analysis of perioperative morbidity and mortality in shoulder arthroplasty patients with preexisting alcohol use disorders. *Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al].* 2014.
8. Kheterpal S, Tremper KK, Heung M, et al. Development and validation of an acute kidney injury risk index for patients undergoing general surgery: results from a national data set. *Anesthesiology* 2009;110(3):505-515.
9. Zhou X, Zhang Y, Teng Y, et al. Predictors of postoperative acute kidney injury in patients undergoing hip fracture surgery: A systematic review and meta-analysis. *Injury.* 2020.
10. Spolverato G, Ejaz A, Hyder O, Kim Y, Pawlik TM. Failure to rescue as a source of variation in hospital mortality after hepatic surgery. *The British journal of surgery.* 2014;101(7):836-846.
11. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer.* 2014;120(7):1035-1041.
12. Rivard PE, Elixhauser A, Christiansen CL, Shibei Z, Rosen AK. Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Medical care research and review : MCRR.* 2010;67(3):321-341.
13. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Medical care.* 2006;44(9):850-861.
14. Rosen AK, Singer S, Shibei Z, Shokeen P, Meterko M, Gaba D. Hospital safety climate and safety outcomes: is there a relationship in the VA? *Medical care research and review : MCRR.* 2010;67(5):590-608.
15. Hawkins AT, Smith AD, Schaumeier MJ, de Vos MS, Hevelone ND, Nguyen LL. The effect of surgeon specialization on outcomes after ruptured abdominal aortic aneurysm repair. *J Vasc Surg.* 2014;60(3):590-596.
16. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality & safety in health care.* 2003;12 Suppl 2:ii58-63.
17. Ricciardi R, Roberts PL, Read TE, Hall JF, Marcello PW, Schoetz DJ. Which adverse events are associated with mortality and prolonged length of stay following colorectal surgery? *J Gastrointest Surg.* 2013;17(8):1485-1493.
18. Bensley RP, Curran T, Hurks R, et al. Open repair of intact thoracoabdominal aortic aneurysms in the American College of Surgeons National Surgical Quality Improvement Program. *J Vasc Surg.* 2013;58(4):894-900.

19. Hobson C, Ozrazgat-Baslanti T, Kuxhausen A, et al. Cost and Mortality Associated With Postoperative Acute Kidney Injury. *Ann Surg.* 2014.

20. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2017;32(6):583-590.

21. Friedman B, Encinosa W, Jiang HJ, Mutter R. Do patient safety events increase readmissions? *Medical care.* 2009;47(5):583-590.

22. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Medical care.* 2013;51(1):37-44.

23. Brown JR, Parikh CR, Ross CS, et al. Impact of perioperative acute kidney injury as a severity index for thirty-day readmission after cardiac surgery. *The Annals of thoracic surgery.* 2014;97(1):111-117.

24. 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.

25. Bath J, Dombrovskiy VY, Vogel TR. Impact of Patient Safety Indicators on readmission after abdominal aortic surgery. *J Vasc Nurs.* 2018;36(4):189-195.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite Measure)

Measure Title: PSI 11 Postoperative Respiratory Failure Rate (Component Measure)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome:** [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process**: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency**: [6] evidence not required for the resource use component.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome:** [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.

- **Intermediate clinical outcome**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process**: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency**: [6] evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- **Process measures incorporating Appropriate Use Criteria:** See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

16. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

17. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

18. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

Respiratory failure—usually defined as unplanned intubation or prolonged ventilation—is considered to be the most serious of the postoperative respiratory complications because it represents the "end stage" of several types of pulmonary complications (e.g., pneumonia, aspiration, pulmonary edema, ARDS) and it often results in prolonged morbidity, mortality, and associated costs. Healthcare facilities can decrease postoperative respiratory failure rates by adopting and following guidelines for assessing perioperative pulmonary risk and implementing recommended preventive strategies for high-risk patients. Careful management of blood products and fluid resuscitation in the perioperative setting may reduce the risk of postoperative respiratory failure due to adult respiratory distress syndrome (ARDS).

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	*
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions	*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	*
Provide all other grades and definitions from the evidence grading system	*
Grade assigned to the **recommendation** with definition of the grade	*
Provide all other grades and definitions from the recommendation grading system	*
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	*

Systematic Review	Evidence
Estimates of benefit and consistency across studies	*
What harms were identified?	*
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with process of care

Numerous studies have identified associations between specific intraoperative risk factors and postoperative respiratory failure.[1-5] 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 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).[1] Hughes et al. identified intraoperative risk factors for the postoperative development of ARDS among 89 patients admitted to the ICU with postoperative respiratory failure. 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).[2] 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).[3] Blum et al. found that among 50,367 patient admissions for common adult surgical procedures between 2004 and 2009, the number of different anesthetics administered during the admission was associated with higher risk of ARDS (OR=1.37).[1] 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.[4] 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 postoperative respiratory failure (personal communication; manuscript under review).

Two studies describe quality improvement interventions that resulted in decreased rates of acute respiratory failure.[6,7] 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.[6] 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.[7] 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, include acute respiratory inadequacy.[8] 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 nasogastric tube decompression 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.[9]

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 et al. 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 0.846 to 0.926).[10] 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 VA and NIS data from 2003 through 2004 (n=116 VA hospitals, n=992 community non-Federal hospitals), Rivard et al. reported lower risk-adjusted rates of postoperative respiratory failure 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).[11] Another study involving 4,581 staff surveys from 30 VA hospitals (2005-2006) by Rosen et al. found that

there was no association between hospital safety climate (overall or for various climate dimensions) and individual hospital-level PSIs, including postoperative respiratory failure.[12]

Association with other outcomes

Several studies found that postoperative respiratory failure is associated with longer length of stay.[10,13-15] In a multivariable analysis of NIS data from 2002-2010, Rahman et al. found that length of stay was significantly longer for patients with postoperative respiratory failure (median 8.0 days) compared to those without respiratory failure (median 4.0 days, p<0.0001).[10] 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).[14] 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 postoperative pulmonary complications (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).[15]

Several studies also found that that postoperative respiratory failure is associated with higher 30-day readmission rates.[13,16,17] In three studies included in a recent literature review by Sabate et al., the estimated increased costs in U.S. dollars associated with postoperative respiratory failure 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.[13] 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),[17] with an adjusted odds ratio of 1.39 (95% CI 1.25 to 1.54). 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).[16] 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).[18]

Four different population-based studies have demonstrated that postoperative respiratory failure 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 acute respiratory failure had significantly greater in-hospital mortality than those who did not develop this complication (5.69% versus 0.04%, p<0.01).[19] Based on an analysis of data from 165,600 senior patients undergoing non-emergent major general surgeries from the ACS NSQIP dataset, Gajdos et al. found that reintubation had one of the highest failure-to-rescue rates among all postoperative complications (25.6%).[14] 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).[20] 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).[21] 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.[22]

Population group disparities

Table 8 presents population group disparities for PSI 11 Postoperative Respiratory Failure Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 8. PSI 11 Postoperative Respiratory Failure Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	19662	4.781	5.820
White	920041	5.507	5.198
Black	78549	8.453	6.046
Other	11585	5.697	4.831
Asian	9340	5.782	4.749
Hispanic	14199	7.465	5.687
North American Native	6124	7.348	5.748
Gender	*	*	*
Female	605665	5.006	5.197
Male	453835	6.751	5.386
Age	*	*	*
<50	39287	7.865	5.757
50-54	27247	7.487	5.417
55-59	42943	7.778	5.431
60-64	55682	7.669	5.192
65-69	307397	4.447	5.242

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
70-74	262105	5.135	5.305
75-79	180021	5.916	5.212
80-84	95877	6.394	5.217
85-89	39029	8.840	5.424
90 plus	9912	8.676	4.986

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review results presented below constitute the most recent update, conducted in August 2020. Search terms included relevant MeSH terms (respiratory distress syndrome, adult) and keywords (post-operative respiratory failure, postoperative respiratory failure, postoperative acute respiratory, acute respiratory failure, respiratory distress syndrome, ARDS, reintubation, prolonged intubation, delayed extubation). Studies focused on early extubation or immediate extubation were excluded, as were those focused on obstetric, peripartum or neonatal care. We combined this clinical search string with MeSH terms (patient admission) and keywords (hospitals, patient admission, inpatient) to identify studies examining inpatient care. Search was limited to English publications. We also tested more inclusive search strings. To provide the most up-to-date evidence, we summarize below the most recent evidence.

1a.4.3. Provide the citation(s) for the evidence.

1. 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.

2. Hughes CG, Weavind L, Banerjee A, Mercaldo ND, Schildcrout JS, Pandharipande PP. Intraoperative risk factors for acute respiratory distress syndrome in critically ill patients. *Anesth Analg.* 2010;111(2):464-467.
3. Shalev D, Kamel H. Risk of Reintubation in Neurosurgical Patients. *Neurocritical care.* 2014.
4. Attaallah AF, Vallejo MC, Elzamzamy OM, Mueller MG, Eller WS. Perioperative risk factors for postoperative respiratory failure. *J Perioper Pract.* 2019;29(3):49-53.
5. Chandler D, Mosieri C, Kallurkar A, et al. Perioperative strategies for the reduction of postoperative pulmonary complications. *Best Pract Res Clin Anaesthesiol.* 2020;34(2):153-166.
6. Braddock CH, 3rd, Szaflarski N, Forsey L, Abel L, Hernandez-Boussard T, Morton J. The TRANSFORM Patient Safety Project: A Microsystem Approach to Improving Outcomes on Inpatient Units. *J Gen Intern Med.* 2014.
7. Cassidy MR, Rosenkranz P, McCabe K, Rosen JE, McAneny D. I COUGH: reducing postoperative pulmonary complications with a multidisciplinary patient care program. *JAMA surgery.* 2013;148(8):740-745.
8. Guimaraes MM, El Dib R, Smith AF, Matos D. Incentive spirometry for prevention of postoperative pulmonary complications in upper abdominal surgery. *Cochrane Database Syst Rev.* 2009(3):CD006058.
9. Lawrence VA, Cornell JE, Smetana GW. Strategies to reduce postoperative pulmonary complications after noncardiothoracic surgery: systematic review for the American College of Physicians. *Ann Intern Med.* 2006;144(8):596-608.
10. 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.
11. Rivard PE, Elixhauser A, Christiansen CL, Shibei Z, Rosen AK. Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Medical care research and review : MCRR.* 2010;67(3):321-341.
12. Rosen AK, Singer S, Shibei Z, Shokeen P, Meterko M, Gaba D. Hospital safety climate and safety outcomes: is there a relationship in the VA? *Medical care research and review : MCRR.* 2010;67(5):590-608.
13. Sabate S, Mazo V, Canet J. Predicting postoperative pulmonary complications: implications for outcomes and costs. *Case reports in anesthesiology.* 2014;27(2):201-209.
14. Gajdos C, Kile D, Hawn MT, Finlayson E, Henderson WG, Robinson TN. Advancing age and 30-day adverse outcomes after nonemergent general surgeries. *J Am Geriatr Soc.* 2013;61(9):1608-1614.
15. Marda M, Pandia MP, Rath GP, Bithal PK, Dash HH. Post-operative pulmonary complications in patients undergoing transoral odontoidectomy and posterior fixation for craniovertebral junction anomalies. *Journal of anaesthesiology, clinical pharmacology.* 2013;29(2):200-204.
16. Lawson EH, Hall BL, Louie R, et al. Association between occurrence of a postoperative complication and readmission: implications for quality improvement and cost savings. *Ann Surg.* 2013;258(1):10-18.

17. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Medical care*. 2013;51(1):37-44.

18. Bath J, Dombrovskiy VY, Vogel TR. Impact of Patient Safety Indicators on readmission after abdominal aortic surgery. *J Vasc Nurs.* 2018;36(4):189-195.

19. Masoomi H, Reavis KM, Smith BR, Kim H, Stamos MJ, Nguyen NT. Risk factors for acute respiratory failure in bariatric surgery: data from the Nationwide Inpatient Sample, 2006-2008. *Surg Obes Relat Dis.* 2013;9(2):277-281.

20. Rahmanian PB, Kroner A, Langebartels G, Ozel O, Wippermann J, Wahlers T. 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.* 2013;17(2):319-326; discussion 326-317.

21. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2017;32(6):583-590.

22. Bensley RP, Curran T, Hurks R, et al. Open repair of intact thoracoabdominal aortic aneurysms in the American College of Surgeons National Surgical Quality Improvement Program. *Journal of vascular surgery.* 2013;58(4):894-900.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite Measure)

Measure Title: PSI 12 Perioperative Pulmonary Embolism or Deep Vein Thrombosis Rate (Component)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome:** [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process**: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency:** [6] evidence not required for the resource use component.
- For measures derived from *patient reports*, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

19. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

20. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

21. Measures of efficiency combine the concepts of resource use **and** quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

Deep vein thrombosis (DVT) is the formation of a blood clot in a deep vein—usually in the leg or pelvic veins. The most serious complication of a DVT is that the clot dislodges and travels to the lungs, becoming a pulmonary embolus (PE). Venous thromboembolism (VTE), which encompasses both DVT and PE, is

common in the perioperative setting, especially after high-risk operations, and can be deadly. Clinical trials have demonstrated that mechanical and pharmacologic interventions can substantially reduce the risk of perioperative VTE among moderate and high-risk surgical patients, especially when these interventions are initiated before or immediately after surgery and continued until or after discharge. Case control studies have demonstrated that early ambulation after surgery can further reduce the risk of perioperative VTE among high-risk surgical patients who receive appropriate mechanical or pharmacologic prophylaxis. Effective and safe prophylactic measures are now available for most high-risk patients, and numerous evidence-based guidelines have been published for the prevention of VTE (most notably by the American College of Chest Physicians and the American Academy of Orthopedic Surgeons).

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	*
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	*
Provide all other grades and definitions from the evidence grading system	*
Grade assigned to the **recommendation** with definition of the grade	*
Provide all other grades and definitions from the recommendation grading system	*
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	*
Estimates of benefit and consistency across studies	*

Systematic Review	Evidence
What harms were identified?	*
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with process of care

A recent systematic review including 19 studies of 11,430 patients concluded that primary prophylaxis in ambulatory cancer patients can significantly reduce the risk of VTE.[1] Several recent studies examined the impact of efforts to improve VTE prophylaxis adherence, tracking changes in incidence over time as processes improved. Most of these studies reported favorable results, with the notable exception of cancer patients. For example, implementation of "mandated risk assessment" with computerized DVT prophylaxis order entry at a tertiary cancer center increased use of prophylaxis without reducing VTE incidence,[2] whereas similar protocols reduced the incidence of postoperative VTE on an vascular surgery service from 1.49% to 0.38%,[3] at a large Russian medical center from 0.88% to 0.42%,[4] and at a large medical center in Abu Dhabi from 0.9-3.1% to 0.1-0.2%.[5] Nelson et al. (2015) analyzed 2006-2011 surgical registry data on colorectal surgery from Washington and reported that use of in-hospital postoperative VTE chemoprophylaxis increased from 59.6% to 91.4%, but 90-day VTE rates did not decrease.[6] Heslin et al. reported that among 12 surgical services in a single institution the most common contributing factor for PSI 12 was "failure to follow protocol," but they did not report the impact of improved adherence on PSI 12.[7] Hussey et al tested an alpha version of the AHRQ QI Toolkit in a one-year quality improvement initiative at an academic medical center. After the electronic medical record was revised so that DVT prophylaxis would be a mandatory part of the order set, PSI 12 rates decreased from 20.7 to 15.9.[8] A similar clinical decision support intervention at the University of Pennsylvania was associated with increased use of "recommended" prophylaxis (from 32.3% to 60.0%) and a concurrent drop in PSI 12 rates from 21.8 to 17.3.[9] The University of California recently reported that a five-campus collaborative effort to improve VTE risk stratification and prophylaxis achieved a 23.8% relative reduction in the incidence of PSI 12 in 2014 relative to 2011.[10] A similar program at Boston Medical Center, which also included an emphasis on early ambulation, was associated with an 84% decrease in DVT incidence (from 1.9% to 0.3%) and a 55% decrease in PE incidence (from 1.1% to 0.5%), lowering the observed-to-expected VTE ratio from 3.41 to 0.94.

In a series of studies from The John Hopkins, use of risk-appropriate VTE prophylaxis in surgical patients increased from 26% (42 of 161) at baseline to 68% (178 of 262) within 12 months, and to 85% after implementation of computer-based "smart order sets."[14] A retrospective review of 92 patients diagnosed with hospital-acquired VTE found that only 43 (47%) received defect-free care, while 49 (53%) had potentially preventable VTE.[11] On the trauma service, 56.0% of residents prescribed "optimal, risk-appropriate" VTE prophylaxis, while attending physicians had a compliance rate of 74.2% (interquartile range, 72.6%-77.3%), indicating that resident practice variation may be an important contributor to VTE events at teaching hospitals.[12] Lau et al. (2015) reported that a performance feedback scorecard with individual peer-to-peer coaching increased the percentage of these residents providing defect-free care from 45% to 78% and reduced the incidence of postoperative VTE from 0.81% to 0.38-0.39%.[12]

AHRQ's Evidence-based Practice Review on Patient Safety summarized the state of the field: "Evidence generally finds that the use of aspirin following these surgical procedures – either as the sole prophylaxis in combination with other pharmacologic agents or in conjunction with mechanical prophylaxis – is equivalent to other agents or has a better safety profile."[13] An older report from the AHRQ's Evidence-based Practice Review on Patient Safety noted that many hospitalized patients are not given risk-appropriate VTE prophylaxis. A 2008 study across 32 countries found that only 59% of at-risk surgical and 40% of at-risk medical patients received guideline-recommended VTE prophylaxis, and a 2004 United States registry study found that only 42% of patients diagnosed with DVT during a hospitalization had received prophylaxis…"[14] Similar findings have been reported from Europe[15] and from 28 Veterans Health Administration hospitals, where "accounting for contraindications and early VTE occurrence, a total of 78% of cases [with PSI 12] and 80% of controls [without PSI 12] were appropriately managed."[16]

Delayed Ambulation

Based on observational data from case control studies and longitudinal intervention studies, delayed ambulation is an independent risk factor for VTE after orthopedic surgery, even accounting for appropriate pharmacologic prophylaxis. In a case-control study of patients undergoing total knee arthroplasty (TKA) in 15 teaching hospitals, among PSI 12 cases with an objectively documented acute VTE within 9 days of surgery (N=130) and randomly selected controls (N=463), only 68% ambulated on day 1 or 2 after surgery despite all patients receiving thromboprophylaxis (pharmacologic in 80%, mechanical alone in 20%). Factors significantly associated with VTE (after adjusting for age, sex, history of VTE, and BMI) were bilateral TKA (OR=4.2; 95% CI: 1.9-9.1), receipt of pharmacological prophylaxis (OR=0.5; 95% CI: 0.3-0.8), and ambulation by postoperative day 2 (OR=0.3; 95% CI: 0.1-0.9).[17] In an earlier case control study based on a sampling frame with 25,388 Medicare fee-for-service beneficiaries who underwent unilateral total hip arthroplasty (THA) in any nonfederal hospital in California, White et al. compared processes of care between 297 randomly selected cases with VTE within 3 months after surgery and 592 randomly selected controls. Factors independently associated with VTE included initial ambulation before day 2 after surgery (OR=0.7; 95% CI 0.5–0.9), use of pneumatic compression (among patients with body-mass index <25; OR=0.3; 95% CI 0.2–0.6), and use of warfarin after discharge (OR=0.6; 95% CI 0.4–1.0).[18] These studies suggest a population fraction of post-arthroplasty VTE attributable to delayed ambulation of at least 10% and perhaps over 40%.

Two studies have reported single-center results of prospectively implementing early ambulation postoperative care protocols. Chandrasekaran et al. found that getting patients out of bed or walking for at least 15–30 minutes twice on the first day after TKA significantly reduced the odds of asymptomatic or symptomatic VTE (OR=0.35; 95% CI: 0.13-0.94) compared with the previous practice of confining patients to bed on that day.[19] Similarly, Pearse et al. implemented a treatment protocol that involved showering and walking up to 30 meters within 24 hours after TKA, and observed a substantial reduction in the odds of

asymptomatic or symptomatic DVT (OR=0.04; 95% CI 0.004-0.30).[20] These findings are supported by several cohort studies summarized in a recent structured review.[21]

Patient and Clinical Risk Factors

Studies have shown variation in PE/DVT by procedure type, suggesting the importance of risk adjustment.[22-25] Total operative time is also associated with increased VTE risk. Kim et al. (2015) reported that the risk of VTE in NSQIP data increased in a stepwise manner with the procedure standardized duration of general anesthesia time.[26] These findings were confirmed by Daley et al. (2015), using a measure of whether total operative time exceeded the upper 95% confidence limit of its expected value.[27]

Several studies have examined the association between patient characteristics and rates of pulmonary embolism and deep vein thrombosis. Associations between PSI 12 and patient characteristics have been found for black race (for post-surgical DVT but not PE),[28] gender,[23,12] age,[9,29] obesity,[14] and select comorbidities (postoperative infection or stroke,[8] disseminated cancer,[8] dependent functional status,[8] return to operating room,[8] preoperative hyponatremia,[13] irritable bowel disease,[15] and congestive heart failure and cancer.[9] Other preoperative risk factors for VTE were identified in studies included: age, ASA risk classification (for colorectal surgery), white race (for esophageal surgery), body mass index (for hysterectomy and colorectal and bariatric surgery), cancer (for craniotomy and hysterectomy) and disseminated cancer (for colorectal surgery), chronic steroid use, emergent or non-elective surgery, open (versus laparoscopic) surgery (for colorectal and bariatric surgery), duration of pre-surgical hospitalization, preoperative sepsis, previous cardiac surgery, weight loss, hypoalbuminemia (for colorectal surgery), history of prior VTE, operation for inflammatory disease (for colorectal surgery), transfer from acute care hospital (for craniotomy), dependent functional status (for craniotomy), or individual comorbidities such as peripheral vascular disease and prior stroke (for craniotomy).[6,30-39] Risk models have been developed and validated for VTE; Obi et al. (2015) and Hachey et al. (2016) validated the Caprini VTE risk assessment model among critically ill surgical patients and after lung cancer resection.[40,41]

Some of the identified risk factors are at least partially under providers' control, and may account for some of the observed hospital-level variation in PSI 12 rates (e.g., pre-surgical days, duration of general anesthesia or surgery, open versus laparoscopic approach, and postoperative complications such as prolonged mechanical ventilation and unplanned reintubation). Surgical duration is an especially noteworthy factor because of its association with resident involvement in surgery.

Association with hospital and health system characteristics

Several studies have examined the influence of various hospital and health system characteristics on the rate of postoperative PE and DVT. One study demonstrated that hospitals with higher percentages of registered nurses with baccalaureate or higher degrees had lower rates of PSI 12,[42] while studies **were inconclusive** regarding the impact of hospital factors such as being within the VA healthcare system,[43-45] teaching status,[44,46] **bed size,[46] location,[46] nurse staffing hours,[43] safety climate[47] and the** implementation of duty-hour regulations.[48] Another study found lower rates of postoperative PE and DVT at low procedure volume hospitals (compared to high-volume hospitals), rural hospitals (compared to urban hospitals), and non-teaching hospitals (compared to teaching hospitals), but statistical test values were not provided.[19]

Association with other outcomes

Cases from the HCUP Nationwide Inpatient Sample that were flagged by this PSI in 2000 had 6.6% excess mortality, 5.4 days of excess hospitalization, and $21,709 in excess hospital charges, relative to carefully matched controls that were not flagged.[49] This finding was confirmed in the Veterans Affairs hospital system, where cases that were flagged by this PSI in 2001 had 6.1% excess mortality, 4.5-5.5 days of excess hospitalization, and $7,205-9,064 in excess hospital costs, relative to carefully matched controls that were not flagged[43]. Carey and Stefos re-estimated the financial impact of each PSI 12 event in the VA system in 2007 as $17,453-18,935, using more sophisticated cost accounting and econometric methods.[50] In another study based on HCUP SID from seven states in 2004 that permit linkage of serial hospitalizations, this indicator was associated with risk ratios of 1.35 for inpatient death, 1.28 for readmission within three months, and 1.25 for readmission within one month (after adjusting for age, gender, payer, comorbidities, specific surgical DRGs, and APR-DRG severity levels).[51] Similarly, in a multivariable analysis of Veterans Health Administration data from 2003-2007, hospitalizations with a PSI 12 event were 33% more likely to result in a readmission within 30 days (OR 1.33; 95% CI 1.23-1.44), after adjusting for age, sex, comorbidities, and other PSI events.[52]

Several other studies have focused on narrower clinical cohorts, with similar results. Bohensky et al. examined cost and length of stay (LOS) following complications in 139,031 knee arthroscopy cases in the Victorian Admitted Episodes Dataset (2000 to 2009).[53] VTE events were the most common complication (0.3%) and the cumulative excess 30-day cost of VTE was $3227 (95% CI $3211-3244). Patients who experienced VTE also had longer median LOS (6 days vs. 1 day, p<0.01) than those without VTE. Ramanan et al. used 2007-2009 National Surgical Quality Improvement Program (NSQIP) data on patients undergoing vascular surgery to show that VTE events increased overall mortality risk among patients with DVT (1.5% to 6.2%) or PE (1.5% to 5.7%), compared to those without VTE.[22] Using data from the NSQIP Semi Annual Reports for 197 US and Canadian hospitals (2007-2008), Borgi et al. demonstrated that VTE events were positively and statistically significantly associated with postoperative mortality (regression slope 0.393; 95% CI 0.235 to 0.551, p<0.0001).[54] In an analysis of Medicare claims data for patients undergoing any of 6 cancer resections in 2005-2009, Short et al. found that after adjusting for patient factors (age, sex, race, income), hospital factors (hospital volume, surgeon volume, surgeon specialty designation, hospital resources, patient characteristics) and tumor factors (tumor stage, site), costs increased significantly in association with postoperative VTE for all six types of surgery (p<0.001).[55] Based on an analysis of the 501,908 hospitalizations involving a brain tumor in the NIS between 2002 and 2010, Rahman et al. (2013) found that patients with postoperative DVT or PE had significantly longer length-of-stay, on average, than patients without these complications (10.4 vs 6.3 days and 8.8 vs. 6.4 days, respectively; p < 0.0001 for both).[29] 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 12 event were associated with an additional 2.83 hospital days compared to patients without a PSI 12 event (p<0.001).[56]

Population group disparities

Table 9 presents population group disparities for PSI 12 Perioperative PE or DVT Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 9. PSI 12 Perioperative PE or DVT Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	37392	3.370	3.612
White	2122522	3.526	3.509
Black	233166	5.237	4.980
Other	30971	3.390	3.252
Asian	29215	2.909	2.760
Hispanic	46115	3.643	3.797
North American Native	18093	2.542	2.418
Gender	*	*	*
Female	1307262	3.593	3.635
Male	1210212	3.750	3.639
Age	*	*	*
<50	107216	3.516	3.911
50-54	70230	3.303	3.646
55-59	109865	3.632	3.904
60-64	142036	3.816	3.880
65-69	586155	3.337	3.582
70-74	528720	3.567	3.555
75-79	411071	3.985	3.673
80-84	284379	3.847	3.646
85-89	178740	3.989	3.476
90 plus	99062	4.018	3.632

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review was conducted in August 2020. Search terms included relevant MeSH terms (Venous Thromboembolism or VTE, Pulmonary embolus (PE) or embolism, DVT or Thrombosis) with MeSH terms (patient admission, hospitals, inpatient, patient safety, AHRQ) to identify studies examining quality of inpatient care. The search was limited to English-language publications.

1a.4.3. Provide the citation(s) for the evidence.

1. Xin Z, Liu F, Du Y, et al. Primary prophylaxis for venous thromboembolism in ambulatory cancer patients: a systematic review and network meta-analysis. *Ann Palliat Med.* 2020.
2. Kukar M, Asaro J, Aquino A, Groman A, Skitzki J, Kane JM. Incidence of Venous Thromboembolic Events in Mandated Risk Assessment versus Optional DVT Prophylaxis Era at a Large Tertiary Cancer Center. *Am Surg.* 2015;81(9):893-898.
3. Durinka JB, Hecht TE, Layne AJ, et al. Aggressive venous thromboembolism prophylaxis reduces VTE events in vascular surgery patients. *Vascular.* 2016;24(3):233-240.
4. Nazarenko GI, Kleymenova EB, Payushik SA, Otdelenov VA, Sychev DA, Yashina LP. Decision support systems in clinical practice: The case of venous thromboembolism prevention. *The International journal of risk & safety in medicine.* 2015;27 Suppl 1:S104-105.
5. Nimeri AA, Gamaleldin MM, McKenna KL, Turrin NP, Mustafa BO. Reduction of Venous Thromboembolism in Surgical Patients Using a Mandatory Risk-Scoring System: 5-Year Follow-Up of an American College of Surgeons National Surgical Quality Improvement Program. *Clin Appl Thromb Hemost.* 2017;23(4):392-396.
6. Nelson DW, Simianu VV, Bastawrous AL, et al. Thromboembolic Complications and Prophylaxis Patterns in Colorectal Surgery. *JAMA surgery.* 2015;150(8):712-720.
7. Heslin MJ, Taylor B, Hawn MT, et al. A 100% departmental mortality review improves observed-to-expected mortality ratios and University HealthSystem Consortium rankings. *J Am Coll Surg.* 2014;218(4):554-562.

8. Hussey PS, Burns RM, Weinick RM, Mayer L, Cerese J, Farley DO. Using a hospital quality improvement toolkit to improve performance on the AHRQ quality indicators. *Jt Comm J Qual Patient Saf.* 2013;39(4):177-184.

9. Umscheid CA, Hanish A, Chittams J, Weiner MG, Hecht TEH. Effectiveness of a novel and scalable clinical decision support intervention to improve venous thromboembolism prophylaxis: a quasi-experimental study. *BMC Medical Informatics and Decision Making.* 2012;12(1):92.

10. UC Center for Health Quality and Innovation (CHQI). Report: UC Center for Health Quality and Innovation. Review and Assessment of Grant Impact to Date. https://www.universityofcalifornia.edu/sites/default/files/uc_chqi_report_8.27.14.pdf Published 2014. Accessed September 17, 2020.

11. Haut ER, Lau BD, Kraus PS, et al. Preventability of Hospital-Acquired Venous Thromboembolism. *JAMA surgery.* 2015;150(9):912-915.

12. Lau BD, Streiff MB, Pronovost PJ, Haider AH, Efron DT, Haut ER. Attending Physician Performance Measure Scores and Resident Physicians' Ordering Practices. *JAMA surgery.* 2015;150(8):813-814.

13. Hall KK, Shoemaker-Hunt S, Hoffman L, et al. In: *Making Healthcare Safer III: A Critical Analysis of Existing and Emerging Patient Safety Practices.* Rockville (MD): Agency for Healthcare Research and Quality (US); 2020.

14. Shekelle PG, Wachter RM, Pronovost PJ, et al. Making health care safer II: an updated critical analysis of the evidence for patient safety practices. *Evid Rep Technol Assess (Full Rep).* 2013(211):1-945.

15. Markovic-Denic L, Zivkovic K, Lesic A, Bumbasirevic V, Dubljanin-Raspopovic E, Bumbasirevic M. Risk factors and distribution of symptomatic venous thromboembolism in total hip and knee replacements: prospective study. *Int Orthop.* 2012;36(6):1299-1305.

16. Borzecki AM, Cowan AJ, Cevasco M, et al. Is development of postoperative venous thromboembolism related to thromboprophylaxis use? A case-control study in the Veterans Health Administration. *Jt Comm J Qual Patient Saf.* 2012;38(8):348-358.

17. Sadeghi B, Romano PS, Maynard G, et al. Mechanical and suboptimal pharmacologic prophylaxis and delayed mobilization but not morbid obesity are associated with venous thromboembolism after total knee arthroplasty: a case-control study. *J Hosp Med.* 2012;7(9):665-671.

18. White RH, Gettner S, Newman JM, Trauner KB, Romano PS. Predictors of rehospitalization for symptomatic venous thromboembolism after total hip arthroplasty. *N Engl J Med.* 2000;343(24):1758-1764.

19. Chandrasekaran S, Ariaretnam SK, Tsung J, Dickison D. Early mobilization after total knee replacement reduces the incidence of deep venous thrombosis. *ANZ J Surg.* 2009;79(7-8):526-529.

20. Pearse EO, Caldwell BF, Lockwood RJ, Hollard J. Early mobilisation after conventional knee replacement may reduce the risk of postoperative venous thromboembolism. *J Bone Joint Surg Br.* 2007;89(3):316-322.

21. Zhang ZH, Shen B, Yang J, Zhou ZK, Kang PD, Pei FX. Risk factors for venous thromboembolism of total hip arthroplasty and total knee arthroplasty: a systematic review of evidences in ten years. *BMC Musculoskelet Disord.* 2015;16:24.

22. Ramanan B, Gupta PK, Sundaram A, et al. In-hospital and postdischarge venous thromboembolism after vascular surgery. *J Vasc Surg.* 2013;57(6):1589-1596.

23. Assareh H, Chen J, Ou L, Hollis SJ, Hillman K, Flabouris A. Rate of venous thromboembolism among surgical patients in Australian hospitals: a multicentre retrospective cohort study. *BMJ Open.* 2014;4(10):e005502.

24. White RH, Zhou H, Romano PS. Incidence of symptomatic venous thromboembolism after different elective or urgent surgical procedures. *Thromb Haemost.* 2003;90(3):446-455.

25. Aziz F, Patel M, Ortenzi G, Reed AB. Incidence of Postoperative Deep Venous Thrombosis Is Higher among Cardiac and Vascular Surgery Patients as Compared with General Surgery Patients. *Ann Vasc Surg.* 2015;29(4):661-669.

26. Kim HK, Jung H, Cho J, Huh S, Lee JM, Kim YW. Therapeutic outcomes and thromboembolic events after treatment of acute arterial thromboembolism of the upper extremity. *Ann Vasc Surg.* 2015;29(2):303-310.

27. Daley BJ, Cecil W, Clarke PC, Cofer JB, Guillamondegui OD. How slow is too slow? Correlation of operative time to complications: an analysis from the Tennessee Surgical Quality Collaborative. *J Am Coll Surg.* 2015;220(4):550-558.

28. Adelani MA, Archer KR, Song Y, Holt GE. Immediate complications following hip and knee arthroplasty: does race matter? *J Arthroplasty.* 2013;28(5):732-735.

29. 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.

30. Jamal MH, Corcelles R, Shimizu H, et al. Thromboembolic events in bariatric surgery: a large multi-institutional referral center experience. *Surg Endosc.* 2015;29(2):376-380.

31. Moghadamyeghaneh Z, Hanna MH, Carmichael JC, Nguyen NT, Stamos MJ. A nationwide analysis of postoperative deep vein thrombosis and pulmonary embolism in colon and rectal surgery. *J Gastrointest Surg.* 2014;18(12):2169-2177.

32. Martin JT, Mahan AL, Ferraris VA, et al. Identifying Esophagectomy Patients at Risk for Predischarge Versus Postdischarge Venous Thromboembolism. *The Annals of thoracic surgery.* 2015;100(3):932-938; discussion 938.

33. Kimmell KT, Jahromi BS. Clinical factors associated with venous thromboembolism risk in patients undergoing craniotomy. *J Neurosurg.* 2015;122(5):1004-1011.

34. Hoh DJ, Rahman M, Fargen KM, Neal D, Hoh BL. Establishing standard hospital performance measures for cervical spinal trauma: a Nationwide In-patient Sample study. *Spinal Cord.* 2016;54(4):306-313.

35. Bekelis K, Kalakoti P, Nanda A, Missios S. A Predictive Model of Unfavorable Outcomes After Benign Intracranial Tumor Resection. *World Neurosurg.* 2015;84(1):82-89.

36. Swenson CW, Berger MB, Kamdar NS, Campbell DA, Jr., Morgan DM. Risk factors for venous thromboembolism after hysterectomy. *Obstet Gynecol.* 2015;125(5):1139-1144.

37. Wang L, Pryor AD, Altieri MS, et al. Perioperative rates of deep vein thrombosis and pulmonary embolism in normal weight vs obese and morbidly obese surgical patients in the era post venous thromboembolism prophylaxis guidelines. *Am J Surg.* 2015;210(5):859-863.

38. Greaves SW, Holubar SD. Preoperative Hospitalization Is Independently Associated With Increased Risk for Venous Thromboembolism in Patients Undergoing Colorectal Surgery: A National Surgical Quality Improvement Program Database Study. *Dis Colon Rectum.* 2015;58(8):782-791.

39. Zhang L, Cao H, Chen Y, Jiao G. Risk factors for venous thromboembolism following spinal surgery: A meta-analysis. *Medicine (Baltimore).* 2020;99(29):e20954.

40. Obi AT, Pannucci CJ, Nackashi A, et al. Validation of the Caprini Venous Thromboembolism Risk Assessment Model in Critically Ill Surgical Patients. *JAMA surgery.* 2015;150(10):941-948.

41. Hachey KJ, Hewes PD, Porter LP, et al. Caprini venous thromboembolism risk assessment permits selection for postdischarge prophylactic anticoagulation in patients with resectable lung cancer. *J Thorac Cardiovasc Surg.* 2016;151(1):37-44.e31.

42. Blegen MA, Goode CJ, Park SH, Vaughn T, Spetz J. Baccalaureate education in nursing and patient outcomes. *J Nurs Adm.* 2013;43(2):89-94.

43. Rivard PE, Elixhauser A, Christiansen CL, Shibei Z, Rosen AK. Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Med Care Res Rev.* 2010;67(3):321-341.

44. Rivard PE, Christiansen CL, Zhao S, Elixhauser A, Rosen AK. Advances in Patient Safety: Is There an Association Between Patient Safety Indicators and Hospital Teaching Status? In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 2: Culture and Redesign).* Rockville (MD)2008.

45. Chen Q, Hanchate A, Shwartz M, et al. Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicator Rates Among Veteran Dual Users. *Am J Med Qual.* 2013;29(4):335-343.

46. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Med Care.* 2006;44(9):850-861.

47. Rosen AK, Singer S, Shibei Z, Shokeen P, Meterko M, Gaba D. Hospital safety climate and safety outcomes: is there a relationship in the VA? *Med Care Res Rev.* 2010;67(5):590-608.

48. Shelton J, Kummerow K, Phillips S, et al. Patient safety in the era of the 80-hour workweek. *J Surg Educ.* 2014;71(4):551-559.

49. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Qual Saf Health Care.* 2003;12 Suppl 2(Suppl 2):ii58-63.

50. Carey K, Stefos T. Measuring the cost of hospital adverse patient safety events. *Health Econ.* 2011;20(12):1417-1430.

51. Friedman B, Encinosa W, Jiang HJ, Mutter R. Do patient safety events increase readmissions? *Med Care.* 2009;47(5):583-590.

52. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Med Care.* 2013;51(1):37-44.

53. Bohensky MA, Ademi Z, deSteiger R, et al. Quantifying the excess cost and resource utilisation for patients with complications associated with elective knee arthroscopy: a retrospective cohort study. *The Knee.* 2014;21(2):491-496.

54. Borgi J, Rubinfeld I, Ritz J, Jordan J, Velanovich V. The differential effects of intermediate complications with postoperative mortality. *Am Surg.* 2013;79(3):261-266.

55. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer.* 2014;120(7):1035-1041.

56. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *Am J Med Qual.* 2017;32(6):583-590.

57. Kaafarani HM, A.M. B, K.M. I, et al. Validity of selected Patient Safety Indicators: opportunities and concerns. *J Am Coll Surg* 2011;212(6):924-934.

58. Mull HJ, Borzecki AM, Loveland S, et al. Detecting adverse events in surgery: comparing events detected by the Veterans Health Administration Surgical Quality Improvement Program and the Patient Safety Indicators. *Am J Surg.* 2014;207(4):584-595.

59. Cima RR, Lackore K, Nehring S, et al. How best to measure surgical quality? Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicators (AHRQ-PSI) and the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) postoperative adverse events at a single institution. *Surgery.* 2011;150(5):943-949.

60. Koch CG, Li L, Hixson E, Tang A, Phillips S, Henderson JM. What are the real rates of postoperative complications: elucidating inconsistencies between administrative and clinical data sources. *Journal of the American College of Surgeons.* 2012;214(5):798-805.

61. Sadeghi B, White RH, Maynard G, et al. Improved Coding of Postoperative Deep Vein Thrombosis and Pulmonary Embolism in Administrative Data (AHRQ Patient Safety Indicator 12) After Introduction of New ICD-9-CM Diagnosis Codes. *Med Care.* 2013.

62. Bilimoria KY, Chung J, Ju MH, et al. Evaluation of surveillance bias and the validity of the venous thromboembolism quality measure. *Jama.* 2013;310(14):1482-1489.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite)

Measure Title: PSI 13 Postoperative Sepsis Rate (Component)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome**: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process:** [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency**: [6] evidence not required for the resource use component.
- For measures derived from **patient reports**, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- **Process measures incorporating Appropriate Use Criteria**: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

22. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

23. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

24. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

Prevention of postoperative sepsis requires consideration of patient level risk factors along with review of process measures aimed at neutralizing the threat of bacterial, viral and fungal contamination posed by healthcare staff, the operating room environment and the patient's endogenous skin flora. Appropriate interventions include preoperative disinfection, clipping instead of shaving of hair, use of appropriate surgical attire, skin preparation of both patient and surgeon, timely prophylactic antibiotic therapy as appropriate based on surgery type, and the maintenance of gut function. Other important aspects of care

include monitoring for early signs of infection, maintaining normal blood glucose levels in patients with diabetes, limiting operative traffic, and temperature control (patient and operating room). In combination, these processes can decrease the size of the pathogen innoculum at the surgical site and/or alter the operative site so it is less hospitable to the growth of bacteria and other pathogens. By reducing the incidence of surgical site infections, these processes also help to prevent postoperative sepsis.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the *systematic review of the body of evidence* that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	*
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	*
Provide all other grades and definitions from the evidence grading system	*
Grade assigned to the **recommendation** with definition of the grade	*
Provide all other grades and definitions from the recommendation grading system	*
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	*
Estimates of benefit and consistency across studies	*
What harms were identified?	*

Systematic Review	Evidence
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with process of care

A recent systematic review including 193 studies and over 30 million patients identified several surgery-specific risk factors for post-operative sepsis, including emergency surgery (OR, 3.38, 95% CI, 2.29-4.98, p<0.001), peri-operative blood transfusion (OR, 1.90, 95% CI, 1.57-2.05, p<0.001), delay of surgery (OR, 1.50, 95% CI, 1.25-1.79, p<0.001), inpatient hospital stay (OR, 2.31, 95% CI, 1.27-4.20, p<0.001) and open surgery (OR, 1.80, 95% CI, 1.57-2.05, p<0.001).[1] Studies have examined the association between postoperative sepsis and hospital safety climate and processes of care. Vogel et al examined NIS data (2003-2007) on patients who developed postoperative infectious complications following three high volume elective major surgical procedures (coronary artery bypass graft, colon resections and lung resections).[2] Using multivariable analysis, they found that surgical delay (as measured by time from admission to elective surgery in days) was associated with sepsis after all three procedures (p < 0.0001). In another study, Rosen et al. used the PSIs, including PSI 13 (version 3.1a), to explore the potential relationship between safety climate, as measured through more than 4500 responses to the Patient Safety Climate in Healthcare Organizations survey, and hospital safety performance.[3] Among the 30 VA hospitals that participated in the survey, postoperative sepsis was not significantly associated with any of the 11 dimensions of patient safety culture (p>0.10 for all comparisons). The authors noted that due to the small sample size, the relatively low rate of PSIs among VA hospitals, and narrow variation across hospitals in patient safety culture, statistical power to detect associations was limited for most PSIs, including PSI 13.

Association with hospital and health system characteristics

Multiple studies have explored the association between postoperative sepsis and health system characteristics, such as hospital teaching status, magnet designation, and public versus private status. A recent systematic review by Plaeke et al. identified an association between large-sized hospitals (OR 1.38, 95% CI, 1.12-1.70, p=0.003) and postoperative sepsis.[1] For example, three studies examined the relationship between nursing staff characteristics and postoperative sepsis. Based on data from the 2005 University HealthSystems Consortium (UHC) operational and clinical databases, Goode and colleagues found that registered nurse staff mix was significantly associated with lower PSI 13 rates after adjusting for nurse staffing hours and hospital case mix ($p<0.05$).[4] Surprisingly, PSI 13 rates were higher at designated Magnet hospitals than at non-Magnet hospitals (observed-to-expected ratio 1.83 versus 1.20; $p<0.05$).[4] Using VA and Nationwide Inpatient Sample (NIS) data from 2003 through 2004, Rivard and colleagues found that rates of PSI 13 (version 2.1) were not significantly associated with nurse staffing hours in either VA or NIS hospitals.[5] Unruh and Zhang used latent growth curve models to examine the relationship between changes in registered nurse (RN) full-time employees (FTEs), registered nurses per adjusted patient day (RN/APD), and PSIs (version 2.1) in Florida hospitals from 1996 through 2004.[6] Postoperative sepsis had strong evidence of sensitivity to nursing care, in that higher baseline RN FTE levels were significantly associated with lower levels of postoperative sepsis. Increases over time in RN/APD were associated with decreased rates of postoperative sepsis.

Multiple studies have also assessed the impact of hospital volume and bed size on PSI 13. Analyzing NIS data on all hospitalizations between 2002 and 2010 that involved coiling or clipping cerebral aneurysms, two studies found that hospital bed size was not associated with PSI 13.[7,8] Likewise, using VA and NIS data from 2003 through 2004, Rivard and colleagues found that rates of PSI 13 (version 2.1) were not significantly associated with bed size in either VA or NIS hospitals.[5] Using the VA Patient Treatment File data, Rosen et al. examined risk-adjusted PSIs (version 2.1) for acute care VA hospitals for fiscal years 2001 to 2004 and found that hospital volume was positively associated with PSI 13 rate (1.90, 95% CI 0.26 to 3.54)(model R^2 0.11).[9] Vogel et al found the opposite relationship between PSI 13 and hospital volume when analyzing CMS data (2005-2007) on Medicare beneficiaries aged 65 years and older with non-ruptured abdominal aortic aneurysms who underwent elective endovascular aortic aneurysm repair (EVAR) or open aortic repair (OAR).[10] Patients in low volume (LV) centers were more likely to develop sepsis than patients in high volume (HV) centers after both EVAR (OR 1.45; 95% CI 1.04 to 2.03) and OAR (OR 1.36; 95% CI 1.11 to 1.68). After adjusting for patient age, gender, race, and comorbidities, the likelihood of developing postoperative sepsis remained significantly greater in LV hospitals than in HV centers for both EVAR (OR 1.44; 95% CI 1.03 to 2.01) and OAR (OR 1.33; 95% CI 1.07 to 1.64). The authors estimated that 111 cases of sepsis (36 after EVAR and 75 after OAR) may have been avoided if all patients were treated at HV hospitals, and that these potentially preventable cases of postoperative sepsis accounted for $6.7 million in extra hospital charges after EVAR and $15.1 million in extra charges after OAR surgery.

Association with other outcomes

Several studies have examined the relationship between postoperative sepsis and subsequent outcomes, including length of stay in the hospital, mortality, and readmissions. For example, Ramanathan et al.

examined data on surgical patients hospitalized between 2011 and 2012 at an academic medical center and found that hospitalizations with PSI 13 (version 3.1) were associated with a 33.3 day mean hospital length of stay, 100% included an intensive care unit (ICU) stay, and 38.5% died in hospital.[11] Rosen et al.[12] reported that hospitalizations with postoperative sepsis had a significantly higher all-cause readmission rate (19.2%) than hospitalizations without PSI 13 (13.0%; $p < 0.0001$). In a multivariable analysis controlling for age, sex, comorbidities, and other PSI events, hospitalizations with a PSI 13 event had 32% higher odds of having a subsequent readmission (OR 1.32; 95% CI 1.12 to 1.57). In another study based on an analysis of the 501,908 hospitalizations involving a brain tumor in the NIS between 2002 and 2010, Rahman et al found that patients with postoperative sepsis had significantly longer lengths of stay ($p < 0.0001$) than similar patients without postoperative sepsis.[13]

Zhan and Miller used AHRQ PSI software on 7.45 million discharges in the HCUP NIS (2000) and found that patients who experienced PSI 13 had higher mean (SD) unadjusted length of stay (250.10 [0.48] vs. 70.20 [0.01]), mean charges $113,708 [2486] vs. $32,328 [72]), and mortality (240.87 [0.85] vs. 10.12 [0.02]) than patients who did not experience a PSI 13 event.[14] However, statistical differences of these comparisons were not reported. Liu and colleagues analyzed NIS (2003-2007) data on patients hospitalized with and without sepsis to examine the association between sepsis and patient health.[15] They found that nearly all PSI (version 4.2) rates were higher among patients with sepsis compared with patients without sepsis, and that among those with sepsis, most PSI rates increased as sepsis severity increased.

Bath et al. used Medicare data (MedPAR) from 2009 to 2012 and found that the likelihood of 30-day readmission among patients undergoing abdominal aortic aneurysm repair was greater among patients with postoperative sepsis (OR=1.53, $p<0.001$).[16] 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 13 event were associated with an additional 2.42 hospital days compared to patients without a PSI 13 event ($p<0.001$), as well as a significantly increased risk of in-hospital mortality (OR=6.62; $p<0.001$).[17] A study by Rhee et al compared 11,534 hospitalizations with hospital-onset sepsis and 83,620 hospitalizations with community-onset sepsis occurring across 136 hospitals from 2009 to 2015, and found that patients with hospital-onset sepsis had longer hospital length of stay (19 days vs. 8 days), were admitted to the ICU more often (60.7% vs. 44.1%), had longer ICU length of stay (6 vs. 4 days) and were at greater risk for in-hospital mortality (odds ratio, 2.5).[18]

Population group disparities

Table 10 presents population group disparities for PSI 13 Postoperative Sepsis Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 10. PSI 13 Postoperative Sepsis Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Unknown	22762	3.207	3.661
White	1110970	4.482	4.539
Black	93991	6.830	5.267
Other	14386	8.272	6.837
Asian	11960	7.860	5.898
Hispanic	16582	6.513	5.188
North American Native	7090	6.065	5.571
Gender	*	*	*
Female	689453	3.664	4.759
Male	588287	6.004	4.698
Age	*	*	*
<50	40869	6.288	4.674
50-54	29493	5.900	4.744
55-59	48271	6.691	5.472
60-64	64626	5.663	4.715
65-69	353019	3.782	4.708
70-74	312064	4.332	4.625
75-79	225227	4.844	4.428
80-84	128745	5.491	4.630
85-89	58281	6.228	4.680
90 plus	17145	5.249	4.517

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review results presented below constitute the most

recent update, conducted in August 2020. Search terms included relevant MeSH terms (sepsis) and keywords (sepsis, SIRS, septic). We combined this clinical search string with MeSH terms (postoperative complications, iatrogenic disease, quality indicators) to identify studies examining postoperative complications or quality measures. Search was limited to English publications. We also tested more inclusive search strings. To provide the most up-to-date evidence, we summarize below the most recent evidence.

1. Plaeke P, De Man JG, Coenen S, Jorens PG, De Winter BY, Hubens G. Clinical- and surgery-specific risk factors for post-operative sepsis: a systematic review and meta-analysis of over 30 million patients. *Surg Today.* 2020;50(5):427-439.
2. Vogel TR, Dombrovskiy VY, Lowry SF. In-hospital delay of elective surgery for high volume procedures: the impact on infectious complications. *J Am Coll Surg.* 2010;211(6):784-790.
3. Rosen AK, Singer S, Shibei Z, Shokeen P, Meterko M, Gaba D. Hospital safety climate and safety outcomes: is there a relationship in the VA? *Med Care Res Rev.* 2010;67(5):590-608.
4. Goode CJ, Blegen MA, Park SH, Vaughn T, Spetz J. Comparison of patient outcomes in Magnet(R) and non-Magnet hospitals. *J Nurs Adm.* 2011;41(12):517-523.
5. Rivard PE, Elixhauser A, Christiansen CL, Shibei Z, Rosen AK. Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Med Care Res Rev.* 2010;67(3):321-341.
6. Unruh LY, Zhang NJ. Nurse staffing and patient safety in hospitals: new variable and longitudinal approaches. *Nurs Res.* 2012;61(1):3-12.
7. Fargen KM, Neal D, Rahman M, Hoh BL. The prevalence of patient safety indicators and hospital-acquired conditions in patients with ruptured cerebral aneurysms: establishing standard performance measures using the Nationwide Inpatient Sample database. *J Neurosurg.* 2013;119(6):1633-1640.
8. Fargen KM, Rahman M, Neal D, Hoh BL. Prevalence of patient safety indicators and hospital-acquired conditions in those treated for unruptured cerebral aneurysms: establishing standard performance measures using the Nationwide Inpatient Sample database. *J Neurosurg.* 2013;119(4):966-973.
9. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Med Care.* 2006;44(9):850-861.
10. Vogel TR, Dombrovskiy VY, Graham AM, Lowry SF. The impact of hospital volume on the development of infectious complications after elective abdominal aortic surgery in the Medicare population. *Vasc Endovascular Surg.* 2011;45(4):317-324.
11. Ramanathan R, Leavell P, Wolfe LG, Duane TM. Agency for Healthcare Research and Quality patient safety indicators and mortality in surgical patients. *Am Surg.* 2014;80(8):801-804.
12. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Med Care.* 2013;51(1):37-44.
13. 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.

14. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality & safety in health care*. 2003;12 Suppl 2:ii58-63.

15. Liu V, Turk BJ, Rizk NW, Kipnis P, Escobar GJ. The association between sepsis and potential medical injury among hospitalized patients. *Chest.* 2012;142(3):606-613.

16. Bath J, Dombrovskiy VY, Vogel TR. Impact of Patient Safety Indicators on readmission after abdominal aortic surgery. *J Vasc Nurs.* 2018;36(4):189-195.

17. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *Am J Med Qual.* 2017;32(6):583-590.

18. Rhee C, Wang R, Zhang Z, Fram D, Kadri SS, Klompas M. Epidemiology of Hospital-Onset Versus Community-Onset Sepsis in U.S. Hospitals and Association With Mortality: A Retrospective Analysis Using Electronic Clinical Data. *Crit Care Med.* 2019;47(9):1169-1176.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0531 (Composite)

Measure Title: PSI 14 Postoperative Wound Dehiscence Rate (Component)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome**: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process**: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency**: [6] evidence not required for the resource use component.
- For measures derived from **patient reports,** evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- **Process measures incorporating Appropriate Use Criteria**: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

25. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

26. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

27. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

Wound dehiscence can be caused by inadequate undermining of the wound during surgery; excessive tension on the wound edges caused by lifting, straining, or excessive wound length; or the wound being located on a highly mobile or high-tension area. Prevention of wound dehiscence focuses primarily on control of patient level factors and technical factors under the control of the surgeon. Technical factors may be associated with surgical technique, incisional factors, and those associated with suture. A well-

planned incision should provide ready access to anticipated pathology and provide adequate exposure but allow for extension if the scope of operation needs to be expanded. The incision should interfere minimally with function by preserving important structures and heal with adequate strength to reduce the risk of wound disruption. A major cause of wound separation is failure of suture to remain anchored in the fascia, suture breakage, knot failure, and excessive stitch interval which allows protrusion of viscera. Additional postoperative processes that can help prevent dehiscence include: preventing undue stress on the wound edges and facilitating healing through adequate nutrition, control of diabetes, and the avoidance of medications that may delay wound healing.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the *systematic review of the body of evidence* that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	*
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions	*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	*
Provide all other grades and definitions from the evidence	*
Grade assigned to the **recommendation** with definition of the grade	*
Provide all other grades and definitions from the recommendation grading system	*
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	*
Estimates of benefit and consistency across studies	*
What harms were identified?	*
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Association with process of care

During the preoperative workup, patients should be evaluated for risk for factors for postoperative wound dehiscence using a validated tool.[1,2] Patients with wound dehiscence are more likely to have received care that departed from professionally recognized standards. For example, Hannan et al. reported that cases with a secondary diagnosis of wound disruption were 3.0 times more likely to have received care that departed from professionally recognized standards than cases without that code (4.3% versus 1.7%), after adjusting for patient demographic, geographic, and hospital characteristics.[1] In the case of abdominal wound dehiscence, in the majority of cases (up to 95 percent),[2-4] the sutures and knots are intact, but the suture has pulled through the fascia. This is usually the result of fascial necrosis from the surgeon placing the sutures too close to the edge or from the wound being under too much tension.

Two additional studies examined the association between postoperative wound dehiscence **and** hospital safety climate and processes of care. Rosen et al. (2010) used the PSIs, including PSI 14 (version 3.1a), to explore the potential relationship between safety climate, as measured through more than 4500 responses to the Patient Safety Climate in Healthcare Organizations survey, and hospital safety performance.[3] They found that among the 30 Veteran's Health Administration hospitals that participated in the survey, the rate of postoperative wound dehiscence was not significantly associated with any of the 11 dimensions of patient safety culture included in the analysis ($p>0.10$ for all comparisons) after adjusting for major teaching status, metropolitan area and nurse-staffing ratio. The relationship between the indicator rate and patient safety culture dimensions remained non-significant when senior managers and frontline staff were analyzed separately. The authors note that due to the small sample size, the relatively low rate of PSIs among VA hospitals, and narrow variation across hospitals in patient safety culture, statistical power to detect associations was limited for most PSIs, including PSI 14. Chen et al. (2013a) reviewed PSI 14 (version 3.1a), in relation to hospital processes of care. Using VA data, 28 out of 158 VA hospitals were selected for a stratified sample based upon observed-to-expected PSI 14 cases.[4] The study found differences in the surgical techniques (incision type and closure techniques), however physician review suggested that the techniques were not an indication of poor quality of care. Overall, postoperative wound dehiscence cases were not determined to be indicative of examined hospital processes of care, though it should be noted that about 25% of the cases and controls were missing data from their charts.

Association with hospital and health system characteristics

Multiple studies have explored the association between health system characteristics and the prevalence of postoperative wound dehiscence. Using fiscal year 2004 data from the Veterans Health Administration (VA) and calendar year 2003 data from the Healthcare Cost and Utilization Project (HCUP) National Inpatient Sample (NIS) (n=116 VA hospitals, n=992 community non-Federal hospitals), Rivard and colleagues[5] found that the risk-adjusted rate of PSI 14 was 4.80 per 1000 (95% CI 3.41 to 6.19) in VA hospitals and 1.55 per 1000 (95% CI 1.19 to 1.90) in non-VA hospitals PSI 14 rates were not significantly associated with nurse staffing hours in either VA or NIS hospitals. Among VA hospitals, both major and minor teaching hospitals had lower risk-adjusted PSI 14 rates than nonteaching hospitals (major OR 0.53, 99% CI 0.19 to 1.51; minor OR 0.55, 99% CI 0.18-1.67). Among nonfederal hospitals, both major and minor teaching hospitals had higher risk-adjusted PSI 14 rates than nonteaching hospitals (major OR 1.58, 99% CI 1.01 to 2.48; minor OR 1.40, 95% 0.96 to 2.05). None of these findings, however, was statistically significant. The same authors[6] confirmed that PSI 14 rates across NIS hospitals, but not across VA hospitals, were significantly higher in minor (OR 1.35, 95% CI 1.02 to 1.79) and major (OR 1.41, 95% CI 1.02 to 1.95) teaching hospitals than in nonteaching hospitals. Rates of this indicator were significantly lower at large hospitals (OR 0.70, 95% CI 0.53 to 0.94) than at small hospitals in the NIS, but not in the VA system. A contemporaneous VA study using a longer period of data (2001-2004) reported that after adjusting for hospital and patient characteristics using a generalized linear model, hospital bed size (-2.90, 95% CI -5.71 to -0.10, $p<0.05$) and mean quality improvement implementation score (-15.54, 95% CI -28.38 to -2.70, $p<0.05$) were negatively associated with PSI 14 (version 2.1) rates (model R^2 0.12).[7]

In 2013, Chen et al. (2013b) compared PSI 14 rates (version 3.1a) among veteran dual users (who were treated in both VA and private sector hospitals) over the period 2002-2007 and found that the odds of experiencing PSI 14 were significantly higher for veteran dual users hospitalized in the VA than in the private sector through Medicare (OR 2.23; 95% CI 1.60 to 3.10).[8] After adjustment for age, sex and 27 comorbidities, the PSI 14 rate among VA hospitalizations was 4.4 per 1000 (95% CI 3.8 to 4.9) versus 1.8 per 1,000 in the private sector (95% CI 1.3 to 2.3).

In an analysis of patient-level Medicare claims data for patients undergoing any of 6 cancer resections for the years 2005-2009, Short et al. found that the postoperative wound dehiscence rate was higher at high procedure volume hospitals than at low-volume hospitals (0.25% vs 0.22%), at rural hospitals than urban hospitals (0.39% vs 0.26%), and at non-teaching hospitals than teaching hospitals (0.30% vs 0.23%) (statistical values not provided).[9]

Association with other outcomes

Other studies have examined the relationship between postoperative wound dehiscence and outcomes including length of stay in the hospital, mortality, and readmissions. For example, Rosen et al. (2013), examined whether PSI events, experienced within index hospitalizations, increased the likelihood of readmission within VA hospitals.[10] Of the 1,807,488 index medical and surgical hospitalizations, there were a total of 262,026 readmissions. Postoperative wound dehiscence resulted in significantly higher rates of

all-cause readmissions (20.0%) compared to those hospitalizations without an event (11.5%; p<0.0001). In a multivariate analysis using AHRQ comorbidity software (version 3.5) - controlling for age, sex, comorbidities, and other PSI events - hospitalizations with a PSI 14 event were 61% more likely to result in subsequent readmissions (OR 1.61; 95% CI 1.27 to 2.05).

Ramanathan et al (2014) retroactively examined data on surgical patients hospitalized between 2011 and 2012 at an academic medical center and found that hospitalizations that included a postoperative wound dehiscence (PSI 14, version 3.1) were associated with a 108.5 day mean hospital length of stay; 100% included an intensive care unit stay, and 0% died in hospital.[11] Of those with an ICU stay, the mean ICU length of stay was 24.0 days.

Chen et al. (2013a) reviewed PSI 14 (version 3.1a) in relation to hospital processes of care using VA hospital Patient Treatment Files for 28 VA hospitals and found that length of stay (LOS) was significantly longer for cases than controls (cases mean LOS 41 days, standard deviation [SD] 43.2; controls LOS 24, SD 36.2; p<0.01).[4] Another study by Zhan and Miller used AHRQ PSI software on 7.45 million discharges in the HCUP NIS for the year 2000 and report that those with PSI 14 had a higher mean (SD) unadjusted length of stay (220.32 [0.61] vs. 60.72 [0.014]), charges ($93,022 [3,336] vs. $22,623 [75]), and percent mortality (330.66 [10.16] vs. 160.53 [0.026]) than those who did not experience a PSI 14 event.[12] However, statistical significance was not reported. The findings related to length of stay from these two studies conflict with those of another study of 501,908 hospitalizations involving a brain tumor in the NIS (2002-2010) by Rahman et al, which reported no association between postoperative wound dehiscence and length of stay.[13]

In an analysis of patient-level Medicare claims data for patients undergoing any of 6 cancer resections for the years 2005-2009, Short et al. found that after adjusting for patient (age, sex, race, income), hospital (hospital volume, surgeon volume, surgeon specialty designation, hospital resources, patient characteristics) and tumor factors (tumor stage, site), costs increased significantly, by more than 40%, for postoperative wound dehiscence among 4 of the 6 types of cancer resection patients (p < 0.001).[9]

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 14 event were associated with an additional 3.09 hospital days compared to patients without a PSI 14 event (p<0.001), as well as a significantly increased risk of in-hospital mortality (OR=72.56; p<0.001).[14]

Population group disparities

Table 11 presents population group disparities for PSI 14 Postoperative Wound Dehiscence Rate for 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 11. PSI 14 Postoperative Wound Dehiscence Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	7925	0.379	0.342
White	457827	0.867	0.806
Black	50179	0.917	0.877
Other	7339	1.090	1.275
Asian	7080	0.847	1.045
Hispanic	11377	0.791	1.070
North American Native	3985	0.502	0.475
Gender	*	*	*
Female	286226	0.528	0.795
Male	259486	1.233	0.824
Age	*	*	*
<50	26871	0.558	0.751
50-54	16277	0.922	0.647
55-59	24677	0.851	0.769
60-64	30726	1.497	0.961
65-69	131633	0.858	0.847
70-74	115036	0.861	0.776
75-79	87627	0.890	0.851
80-84	59928	0.734	0.736
85-89	35746	0.671	0.747
90 plus	17191	0.931	1.032

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review results presented below constitute the most recent update, conducted in August 2020. Search terms included relevant MeSH terms (surgical wound dehiscence). We combined this clinical search string with MeSH terms (postoperative complications) and keywords (patient safety) to identify studies examining inpatient care. Search was limited to English publications. We also tested more inclusive search strings.

1a.4.3. Provide the citation(s) for the evidence.

1. van Ramshorst GH, Nieuwenhuizen J, Hop WC, et al. Abdominal wound dehiscence in adults: development and validation of a risk model. *World J Surg.* 2010;34(1):20-27.
2. Kenig J, Richter P, Lasek A, Zbierska K, Zurawska S. The efficacy of risk scores for predicting abdominal wound dehiscence: a case-controlled validation study. *BMC Surg.* 2014;14:65.
3. Rosen AK, Singer S, Shibei Z, Shokeen P, Meterko M, Gaba D. Hospital safety climate and safety outcomes: is there a relationship in the VA? *Medical care research and review : MCRR.* 2010;67(5):590-608.
4. Chen Q, Borzecki AM, Cevasco M, et al. Examining the relationship between processes of care and selected AHRQ patient safety indicators postoperative wound dehiscence and accidental puncture or laceration using the VA electronic medical record. *Am J Med Qual.* 2013;28(3):206-213.
5. Rivard PE, Christiansen CL, Zhao S, Elixhauser A, Rosen AK. Advances in Patient Safety

   Is There an Association Between Patient Safety Indicators and Hospital Teaching Status? In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 2: Culture and Redesign).* Rockville (MD): Agency for Healthcare Research and Quality (US); 2008.
6. Rivard PE, Elixhauser A, Christiansen CL, Shibei Z, Rosen AK. Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Medical care research and review : MCRR.* 2010;67(3):321-341.
7. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Medical care.* 2006;44(9):850-861.
8. Chen Q, Hanchate A, Shwartz M, et al. Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicator Rates Among Veteran Dual Users. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2013;29(4):335-343.
9. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer.* 2014;120(7):1035-1041.
10. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Medical care.* 2013;51(1):37-44.

11. Ramanathan R, Leavell P, Wolfe LG, Duane TM. Agency for Healthcare Research and Quality patient safety indicators and mortality in surgical patients. *Am Surg.* 2014;80(8):801-804.

12. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality & safety in health care.* 2003;12 Suppl 2:ii58-63.

13. 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.

14. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2017;32(6):583-590.

Measure Number (*if previously endorsed*)**:** 0531 (Composite)

Measure Title: PSI 15 Unrecognized Abdominopelvic Accidental Puncture or Laceration Rate (Component)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Patient Safety and Adverse Events Composite (PSI 90; NQF #0531)

Date of Submission: 10/28/2020

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome:** [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- **Intermediate clinical outcome**: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process:** [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- **Structure:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- **Efficiency:** [6] evidence not required for the resource use component.
- For measures derived from ***patient reports***, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- **Process measures incorporating Appropriate Use Criteria:** See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

28. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

29. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

30. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

31. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*) Outcome

☒ Outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health- related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☐ Process:

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

This measure identifies surgical injuries to an abdominopelvic organ (e.g., bowel, bladder, liver, spleen, diaphragm, kidney) that were unintended and presumptively not recognized and treated at the time of occurrence. This definition would be met if, for example, a surgeon errantly creates a full-thickness injury of the small intestine with a cautery device or scissors while dissecting adhesions AND does not recognize the injury until reoperation a few days later for intra-abdominal sepsis. The rationale for this measure is that these injuries have major adverse consequences for patients and are usually preventable. Physicians can prevent these injuries either by avoiding the accidental puncture or laceration with more careful technique or, if that is not possible, by promptly identifying and treating the accidental puncture or laceration at the time it occurs.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

NA

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

NA

1a.3. SYSTEMATIC REVIEW (SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the

evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

NA

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Systematic Review	Evidence
Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	*
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions	*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	*

Systematic Review	Evidence
Provide all other grades and definitions from the evidence grading system	*
Grade assigned to the **recommendation** with definition of the grade	*
Provide all other grades and definitions from the recommendation grading system	*
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	*
Estimates of benefit and consistency across studies	*
What harms were identified?	*
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	*

*cell intentionally left blank

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Much of the available literature regarding PSI 15 concerns the pre-2015 version of the indicator, rather than the reformulated version, which focuses on abdominopelvic injuries that were not treated at the time of occurrence. Although the literature summarized below may provide some relevant information, reviewers should be cognizant that the current version of PSI 15 focuses on a more homogeneous and consequential subset of events.

Association with process of care

Two studies examined the association between hospital safety climate and accidental puncture and laceration. Rosen et al. used PSI 15 (version 3.1a), to explore the potential relationship between safety climate, as measured through more than 4500 responses to the Patient Safety Climate in Healthcare Organizations survey, and hospital safety performance.[1] Among the 30 VA hospitals that participated in the survey, the rate of accidental puncture or laceration was not significantly associated with any of the 11 dimensions of patient safety culture included in the analysis ($p>0.10$ for all comparisons). Analyses were adjusted for major teaching status, metropolitan area and nurse-staffing ratio. The relationship between PSI 15 rates and patient safety

culture dimensions remained non-significant when senior managers and frontline staff were analyzed separately. The authors note that due to the small sample size, the relatively low rate of PSIs among VA hospitals, and narrow variation across hospitals in patient safety culture, statistical power to detect associations was limited.

In a study testing construct validity using an implicit process measure of quality created through the hospital accreditation review process, smoothed rates of PSI 15 among 2,116 hospitals surveyed by The Joint Commission were significantly ($p<0.01$) associated with summary evaluation scores, in the expected direction.[2] Chen et al. reviewed PSI 15 (version 3.1a) in relation to hospital processes of care using data from 28 VA hospitals, but was unable to confirm any significant associations between examined processes of care and incidence of PSI 15 (although the study was likely underpowered, as it included only 112 PSI 15 events).[3]

Other studies have examined the role played by factors such as procedure timing, physician ranking, and duty-hour regulations. For example, Shelton et al., analyzed 376 million patient discharges from NIS between 1998-2007 to evaluate the effect of the 2003 U.S. implementation of duty-hour regulations, limiting resident work hours to 80 per week, within teaching and non-teaching hospitals; non-teaching hospitals served as the control group.[4] They found that the rates of accidental puncture or laceration prior to implementation were 30.27 and 42.27, per 10,000 discharges, in non-teaching and teaching hospitals, respectively. Rates of accidental puncture or laceration were not significantly altered after the implementation of the duty-hour regulations (non-teaching 28.62, teaching 24.65 per 10,000 discharges). Another study by Chen and colleagues (2013a) examined whether PSI 15 events are affected by hospital processes of care such as timing of procedure and physician ranking.[5] Using VA administrative data from October 2002-September 2007, AHRQ PSI software (version 3.1a), and medical chart review, the authors identified 95 matched case-control pairs for PSI 15. There were no significant differences found for operating room procedures performed during the weekend (n=3, 3.9%; n=4, 4.4%) or at night (n=3, 3.9%; n=7, 7.8%) between cases and controls, respectively. The authors also found no association between physician rank and PSI 15 within the matched pairs – attending physicians (cases n=33, 42.3%, controls n=33, 36.7%) and trainees (cases n=41, 52.6%; controls n=53, 58.9%) had similar rates of PSI 15 events.

Association with hospital and health system characteristics

Multiple studies have explored the association between health system characteristics and the occurrence of accidental punctures or lacerations. In 2010, Rivard and colleagues compared the relationship between PSI 15 (version 2.1) rates and various hospital characteristics in Veteran's Health Association (VA) vs. community non-Federal hospitals.[6] Using VA and Healthcare Cost and Utilization Project (HCUP) Nationwide Inpatient Sample (NIS) data from 2003 through 2004 (n=116 VA hospitals, n=992 community non-Federal hospitals), they found that the risk-adjusted rate of PSI 15 was 3.93 per 1000 (95% CI 3.31 to 4.54) in VA hospitals and 3.29 per 1000 (95% CI 3.03 to 3.56) in non-VA hospitals (from the NIS dataset). Rates were significantly higher among minor teaching hospitals, compared to nonteaching hospitals, in both samples (VA OR 2.12, 95% CI 1.11 to 4.08; NIS OR 1.23, 95% CI 1.04 to 1.47). Compared to nonteaching status, major teaching status was significantly associated with higher rates of accidental puncture and laceration, compared to nonteaching hospitals, in NIS hospitals only (OR 1.36, 95% CI 1.11 to 1.66). Another study using fiscal year 2004 data from the VA and calendar year 2003 NIS (PSI version 2.1)[7] reported that for both VA and NIS data, PSI 15 rates were lower in non-teaching hospitals than in either major or minor teaching hospitals. However, a contemporaneous VA study using a longer period of data (2001-2004) found no hospital characteristics (e.g. bed size, teaching status, location) that were associated with PSI 15 (version 2.1) across acute care VA hospitals.[8]

Chen et al (2013b) compared PSI 15 (version 3.1a) rates among 266,203 veteran dual users (who were treated in both VA and private sector hospitals) over the period 2002-2007 and found that the odds of experiencing PSI 15 were significantly higher for veteran dual users hospitalized in the VA than in the private sector through Medicare (OR 1.30, 95% CI 1.20 to 1.42).[5] After adjustment for age, sex and 27 comorbidities, the PSI 15 rate among VA hospitalizations was 5.6 per 1000 (95% CI 5.3 to 5.8) versus 3.9 per 1000 among Medicare hospitalizations (95% CI 3.7 to 4.1).

Basu & Friedman analyzed the HCUP database focusing on 3,481,086 senior Medicare patients in Florida hospitals and found that Health Maintenance Organization (HMO) patients had more PSI 15 events than Fee for Service (FFS) patients (OR = 1.155, $p<0.05$).[9] In an analysis of Medicare claims data for patients undergoing any of 6 cancer resections in 2005-2009, Short et al. found that the accidental puncture or laceration rate was higher at high procedure volume hospitals than at low-volume hospitals (2.17% vs 2.10%), at urban hospitals than rural hospitals (2.11% vs 1.89%), and at teaching hospitals than non-teaching hospitals (2.22% vs 1.98%).[10]

Association with other outcomes

Multiple studies have examined the relationship between PSI 15 and outcomes including readmissions and cost. Rosen et al., examined whether PSI events, experienced within index hospitalizations, increased the likelihood of readmission within VA hospitals.[11] Of the 1,807,488 index medical and surgical hospitalizations, there were a total of 262,026 readmissions. Accidental puncture or laceration resulted in significantly higher rates of all-cause readmissions (15.3%) compared to those hospitalizations without an event (14.6%; $p<0.0001$). In a multivariate analysis using AHRQ comorbidity software (version 3.5) - controlling for age, sex, comorbidities, and other PSI events - hospitalizations with a PSI 15 event were not significantly more likely to result in subsequent readmissions (OR 1.07; 95% CI 0.99 to 1.15). By contrast, an earlier study of 1,409,547 adults from about 1,080 hospitals in 7 geographically dispersed states (California, Florida, Missouri, New York, Tennessee, Utah, and Virginia) for surgical procedures in 2004, assembled from the State Inpatient Databases (SID) of the Healthcare Cost and Utilization Project, reported that PSI 15 events were independently and significantly ($p<0.01$) associated with inpatient deaths, 90-day readmissions, and 30-day readmissions (odds ratios 1. 52, 1. 16, and 1. 25, respectively).[12] These results were adjusted through multinomial logistic regression for age, gender, severity levels in the All Patient Refined DRG classification system from 3M Health Information Systems, coding system, payer group, presence of specific comorbid conditions, and specific surgical DRGs. It is not clear why the VA and SID results differed, but under-ascertainment of readmissions in VA data (because of the availability of non-VA hospitals) represents one possibility.

Analyses of financial outcomes such as length of stay and costs (or charges) have been more consistent. Cases from the 2000 Nationwide Inpatient Sample that were flagged by PSI 15 had 2.2% excess mortality ($p<0.001$), 1.34 days of excess hospitalization ($p<0.001$), and $8,271 in excess hospital charges ($p<0.001$), relative to carefully matched controls that were not flagged. These differences were robust to the specific adjustment approach that was used (i.e., propensity matching versus multivariable logistic modeling), but they exceeded corresponding estimates without adjustment.[13] This finding was confirmed in the Veterans Affairs hospital system, where cases that were flagged by this PSI had 3.2% excess mortality, 1.4-3.1 days of excess hospitalization, and $3,359-6,880 in excess hospital costs, relative to carefully matched controls that were not flagged.[14]

In an analysis of patient-level Medicare claims data for patients undergoing any of 6 cancer resections for the years 2005-2009, Short et al. found that after adjusting for patient (age, sex, race, income), hospital (hospital

volume, surgeon volume, surgeon specialty designation, hospital resources, patient characteristics) and tumor factors (tumor stage, site), costs increased significantly, by 15% to 21% for accidental puncture or laceration among 5 of the 6 types of cancer resection patients (p<0.001).[10]

Based on an analysis of the 501,908 hospitalizations involving a brain tumor in the NIS between 2002 and 2010, Rahman et al found that patients with accidental punctures or lacerations had significantly longer length of stay (LOS) (p < 0.0001) than patients without this indicator.[15] Ramanathan et al retroactively examined data on surgical patients hospitalized between 2011 and 2012 at an academic medical center and found that those hospitalizations that included an accidental puncture or laceration (PSI 15, version 3.1) were associated with a 17.4 day mean hospital LOS, 45.6% included an intensive care unit stay, and 5.9% died in hospital.[16] Of those with an ICU stay, the mean intensive care unit (ICU) LOS was 9.5 days.

In a retrospective study using data collected from a single-hospital department of colorectal surgery, Kin et al. found that accidental puncture or laceration cases had more diagnoses of enterocutaneous fistula (11% vs 2%, p < 0.001), reoperative cases (91% vs 61%, p < 0.001), open surgery (96% vs 77%, p < 0.001), longer operative times (186 vs 146 minutes, p = 0.001), and increased length of stay (10 vs 7 days, p = 0.002).[17]

Bath et al. used Medicare data (MedPAR) from 2009 to 2012 and found that the likelihood of 30-day readmission among patients undergoing abdominal aortic aneurysm repair was greater among patients with unrecognized abdominopelvic puncture or laceration (OR=1.40, p=0.009).[18] Gray et al. retrospectively examined 57,000 inpatient discharges at six hospitals between July 2012 and June 2014 and found that hospitalizations involving an accidental puncture or laceration were associated with an additional 1.35 hospital days compared to patients without an event (p<0.001), as well as a significantly increased risk of 30-day readmission (OR=4.29; p<0.001) and in-hospital mortality (OR=2.59; p<0.001).[19] Bohnen et al used ACS-NSQIP data from 2007 to 2012 and found that in multivariable analyses, PSI 15 events were independently associated with increased 30-day mortality (OR=3.19; p=0.002), 30-day morbidity (OR=2.68; p<0.001) and prolonged postoperative length of stay (OR=1.85; p=0.001).[20] Using the same dataset, Nandan et al found that in multivariable analyses, PSI 15 events were associated with an increase in readmission rates (OR=2.17; p=0.008).

Population group disparities

Table 12 presents population group disparities for PSI 15 Unrecognized Abdominopelvic Accidental Puncture or Laceration Rate from 3,254 measured entities by applying the CMS PSI v10.0 software to a full year of Medicare FFS claims data from July 1, 2017 to June 30, 2018.

Table 12: PSI 15 Rate Disparities

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Race	*	*	*
Unknown	20076	1.295	1.280
White	1251324	1.199	1.152
Black	196235	1.289	1.517
Other	23021	1.651	1.822

Population-Based Disparity Factor	N (beneficiaries)	Observed Rate per 1,000	Adjusted Rate per 1,000
Asian	25219	1.546	1.832
Hispanic	38350	1.173	1.495
North American Native	11197	1.429	1.431
Gender	*	*	*
Female	782040	1.362	1.215
Male	783382	1.088	1.224
Age	*	*	*
<50	91024	1.077	1.335
50-54	50348	1.231	1.124
55-59	75367	0.942	1.083
60-64	94531	1.164	1.173
65-69	312823	1.260	1.172
70-74	289356	1.369	1.297
75-79	244653	1.328	1.204
80-84	191918	1.230	1.159
85-89	136689	1.192	1.356
90 plus	78713	0.788	1.269

Source: CMS Medicare Fee-for-Service Data, 7/2017 – 6/2018

*cell intentionally left blank

For additional information on disparities see the PSI 90 Composite (NQF #0531) Measure Information Form Section 1b.4.

1a.4.2 What process was used to identify the evidence?

Formal environmental scans of the literature, including routine PubMed searches are performed to continually update evidence. The current evidence review results presented below constitute the most recent update, conducted in August 2020. Search terms included relevant keywords (accidental puncture, laceration). Search was limited to English publications. We also tested more inclusive search strings.

1a.4.3. Provide the citation(s) for the evidence.

1. Rosen AK, Singer S, Shibei Z, Shokeen P, Meterko M, Gaba D. Hospital safety climate and safety outcomes: is there a relationship in the VA? *Medical care research and review : MCRR.* 2010;67(5):590-608.
2. Miller MR, Pronovost P, Donithan M, et al. Relationship between performance measurement and accreditation: implications for quality of care and patient safety. *Am J Med Qual.* 2005;20(5):239-252.
3. Chen Q, Borzecki AM, Cevasco M, et al. Examining the relationship between processes of care and selected AHRQ patient safety indicators postoperative wound dehiscence and accidental puncture or laceration using the VA electronic medical record. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2013;28(3):206-213.
4. Shelton J, Kummerow K, Phillips S, et al. Patient safety in the era of the 80-hour workweek. *Journal of surgical education.* 2014;71(4):551-559.
5. Chen Q, Hanchate A, Shwartz M, et al. Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicator Rates Among Veteran Dual Users. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2013;29(4):335-343.
6. Rivard PE, Elixhauser A, Christiansen CL, Shibei Z, Rosen AK. Testing the association between patient safety indicators and hospital structural characteristics in VA and nonfederal hospitals. *Medical care research and review : MCRR.* 2010;67(3):321-341.
7. Rivard PE, Christiansen CL, Zhao S, Elixhauser A, Rosen AK. Advances in Patient Safety: Is There an Association Between Patient Safety Indicators and Hospital Teaching Status? In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 2: Culture and Redesign).* Rockville (MD)2008.
8. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Medical care.* 2006;44(9):850-861.
9. Basu J, Friedman B. Adverse events for hospitalized medicare patients: is there a difference between HMO and FFS enrollees? *Social work in public health.* 2013;28(7):639-651.
10. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer.* 2014;120(7):1035-1041.
11. Rosen AK, Loveland S, Shin M, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Medical care.* 2013;51(1):37-44.
12. Friedman B, Encinosa W, Jiang HJ, Mutter R. Do patient safety events increase readmissions? *Medical care.* 2009;47(5):583-590.
13. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality & safety in health care.* 2003;12 Suppl 2:ii58-63.
14. Rivard PE, Luther SL, Christiansen CL, et al. Using Patient Safety Indicators to Estimate the Impact of Potential Adverse Events on Outcomes. *Medical Care Research and Review.* 2008;65(1):67-87.
15. 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.
16. Ramanathan R, Leavell P, Wolfe LG, Duane TM. Agency for Healthcare Research and Quality patient safety indicators and mortality in surgical patients. *Am Surg.* 2014;80(8):801-804.
17. Kin C, Snyder K, Kiran RP, Remzi FH, Vogel JD. Accidental puncture or laceration in colorectal surgery: a quality indicator or a complexity measure? *Dis Colon Rectum.* 2013;56(2):219-225.

18. Bath J, Dombrovskiy VY, Vogel TR. Impact of Patient Safety Indicators on readmission after abdominal aortic surgery. *J Vasc Nurs.* 2018;36(4):189-195.

19. Gray DM, 2nd, Hefner JL, Nguyen MC, Eiferman D, Moffatt-Bruce SD. The Link Between Clinically Validated Patient Safety Indicators and Clinical Outcomes. *American journal of medical quality : the official journal of the American College of Medical Quality.* 2017;32(6):583-590.

20. Bohnen JD, Mavros MN, Ramly EP, et al. Intraoperative Adverse Events in Abdominal Surgery: What Happens in the Operating Room Does Not Stay in the Operating Room. *Ann Surg.* 2017;265(6):1119-1125.

- 1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

Not applicable (composite measure)

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

Table 1. Medicare Fee-for-Service Reference Population Rate and Distribution of Hospital Performance on PSI 90 (Patient Safety Composite)

Year	N	Mean	SD	min	p10	p25	Med	p75	p90	max
2016-2017	3212	0.995	0.174	0.567	0.842	0.906	0.970	1.036	1.181	5.326
2017-2018	3212	0.994	0.166	0.530	0.845	0.907	0.970	1.029	1.178	3.791
2018-2019	3212	0.996	0.161	0.629	0.849	0.913	0.971	1.032	1.174	2.588

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

Abbreviations: SD=standard deviation; p=percentile

Table 2. Distribution of Hospital Performance on PSI 90 Specific Indicators

	2017	2017	2018	2018	2019	2019
Component	N	Mean	N	Mean	N	Mean
PSI 03	3207	0.593	3208	0.618	3208	0.580
PSI 06	3211	0.257	3211	0.243	3210	0.221
PSI 08	3209	0.108	3209	0.106	3207	0.100
PSI 09	3086	2.408	3083	2.532	3075	2.481
PSI 10	2970	1.323	2981	1.368	2960	1.337
PSI 11	2961	6.801	2973	5.755	2950	4.876

PSI 12	3086	3.784	3083	3.670	3077	3.572
PSI 13	2964	4.702	2966	4.743	2943	4.491
PSI 14	3036	0.977	3035	0.865	3021	0.879
PSI 15	3112	1.273	3104	1.226	3086	1.180

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

For PSI 90, multiple denominator populations are aggregated into a composite score at the facility level. Since the composite score is not estimated at the patient level or the population group level, disparities at these levels cannot be calculated.

To satisfy the NQF requirement, we are reporting our analysis of population group disparities for the individual component measures in PSI 90. These analyses were performed by applying the CMS PSI v10.0 software to the most recent full year of Medicare FFS claims data for the reference population for that software version (7/1/2017-6/30/2018). PSI 90 Component Indicator Rate Disparities for 9,619,208 hospital discharges from 3,254 measured entities are presented in Evidence Attachment Tables 3-12. We explore disparities by sex, age group, and race/ethnicity (see Evidence Attachment Tables 3-12). Age group disparities should be interpreted cautiously because Medicare FFS beneficiaries under 65 years of age are more likely to be disabled, and less likely to be healthy, than the general US population of similar age. We also used same-month Medicaid eligibility status, identified in the Medicare Beneficiary Summary File, as a proxy for socioeconomic status, in accord with recommendations from NQF and the Assistant Secretary for Planning and Evaluation (ASPE).

In summary, disparities exist for the PSI 90 component measures, but there is no consistent pattern across these components. This finding is not surprising as the PSI 90 component measures focus on hospital-acquired complications of care. For example, men have at least 20% higher observed rates than women for PSI 03 (Pressure UIcer), PSI 09 (Postoperative Hemorrhage or Hematoma), PSI 10 (Postoperative Acute Kidney Injury Requiring Dialysis), PSI 11 (Postoperative Respiratory Failure), PSI 13 (Postoperative Sepsis), and PSI 14 (Postoperative Wound Dehiscence). Men have at least 20% lower observed rates than women for PSI 06 (Iatrogenic Pneumothorax), PSI 08 (In-hospital Fall with Hip Fracture), and PSI 15 (Unrecognized Accidental Puncture or Laceration). All of the PSI risk-adjustment models include sex, age groups, and sex-age interactions. Therefore, the observed disparities across age and sex categories greatly diminish or disappear after risk-adjustment, as intended.

Across racial-ethnic categories, the Medicare FFS data show at least 25% higher adjusted rates among Black patients, relative to White patients, for only three PSIs: PSI 03 (Pressure Injury), PSI 12 (Perioperative Deep Vein Thrombosis or Pulmonary Embolism), and PSI 15 (Unrecognized Accidental Puncture or Laceration). Comparing Hispanic patients with White patients, Hispanics had at least 20% higher adjusted PSI rates only for PSI 14 (Postoperative Wound Dehiscence) and PSI 15 (Unrecognized Accidental Puncture or Laceration). For all other PSI 90 component measures, Black and Hispanic patients had lower or similar adjusted rates, when compared with White patients.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

Not applicable, see above.

1c. Composite Quality Construct and Rationale

1c.1. A composite performance measure is a combination of two or more component measures, each of which individually reflects quality of care, into a single performance measure with a single score.

For purposes of NQF measure submission, evaluation, and endorsement, the following will be considered composites:

- Measures with two or more individual performance measure scores combined into one score for an accountable entity.
- Measures with two or more individual component measures assessed separately for each patient and then aggregated into one score for an accountable entity:
  - all-or-none measures (e.g., all essential care processes received, or outcomes experienced, by each patient);

1c.1. Please identify the composite measure construction: two or more individual performance measure scores combined into one score

1c.2. Describe the quality construct, including:

- the overall area of quality
- included component measures and
- the relationship of the component measures to the overall composite and to each other.

CMS Patient Safety Indicator (PSI) 90, also known as the Patient Safety and Adverse Events Composite, combines information about 10 common patient safety events that may occur in hospitalized patients. It was created to provide a simple and transparent single metric that can be used to better understand, communicate, and track patient safety in U.S. hospitals. The underlying concept, as described by the Institute of Medicine, is that safety is "freedom from accidental injury"(1) and that safe care "involves making evidence-based clinical decisions to maximize the health outcomes of an individual and to minimize the potential for harm"(2). This concept is closely linked to CMS's priority to implement quality initiative assuring quality healthcare for Medicare Beneficiaries using tools to achieve effective, safe, efficient, patient-centered, equitable, and timely care. Section 3008 of the Affordable Care Act established the Hospital-Acquired Condition (HAC) Reduction Program to encourage hospitals to reduce HACs. Beginning with Fiscal Year (FY) 2015 discharges (i.e. October 1, 2014), the HAC Reduction Program requires the Secretary of Health and Human Services (HHS) to adjust payments. As set forth in the Affordable Care Act, CMS may reduce payments for the worst-performing 25 percent of hospitals by up to one percent, and publicly report hospitals' measure scores, domain scores, and HAC Reduction Program data.

CMS PSI 90 combines the smoothed indirectly standardized morbidity ratios (observed/expected ratio) from 10 component indicators: PSI 03 Pressure Ulcer, PSI 06 Iatrogenic Pneumothorax, PSI 08 In-Hospital Fall with Hip Fracture, PSI 09 Postoperative Hemorrhage or Hematoma, PSI 10 Postoperative Acute Kidney Injury Requiring Dialysis, PSI 11 Postoperative Respiratory Failure, PSI 12 Perioperative Pulmonary Embolism or Deep Vein Thrombosis, PSI 13 Postoperative Sepsis Rate, PSI 14 Postoperative Wound Dehiscence, and PSI 15 Unrecognized Abdominopelvic Accidental Puncture or Laceration Rate.

From a conceptual perspective, the CMS PSI 90 composite should reflect the likelihood of harm associated with a wide range of potentially preventable adverse events. Within this conceptualization, each PSI is an individual predictor of an important and relevant aspect of harm. Thus, the likelihood of harm is expressed as the probability of a potentially preventable adverse event (such as postoperative sepsis) times the average net severity of harm associated with that event. In this conceptualization, CMS PSI 90 is modeled as a heterogeneous, formative index, meaning that the composite is "formed from" a set of measured components, each of which reflects different but overlapping aspects of care. The use of administrative data provides an

inexpensive and fairly comprehensive approach to measuring a wide range of elements of the construct of harm while each component indicator contributes a unique aspect of harm.

The final weight for each component measure is the product of harm weights and volume weights (numerator weights). Harm weights are calculated by multiplying empirical estimates of probability of excess harms associated with the patient safety event by the corresponding utility weights linked to each of the harms (1-disutility). Disutility is the measure of the severity of the adverse events associated with each of the harms (i.e.., outcome severity, or at least preferred states from the patient perspective). The harm weights are calculated using linked claims data for two years of Medicare Fee for Service beneficiaries. Volume weights, the second part of the final weight, are calculated based on the number of safety-related events for the component indicators in the fee-for-service reference population.

For more information, see https://www.qualitynet.org/inpatient/measures/psi/resources.

1. Kohn LT, Corrigan JM, Donaldson MS, eds., Institute of Medicine. To err is human: Building a safer health system. Washington, DC: National Academies Press, 1999.
2. Aspden P, Corrigan JM, Wolcott J, Erickson SM, eds., Institute of Medicine. Patient safety: achieving a new standard for care. Washington, DC: National Academies Press; 2004.

1c.3. Describe the rationale for constructing a composite measure, including how the composite provides a distinctive or additive value over the component measures individually.

The CMS PSI 90 composite measure was developed to summarize patient safety across multiple indicators to monitor performance over time or across regions and populations using a methodology that can be applied at the national, regional, State and provider level. Practically, a composite was constructed to increase statistical precision due to an increase in the effective sample size and to address the issue of competing priorities where more than one component measure may be important; and to assist consumers in selecting hospitals, providers in allocating resources to reduce patient safety events, and payers in assessing performance.

1c.4. Describe how the aggregation and weighting of the component measures are consistent with the stated quality construct and rationale.

The CMS PSI 90 composite is conceptualized as a formative composite, meaning that the indicator is formed from a set of component indicators, and in this case, these indicators assess different harms and aspects of patient safety.

Formative composites require deliberate selection of weights that best support the decision-making purpose of the composite. Composite measures typically provide a more valid signal if more "important" components are weighted more heavily than less "important" components. A variety of weighting methods exist. Previously, the CMS PSI 90 composite weighted each component according to the number of “opportunities” to provide the optimal process of care or experience the optimal outcome (based on the concept that relatively rare events become more important to the extent that more patients are at risk of experiencing them).

The CMS PSI 90 composite reflects the “redesigned” PSI 90 that was submitted to and endorsed by NQF in 2015, in direct response to feedback from the NQF Patient Safety Standing Committee, NQF members, and many other stakeholders. The measure reflects an approach which is based on (1) soliciting patients’ or clinicians’ judgments about the relative importance of each component (based on the concept that some events are more important, from the clinical or public health perspective, than others of equal frequency) and (2) using more complex statistical and empirical methods to estimate the relative importance of each component (based on the concept that relative importance can be estimated from a causal model in which adverse events are a final common pathway leading to death, prolonged hospital stay, or other undesirable outcomes). The CMS PSI 90 composite retains the “formative” construct, with weighting based on empirical estimates of importance (versus a “reflective” construct based on an underlying unobserved construct of patient safety). The formative construct approach was deemed preferable during the redesign because (1) it is historically consistent with how the PSIs were developed and how PSI 90 was conceived; (2) it retains the conceptual advantages of a single composite, whereas applying item response theory might require division

into multiple composites; and (3) it is driven by stronger theory; i.e., decision-making by providers, consumers, and other stakeholders should be driven by the objective of reducing net harm and increasing utility.

The goals of this formative composite is to assess and improve safety (freedom from harm) for a population that may be at risk for a variety of different adverse events, each of which may have different causes and potential mechanisms. The composite must balance the competing risks of these different events, based (in the case of PSI 90) on the perspective of patients' experience of inpatient care. Thus, the composite is designed in a manner that not only enhances reliability, but also reflects competing importance for specific cohorts.

Article I. 2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

Critical Care, Gastrointestinal (GI), Renal, Respiratory, Surgery, Surgery : Cardiac Surgery, Surgery : General Surgery, Surgery : Perioperative and Anesthesia, Surgery : Thoracic Surgery, Surgery : Vascular Surgery

De.6. Non-Condition Specific*(check all the areas that apply):*

Safety, Safety : Complications, Safety : Healthcare Associated Infections

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

Populations at Risk

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

https://www.qualitynet.org/inpatient/measures/psi/resources

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

Attachment**:** PSI_90_v10.xlsx

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

Yes

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

Since the last NQF measure update in October 2018 (which referenced AHRQ v6.0 [2017]), all CMS v10.0 component measure specifications are now using ICD-10-CM diagnosis codes and/or ICD-10-PCS procedure codes. State-of-the-art conversion processes were used to convert CMS PSI 90 and its component measures from ICD-9-CMS to ICD-10-CMS/PCS (1, 2). Other changes to the component measures are based on user comments and suggestions, as well as periodic literature scans by the CMS and AHRQ. These changes are as follows:

PSI 03 (Pressure Ulcer):

- Restricted denominator exclusion to qualifying discharges with a principal diagnosis code for pressure ulcer stage III or IV (or unstageable) instead of excluding discharges with a principal diagnosis code of pressure ulcer (any stage). A change in ICD-9-CM coding guidance established that "If a patient is admitted to an inpatient hospital with a pressure ulcer at one stage and it progresses to a higher stage, two separate codes should be assigned: one code for the site and stage of the ulcer on admission and a second code for the same ulcer site and the highest stage reported during the stay." This change allows PSI 03 to capture pressure injuries that are documented as stage 1 or 2 at admission (present on admission, or POA) but progress to stage 3 or 4 after admission.
- Removed denominator exclusions for the following procedures and conditions in ICD-10: pedicle graft procedures, hemiplegia or similar plegias, spina bifida or anoxic encephalopathy, and major skin disorders. Before POA reporting was required, these conditions and procedures potentially associated with pressure ulcers were assumed to indicate that the pressure injury was POA. Therefore, exclusions for these conditions and procedures served as a means of removing events that might not be attributable to hospitals. However, now that POA status is required, these exclusions are redundant and lead to undercounting of hospital-acquired pressure ulcers.
- Removed denominator exclusion for patients admitted from acute hospitals or SNFs/ICFs. Before POA reporting was required, these conditions and procedures potentially associated with pressure ulcers were assumed to indicate that the pressure injury was POA. Therefore, exclusions for these conditions and procedures served as a means of removing events that might not be attributable to hospitals. However, now that POA status is required, these exclusions are redundant and lead to undercounting of hospital-acquired pressure ulcers.
- Added denominator exclusions for diagnosis codes for severe burns (>20% body surface area, BSA) and exfoliative disorders of the skin (>20% BSA). Patients with severe burns are at an increased risk for skin breakdown and already receive intensive skin care as a result of their burn-related injury. Despite best efforts, progression to stage III or IV pressure ulcers may be largely unpreventable, which is inconsistent with the intent of PSI 03 to capture preventable hospital-acquired pressure ulcers. The same logic applies to exfoliative disorders involving large areas of skin surface, such as Stevens-Johnson Syndrome.

PSI 06 (Iatrogenic Pneumothorax)

- No material changes

PSI 08 (In Hospital Fall with Hip Fracture)

- Revised denominator statement to include medical DRGs (in addition to surgical DRGs). The complication can occur in both medical and surgical patients. Previously medical patients were excluded due to concerns of capturing fractures present on admission, but present on admission data allows for dropping this criterion.

- Removed denominator exclusion where the first or only operating room procedure is hip fracture repair. With the inclusion of "present on admission" criteria it is no longer necessary to focus on surgical patients to avoid false positives. PSI 08 now includes patients whose only operating room procedure was a hip fracture repair.
- Removed denominator exclusion for with diagnosis codes for self-inflicted injury. Exclusion of self-inflicted injuries was removed because self-inflicted harm could be better addressed with risk-adjustment rather than exclusion. Hospitals should be expected to make efforts to prevent patient self-inflicted harm. Self-inflicted harm is extremely unlikely to result in a hip fracture.
- Removed denominator exclusion for MDC 8 (diseases and disorders of the musculoskeletal system and connective tissue). When the denominator was expanded to medical and surgical patients, this exclusion had the unintended effect of removing patients who were admitted for a medical condition assigned to MDC 08, fell, and sustained a hip fracture. Hospitals may be expected to prevent falls with hip fracture in these patients.

PSI 09 (Perioperative Hemorrhage or Hematoma Rate)

- Removed denominator exclusion for patients in whom the only operating room procedure is a procedure potentially related to treatment of perioperative hemorrhage or hematoma unless there is any secondary ICD-10-CM diagnosis code for perioperative hemorrhage or hematoma. This change was necessitated by the fact that ICD-10-PCS procedure codes do not incorporate any diagnostic information; therefore, the same procedure may be performed to drain an abscess or a hematoma. The PSI software was rewritten to narrow the exclusion to patients who had (for example) a drainage or extirpation procedure for hemorrhage or hematoma, with no other major operations.
- Added denominator exclusion for diagnosis codes for coagulation disorders. Antineoplastic chemotherapy induced pancytopenia and other disorders impacting coagulation were added to the definition of platelet disorders for the purpose of excluding patients in the ICD-10 version of PSI 09. As an antiplatelet disorder, patients with antineoplastic chemotherapy induced pancytopenia have a higher risk for a PSI 09 event and should consequently be excluded from the measure. Other disorders can decrease coagulation.

PSI 10 (Postoperative Acute Kidney Injury Requiring Dialysis Rate)

- Added denominator exclusion for diagnosis code present on admission for solitary kidney disease and any procedure code for partial nephrectomy. In the setting of a solitary kidney, partial nephrectomy is expected to lead to significant compromise of renal function, potentially requiring temporary or permanent renal replacement therapy.

PSI 11 (Postoperative Respiratory Failure)

- Revised numerator statement to include only secondary procedure codes for reintubation or mechanical ventilation (not principal procedure codes) occurring one or more days after the first major operating room procedure code. The principal procedure is defined as the procedure most closely related to the principal diagnosis; the target population for PSI 11 consists of patients who are not in respiratory failure at admission (and therefore would not have a principal procedure of intubation or mechanical ventilation).
- Revised denominator exclusion to include any procedure codes for facial surgery, not limited to those including a diagnosis code for craniofacial anomalies. These ICD-10 denominator exclusions were restricted to those that involve an inherent risk of airway compromise, with input from a general surgeon and an otolaryngologist. More specific procedure codes in ICD-10-PCS permit a more tailored denominator exclusion based on the procedures that involve airway compromise requiring extended intubation.

PSI 12 (Perioperative Pulmonary Embolism or Deep Vein Thrombosis)

- Revised numerator statement to limit to proximal deep vein thrombosis. This change was based on emerging evidence of "overdiagnosis" of distal vein thrombosis. Users raised concerns about the impact of inter-hospital variation in postoperative surveillance and diagnostic testing on the rate of PSI 12; this variation was linked to the observation that major teaching hospitals often have higher PSI 12 rates than

minor teaching and non-teaching hospitals. Routine use of ultrasound for postoperative surveillance is not an evidence-based practice, and is not endorsed in clinical practice guidelines (e.g., American College of Chest Physicians). It appears that many of the distal thromboses discovered through routine surveillance would never have caused symptoms, although they do require observation due to the risk of embolization.

- Revised denominator exclusion to exclude cases with a principal diagnosis code (or secondary diagnosis code present on admission) for proximal deep vein thrombosis. This change was linked to the preceding change; all PSI specifications exclude patients admitted with the complication in question.
- Revised denominator exclusion for cases where the only operating room procedure was interruption of vena cava. This change modified the previous exclusion so that cases are excluded only if they are the only operating room procedure, instead of the principal procedure. The principal procedure is defined as the procedure most closely related to the principal diagnosis, which is not relevant to the intent of this exclusion.
- Added denominator exclusion for cases where a procedure for pulmonary arterial thrombectomy occurs before or on the same day as the first operating room procedure or as the only operating room procedure. Pulmonary arterial thrombectomy procedures should not qualify a patient as a surgical patient if no other OR procedures were performed prior to the thrombectomy, because the thrombectomy was presumably performed to treat a pulmonary embolism. Therefore, failure to exclude thrombectomy procedures from the denominator may lead to false positives for PSI 12 events. (Such an exclusion could not be implemented in ICD-9 due to lack of specific codes for pulmonary arterial thrombectomy.)

PSI 13 (Sepsis)

No material changes

PSI 14 (Postoperative Wound Dehiscence Rate)

- Revised numerator statement to include cases involving both procedure codes for repair of the abdominal wall and diagnosis codes for disruption of internal surgical wound. This change was necessitated by the conversion from ICD-9-CM to ICD-10-CM/PCS; the latter code set has no specific procedure codes for reclosure of a postoperative disruption of the abdominal wall. This concept can only be captured using a combination of diagnosis codes (for surgical wound disruption) and procedure codes (for repair of the abdominal wall).
- Revised denominator exclusion to exclude cases where the procedure for abdominal wall closure occurs on or before the day of the first open or laparoscopic abdominopelvic surgery procedure. This type of denominator exclusion applies across all surgical PSIs, to ensure that the index procedure actually preceded the PSI-triggering reparative procedure.

PSI 15 (Unrecognized Abdominopelvic Accidental Puncture or Laceration Rate)

No material changes

(1) Agency for Healthcare Research and Quality. AHRQ ICD-10-CM/PCS Conversion Project. November 15, 2013. Available at: https://www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2013/C.14.10.D001_REVISED.pdf. Accessed June 22, 2020.

(2) Utter GH, Cox GL, Atolagbe OO, et al. Conversion of the Agency for Healthcare Research and Quality's Quality Indicators from ICD-9-CM to ICD-10-CM/PCS: The Processes, Results, and Implications for Users. Health Services Research;53(5). https://doi.org/10.1111/1475-6773.12981

For further details regarding the original conceptual framework underlying CMS Medicare PSI 90, and the methods used for utility assessment and harm weighting, please refer to the supplemental materials submitted by AHRQ as part of the last cycle of NQF review at https://www.qualityforum.org/QPS/0531.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

PSI 03: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for pressure ulcer stage III or IV (or unstageable).

PSI 06: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for iatrogenic pneumothorax.

PSI 08: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for hip fracture.

PSI 09: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with: any secondary ICD-10-CM diagnosis codes for perioperative hemorrhage or hematoma and any-listed ICD-10-CM procedure codes for treatment of hemorrhage or hematoma (Note: The ICD-10-CM specification is limited to postoperative hemorrhage or hematoma).

PSI 10: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for acute renal failure and any-listed ICD-10-CM procedure codes for dialysis.

PSI 11: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with either: any secondary ICD-10-CM diagnosis code for acute respiratory failure; or any-listed ICD-10-CM procedure codes for a mechanical ventilation for 96 consecutive hours or more that occurs zero or more days after the first major operating room procedure code (based on days from admission to procedure); or any-listed ICD-10-CM procedure codes for a mechanical ventilation for less than 96 consecutive hours (or undetermined) that occurs two or more days after the first major operating room procedure code (based on days from admission to procedure); or any-listed ICD-10-CM procedure codes for a reintubation that occurs one or more days after the first major operating room procedure code (based on days from admission to procedure).

PSI 12: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with a secondary ICD-10-CM diagnosis code for proximal deep vein thrombosis or a secondary ICD-10-CM diagnosis code for pulmonary embolism.

PSI 13: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for sepsis.

PSI 14: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any-listed ICD-10-PCS procedure codes for repair of the abdominal wall and any-listed ICD-10-CM diagnosis code for disruption of internal surgical wound

PSI 15: Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with any secondary ICD-10-CM diagnosis codes for accidental puncture or laceration during a procedure and second abdominopelvic operation >=1 day after an index abdominopelvic operation.

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

See attached technical specifications for complete list of numerator details, which are also available at:

https://www.qualitynet.org/inpatient/measures/psi/resources and
https://www.qualitynet.org/files/5ebeeee9641cb00023dd1f96?filename=2019_PSI_TechSpecs_Excel.zip

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

PSI 03: Surgical or medical discharges, for patients ages 18 years and older. Surgical and medical discharges are defined by specific MS-DRG codes.

PSI 06: Surgical and medical discharges, for patients ages 18 years and older. Surgical and medical discharges are defined by specific MS-DRG codes.

PSI 08: Discharges, for patients ages 18 years and older, in a medical DRG or in a surgical DRG, with any listed ICD-10-PCS procedure codes for an operating room procedure.

PSI 09: Surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Surgical discharges are defined by specific MS-DRG codes.

PSI 10: Elective surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Elective surgical discharges are defined by specific MS-DRG codes with admission type recorded as elective.

PSI 11: Elective surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Elective surgical discharges are defined by specific MS-DRG codes with admission type recorded as elective.

PSI 12: Surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Surgical discharges are defined by specific MS-DRG codes.

PSI 13: Elective surgical discharges, for patients ages 18 years and older, with any-listed ICD-10-PCS procedure codes for an operating room procedure. Elective surgical discharges are defined by specific MS-DRG codes with admission type recorded as elective.

PSI 14: Discharges, for patients ages 18 years and older, with any-listed ICD-10-CM procedure codes for abdominopelvic surgery, open approach, or with any-listed ICD-10-PCS procedure codes for abdominopelvic surgery, other than open approach.

PSI 15: Surgical and medical discharges, for patients ages 18 years and older, with any ICD-10-PCS procedure code for an abdominopelvic procedure

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The attached technical specifications and appendices include a complete list of denominator codes and details, which are also available at:
https://www.qualitynet.org/files/5ebeeee9641cb00023dd1f96?filename=2019_PSI_TechSpecs_Excel.zip

PSI 03: See PSI Appendix B - Medical Discharge MS-DRGs and Appendix C - Surgical Discharge MS-DRGs for the full list of codes

PSI 06: See PSI Appendix C - Surgical Discharge MS-DRGs for the full list of codes

PSI 08: See PSI Appendix A - Operating Room Procedure Codes, Appendix B - Medical Discharge MS-DRGs and Appendix C - Surgical Discharge MS-DRGs for the full list of codes, and Appendix E - excluded Trauma Diagnosis Codes

PSI 09: See PSI Appendix A - Operating Room Procedure Codes and Appendix C - Surgical Discharge MS-DRGs for the full list of codes

PSI 10: See PSI Appendix A - Operating Room Procedure Codes and Appendix C - Surgical Discharge MS-DRGs for the full list of codes

PSI 11: See PSI Appendix A - Operating Room Procedure Codes and Appendix C - Surgical Discharge MS-DRGs for the full list of codes

PSI 12: See PSI Appendix A - Operating Room Procedure Codes and Appendix C - Surgical Discharge MS-DRGs for the full list of codes

PSI 13: See PSI Appendix A - Operating Room Procedure Codes and Appendix C - Surgical Discharge MS-DRGs for the full list of codes

PSI 14: see attached technical specifications for the full list of codes

PSI 15: see attached technical specifications plus Appendix B - Medical Discharge MS-DRGs and Appendix C - Surgical Discharge MS-DRGs for the full list of codes

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

PSI 03:

- Length of stay of less than 3 days
- Principal ICD-10-CM diagnosis code for pressure ulcer stage III or IV (or unstageable)
- All secondary ICD-10-CM diagnosis codes for pressure ulcer III or IV (or unstageable) present on admission. If more than one diagnosis of pressure ulcer is present, all diagnoses must be present on admission for the discharge to be excluded
- Any listed ICD-10-CM diagnosis code for severe burns (>20% body surface area)
- Any listed ICD-10-CM diagnosis code for exfoliative disorders of the skin (>20% body surface area)
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 06:

- Principal ICD-10-CM diagnosis code for iatrogenic pneumothorax
- Any secondary ICD-10-CM diagnosis code for iatrogenic pneumothorax present on admission, among patients qualifying for the numerator
- Any listed ICD-10-CM diagnosis codes for specified chest trauma (rib fractures, traumatic pneumothorax and related chest wall injuries)
- Any listed ICD-10-CM diagnosis codes for pleural effusion
- Any listed ICD-10-PCS procedure codes for thoracic surgery
- Any listed ICD-10-CM procedure codes for cardiac procedure;
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 08:

- Principal ICD-10-CM diagnosis code for hip fracture
- Any secondary ICD-10-CM diagnosis code for hip fracture present on admission, among patients otherwise qualifying for the numerator
- Principal ICD-10-CM diagnosis code for seizure
- Principal ICD-10-CM diagnosis code for syncope
- Principal ICD-10-CM diagnosis code for stroke and occlusion of arteries
- Principal ICD-10-CM diagnosis code for coma

- Principal ICD-10-CM diagnosis code for cardiac arrest
- Principal ICD-10-CM diagnosis code for poisoning
- Principal ICD-10-CM diagnosis code for trauma
- Principal ICD-10-CM diagnosis code for delirium and other psychoses
- Principal ICD-10-CM diagnosis code for anoxic brain injury
- Any listed ICD-10-CM diagnosis codes for metastatic cancer
- Any listed ICD-10-CM diagnosis codes for lymphoid malignancy
- Any listed ICD-10-CM diagnosis codes for bone malignancy
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 09:

- Principal ICD-10-CMS diagnosis code for perioperative hemorrhage or postoperative hematoma
- Any secondary ICD-10-CM diagnosis present on admission for perioperative hemorrhage or postoperative hematoma, among discharges that otherwise qualify for the numerator
- The only operating room procedure is for treatment of perioperative hemorrhage, or hematoma and with any secondary ICD-10-CM diagnosis codes for perioperative hemorrhage or hematoma
- Treatment of postoperative hemorrhage or hematoma occurs one day or more before the first operating room procedure, and with any secondary ICD-10-CM diagnosis codes for postoperative hemorrhage or hematoma
- With any listed ICD-10-CM diagnosis codes for coagulation disorders
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 10:

- Principal ICD-10-CM diagnosis code for acute renal failure, cardiac arrest, cardiac dysrhythmia, shock or chronic kidney failure
- Any secondary ICD-10-CM diagnosis code for acute kidney failure, cardiac arrest, cardiac dysrhythmia, shock or chronic kidney failure, present on admission, among patients otherwise qualifying for the numerator
- Any dialysis procedure that occurs before or on the same day as the first operating room procedure
- Any dialysis access procedure occurring before or on the same day as the first operating room procedure
- Principal ICD-10-CM (or secondary diagnosis present on admission) for urinary tract obstruction
- Any ICD-10-CM diagnosis code present on admission for solitary kidney disease and any ICD-10-PCS procedure code for partial nephrectomy
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 11:

- Principal ICD-10-CM diagnosis code for acute respiratory failure
- Any secondary ICD-10-CM diagnosis code for respiratory failure present on admission, among patients otherwise qualifying for the numerator

- Only operating room procedure is tracheostomy
- Procedure for tracheostomy occurs before the first operating room procedure
- Any listed ICD-10-CM diagnosis codes for neuromuscular disorder
- Any listed ICD-10-PCS procedure codes for laryngeal or pharyngeal, nose, mouth pharynx or facial surgery
- Any listed ICD-10-CM procedure codes for esophageal resection
- Any listed ICD-10-CM procedure codes for lung cancer
- Any listed ICD-10-CM diagnosis codes for degenerative neurological disorder
- Any listed ICD-10-CM procedure codes for lung transplant
- MDC 4 (diseases/disorders of respiratory system);
- MDC 5 (diseases/disorders of circulatory system);
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 12:

- Principal ICD-10-CM diagnosis code for proximal deep vein thrombosis (DVT) or pulmonary embolism (PE),
- Any secondary ICD-10-CM diagnosis code for DVT or PE present on admission, among patients otherwise qualifying for the numerator
- Procedure for interruption of vena cava occurs before or on the same day as the first operating room procedure
- Only operating room procedure was interruption of vena cava
- Any listed ICD-10-CM diagnosis code for acute brain or spinal injury present on admission
- Any listed ICD-10-PCS procedure code for extracorporeal membrane oxygenation (ECMO)
- Procedure for pulmonary arterial thrombectomy occurs before or on the same day as the first operating room procedure
- Only operating room procedure was for pulmonary arterial thrombectomy
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 13:

- Principal ICD-10-CM diagnosis code for sepsis or infection
- Any secondary ICD-10-CM diagnosis code for sepsis or infection present on admission, among patients otherwise qualifying for the numerator
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 14:

- Procedure for abdominal wall reclosure occurs on or before the day of the first open abdominopelvic surgery procedure, if any, and the day of the first laparoscopic abdominopelvic surgery procedure, if any
- Any listed ICD-10-CM diagnosis codes or any-listed ICD-10-PCS procedure codes for immunocompromised state
- Principal ICD-10-CM diagnosis code for disruption of internal operation wound

- Any secondary ICD-10-CM diagnosis code for disruption of internal operation wound present on admission
- Length of stay less than two (2) days-MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

PSI 15:

- Principal ICD-10-CM diagnosis code for accidental puncture or lacerations during a procedure
- Any secondary ICD-10-CM diagnosis code for accidental puncture or laceration during a procedure, among patients otherwise qualifying for the numerator
- MDC 14 (pregnancy, childbirth, and puerperium)
- Missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), or principal diagnosis (DX1=missing)

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

PSI 03: For a complete list of excluded codes, see attached technical specifications

PSI 06: For a complete list of excluded codes, see attached technical specifications

PSI 08: For a complete list of excluded codes, see attached technical specifications

PSI 09: For a complete list of excluded codes, see attached technical specifications

PSI 10: For a complete list of excluded codes, see attached technical specifications

PSI 11: For a complete list of excluded codes, see attached technical specifications

PSI 12: For a complete list of excluded codes, see attached technical specifications

PSI 13: For a complete list of excluded codes, see attached technical specifications and PSI Appendix D – Infection Diagnosis Codes

PSI 14: For a complete list of excluded codes, see attached technical specifications and PSI Appendix F – Immunocompromised State Diagnosis and Procedure Codes

PSI 15: For a complete list of excluded codes, see attached technical specifications

Excluded codes for all components are also available at:
https://www.qualitynet.org/inpatient/measures/psi/resources and
https://www.qualitynet.org/files/5ebeeee9641cb00023dd1f96?filename=2019_PSI_TechSpecs_Excel.zip

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

Not applicable.

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

Statistical risk model

If other:

S.12. Type of score:

Other (specify):

If other: Observed to expected ratio (component measures); Weighted average of the smoothed (empirical Bayes shrinkage) risk standardized observed to expected ratios (composite)

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Lower score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

For each component: The observed rate is the number of discharge records where the patient experienced the adverse event divided by the number of discharge records at risk for the event. The expected rate is a comparative rate that incorporates information about a reference population that is not part of the user's input dataset – what rate would be observed if the expected level of care observed in the reference population and estimated with risk adjustment regression models, were applied to the mix of patients with demographic and comorbidity distributions observed in the user's dataset? The expected rate is calculated only for risk-adjusted indicators.

The expected rate is estimated for each person using a generalized estimating equations (GEE) approach to account for correlation at the hospital or provider level.

The risk-adjusted rate is a comparative rate that also incorporates information about a reference population that is not part of the input dataset – what rate would be observed if the level of care observed in the user's dataset were applied to a mix of patients with demographics and comorbidities distributed like the reference population? The risk adjusted rate is calculated using the indirect method as observed rate divided by expected rate multiplied by the reference population rate. The smoothed rate is the weighted average of the risk-adjusted rate from the user's input dataset and the rate observed in the reference population; the smoothed rate is calculated with a shrinkage estimator to result in a rate near that from the user's dataset if the provider's rate is estimated in a stable fashion with minimal noise, or to result in a rate near that of the reference population if the variance of the estimated rate from the input dataset is large compared with the hospital-to-hospital variance estimated from the reference population. Thus, the smoothed rate is a weighted average of the risk-adjusted rate and the reference population rate, where the weight is the signal-to-noise ratio. In practice, the smoothed rate brings rates toward the mean, and tends to do this more so for outliers (such as rural hospitals)

The composite measure is a weighted average of the smoothed (empirical Bayes shrinkage) indirectly standardized morbidity ratios (observed to expected ratios) of the component indicators. The final weight for each component is based on two concepts: the volume of the adverse event and the harm associated with the adverse event.

The volume weights were calculated based on the number of safety-related events for the component indicators in the fee-for-service reference population. The harm weights were calculated by multiplying empirical estimates of the probability of excess harms associated with each patient safety event by the corresponding utility weights (1–disutility). Disutility is the measure of the severity of the adverse events associated with each of the harms (i.e., outcome severity, or least preferred states from the patient perspective). These excess harm probabilities were estimated by comparing patients with a safety-related event to very similar, otherwise eligible patients without that safety-related event over up to 1 year after the discharge during which the index event happened. Linked claims data for 2 years of Medicare Fee for Service beneficiaries (2016 - 2018) were used for this analysis. To account for confounders in estimating the marginal impact of each PSI on the risk of excess harms, inverse probability propensity weighting with indicator- and harm-specific propensity models were calculated that included age, sex, racial/ethnic categories, Medicaid eligibility, point of origin, modified Medicare Severity–Diagnosis-Related Group categories, Elixhauser comorbidities, and co-occurring PSIs.

CMS PSI 90 results center on 1.0 to improve interpretability. This means that the CMS PSI 90 composite value of the entire population of the input data equals 1.0. Hospital-level CMS PSI 90 results can be compared with 1.0. Adjusting for case mix, a CMS PSI 90 composite value less than 1.0 indicates a value better than the average of the reference population; likewise, a CMS PSI 90 composite value greater than 1.0 indicates a value worse than the average of the reference population.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

Not applicable.

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

Not applicable.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Claims

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

While the measure is tested and specified using fee-for-service data from the Centers for Medicare and Medicaid Services (CMS) (see section 1.1 and 1.2 of the measure testing form), the measure specifications and software are specified to be used with any ICD-10-CM-coded administrative billing/claims/discharge dataset with Present on Admission (POA) information.

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Facility

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Inpatient/Hospital

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

CMS PSI 90 was developed to provide a simple and transparent single metric that can be used to better understand, communicate and track patient safety in US hospitals. The indicator is comprised of eleven component PSIs which are calculated using readily available, routinely collected administrative data. The composite is conceptualized as a formative composite, meaning that the indicator is formed from a set of component indicators that assess different harms and aspects of patient safety.

CMS PSI 90 is a combination of the reliability-adjusted (smoothed), risk-standardized observed-to-expected ratio for each composite. CMS PSI 90 weights reflect a potential harm-based approach and are based on three components 1) excess harm associated with the PSI, 2) the estimated preferences for health states reflected

by these harms and 3) the volume of the PSI complication. Below we describe the methods used to quantify each portion of the harm based weights and the calculation method for the weights.

The excess harms for each component PSI were estimated using the CMS Inpatient and Outpatient Standard Analytic File (SAF) and the Denominator file. Potential harms, such as mortality, readmissions and additional treatments, were identified by a team of physicians and nurses for each PSI using literature review, environmental scan and clinical judgement. These harms were specified using variables available in the CMS dataset.

We estimated the average excess number of harmful outcomes associated with the occurrence of the PSI event using a separate cohort sample for each component indicator based on the patient records of patients qualifying for that PSI denominator. Index events included observations with the PSI, and the comparison group was those without the PSI. To account for potential confounding between the risk factors for developing a PSI and the risk factors for developing the harms independent of the PSI, we utilized propensity weighting using the risk models for each PSI. We used an inverse probability of treatment weighting approach to estimate the average treatment effect on the treated (ATT) or those with the PSI event. We followed patients for up to 1 year. Separate regression models were fit for each harm outcome. Linear probability models were used for binary outcomes (e.g. mortality, readmission) and a linear model was used to model length of stay.

To assign a relative value for decrements to the quality of life for each PSI event and its sequelae, we adopted the utility scale. A health utility refers to an individual's preference for a specific health state on a scale of 0 to 1, where 0 is equivalent to death and 1 denotes perfect health. These utility values weight different health states according to their relative valuation. They are widely used in health economic analysis (e.g., calculation of quality-of life years saved by a treatment) because they represent stable and assessable population values.

Because adequate utility values were not available in the literature for each health state and because our primary goal was to understand the relative disutility of each harm, we utilized a two-pronged approach in which we elicited relative rankings of each harm from clinicians and then fitted these rankings to known literature-based disutilities. The advantage of the utility approach is that it adopts a commonly used scale from 0-1 that can be converted to a harm scale (1-utility) to weight the relative quality-of-life effect on patients of a diverse set of PSI-related harms. Insignificant events to a patient are not given any weight since there is no disutility. Finally, average utility values represent a relative preference for one health state versus another at a group level, the appropriate analytic level for a quality indicator composite. Relative rankings of utilities are robust at the population level, regardless of utility assessment method chosen.

For each component indicator in the CMS PSI 90 composite, two sets of values need to be computed or estimated in order to calculate the final weights. The first is the excess risk of the outcomes (risk difference) that may occur in association with the indicator patient safety event. The second is the set of numerator weights, which are calculated from the volume (count) of component events in the CMS fee-for-service (FFS)reference population. Please see the testing attachment for additional details on the calculation of final weights including a formula for that calculation.

2. Validity – See attached Measure Testing Submission Form

PSI90-Composite-Testing-Attachment-508-FINAL.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include

information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

Yes - Updated information is included

- Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 531

Composite Measure Title: Patient Safety Indicator (PSI) 90: Patient Safety and Adverse Events Composite

Date of Submission: July 31, 2020

Composite Construction:

☒ Two or more individual performance measure scores combined into one score

☐ All-or-none measures (e.g., all essential care processes received or outcomes experienced by each patient)

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect **of testing**,*(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.*

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for different components in the composite, indicate the component after the checkbox. If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.17***)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☒ claims	☒ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	Measure Tested with Data From:
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other:	☐ other:

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

Most analyses were completed using Medicare Fee-for-Service (FFS) discharge data from Inpatient Prospective Payment System (IPPS) hospitals, including hospitals in Maryland and excluding Veteran's Administration Hospitals, from July 1, 2016 – June 30, 2019. The files included monthly inpatient (Part A) claims files[1] (Research Identifiable Files, or RIF) and Medicare Beneficiary Summary Files[2] from 2016 to 2019. These files contain diagnosis codes (ICD-10-CM), procedure codes (ICD-10-PCS), dates of service, cost and revenue codes, provider identifiers and beneficiary information. The final dataset included 13,611,933 individuals, 3345 hospitals, and 28,745,550 hospital stays.

Confirmatory testing was completed using selected Healthcare Cost and Utilization Project (HCUP) State Inpatient Databases (SID), because these data supported NQF's previous endorsement of PSI 90. HCUP is a family of health care databases and related software tools and products developed through a Federal-State-Industry partnership and sponsored by the Agency for Healthcare Research and Quality (AHRQ).[3] The HCUP SID contain all inpatient discharge abstracts from nonfederal acute care hospitals in participating States, translated into a uniform format to facilitate multi-State comparisons and analyses. For the analyses presented here, we used 2016-2017 data from Arizona, Florida, Kentucky, Nevada, Maryland, Washington, Maine, Minnesota and Nebraska, which include 13,390,121 hospital discharges. These data sets were selected because they had all necessary data elements (i.e., "present on admission" flag for each diagnosis code, procedure dates, de-identified hospital identifiers), were readily available through the HCUP Central Distributor, and offered geographic diversity at a reasonable cost.

1.3. What are the dates of the data used in testing?

July 1, 2016 – June 30, 2019 CMS data; January 1, 2016 – December 31, 2017 HCUP data

---

[1] Centers for Medicare and Medicaid Services (CMS) Research Identifiable Data Files (RIF). https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/IdentifiableDataFiles. Accessed June 15, 2020.

[2] Centers for Medicare and Medicaid Services (CMS) Master Beneficiary Summary File (MBSF) limited data set (LDS). https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/MBSF-LDS. Accessed June 15, 2020.

[3] HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2008-2012. Agency for Healthcare Research and Quality, Rockville, MD. http://www.hcup-us.ahrq.gov/sidoverview.jsp. (AHRQ QI Software Version 6.0 alpha)

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	Measure Tested at Level of:
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☐ other:	☐ other:

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

Table 1. Number and Descriptive Characteristics of Hospitals Included in Testing and Analysis, Medicare Fee-for-Service 2016-2019 and HCUP SID 2016-2017 (9 States)

Hospital Category	Medicare FFS Data (7/1/2016-6/30/2017)	Medicare FFS Data (7/1/2017-6/30/2018)	Medicare FFS Data (7/1/2018-6/30/2019)	HCUP All-Payer SID Data (1/1/2016-12/31/2016)	HCUP All-Payer SID Data (1/1/2017-12/31/2017)
Investor Owned; <100 beds	304	307	300	127	124
Investor Owned; >100 beds	460	455	448	118	117
Not-for-Profit (Rural); <100 beds	407	392	381	114	114
Not-for-Profit (Rural); >100 beds	236	236	236	23	23
Not-for-Profit (Urban); <100 beds	332	333	335	101	101
Not-for-Profit (Urban); 100-299 beds	781	779	769	113	120
Not-for-Profit (Urban); >300 beds	751	752	750	97	98
Total	**3,271**	**3,254**	**3,219**	**693**	**697**

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

HCUP all-payer discharges (1/1/2016-12/31/2017) for AZ, FL, KY, MD, ME, MN, NE, NV, WA, processed with AHRQ v2019 software.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

Table 2. Number of Patients Included in Testing and Analysis, by PSI 90 Component Indicator, Medicare Fee-for-Service 2016-2019 and HCUP SID 2016-2017 (9 States)

PSIs	Medicare FFS Data (7/1/2016-6/30/2017)	Medicare FFS Data (7/1/2017-6/30/2018)	Medicare FFS Data (7/1/2018-6/30/2019)	HCUP All-Payer SID Data (1/1/2016-12/31/2016)	HCUP All-Payer SID Data (1/1/2017-12/31/2017)
PSI 03	6,828,538	6,715,206	6,535,165	3,182,123	3,199,669
PSI 06	8,993,318	8,861,590	8,643,702	4,699,253	4,748,375
PSI 08	7,894,667	7,792,221	7,564,793	3,935,755	3,982,230
PSI 09	2,401,432	2,350,443	2,310,743	1,278,901	1,277,930
PSI 10	1,347,526	1,305,634	1,251,666	692,952	692,588
PSI 11	1,099,250	1,056,932	1,006,536	589,062	587,112
PSI 12	2,568,511	2,511,226	2,461,097	1,360,406	1,353,880
PSI 13	1,316,120	6,715,206	1,222,407	669,705	670,366
PSI 14A	270,185	262,204	258,029	167,207	165,728
PSI 14B	286,560	282,060	281,217	195,140	190,897
PSI 15	1,589,209	1,561,479	1,537,324	868,142	868,733

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

HCUP all-payer discharges (1/1/2016-12/31/2017) for AZ, FL, KY, MD, ME, MN, NE, NV, WA, processed with AHRQ v2019 software.

Notes: PSI 14A reflects postoperative wound dehiscence with an open approach and PSI 14B reflects non-open approach.

Table 3A. Descriptive Characteristics of Patients Included in Testing and Analysis, by PSI 90 Component Indicator, Medicare Fee-for-Service Population (2016-2019)

PSI	Male (%)	Female (%)	Mean age	Median age	Age (SD)	White (%)	Black (%)	Hispanic (%)	API (%)	Native Amer (%)	Other (%)
PSI 03	45.7	54.4	73.1	74.0	13.3	80.4	13.0	2.3	1.5	0.7	1.2
PSI 06	45.9	54.1	72.7	74.0	13.3	80.5	12.9	2.3	1.4	0.8	1.2
PSI 08	46.4	53.6	72.5	73.0	13.1	80.4	13.0	2.3	1.4	0.8	1.2
PSI 09	47.1	52.9	71.5	72.0	11.1	84.6	9.3	1.8	1.1	0.7	1.2
PSI 10	45.6	54.4	70.8	71.0	9.5	87.3	7.3	1.3	0.9	0.6	1.1
PSI 11	42.5	57.5	70.1	70.0	9.6	86.9	7.5	1.3	0.9	0.6	1.1

PSI	Male (%)	Female (%)	Mean age	Median age	Age (SD)	White (%)	Black (%)	Hispanic (%)	API (%)	Native Amer (%)	Other (%)
PSI 12	47.8	52.2	71.5	72.0	11.2	84.4	9.4	1.8	1.1	0.7	1.2
PSI 13	45.8	54.2	70.8	71.0	9.5	87.0	7.5	1.3	0.9	0.6	1.1
PSI 14A	45.3	54.7	71.1	71.0	10.8	85.1	9.1	1.6	1.1	0.6	1.2
PSI 14B	49.4	50.7	70.7	71.0	11.7	82.9	9.5	2.5	1.5	0.8	1.5
PSI 15	49.8	50.2	71.5	72.0	12.4	80.1	12.6	2.4	1.6	0.7	1.4

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

Notes: PSI 14A reflects postoperative wound dehiscence with an open approach and PSI 14B reflects non-open approach.

Abbreviations: SD=standard deviation; API=Asian or Pacific Islander

Table 3B. Descriptive Characteristics of Patients Included in Testing and Analysis, by PSI 90 Component Indicator, HCUP SID Population (2016) from 9 States

PSI	Male (%)	Female (%)	Mean age	Median age	Age (SD)	White (%)	Black (%)	Hispanic (%)	API (%)	Native Amer (%)	Other (%)
PSI 03	48.6	51.4	63.3	65.0	18.3	70.9	12.5	9.4	1.3	0.8	5.0
PSI 06	48.6	51.4	61.8	64.0	18.5	70.7	12.5	9.8	1.3	0.8	5.0
PSI 08	48.1	51.9	62.7	65.0	17.8	71.2	12.3	9.8	1.3	0.8	4.8
PSI 09	48.3	51.7	61.7	64.0	16.1	73.9	9.1	9.1	1.2	0.7	6.0
PSI 10	44.9	55.1	62.4	64.0	14.1	77.4	7.6	6.3	1.1	0.5	7.1
PSI 11	42.2	57.8	61.4	63.0	14.1	77.1	7.9	6.4	1.1	0.5	7.0
PSI 12	49.0	51.0	62.0	64.0	16.0	73.8	9.1	9.1	1.3	0.7	6.0
PSI 13	44.9	55.1	62.5	64.0	14.1	77.2	7.7	6.3	1.1	0.5	7.2
PSI 14A	42.9	57.1	59.5	61.0	16.1	70.3	11.3	9.7	1.4	0.7	6.6
PSI 14B	45.8	54.2	57.8	59.0	17.0	68.6	10.1	13.5	1.6	0.9	5.3
PSI 15	47.6	52.4	60.6	62.0	17.2	69.4	11.6	11.2	1.5	0.8	5.5

Source: HCUP all-payer discharges (1/1/2016-12/31/2016) for AZ, FL, KY, MD, ME, MN, NE, NV, WA, processed with AHRQ v2019 software.

Notes: Race set to WHITE for all encounters for Nebraska and Maine as no RACE provided in STATE SID data. Derived AGE for Maine as "age group" is reported instead of AGE.

PSI 14A reflects postoperative wound dehiscence with an open approach and PSI 14B reflects non-open approach.

Abbreviations: SD=standard deviation; API=Asian or Pacific Islander

Table 3C. Descriptive Characteristics of Patients Included in Testing and Analysis, by PSI 90 Component Indicator, HCUP SID Population (2017) from 9 States

PSI	Male (%)	Female (%)	Mean age	Median age	Age (SD)	White (%)	Black (%)	Hispanic (%)	API (%)	Native Amer (%)	Other (%)
PSI 03	49.0	51.0	63.4	66.0	18.4	71.8	12.7	9.5	1.4	0.7	4.0
PSI 06	49.0	51.0	62.0	64.0	18.5	71.6	12.7	9.8	1.4	0.7	3.9
PSI 08	48.5	51.5	63.1	65.0	17.8	72.0	12.5	9.7	1.4	0.7	3.7
PSI 09	48.6	51.5	62.1	64.0	15.9	75.0	9.2	9.2	1.3	0.6	4.7
PSI 10	45.2	54.8	62.8	65.0	14.0	78.7	7.7	6.4	1.2	0.4	5.6
PSI 11	42.5	57.5	61.8	64.0	14.0	78.5	8.0	6.5	1.2	0.4	5.4
PSI 12	49.2	50.9	62.3	64.0	15.9	74.9	9.2	9.2	1.4	0.6	4.7
PSI 13	45.2	54.8	62.9	65.0	14.0	78.5	7.8	6.4	1.3	0.4	5.6
PSI 14A	43.5	56.5	59.9	61.0	15.9	71.6	11.2	9.6	1.6	0.5	5.6
PSI 14B	46.0	54.0	58.0	59.0	17.0	69.1	10.3	13.6	1.8	0.7	4.4
PSI 15	48.1	51.9	60.9	63.0	17.1	70.3	11.7	11.2	1.7	0.7	4.5

Source: HCUP all-payer discharges (1/1/2017-12/31/2017) for AZ, FL, KY, MD, ME, MN, NE, NV, WA, processed with AHRQ v2019 software.

Notes: Race set to WHITE for all encounters for Nebraska and Maine as no RACE provided in STATE SID data. Derived AGE for Maine as "age group" is reported instead of AGE.

PSI 14A reflects postoperative wound dehiscence with an open approach and PSI 14B reflects non-open approach.

Abbreviations: SD=standard deviation; API=Asian or Pacific Islander

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

The data sets described above were used for all aspects of testing.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

We analyzed PSI 90 rates among dual-eligible (Medicaid plus Medicare) beneficiaries and racial/ethnic minority beneficiaries. The FFS data do not include other relevant information at the individual level. In addition, we categorized hospitals by the percentage of dual-eligible patients and the percentage of minority

patients among all of their Medicare FFS admissions. We do not currently have access to 9-digit zip code data necessary for geocoding to census tracts. We do not have patient-reported data.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

***Note**: Current guidance for composite measure evaluation states that reliability must be demonstrated for the composite performance measure score.*

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. Describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

Component Reliability: For component PSIs, we calculated measure reliability using estimated intraclass correlation coefficients (ICC), a type of signal-to-noise analysis.[4] Signal-to-noise analysis seeks to distinguish the true variation in measure scores across providers (signal) from measurement error (noise). For each hospital, reliability (also known as the signal-to-noise ratio) is defined as:

$$Reliability = \frac{signal\ variance}{signal\ variance + noise\ variance}$$

We define noise variance and signal variance as follows:

- **Noise variance**: the conditional variance of a measure (in this case, the observed hospital-level risk-adjusted rate), given the true risk-adjusted hospital rate, where the conditional variance is due to sampling error within each hospital.
- **Signal variance:** the between-hospital variance in the true value of the measures (that is, variation due to hospital performance).

We calculated the noise variance for each PSI 90 component as the sampling variance of the risk-adjusted rates, assuming each discharge follows a Bernoulli distribution, with the probability of an adverse event being estimated from the population risk-adjustment model. The calculation of the signal variance for each component measure assumes the following implicit two-stage model:

- Stage 1: True risk-adjusted hospital rates are distributed approximately normal (reference population rate, signal variance)
- Stage 2: Sampled risk-adjusted hospital rates, given true risk-adjusted hospital rates, are distributed approximately normal (true risk-adjusted hospital rate, noise variance)

To estimate the signal variance, we implemented an estimation procedure similar to the method proposed by Carl Morris.[5] We used an iterative algorithm because the noise variance for different hospitals varies substantially. The same iterative method is used in CMS v10.0 software to calculate smoothed PSI rates.

[4] Adams, John L. The Reliability of Provider Profiling: A Tutorial. Santa Monica, CA: RAND Corporation, 2009. http://www.rand.org/pubs/technical_reports/TR653.html

[5] Morris, C. "Parametric Empirical Bayes Inference: Theory and Applications." *Journal of the American Statistical Association,* vol. 78, no. 381, March 1983, pp. 47–55.

Composite Reliability: The signal-to-noise reliability approach does not apply to PSI 90 as a composite measure, because PSI 90 is a weighted average of risk-adjusted, reliability-adjusted (smoothed) component measures. In other words, each hospital's own signal-to-noise reliability is used as a shrinkage parameter to determine how far to shrink that hospital's estimate toward the national reference mean of 1.0. Through this process, noise variance is essentially removed. Therefore, we apply split-half and test-retest approaches to estimate the reliability of smoothed measures such as PSI 90.

For hospital h in subsample t where each hospital subsample is based on summarizing performance across a varying number of denominator-eligible cases (n_{ht}), we assumed that the smoothed and risk-adjusted performance measure for hospital h and subsample t (Y_{ht}) follows a simple two-level model:

$$Y_{ht} = \mu + \alpha_h + \varepsilon_{ht}$$

where the hospital effects (α_h) are sampled from a normal distribution with mean 0 and variance of hospital effects (σ_b^2) and the residual errors (ε_{ht}) are independently sampled from a normal distribution with mean 0 and variance:

$$\sigma_e^2/n_{ht}$$ [6]

The subsamples here could come from different calendar periods or from randomly generated subsamples (e.g. split-halves) of patients, stratified by hospital. In the split-half approach, we set T=2 without replacement, resulting in two records per hospital based on all-inclusive and mutually exclusive subsamples. In the test-retest approach, we "tested" using the publicly reported v10 data period, 7/1/2016-6/30/2018, and "retested" using the subsequent year of non-overlapping data. Note that the specification of the residual error variance assumes that, conditional on hospital random effects, the variance is inversely proportional to the sample size used to form the hospital-subsample estimate.

We used SAS PROC NLMIXED to analyze the dataset where the units of analysis are hospital subsample estimates. This allowed us to specify a two-level random effects model (hospital subsamples nested within hospital) to properly account for the between-observation variation in denominator sizes, so that we could obtain maximum likelihood estimates of the variance components, including the between hospital variance component (σ_b^2) and the error variance component (σ_e^2). These estimates were then used in a "plug-in" estimator of the classical intracluster correlation coefficient (ICC):

$$ICC(n) = \sigma_b^2 / [(\sigma_b^2 + (\sigma_e^2 / n)] = nR / (nR + 1),$$

where $R = \sigma_b^2 / \sigma_e^2$, which is the ratio of the between-hospital variance component (σ_b^2) over the error variance component (σ_e^2), and n is a hospital's denominator-eligible sample size.

Weighted averaging of multiple component PSIs (each of which is separately risk-adjusted and reliability-adjusted) helps to ensure the validity of the distributional assumption for PSI 90. By design, hospital-level PSI 90 values are centered around 1 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). [6] Because this ICC depends only on the ratio of between-hospital to within-hospital estimated variance components, and the number of denominator-eligible cases at each hospital, we can estimate reliability as a function of the hospital's

[6] The methods for constructing PSI 90 are fully described in: Quality Indicator Empirical Methods, prepared for Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. Revised September 2019. Available at: https://www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2019/Empirical_Methods_2019.pdf, accessed May 29, 2020.

subsample size, using an application of the Spearman-Brown prophecy formula.[7] We applied this methodology to hospital subsamples that were formed by randomly dividing the two years of patient data in the current v10 reporting period (July 1, 2016 through June 30, 2018)[8] from each hospital into two, then executing the PSI software separately on each split-half, to yield two estimates per hospital. We repeated this exercise after adding a subsequent year of data (July 1, 2018 through June 30, 2019) to assess test-retest reliability.

2a2.3. What were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

[7] 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

[8] https://www.qualitynet.org/inpatient/measures/psi/resources, accessed May 29, 2020.

Component Reliability:

Table 4. Weighted Mean Signal-to-Noise Reliability for PSI 90 Component Measures across Hospitals in Medicare Fee-for-Service Data, 2016-2018 (Comparing Current CMS v10.0 with Previous CMS v9.0)

PSIs	CMS v10.0 signal-to-noise reliability (N)	CMS v10.0 signal-to-noise reliability (Weighted mean)	CMS v9.0 signal-to-noise reliability (Weighted mean)
PSI 03	3,294	0.777	0.784
PSI 06	3,305	0.400	0.388
PSI 08	3,303	0.152	0.208
PSI 09	3,130	0.469	0.485
PSI 10	3,008	0.489	0.509
PSI 11	2,998	0.652	0.654
PSI 12	3,131	0.610	0.609
PSI 13	2,994	0.554	0.567
PSI 14	3,060	0.167	0.261
PSI 15	3,152	0.443	0.443

Source: CMS v10.0 findings were generated by UC Davis through analysis of Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018). CMS v9.0 represents analysis of Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) through CMS v9.0, reported in the CVP/Mathematica Scientific Acceptability Report, December 2019.

These reliability estimates are based on risk-adjusted measures; after smoothing or shrinkage, reliability estimates exceed 0.99.

Table 5. Distribution of Signal-to-Noise Reliability for PSI 90 Component Measures across Hospitals in Medicare Fee-for-Service Data, 2016-2018 (Using Current CMS v10.0)

PSIs	Median (50th percentile)	25th percentile	75th percentile	% Hospitals ≥0.4
PSI 03	0.668	0.347	0.829	71.3
PSI 06	0.132	0.039	0.283	11.6
PSI 08	0.036	0.012	0.077	0.03
PSI 09	0.113	0.035	0.288	13.4
PSI 10	0.079	0.024	0.241	11.5
PSI 11	0.39	0.180	0.616	44.3
PSI 12	0.341	0.153	0.564	41.1
PSI 13	0.178	0.063	0.387	21.6
PSI 14	0.026	0.010	0.059	0.1

PSIs	Median (50th percentile)	25th percentile	75th percentile	% Hospitals ≥0.4
PSI 15	0.148	0.054	0.297	13.0

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with CMS v10.0 software.

These reliability estimates are based on risk-adjusted measures; after smoothing or shrinkage, reliability estimates exceed 0.99.

Composite Reliability:

Table 6. Split Sample PSI 90 Reliability at Hospital Level in Medicare Fee-for-Service Data, 2016-2018

Reliability Assessment	24 months of data	36 months of data
Hospitals meeting 3 case minimum	3,305	3,305
Median Intracluster Correlation Coefficient (ICC)	0.74	0.81
% Hospitals meeting ICC≥0.6	67%	76%
% Hospitals meeting ICC≥0.4	83%	89%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with CMS v10.0 software.

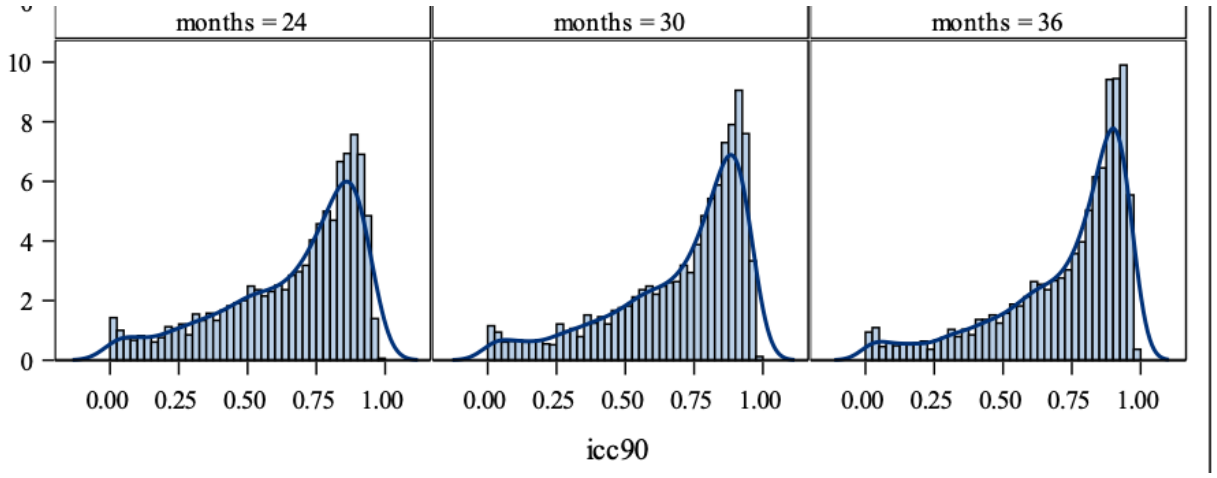
Table 7. Test-Retest PSI 90 Consistency at Hospital Level in Medicare Fee-for-Service Data, 2016-2019

Reliability Assessment	24 months of data	36 months of data
Hospitals meeting 3 case minimum	3,305	3,305
Median Intracluster Correlation Coefficient (ICC)	0.61	0.70
% Hospitals meeting ICC≥0.6	51%	62%
% Hospitals meeting ICC≥0.4	72%	81%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

In this analysis, the third year of data was "held out" to assess test-retest reliability.

Figure 1. Distribution of Split Sample Intracluster Correlation Coefficients for PSI 90 by Duration of Data Collection Period (Percentage of hospitals is shown on y axis)



Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

Table 8. Test-Retest PSI 90 Consistency at Hospital Level in All-Payer State Inpatient Data (2016-2017), Using 12 Months of Data to Estimate ICC (C,1)

State	# Hospitals	Overall ICC	95% CI, Lower Bound	95% CI, Upper Bound
9 States Combined	900	0.746	0.716	0.774

Source: HCUP inpatient discharge data from (1/1/2016-12/31/2017) processed with AHRQ v2019 software.

Abbreviations: CI=confidence interval; ICC=intracluster correlation coefficient

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Composite: Reliability scores vary from 0 to 1, with a score of 0 indicating that all variation is attributable to measurement error whereas a score of 1 implies that all variation is caused by real differences in performance across accountable entities. As shown in **Table 6**, PSI 90 demonstrates moderate-to-high score reliability at the hospital level, with an overall split half (intracluster correlation coefficient, ICC) reliability estimate of 0.74 based on 24 months of Medicare FFS claims data, and 67% of facilities exceeding ICC=0.6. If even higher reliability were desired, the data period could be increased to 36 months, with split half reliability of 0.81 and 76% of facilities exceeding ICC=0.6 **(Table 6)**. As the reliability distribution in **Figure 1** shows, only 2-3% of hospitals have very low reliability (ICC<0.05). CMS anticipates excluding most of these low-reliability hospitals from public reporting using a minimum volume threshold (e.g., 25 denominator records) and a missing data threshold, as described further below.

An even more rigorous test of reliability at the hospital level is to use a holdout sample from a separate, subsequent time period, a concept known as test-retest reliability. In this approach, the within-hospital signal is diluted by changes over time, as hospitals invest in quality improvement activities and systematically improve their performance at different rates. As shown in **Table 7**, the current 24-month reporting period still meets the reliability standard for hospital-level reporting with a median test-retest ICC of 0.61.

We also analyzed test-retest consistency using all-payer data from several states (instead of two years of Medicare FFS claims data) because PSI 90 was originally submitted to NQF as an all-payer claims-based measure. As shown in **Table 8**, these test-retest ICCs between test year 2016 and retest year 2017 varied across states but averaged to 0.746 (95% confidence interval, 0.716-0.774). This value is almost identical to that provided in AHRQ's 2015 submission for this measure (ICC=0.76), showing that score-level reliability has been consistent over time.

Components: As shown in **Tables 4 and 5**, signal-to-noise reliability varies across the PSI 90 component measures, with more frequent events (i.e., PSI 03, PSI 11, PSI 12, PSI 13) having higher signal-to-noise reliability (weighted mean >0.5) than rare events (i.e., PSI 08, PSI 14, with weighted mean <0.3). Among all the PSI 90 component measures, only PSI 03 would definitively meet a conventional threshold for public reporting (e.g., median >0.6 in **Table 5**, weighted mean >0.7 in **Table 4**) as a standalone measure using 24 months of Medicare FFS claims data, which highlights the importance of PSI 90 as a composite measure that draws statistical strength from all of its component measures.

2b1. VALIDITY TESTING

***Note**: Current guidance for composite measure evaluation states that validity should be demonstrated for the composite performance measure score. If not feasible for initial endorsement, acceptable alternatives include assessment of content or face validity of the composite OR demonstration of validity for each component. Empirical validity testing of the composite measure score is expected by the time of endorsement maintenance.*

2b1.1. What level of validity testing was conducted?

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ Composite performance measure score

☒ **Empirical validity testing**

☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

☒ **Validity testing for component measures** *(check all that apply)*

***Note**: applies to ALL component measures, unless already endorsed or are being submitted for individual endorsement.*

☐ Endorsed (or submitted) as individual performance measures

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ Empirical validity testing of the component measure score(s)

☒ Systematic assessment of face validity of component measure score(s) as an indicator of quality

or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

Face validity refers to the degree to which evidence, clinical judgment, 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.

In 2014-15, a standing workgroup of clinicians with experience relevant to the PSI90 composite and its component indicators was convened to review the measure and provide guidance regarding indicator refinements. Standing workgroup members were solicited via Federal Register notices to ensure that interested parties were allowed equal opportunity for participation. Members could be nominated or self-nominated. Nominees were selected by a stringent rating system that measured each nominee's quality measurement knowledge, quality improvement experience, clinical expertise, written publications related to the use and application of the AHRQ QIs and their knowledge of the NQF measure endorsement process. Workgroup members selected for participation were familiar with the routine updates and maintenance of the AHRQ QIs, relevant literature pertaining to potential enhancements of the AHRQ QIs, methodological changes and refinements, application to the software refinements, and AHRQ QI user needs. Panel composition was designed to ensure a wide variety of quality indicator knowledge and experience.

In 2019, CMS convened a new Technical Expert Panel, following standard processes outlined in the CMS Measures Management System Blueprint,[9] to advise the measure developer on updated specifications and scientific acceptability testing for PSI 90, and to assess the results of this testing. To determine face validity, we obtained input from members of this TEP to determine whether they think the measure as specified will help inform consumers and help providers improve quality. On July 20, 2020, TEP members voted 12-1 in favor of continued use of PSI 90, subject to reassessment as additional validation data and measures become available.

Component Validity: Predictive validity is a type of construct validity that focuses on a measure's ability to predict subsequent outcomes of well-established validity and clinical importance. In this case, we assessed predictive validity based on the estimated marginal effect of each PSI 90 component event on subsequent harms in the Medicare FFS population, after adjusting for the propensity of that PSI (and the occurrence of other PSIs that could contribute to causing the same outcomes). These subsequent harms include:

- Death at hospital discharge, and within 30 and 180 days after discharge (the longer time window was used for PSIs 03, 10, 11, 12, 13, and 14 based on expert input confirmed by empirical analysis of post-event survival curves)
- Readmission to an acute care hospital within 30 days after discharge (and inpatient days during that time window)
- Transfer to long-term acute, inpatient rehabilitation, or post-acute skilled nursing care (and inpatient days in that setting)
- Admission to long-term skilled nursing care (and days in that setting)
- Chronic dialysis (for PSI 10)
- Late complications such as tracheostomy to support long-term mechanical ventilation (PSI 11), osteomyelitis and other deep soft tissue infections (PSI 03), anoxic brain injury or other shock-related complications (PSI 09), extension of thrombosis or anticoagulant-related bleeding (PSI 12), enterocutaneous fistula or incisional hernia (PSI 14), and abscess or fistula (PSI 15).

We calculated excess harm risks for each PSI 90 component using CMS datasets: the Research Identifiable Files (RIF) and the Medicare Master Beneficiary Summary Files (MBSF). These files were used to estimate the average excess number of harmful outcomes associated with each component PSI using a separate cohort for each indicator based on denominator-eligible records. Index hospitalizations with the PSI (numerator event)

[9] https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/MMS-Blueprint, accessed 7/26/2020.

were compared with eligible hospitalizations without the PSI event. To account for potential confounding between the risk factors for developing a PSI and the risk factors for developing the harms independent of the PSI, we weighted observations by the inverse probability of treatment to estimate the "average treatment effect in the treated" for those with the PSI event. We then fit separate regression models for each harm outcome: probability models for binary outcomes and linear models for length of stay.

Composite Validity: Construct validity refers to the extent to which the measure generates estimates that are consistent with a construct or conceptual framework regarding how safe care is produced and defined. For example, convergent validity refers to whether multiple measures of an underlying concept are positively correlated with each other. To assess the convergent validity of the PSIs, we compared PSI results with related measures of patient safety and outcomes at the hospital level, publicly available on https://data.Medicare.gov. Using Spearman rank correlation coefficients, we compared hospital-level PSI rates with rates of complications for hip/knee replacement patients, risk-standardized 30-day readmission rates (e.g., hospital-wide unplanned all-cause readmissions) and health care-associated infection measures from the National Healthcare Safety Network (central line associated bloodstream infection, Clostridium difficele infection, catheter-associated urinary tract infection, surgical site infection, and methicillin-resistant Staphylococcus aureus). Correlations among these measures would support the validity of the PSIs 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 PSIs in terms of the populations and conditions being measured.

We further assessed convergent validity using the results of the Leapfrog Group's Hospital Safety Survey, which is used (in combination with PSI 90 and other measures) to assign Hospital Safety Grades. A key advantage of this latter data source is that hospital respondents are audited and asked to provide documentation to support the accuracy of their survey responses. We hypothesized that hospitals with greater implementation of safe practices would have lower PSI 90 rates than hospitals with less implementation of safe practices.

Known groups validity is a final 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 full-time 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%).

We hypothesized that PSI 90 rates (reflecting risk-adjustment) would be equivalent or lower at teaching hospitals and at hospitals with high nurse staffing and skill mix than at non-teaching hospitals and hospitals with low nurse staffing and skill mix, respectively.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Component Validity:

Table 9. Predictive Validity of PSI 90 Components at the Patient Level, Showing the Average Marginal Effect of Each PSI Event on Subsequent Adverse Outcomes (after Adjusting for Confounding Factors through Inverse Probability Propensity Weighting)

Adverse Outcome (absolute diff in days or %)	PSI 03	PSI 06	PSI 08	PSI 09	PSI 10	PSI 11	PSI 12	PSI 13	PSI 14	PSI 15
Hospital Length of Stay (days)	9.3	4.6	4.5	5.1	11.4	7.1	8.0	12.0	12.2	14.2
30-day readmission	5.0%	0.0%	9.1%	4.7%	6.3%	5.2%	5.8%	4.8%	8.4%	7.3%
Death (30*/180 days)	27.0 %	13.0 %*	7.3%*	4.5%*	32.7%	18.6%	13.4%	28.6%	10.8%	10.1% *
Long-term SNF admission	9.3%	0.0%	25.3%	3.1%	2.9%	6.3%	5.0%	6.5%	10.2%	5.4%
SNF Length of Stay (days)	8.8	0	18.6	2.5	1.8	4.7	3.9	4.9	8.2	4.8
Late complication**	7.4%	--	--	10.7%	--	--	--	--	1.7%	9.4%
Late operation***	4.9%	--	0.12%	--	--	--	--	--	--	--
Late incisional hernia	--	--	--	--	--	--	--	--	1.4%	--
Tracheostomy	--	--	--	--	--	14.0%	--	--	--	--
DVT/PE/bleed (ED visit)	--	--	--	--	--	--	47.6%	--	--	--
Long-term dialysis	--	--	--	--	4.3%	--	--	--	--	--

Source: CMS Inpatient and Outpatient Medicare Fee-For-Service data in the 100% standard analytical files (SAF).

Note: A separate cohort sample was drawn for each component PSI that was defined by that PSI's denominator criteria among inpatient stays in CY 2012. Data from CY 2013 were used for follow-up only, to ensure that a full 365 days of follow-up were available for each observation. Marginal effects were estimated using "average treatment effect in the treated" inverse propensity score weighting based on log odds of expected value from risk-adjustment models (0.72<c<0.91), with addition of sociodemographic factors and co-occurring PSIs, to account for confounding and alternate pathways to the same harm state.

N/A indicates harms that are not relevant to a particular component PSI measure.

** indicates 30-day death*

*** Late complications include osteomyelitis and other deep soft tissue infections for PSI 03 (Pressure Ulcer), anoxic brain injury or other shock-related complications for PSI 09 (Postoperative Hemorrhage or Hematoma), enterocutaneous fistula for PSI 14 (Postoperative Wound Dehiscence), and abscess or fistula for PSI 15 (Unrecognized Abdominopelvic Accidental Puncture/Laceration).*

*** Late operations include flap and graft procedures for PSI 03 (Pressure Ulcer) and reoperations to treat complications of the original repair for PSI 08 (In-hospital Fall with Hip Fracture).*

Abbreviations: DVT=deep vein thrombosis; ED=emergency department; PE=pulmonary embolism; SNF=skilled nursing facility

-- cell intentionally left blank

Composite Validity:

Table 10. Convergent Validity Between PSI 90 and Infection-Related Outcome Measures by Spearman Rank Correlation, Using Different Data Periods

Hospital Compare Measures	Hospitals	PSI 90 – CMS v10, 2016-2018	PSI 90 – CMS v10, 2017-2019
Hip/knee complication rate	2,387	0.149***	0.136***
Central line-associated bloodstream infection (CLABSI)	2,273	0.042*	0.040
Catheter-associated urinary tract infection	2,536	0.047*	0.060**
Clostridium difficile (C. diff) infection	2,946	0.054**	0.060**
Surgical-site infection (SSI) following abdominal hysterectomy/colon procedure	2,425	0.108***	0.104***
Methicillin-resistant Staphylococcus aureus (MRSA) bacteremia	2,131	0.058*	0.072**
Total healthcare-acquired condition (HAC) score	3,188	0.420***	0.347***

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with v10.0 CMS software.

Other measure results retrieved from Hospital Compare data on https://data.medicare.gov/ which was updated on 10/30/2019, including infection data from 1/1/2017-12/31/2018

** p<0.05, ** p<0.005, *** p<0.0005*

Table 11. Convergent Validity Between PSI 90 and 30-day Readmission Measures by Spearman Rank Correlation, Using Different Data Periods

Hospital Compare Measures	Hospitals	PSI 90 – CMS v10, 2016-2018	PSI 90 – CMS v10, 2017-2019
30-day readmission: Acute Myocardial Infarction (AMI)	2,061	0.024	0.016
30-day readmission: Coronary Artery Bypass Graft (CABG)	994	0.058	0.054
30-day readmission: Chronic Obstructive Pulmonary Disease (COPD)	2,836	0.037*	0.043*
30-day readmission: Heart Failure	2,856	0.059**	0.045*
30-day readmission: Hip and Knee	2,460	0.084***	0.096***
30-day readmission: Pneumonia	2,927	0.065***	0.070***
30-day readmission: Hospital-wide	3,140	0.138***	0.145***

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with v10.0 CMS software.

Other measure results retrieved from Hospital Compare data on https://data.medicare.gov/ which was updated on 10/30/2019, including readmission data from 7/1/2015-6/30/2018.

** p<0.05, ** p<0.005, *** p<0.0005*

Table 12. Convergent Validity Between PSI 90 and Leapfrog Survey Safe Practice Scores, Based on Mean PSI 90 Values by Category of Response

Safe Practice Score: Overall Performance	Hospitals (%)	Mean Score (SD)	Relative risk (compared with "fully meets standard")
Fully meets standard	1,493 (49.2%)	1.002 (0.215)	1
Substantial progress	183 (6.0%)	0.983 (0.179)	0.981
Some progress	45 (1.5%)	1.009 (0.193)	1.007
Willing to report	38 (1.3%)	1.033 (0.192)	1.031
Declined to respond	1,274 (42.0%)	0.986 (0.174)	0.984

Source: 2019 (v8.0) Leapfrog Hospital Survey linked with Medicare FFS discharges from IPPS hospitals (7/1/2017-6/30/2019) processed with v10.0 CMS software.

Abbreviation: SD=standard deviation

Table 13. Convergent Validity Between PSI 90 and Leapfrog Survey Safe Practice Scores by Spearman Rank Correlation, Excluding Hospitals that Declined to Respond

Performance on Safe Practice Measures	Hospitals	Mean Score (SD)	PSI 90 – CMS v10, 2017-2019
Culture of safety leadership structures and systems (out of 120 points)	1,759	116.92 (8.46)	0.034
Culture measurement, feedback and interventions (out of 120 points)	1,759	116.47 (12.95)	-0.020
Risks and hazards (out of 100 points)	1,759	97.25 (9.63)	-0.017
Nursing workforce (out of 100 points)	1,759	97.60 (9.09)	-0.021
Hand hygiene (out of 60 points)	1,759	57.22 (7.93)	-0.017

Source: 2019 (v8.0) Leapfrog Hospital Survey linked with Medicare FFS discharges from IPPS hospitals (7/1/2017-6/30/2019) processed with v10.0 CMS software.

Abbreviation: SD=standard deviation

Table 14. Known Groups Validity for PSI 90

Known Groups Category	Hospitals (%)	Mean	SD
Hospital Teaching	--	--	--
Resident FTE/bed ratio = 0	2,564 (79.3%)	0.978	0.168
Resident FTE/bed ratio (0 - 0.25)	471 (14.6%)	1.013	0.194
Resident FTE/bed ratio $\geq$0.25	196 (6.1%)	1.125	0.372

Known Groups Category	Hospitals (%)	Mean	SD
Hospital Nursing*	--	--	--
Nurse FTE/bed ratio <1.0	950 (29.4%)	0.985	0.154
Nurse FTE/bed ratio (1.0 – 2.0)	1,700 (52.6%)	0.995	0.188
Nurse FTE/bed ratio >2.0	581 (18.0%)	0.992	0.258
Nursing Skill Mix**	--	--	--
Low (<0.85)	765 (23.7%)	0.995	0.175
Medium (0.85-0.975)	1,359 (42.0%)	1.002	0.218
High (>0.975)	1,107 (34.3%)	0.976	0.172

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviation: FTE=full-time equivalent; SD=standard deviation

** Spearman rank correlation between nurse FTE/bed ratio and CMS Medicare PSI 90 is -0.042 (p<0.05).*

*** Spearman rank correlation between nursing skill mix percentage and CMS Medicare PSI 90 is -0.071 (p<0.0001)*

- - cell intentionally left blank

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

We assessed each component measure individually to determine whether it was valid to include in the composite measure by calculating the average marginal effects, or absolute difference in event rates, between balanced groups with and without the PSI. We used inverse propensity score weighting to balance the PSI-exposed and unexposed groups on measured characteristics, including all features in the corresponding risk-adjustment model and additional social risk factors (e.g., dual eligibility, race/ethnicity). As shown in **Table 9** above, all component events were independently predictive of hospital length of stay and 30-180 day mortality. For example, the average patient who experienced a stage 3, 4, or unstageable pressure injury (PSI 03) spent an extra 9.3 days in the hospital, and had a 27% higher absolute risk of death within 180 days, than an otherwise identical patient who did not experience PSI 03. All events except for iatrogenic pneumothorax (PSI 06) were also independently predictive of hospital readmissions and nursing home admissions. For example, the average patient who experienced PSI 03 spent an extra 8.8 days in post-acute skilled nursing care, and had a 5.0% higher absolute risk of readmission within 30 days and a 9.3% higher absolute risk of long-term nursing home placement, than an otherwise identical patient who did not experience PSI 03. These findings strongly support the predictive validity of PSI 90 and its components at the patient level.

Construct validity for PSI 90 at the hospital level is moderate. PSI 90 correlates satisfactorily with other, independently collected, NQF-endorsed measures of hospital harms, including hospital-acquired infection standardized morbidity ratios (**Table 10**) and 30-day risk-standardized readmission rates (**Table 11**). Given that PSI 90 is dominated by perioperative and postoperative complications, it is not surprising that these correlations are higher for surgical patients (e.g., surgical site infection rates and hip/knee complication rates in **Table 10**) than for medical patients. As shown in **Table 11**, rank correlations between PSI 90 and hospital-wide readmission rates are especially high at 0.138–0.145 (p<0.0001). (In previous NQF endorsement review, a very similar correlation of 0.11 (p<0.0001) was reported with the Potentially Preventable Readmission Rate,

based on a methodology developed by 3M Health Information Systems.) Although these correlations may seem low, they are consistent across all comparisons and they reflect the fact that each type of hospital performance measure contributes unique information.

To explore process-outcome relationships, we assessed hospital-level correlations between PSI 90 and implementation of evidence-based safe practices, as reported by hospitals in the annual survey conducted by the Leapfrog Group. **Table 12** shows that the 38 hospitals that responded to the Leapfrog survey but reported little progress in implementing Safe Practices performed 3.3% worse on PSI 90, on average, than the 1,676 hospitals that reported full or nearly full implementation, while the 45 hospitals that reported "some" progress performed 0.9% worse. **Table 13** confirms that hospitals that reported greater implementation of NQF-endorsed Safe Practices (with the exception of "Culture of Safety Leadership Structures and Systems") had nonsignificantly lower PSI 90 scores than hospitals that reported less implementation, with rank correlations of -0.017 to -0.021. These analyses are limited by ceiling effects in the Leapfrog survey; for example, the "Culture of Safety Leadership" score has a mean value of nearly 117 of 120, and a standard deviation of less than 9 (**Table 13**).

Analyses of known groups validity support PSI 90 for nursing skill mix, which has been the most consistent nursing-related correlate of hospital outcomes in prior research (see 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):3404-23; also Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K. Nurse-staffing levels and the quality of care in hospitals. *N Engl J Med* 2002;346(22):1715-22). As shown in **Table 14**, hospitals that rely almost entirely on registered nurses, with a skill mix ratio greater than 0.975, had at least 2% better performance on CMS Medicare PSI 90 than hospitals that rely more heavily on licensed practical or vocational nurses (i.e., mean PSI 90 score 0.976 versus 0.995-1.002). As summarized in footnotes to **Table 14**, rank correlations with PSI 90 scores were statistically significant for both nurse staffing ratios ($r=-0.042$, $p<0.05$) and skill mix ($r=-0.071$, $p<0.0001$). Expressed another way (not shown in tables), 28% and 43% of statistical outliers with low PSI 90 scores reported the highest levels of nurse staffing and skill mix, respectively, versus 23% and 31% of statistical outliers with high PSI 90 scores.

It may be surprising that teaching hospitals appear to perform worse on PSI 90 than non-teaching hospitals, given a robust literature on mortality measures showing the opposite relationship. However, PSI 90 is dominated by surgical complications for which trainees may have higher incidence than experienced practitioners. Higher standard deviations for teaching hospitals indicate substantial variation in performance within this subgroup of hospitals.

2b2. EXCLUSIONS ANALYSIS

Note: *Applies to the composite performance measure, as well all component measures unless they are already endorsed or are being submitted for individual endorsement.*

NA ☐ no exclusions — *skip to section 2b4*

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

For component indicators, the current exclusion criteria are intrinsically embedded into the logic of the indicator (in SAS), including all of the risk-adjustment models and other analyses presented here. With a few

minor exceptions, all of these exclusions were present through both previous rounds of NQF endorsement. To test the impact of these exclusions, we removed them one at a time from the logic of each component indicator, and enumerated the marginal and relative (%) increase in the number of numerator and denominator records as a result (and the resulting impact of dropping each exclusion on the observed indicator rate). The composite indicator does not apply additional exclusion criteria beyond those that are part of the component indicator specifications.

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

Table 15a. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Pressure Ulcer (PSI 03)

PSI 03	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**13,477,287**	*	**8,126**	*	**0.603**	*
Principal diagnosis of stage 3, 4 or unstageable PU	13,944	0.1%	117	1.4%	0.611	1.3%
All diagnoses of stage 3, 4 or unstageable PU are present on admission	254,575	1.9%	254,575	3132.8%	19.131	3072.9%
Severe burns (>20% BSA)	784	0.0%	11	0.1%	0.604	0.1%
Exfoliative disorders of the skin (>20% BSA)	91	0.0%	0	0.0%	0.603	0.0%
MDC 14 (obstetrics)	19,818	0.1%	1	0.0%	0.602	-0.1%
LOS <3 days	5,036,196	37.4%	73	0.9%	0.443	-26.5%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: PU – pressure ulcer; MDC – major diagnostic category; LOS – length of stay; BSA – body surface area

*cell intentionally left blank

Table 15b. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Iatrogenic Pneumothorax (PSI 06)

PSI 06	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**17,444,847**	*	**4,421**	*	**0.253**	*
Principal diagnosis of iatrogenic pneumothorax (or secondary diagnosis present on admission)	4,887	0.0%	1,571	35.5%	0.343	35.5%
Chest trauma or pleural effusion or MDC 14 (obstetrics)	690,998	4.0%	4,057	91.8%	0.467	84.5%
Thoracic surgery or lung, cardiac, or diaphragmatic procedure	513,188	2.9%	11,109	251.3%	0.865	241.2%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: MDC – major diagnostic category

*cell intentionally left blank

Table 15c. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, In Hospital Fall with Hip Fracture (PSI 08)

PSI 08	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**15,331,356**	*	**1,660**	*	**0.108**	*
Principal diagnosis of hip fracture	19,262	0.1%	5	0.3%	0.108	0.2%
Secondary diagnosis of hip fracture present on admission	14,425	0.1%	14,425	869.0%	1.048	868.1%
Principal diagnosis of seizure, syncope, stroke, coma, cardiac arrest, poisoning, trauma, delirium, psychosis, anoxic brain injury	1,646,475	10.7%	161	9.7%	0.107	-0.9%
Metastatic, lymphoid, or bone cancer	1,061,048	6.9%	133	8.0%	0.109	1.0%
MDC 14 (obstetrics)	33,360	0.2%	0	0.0%	0.108	-0.2%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: MDC – major diagnostic category

*cell intentionally left blank

Table 15d. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Perioperative Hemorrhage or Hematoma (PSI 09)

PSI 09	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**4,672,895**	*	**11,644**	*	**2.492**	*
Principal diagnosis of postprocedural hemorrhage or hematoma	7,410	0.2%	24	0.2%	2.493	0.0%
Secondary diagnosis of postprocedural hemorrhage or hematoma present on admission	3,080	0.1%	3,080	26.5%	3.149	26.4%
Only OR procedure is for control of hemorrhage or hematoma	84,581	1.8%	1,031	8.9%	2.664	6.9%
Control of hemorrhage or hematoma occurs before first OR procedure	452	0.0%	452	3.9%	2.588	3.9%
MDC 14 (obstetrics)	12,602	0.3%	4	0.0%	2.486	-0.2%
Coagulation disorder	318,445	6.8%	5,165	44.4%	3.368	35.1%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: MDC – major diagnostic category; OR – operating room

*cell intentionally left blank

Table 15e. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Postoperative Acute Kidney Injury Requiring Dialysis (PSI 10)

PSI 10	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**2,601,337**	*	**3,529**	*	**1.357**	*
Principal diagnosis of acute kidney injury	12,730	0.5%	10	0.3%	1.354	-0.2%
Secondary diagnosis of acute kidney injury present on admission	554	0.0%	554	15.7%	1.569	15.7%

PSI 10	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**2,601,337**	*	**3,529**	*	**1.357**	*
Dialysis on or before date of first OR procedure	17,923	0.7%	194	5.5%	1.421	4.8%
Dialysis access procedure on or before date of first OR procedure	1,977	0.1%	1	0.0%	1.356	0.0%
Principal diagnosis of cardiac arrest or dysrhythmia, shock, or chronic kidney disease	10,148	0.4%	31	0.9%	1.363	0.5%
Secondary diagnosis of cardiac arrest or dysrhythmia, shock, or chronic kidney disease on admission	990	0.0%	990	28.1%	1.737	28.0%
Solitary kidney or S/P nephrectomy	432	0.0%	12	0.3%	1.361	0.3%
MDC 14 (obstetrics)	6,942	0.3%	0	0.0%	1.353	-0.3%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: MDC – major diagnostic category; OR – operating room; S/P – status post

*cell intentionally left blank

Table 15f. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Postoperative Respiratory Failure (PSI 11)

PSI 11	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**2,105,347**	*	**12,954**	*	**6.153**	*
Principal diagnosis of postprocedural respiratory failure	0	0.0%	0	0.0%	6.153	0.0%
Secondary diagnosis of postprocedural respiratory failure present on admission	1,297	0.1%	1,297	10.0%	6.765	9.9%
Tracheostomy is only OR procedure	5	0.0%	0	0.0%	6.153	0.0%

PSI 11	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**2,105,347**	*	**12,954**	*	**6.153**	*
Tracheostomy occurs before first OR procedure	0	0.0%	0	0.0%	6.153	0.0%
Neuromuscular disorder	6,518	0.3%	239	1.8%	6.247	1.5%
Laryngeal, pharyngeal, nose, mouth, facial surgery	21,753	1.0%	1,991	15.4%	7.026	14.2%
Esophageal resection	3,913	0.2%	374	2.9%	6.319	2.7%
Lung cancer	806	0.0%	31	0.2%	6.165	0.2%
Degenerative neurologic disorder	50,540	2.4%	1,304	10.1%	6.614	7.5%
Lung transplant	9	0.0%	1	0.0%	6.153	0.0%
MDC 4 (respiratory), MDC 5 (circulatory), MDC 14 (obstetrics)	401,708	19.1%	11,996	92.6%	9.952	61.7%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: MDC – major diagnostic category; OR – operating room

*cell intentionally left blank

Table 15g. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Perioperative Pulmonary Embolism or Deep Vein Thrombosis (PSI 12)

PSI 12	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**4,996,379**	*	**18,788**	*	**3.760**	*
Principal diagnosis of DVT or PE	3,562	0.1%	39	0.2%	3.765	0.1%
Secondary diagnosis of DVT or PE present on admission	25,893	0.5%	25,893	137.8%	8.897	136.6%
IVC filter on or before date of first OR procedure	3,619	0.1%	678	3.6%	3.893	3.5%
PA thrombectomy on or before date of first OR procedure	122	0.0%	14	0.1%	3.763	0.1%
IVC filter or PA thrombectomy is only OR procedure	0	0.0%	0	0.0%	3.760	0.0%

PSI 12	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**4,996,379**	*	**18,788**	*	**3.760**	*
Intracranial or spinal cord trauma	61,687	1.2%	966	5.1%	3.905	3.9%
Extracorporeal membrane oxygenation	4,039	0.1%	200	1.1%	3.797	1.0%
MDC 14 (obstetrics)	12,896	0.3%	4	0.0%	3.751	-0.2%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: MDC – major diagnostic category; OR – operating room; DVT – deep vein thrombosis; PE – pulmonary embolism; IVC – inferior vena cava; PA – pulmonary arterial

*cell intentionally left blank

Table 15h. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Postoperative Sepsis (PSI 13)

PSI 13	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**2,537,482**	*	**12,142**	*	**4.785**	*
Principal diagnosis of sepsis	0	0.0%	0	0.0%	4.785	0.0%
Secondary diagnosis of sepsis present on admission	0	0.0%	0	0.0%	4.785	0.0%
Principal diagnosis of bacterial infection	98,630	3.9%	720	5.9%	4.879	2.0%
Secondary diagnosis of bacterial infection present on admission	3,444	0.1%	3,444	28.4%	6.134	28.2%
MDC 14 (obstetrics)	6,930	0.3%	4	0.0%	4.774	-0.2%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: MDC – major diagnostic category

*cell intentionally left blank

Table 15i. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Postoperative Wound Dehiscence (PSI 14)

PSI 14	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**1,094,111**	*	**1,000**	*	**0.914**	*
Wound closure on or before date of first OR procedure	3,916	0.4%	402	40.2%	1.277	39.7%
Principal diagnosis of disruption of surgical wound or secondary diagnosis of disruption of surgical wound present on admission	4,398	0.4%	105	10.5%	1.006	10.1%
Diagnosis or procedure indicating immuno-compromised state	246,221	22.5%	483	48.3%	1.106	21.1%
LOS < 2 days	105,746	9.7%	0	0.0%	0.833	-8.8%
MDC 14 (obstetrics)	3,514	0.3%	0	0.0%	0.911	-0.3%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: LOS – length of stay; MDC – major diagnostic category; OR – operating room

*cell intentionally left blank

Table 15j. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Observed National Rate per 1000, Unrecognized Abdominopelvic Accidental Puncture or Laceration (PSI 15)

PSI 15	Denominator Count (N)	Denominator Percent Change	Numerator Count (N)	Numerator Percent Change	Rate per 1,000 Dropping Exclusion	Rate per 1,000 Percent Change
Current specification	**3,098,185**	*	**3,910**	*	**1.262**	*
Principal diagnosis of accidental puncture or laceration	3,524	0.1%	5	0.1%	1.262	0.0%
Secondary diagnosis of accidental puncture or laceration present on admission	526	0.0%	526	13.5%	1.432	13.4%
MDC 14 (obstetrics)	7,378	0.2%	4	0.1%	1.260	-0.1%

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

Abbreviations: MDC – major diagnostic category

*cell intentionally left blank

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note:* ***If patient preference is an exclusion****, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

Several exclusions are applied to each indicator because they are part of the inherent design of the Patient Safety Indicators. First, all indicators (**Tables 15a-15j**) exclude cases with a qualifying complication code in the principal position, because these diagnoses are present on admission based on the definition of "principal diagnosis" as "the condition established after study to be chiefly responsible for occasioning the admission of the patient for care." Second, all indicators exclude cases with a qualifying complication code in a secondary position but classified as "present on admission" (POA)(including "clinically undetermined"), because it is not appropriate to attribute such complications to that hospital. For PSI 13 (**Table 15h**), these "principal diagnosis" and "secondary diagnosis present POA" exclusions are applied not just to the numerator-triggering diagnosis of sepsis, but also to associated bacterial infections. For PSI 10 (**Table 15e**), these exclusions are applied not just to the numerator-triggering diagnosis of acute kidney failure, but also to associated conditions such as cardiac arrest or ventricular arrhythmia, shock, or stage 5 chronic kidney disease. Third, obstetric patients (MDC=14) are excluded from all indicators (**Tables 15a-15j**) because these patients are at very low risk for PSI 90 events, and when they are at risk, different and less specific ICD-10-CM codes are typically used. Completely different technical specifications would be needed for the obstetric population. Fourth, the indicators that involve procedures as part of their numerator definitions (**Tables 15d-15g, 15i**) all have exclusions for situations where the numerator-triggering (or numerator-related, for PSIs 11 and 12) procedure is the only operating room procedure or preceded the procedure(s) that qualified the record for the denominator. This last scenario merits exclusion because it suggests that the complication of interest developed preoperatively rather than postoperatively. Although dropping these four types of denominator exclusions would have a substantial quantitative impact, as shown in **Tables 15a-15j**, they cannot be dropped because they help to operationalize each indicator's focus on complications that occurred during that hospital stay and/or after an OR procedure.

Additional exclusions are applied on an indicator-specific basis to reduce the false positive rate or to exclude patients for whom expert panel members agreed that the indicator event is essentially nonpreventable. For example, length of stay exclusions are applied to PSI 03 (**Table 15a**) and PSI 14 (**Table 15i**) because of previous findings (from clinical research corroborated by chart review studies) that these complications almost never arise within that length-of-stay period. Patients with severe burns or exfoliative disorders of the skin, affecting at least 20% of body surface area, are excluded from PSI 03 (**Table 15a**) because they are at very high risk for additional skin injury and it is often infeasible to turn them frequently. Similarly, patients with intracranial or spinal cord trauma are excluded from PSI 12 (**Table 15g**) because they cannot receive pharmacologic thromboprophylaxis, while patients on extracorporeal membrane oxygenation are excluded because they have a very high risk of venous thromboembolism despite pharmacologic thromboprophylaxis. Patients receiving procedures that require traversing the pleural space are excluded from PSI 06 (**Table 15b**) because this indicator focuses on patients in whom the pleural space is inadvertently entered (e.g., during central venous catheterization). Patients requiring prolonged intubation to protect a vulnerable airway during facial, oropharyngeal, or laryngeal surgery are excluded from PSI 11 (**Table 15f**) because this indicator uses prolonged intubation as a numerator trigger suggesting respiratory failure.

None of the PSI exclusions imposes a burden by increasing the complexity of data collection or analysis, because they use standard claims data elements and are embedded within the public-use software. None of the PSI exclusions is based on patient or provider preference. All of the PSI exclusions have been recommended or endorsed by expert panels convened by AHRQ and/or CMS.

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

Note: *Applies to all outcome or resource use component measures, unless already endorsed or are being submitted for individual endorsement.*

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b4.

2b3.1. What method of controlling for differences in case mix is used? *(check all that apply)*

☐ Endorsed (or submitted) as individual performance measures

☐ **No risk adjustment or stratification**

☒ **Statistical risk model with 49 (PSI 14B) - 135 (PSI 03) risk factors**

☐ **Stratification by risk categories**

☐ **Other**

2b3.1.1 If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

Each of the PSI component risk models is shown below in **Tables 16a-16k** with risk factor names and denominator prevalences, coefficient estimates, and 95% confidence intervals surrounding those coefficient estimates. **Tables 16 i-16j** represent risk models for two versions of PSI 14; PSI 14A reflects postoperative wound dehiscence with an open approach and PSI 14B reflects the same complication with a laparoscopic, endoscopic, or percutaneous approach. In each model, risk factors are listed in the following sequence:

- Age categories, generally 5 years in width, using the youngest category as the omitted referent
- Sex categories, using female as the omitted referent
- Two-way age-sex interactions, which allow for different age-outcome relationships among men versus women
- AHRQ (Elixhauser) comorbidities, coded using publicly available HCUP software[10]
- Major Diagnostic Categories (MDCs) based on the body system of the principal diagnosis[11]

[10] https://www.hcup-us.ahrq.gov/toolssoftware/comorbidityicd10/comorbidity_icd10.jsp

[11] MDCs fully described in Table A.3 of the PSI Parameter Estimates: https://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V2019/Parameter_Estimates_PSI_v2019.pdf

- Modified Diagnosis Related Groups (MDRGs), based on aggregation of adjacent MS-DRGs with or without comorbidities and complications [12]
- Admission by transfer in from another hospital

Table 16a. PSI 03 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-6.9393	-7.1295	-6.7491	0.0000
Age_50_54	428,693	0.0493	-0.2124	0.3110	0.7120
Age_55_59	641,627	0.4157	0.1870	0.6444	0.0004
Age_60_64	798,260	0.3956	0.1791	0.6121	0.0003
Age_65_69	2,152,971	0.2864	0.0976	0.4752	0.0030
Age_70_74	2,129,934	0.2864	0.1058	0.4670	0.0019
Age_75_79	1,997,245	0.4518	0.2662	0.6373	0.0000
Age_80_84	1,808,102	0.3704	0.1812	0.5595	0.0001
Age_85_89	1,539,535	0.2810	0.0877	0.4743	0.0044
Age_90Plus	1,217,648	0.2798	0.0701	0.4894	0.0089
MALE	6,152,581	0.1975	-0.0073	0.4024	0.0588
Age_50_54*MALE	220,247	0.1783	-0.1524	0.5091	0.2906
Age_55_59*MALE	329,821	-0.1949	-0.4789	0.0892	0.1787
Age_60_64*MALE	406,944	-0.0610	-0.3306	0.2086	0.6576
Age_65_69*MALE	1,071,254	-0.0236	-0.2579	0.2108	0.8437
Age_70_74*MALE	1,024,245	-0.0118	-0.2379	0.2143	0.9186
Age_75_79*MALE	927,267	-0.1334	-0.3554	0.0886	0.2388
Age_80_84*MALE	788,061	-0.0524	-0.2886	0.1838	0.6639
Age_85_89*MALE	611,629	-0.0275	-0.2729	0.2179	0.8261
Age_90Plus*MALE	398,523	-0.0707	-0.3468	0.2053	0.6155
DM	1,812,405	-0.1696	-0.2588	-0.0804	0.0002
CHF	2,523,499	0.6575	0.5973	0.7178	0.0000
ARTH	608,252	-0.0667	-0.1882	0.0548	0.2819
COAG	947,840	0.1231	0.0549	0.1913	0.0004
DMCX	3,015,976	0.3197	0.2629	0.3765	0.0000
METS	427,596	0.0926	-0.0153	0.2004	0.0926
PARA	652,603	0.9632	0.8930	1.0335	0.0000
HTN_C	8,919,910	-0.3905	-0.4426	-0.3384	0.0000
LIVER	597,668	-0.1205	-0.2198	-0.0212	0.0174
LYMPH	164,994	-0.0966	-0.2820	0.0887	0.3068
LYTES	4,584,880	0.3168	0.2655	0.3680	0.0000
NEURO	1,705,282	0.1905	0.1285	0.2525	0.0000
OBESE	2,211,313	-0.0728	-0.1387	-0.0069	0.0305
PSYCH	579,218	0.1229	0.0122	0.2337	0.0296
TUMOR	432,435	-0.1160	-0.2359	0.0039	0.0580
ULCER	148,708	0.2295	0.0730	0.3859	0.0041
VALVE	970,433	-0.2394	-0.3299	-0.1488	0.0000

[12] MDRGs fully described in Table A.2 of the PSI Parameter Estimates: https://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V2019/Parameter_Estimates_PSI_v2019.pdf

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
ALCOHOL	408,480	-0.1447	-0.2804	-0.0091	0.0365
ANEMDEF	3,415,573	-0.0651	-0.1224	-0.0079	0.0258
BLDLOSS	153,760	0.0389	-0.1474	0.2251	0.6826
DEPRESS	2,071,963	-0.3796	-0.4550	-0.3043	0.0000
HYPOTHY	2,656,060	-0.3659	-0.4340	-0.2979	0.0000
CHRNLUNG	3,723,966	-0.3068	-0.3591	-0.2544	0.0000
PERIVASC	1,218,098	0.0927	0.0272	0.1582	0.0055
PULMCIRC	417,600	0.2096	0.1023	0.3169	0.0001
RENLFAIL	3,554,904	0.1960	0.1391	0.2530	0.0000
WGHTLOSS	1,271,554	1.1675	1.1092	1.2258	0.0000
MDC_1	1,003,014	-1.1050	-1.2696	-0.9404	0.0000
MDC_3	81,467	-1.3694	-1.7045	-1.0342	0.0000
MDC_4	1,911,180	-0.2771	-0.4139	-0.1403	0.0001
MDC_6	1,401,245	-1.1827	-1.3309	-1.0344	0.0000
MDC_7	398,613	-1.6503	-1.8437	-1.4569	0.0000
MDC_8	1,700,694	-0.8931	-1.0159	-0.7704	0.0000
MDC_9	314,500	-0.9360	-1.1818	-0.6902	0.0000
MDC_10	462,721	-1.0667	-1.2840	-0.8494	0.0000
MDC_11	1,048,492	-1.1485	-1.3132	-0.9838	0.0000
MDC_12	41,734	-1.5918	-2.1290	-1.0547	0.0000
MDC_13	47,645	-1.5774	-2.1206	-1.0343	0.0000
MDC_16	185,029	-1.4871	-1.8257	-1.1485	0.0000
MDC_17	109,212	-1.2364	-1.5100	-0.9628	0.0000
MDC_19	184,580	-1.8724	-2.4968	-1.2480	0.0000
MDC_20	91,079	-2.2090	-2.7955	-1.6225	0.0000
MDC_21	171,233	-0.6190	-0.8633	-0.3747	0.0000
MDRG_103	59,807	-0.5754	-1.0346	-0.1163	0.0140
MDRG_104	11,993	1.1051	0.7046	1.5057	0.0000
MDRG_108	20,462	1.0442	0.7353	1.3530	0.0000
MDRG_114	304,059	-0.7314	-0.9924	-0.4705	0.0000
MDRG_118	34,491	-1.6837	-2.6622	-0.7052	0.0007
MDRG_125	59,376	-0.7300	-1.1996	-0.2604	0.0023
MDRG_128	89,378	-0.9362	-1.4010	-0.4713	0.0001
MDRG_401	61,062	-0.5456	-0.8742	-0.2170	0.0011
MDRG_403	79,788	-2.5291	-3.1900	-1.8681	0.0000
MDRG_404	183,778	-1.2082	-1.4220	-0.9944	0.0000
MDRG_405	46,299	-1.5970	-2.0991	-1.0949	0.0000
MDRG_406	27,654	-2.8628	-4.3680	-1.3576	0.0002
MDRG_408	235,550	-1.6209	-1.9457	-1.2962	0.0000
MDRG_409	439,619	-2.7845	-3.1256	-2.4433	0.0000
MDRG_410	437,751	-1.8956	-2.1356	-1.6557	0.0000
MDRG_411	18,860	-2.6824	-4.2322	-1.1326	0.0007
MDRG_413	63,545	-4.4049	-6.6131	-2.1968	0.0001
MDRG_416	124,853	-0.3368	-0.5237	-0.1499	0.0004
MDRG_503	89,127	-0.1032	-0.2941	0.0876	0.2890
MDRG_505	16,186	-0.6012	-1.1274	-0.0750	0.0251

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_507	114,562	-0.6900	-0.9189	-0.4612	0.0000
MDRG_509	33,181	0.4005	0.2007	0.6004	0.0001
MDRG_510	90,824	-1.6854	-2.0745	-1.2963	0.0000
MDRG_511	147,482	-2.3070	-2.7037	-1.9103	0.0000
MDRG_514	114,381	-0.8282	-1.0567	-0.5997	0.0000
MDRG_515	8,361	-1.0460	-1.7876	-0.3044	0.0057
MDRG_517	10,400	-1.2004	-2.0154	-0.3854	0.0039
MDRG_520	213,117	-1.8885	-2.1574	-1.6195	0.0000
MDRG_522	140,375	-2.1538	-2.5533	-1.7543	0.0000
MDRG_524	764,434	-2.0281	-2.2010	-1.8552	0.0000
MDRG_526	2,909	0.1643	-0.6218	0.9504	0.6821
MDRG_527	77,619	-1.6872	-2.1005	-1.2740	0.0000
MDRG_529	45,080	-4.0844	-5.6767	-2.4920	0.0000
MDRG_531	251,190	-2.3571	-2.7099	-2.0043	0.0000
MDRG_533	79,548	-3.9946	-5.4448	-2.5445	0.0000
MDRG_535	97,319	-1.5866	-1.8743	-1.2989	0.0000
MDRG_540	40,248	-1.7141	-2.3417	-1.0865	0.0000
MDRG_541	19,076	-0.2418	-0.6544	0.1708	0.2507
MDRG_542	57,121	-0.2639	-0.4940	-0.0339	0.0245
MDRG_543	30,150	-1.8713	-2.6728	-1.0697	0.0000
MDRG_601	50,755	0.3791	0.0826	0.6756	0.0122
MDRG_602	212,677	0.5156	0.3171	0.7141	0.0000
MDRG_614	39,126	-0.5376	-0.9940	-0.0813	0.0209
MDRG_615	298,309	-0.9529	-1.2229	-0.6830	0.0000
MDRG_619	139,199	-1.2830	-1.7052	-0.8608	0.0000
MDRG_620	253,999	-1.6213	-2.0116	-1.2310	0.0000
MDRG_706	4,882	1.2226	0.4765	1.9686	0.0013
MDRG_806	56,953	-0.6162	-1.1216	-0.1107	0.0169
MDRG_807	516,280	-0.6772	-0.8794	-0.4749	0.0000
MDRG_834	99,202	-1.0851	-1.5124	-0.6577	0.0000
MDRG_839	49,346	-1.0200	-1.6198	-0.4203	0.0009
MDRG_910	202,943	-1.2233	-1.6112	-0.8354	0.0000
MDRG_1007	129,515	-0.6016	-0.9568	-0.2463	0.0009
MDRG_1008	202,363	-0.5768	-0.9040	-0.2495	0.0006
MDRG_1102	13,201	0.7752	0.1965	1.3538	0.0086
MDRG_1104	18,452	0.1046	-0.4557	0.6648	0.7145
MDRG_1109	25,290	0.5966	0.2930	0.9002	0.0001
MDRG_1110	391,682	-0.9325	-1.1454	-0.7195	0.0000
MDRG_1113	327,045	-0.9120	-1.1601	-0.6638	0.0000
MDRG_1604	107,320	-1.2930	-1.9647	-0.6214	0.0002
MDRG_1708	9,268	1.2555	0.7481	1.7629	0.0000
MDRG_1801	178,767	0.3420	0.2359	0.4481	0.0000
MDRG_1807	56,211	0.4054	0.2578	0.5530	0.0000
MDRG_1808	1,202,526	-1.2325	-1.3564	-1.1087	0.0000
MDRG_1915	130,488	-1.6975	-2.7906	-0.6044	0.0023
MDRG_2104	33,368	0.0124	-0.3850	0.4099	0.9511

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_2107	64,592	-1.5039	-2.0944	-0.9133	0.0000
MDRG_2303	56,563	-2.0671	-2.6760	-1.4581	0.0000
MDRG_2406	789	0.8576	-0.0600	1.7752	0.0670
MDRG_2408	7,631	0.1912	-0.3302	0.7127	0.4723
MDRG_7701	4,765	1.7734	1.4741	2.0727	0.0000
TRNSFER	900,717	0.6103	0.5261	0.6944	0.0000

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16b. PSI 06 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-8.9678	-9.2417	-8.6938	0.0000
Age_50_54	575,901	0.1466	-0.2738	0.5670	0.4944
Age_55_59	846,152	0.2961	-0.0629	0.6551	0.1060
Age_60_64	1,031,400	0.4993	0.1680	0.8306	0.0031
Age_65_69	2,949,377	0.6135	0.3270	0.9000	0.0000
Age_70_74	2,831,610	0.6324	0.3440	0.9209	0.0000
Age_75_79	2,563,445	0.6633	0.3757	0.9510	0.0000
Age_80_84	2,260,597	0.7519	0.4660	1.0378	0.0000
Age_85_89	1,876,023	0.6605	0.3736	0.9474	0.0000
Age_90Plus	1,463,804	0.5989	0.3112	0.8867	0.0000
MALE	8,013,134	-0.1792	-0.5495	0.1912	0.3431
Age_50_54*MALE	295,697	0.0189	-0.5521	0.5898	0.9484
Age_55_59*MALE	435,323	-0.1434	-0.6396	0.3527	0.5710
Age_60_64*MALE	525,830	-0.5306	-1.0510	-0.0103	0.0456
Age_65_69*MALE	1,464,382	-0.3621	-0.7643	0.0402	0.0777
Age_70_74*MALE	1,364,214	-0.4828	-0.8867	-0.0789	0.0191
Age_75_79*MALE	1,193,600	-0.2116	-0.6073	0.1841	0.2947
Age_80_84*MALE	990,096	-0.2173	-0.6220	0.1874	0.2926
Age_85_89*MALE	748,832	-0.1689	-0.5716	0.2339	0.4112
Age_90Plus*MALE	479,084	-0.0828	-0.4990	0.3335	0.6967
DM	2,457,820	-0.4884	-0.5904	-0.3864	0.0000
DMCX	3,610,862	-0.4102	-0.5051	-0.3153	0.0000
DRUG	390,137	-0.0639	-0.3306	0.2028	0.6387
PARA	812,454	0.2555	0.1161	0.3950	0.0003
HTN_C	11,543,196	-0.0568	-0.1191	0.0055	0.0739
LIVER	697,298	0.0939	-0.0751	0.2630	0.2760
LYMPH	188,691	-0.3017	-0.6407	0.0374	0.0812
LYTES	5,368,647	0.0647	-0.0045	0.1339	0.0670
NEURO	2,078,347	0.0094	-0.0855	0.1043	0.8456
OBESE	2,768,015	-0.6502	-0.7647	-0.5357	0.0000
PSYCH	734,606	0.0325	-0.1492	0.2141	0.7261
ALCOHOL	497,807	-0.1086	-0.3219	0.1047	0.3184
ANEMDEF	3,930,571	-0.0841	-0.1626	-0.0056	0.0357
BLDLOSS	177,041	-0.0988	-0.4189	0.2213	0.5452

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
DEPRESS	2,609,164	-0.1827	-0.2779	-0.0874	0.0002
HYPOTHY	3,344,457	0.0039	-0.0737	0.0815	0.9218
CHRNLUNG	4,481,843	0.0919	0.0189	0.1650	0.0137
PERIVASC	1,465,865	-0.0090	-0.1137	0.0957	0.8668
PULMCIRC	441,588	0.1242	-0.0689	0.3173	0.2076
RENLFAIL	4,237,449	-0.1419	-0.2228	-0.0611	0.0006
WGHTLOSS	1,326,399	0.7815	0.6919	0.8712	0.0000
MDRG_102	25,009	1.9448	1.5880	2.3017	0.0000
MDRG_103	76,034	1.0590	0.7232	1.3947	0.0000
MDRG_111	80,860	-1.0837	-1.9631	-0.2043	0.0157
MDRG_114	438,859	-0.5115	-0.8105	-0.2125	0.0008
MDRG_116	90,376	-2.6664	-4.5982	-0.7345	0.0068
MDRG_125	84,701	-1.2502	-2.2332	-0.2673	0.0127
MDRG_401	1,531	2.4877	1.3485	3.6269	0.0000
MDRG_402	17,407	1.9506	1.5008	2.4005	0.0000
MDRG_403	107,092	-2.1924	-3.5945	-0.7903	0.0022
MDRG_404	194,986	-0.5151	-0.9244	-0.1059	0.0136
MDRG_405	23,934	2.0638	1.7468	2.3807	0.0000
MDRG_409	577,886	-1.5239	-1.9609	-1.0869	0.0000
MDRG_410	532,478	-1.3912	-1.8007	-0.9817	0.0000
MDRG_415	28,042	0.5888	-0.1008	1.2783	0.0942
MDRG_416	131,506	2.1636	2.0031	2.3241	0.0000
MDRG_504	30,051	4.0963	3.9391	4.2535	0.0000
MDRG_510	121,273	4.1828	4.0901	4.2755	0.0000
MDRG_517	12,909	2.5651	2.1610	2.9693	0.0000
MDRG_520	312,810	-0.6099	-0.9938	-0.2260	0.0018
MDRG_527	112,284	-1.8031	-2.9345	-0.6717	0.0018
MDRG_531	431,669	-0.8796	-1.2384	-0.5208	0.0000
MDRG_533	140,318	-2.0332	-3.1388	-0.9275	0.0003
MDRG_534	94,480	-2.5475	-4.4160	-0.6790	0.0075
MDRG_540	61,652	1.7391	1.4712	2.0071	0.0000
MDRG_542	62,474	1.0421	0.6425	1.4416	0.0000
MDRG_601	28,995	1.6977	1.3468	2.0486	0.0000
MDRG_602	222,646	0.9932	0.7883	1.1981	0.0000
MDRG_613	105,146	-0.6889	-1.3097	-0.0682	0.0296
MDRG_615	413,056	-0.6779	-1.0120	-0.3437	0.0001
MDRG_620	382,412	-1.2219	-1.6575	-0.7864	0.0000
MDRG_701	19,025	2.5986	2.2720	2.9253	0.0000
MDRG_705	100,940	0.2974	-0.1335	0.7282	0.1761
MDRG_710	107,509	-0.9815	-1.7865	-0.1765	0.0169
MDRG_711	73,623	0.2046	-0.3448	0.7540	0.4654
MDRG_803	148,391	0.6330	0.2868	0.9792	0.0003
MDRG_806	86,007	-0.9978	-1.8811	-0.1145	0.0268
MDRG_807	1,015,380	-2.1594	-2.6104	-1.7083	0.0000
MDRG_834	123,374	-1.1259	-1.8649	-0.3868	0.0028
MDRG_910	263,653	-1.8032	-2.5958	-1.0107	0.0000
MDRG_1006	16,866	0.1180	-1.2900	1.5260	0.8695
MDRG_1007	193,031	-0.7022	-1.2092	-0.1952	0.0066
MDRG_1008	310,221	-1.1557	-1.6669	-0.6444	0.0000
MDRG_1103	31,148	2.9087	2.6613	3.1561	0.0000
MDRG_1110	516,018	-0.8801	-1.2079	-0.5522	0.0000
MDRG_1113	433,522	-2.7751	-3.6399	-1.9104	0.0000
MDRG_1118	186,582	-0.3913	-0.8308	0.0482	0.0810

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_1604	162,649	-2.6188	-4.0270	-1.2107	0.0003
MDRG_1801	164,985	1.4625	1.2690	1.6560	0.0000
MDRG_1807	53,271	2.6611	2.4776	2.8445	0.0000
MDRG_1808	1,408,885	0.3504	0.2264	0.4745	0.0000
MDRG_2107	102,944	-0.5964	-1.2947	0.1018	0.0941
MDRG_2303	83,930	-2.6617	-4.6295	-0.6938	0.0080
TRNSFER	1,026,048	0.0838	-0.0221	0.1897	0.1208

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16c. PSI 08 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-9.9985	-10.5228	-9.4743	0.0000
Age_50_54	510,084	0.7705	0.0581	1.4828	0.0340
Age_55_59	757,409	1.0670	0.4340	1.7000	0.0010
Age_60_64	929,797	1.1878	0.5809	1.7948	0.0001
Age_65_69	2,663,640	1.1644	0.6073	1.7214	0.0000
Age_70_74	2,545,816	1.3821	0.8312	1.9330	0.0000
Age_75_79	2,269,167	1.7965	1.2569	2.3360	0.0000
Age_80_84	1,951,860	1.9602	1.4160	2.5045	0.0000
Age_85_89	1,581,625	2.2947	1.7532	2.8362	0.0000
Age_90Plus	1,226,182	2.4635	1.9118	3.0152	0.0000
MALE	7,115,396	0.3595	-0.3382	1.0571	0.3125
Age_50_54*MALE	260,106	-0.4069	-1.4002	0.5864	0.4221
Age_55_59*MALE	388,940	-1.0322	-1.9918	-0.0726	0.0350
Age_60_64*MALE	474,901	-0.5756	-1.4071	0.2559	0.1749
Age_65_69*MALE	1,329,729	-1.0357	-1.7985	-0.2729	0.0078
Age_70_74*MALE	1,235,807	-0.9858	-1.7360	-0.2355	0.0100
Age_75_79*MALE	1,067,050	-0.7246	-1.4491	-0.0001	0.0500
Age_80_84*MALE	867,987	-0.8070	-1.5382	-0.0757	0.0306
Age_85_89*MALE	645,346	-0.8004	-1.5430	-0.0579	0.0346
Age_90Plus*MALE	414,163	-0.8973	-1.6472	-0.1475	0.0190
DM	2,153,452	-0.1918	-0.3603	-0.0234	0.0256
CHF	2,639,203	0.5177	0.3888	0.6467	0.0000
ARTH	697,441	0.1836	-0.0177	0.3848	0.0739
COAG	889,895	0.1549	-0.0421	0.3519	0.1232
DMCX	3,348,070	-0.2011	-0.3431	-0.0591	0.0055
DRUG	318,823	0.1993	-0.1851	0.5837	0.3095
PARA	686,649	0.0921	-0.1568	0.3410	0.4685
HTN_C	9,977,703	-0.1443	-0.2605	-0.0280	0.0150
LIVER	643,704	0.1759	-0.0770	0.4288	0.1728
LYTES	4,796,163	0.3363	0.2183	0.4542	0.0000
NEURO	1,858,374	0.5296	0.3981	0.6611	0.0000
OBESE	2,600,487	-0.4509	-0.6101	-0.2917	0.0000
PSYCH	639,916	-0.0130	-0.3143	0.2884	0.9327
TUMOR	482,527	0.3476	0.1016	0.5935	0.0056

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
ULCER	153,491	-0.2406	-0.7271	0.2459	0.3324
VALVE	977,977	0.2416	0.0792	0.4041	0.0036
ALCOHOL	418,213	0.6098	0.3206	0.8990	0.0000
ANEMDEF	3,540,532	0.2395	0.1160	0.3630	0.0001
BLDLOSS	159,127	0.2732	-0.1298	0.6762	0.1840
DEPRESS	2,277,988	0.1681	0.0391	0.2971	0.0106
HYPOTHY	2,938,632	-0.0482	-0.1677	0.0714	0.4299
CHRNLUNG	4,144,493	0.4246	0.3162	0.5329	0.0000
PERIVASC	1,374,451	-0.0384	-0.2093	0.1324	0.6592
PULMCIRC	401,845	0.5920	0.3781	0.8059	0.0000
RENLFAIL	3,920,745	0.1457	0.0211	0.2703	0.0219
WGHTLOSS	1,171,064	0.7278	0.5788	0.8769	0.0000
MDRG_103	40,137	-1.2430	-2.6244	0.1384	0.0778
MDRG_107	74,110	-1.6138	-2.7480	-0.4797	0.0053
MDRG_111	80,533	-1.9305	-2.9000	-0.9610	0.0001
MDRG_117	69,792	-1.8098	-2.6915	-0.9281	0.0001
MDRG_125	80,092	-2.2414	-3.3736	-1.1092	0.0001
MDRG_401	55,916	-2.7125	-4.6447	-0.7802	0.0059
MDRG_402	38,048	-2.7369	-4.6856	-0.7883	0.0059
MDRG_403	103,654	-3.0085	-4.3871	-1.6299	0.0000
MDRG_404	204,440	-2.7740	-3.4301	-2.1178	0.0000
MDRG_405	23,784	-1.8233	-3.2022	-0.4443	0.0096
MDRG_408	298,364	-3.5068	-4.3858	-2.6278	0.0000
MDRG_409	578,500	-2.3139	-2.7851	-1.8427	0.0000
MDRG_410	537,005	-2.9462	-3.4570	-2.4353	0.0000
MDRG_415	28,900	-2.4601	-4.4057	-0.5145	0.0132
MDRG_416	140,785	-2.7399	-3.6127	-1.8670	0.0000
MDRG_503	91,072	-1.4787	-2.4547	-0.5027	0.0030
MDRG_507	113,483	-1.4665	-2.4383	-0.4948	0.0031
MDRG_509	35,708	-1.0800	-2.0602	-0.0999	0.0308
MDRG_510	25,727	0.0267	-0.6638	0.7171	0.9396
MDRG_511	257,637	-1.6367	-2.2601	-1.0133	0.0000
MDRG_514	138,628	-0.7672	-1.2710	-0.2633	0.0028
MDRG_520	312,009	-2.3424	-2.9410	-1.7439	0.0000
MDRG_522	199,347	-3.8162	-5.7509	-1.8814	0.0001
MDRG_524	963,251	-2.4901	-2.8506	-2.1297	0.0000
MDRG_527	108,195	-3.1225	-4.4967	-1.7483	0.0000
MDRG_529	82,329	-3.4781	-5.4251	-1.5311	0.0005
MDRG_531	429,443	-2.8768	-3.5733	-2.1804	0.0000
MDRG_533	60,463	-3.1460	-5.0891	-1.2028	0.0015
MDRG_534	93,585	-2.0509	-3.1696	-0.9321	0.0003
MDRG_535	124,949	-2.8481	-3.9798	-1.7165	0.0000
MDRG_540	63,043	-1.9255	-3.0500	-0.8010	0.0008
MDRG_542	70,989	-1.4487	-2.4263	-0.4710	0.0037
MDRG_543	34,014	-1.3878	-2.7549	-0.0208	0.0466
MDRG_601	60,092	-1.1748	-1.9698	-0.3798	0.0038
MDRG_602	202,597	-1.7638	-2.3526	-1.1749	0.0000
MDRG_604	38,286	-2.5501	-4.4960	-0.6043	0.0102
MDRG_611	20,459	-0.2745	-1.0772	0.5283	0.5028
MDRG_613	101,491	-3.2598	-4.6377	-1.8820	0.0000
MDRG_615	404,679	-3.1938	-3.8928	-2.4949	0.0000
MDRG_616	26,053	-2.4506	-4.3965	-0.5047	0.0136
MDRG_619	185,020	-4.2436	-6.1800	-2.3072	0.0000

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_620	369,656	-3.8744	-4.9970	-2.7518	0.0000
MDRG_621	140,357	-4.0733	-6.0021	-2.1445	0.0000
MDRG_705	100,871	-1.6784	-2.5484	-0.8085	0.0002
MDRG_708	50,704	-2.8984	-4.8625	-0.9343	0.0038
MDRG_710	107,714	-2.7656	-4.1525	-1.3787	0.0001
MDRG_711	72,158	-2.4751	-3.8524	-1.0979	0.0004
MDRG_801	44,414	-1.9911	-3.9119	-0.0703	0.0422
MDRG_805	36,946	1.1963	0.8220	1.5707	0.0000
MDRG_806	80,488	1.6897	1.4775	1.9020	0.0000
MDRG_807	883,807	0.5290	0.3754	0.6826	0.0000
MDRG_811	14,942	1.4344	1.0382	1.8306	0.0000
MDRG_812	108,058	-3.0681	-4.9887	-1.1475	0.0017
MDRG_826	44,460	-2.5861	-4.5201	-0.6520	0.0088
MDRG_831	29,405	-2.1513	-3.5072	-0.7954	0.0019
MDRG_835	29,058	-2.4331	-4.3756	-0.4906	0.0141
MDRG_838	22,771	-0.4395	-1.2429	0.3640	0.2837
MDRG_901	25,462	-2.1411	-4.0879	-0.1943	0.0311
MDRG_903	22,901	-1.9094	-3.8574	0.0386	0.0547
MDRG_906	16,442	-2.0968	-4.0515	-0.1421	0.0355
MDRG_910	257,222	-3.3771	-4.5028	-2.2513	0.0000
MDRG_911	23,562	-2.4705	-4.4221	-0.5190	0.0131
MDRG_1006	16,182	2.2378	1.8557	2.6200	0.0000
MDRG_1007	187,189	-2.9822	-4.1043	-1.8602	0.0000
MDRG_1008	296,735	-2.3848	-3.0109	-1.7586	0.0000
MDRG_1010	47,677	-1.8508	-2.8339	-0.8677	0.0002
MDRG_1104	22,822	-1.6446	-3.5905	0.3013	0.0976
MDRG_1107	33,311	-0.0040	-0.6012	0.5933	0.9896
MDRG_1110	498,872	-3.1202	-3.6818	-2.5585	0.0000
MDRG_1113	423,902	-3.4219	-4.1125	-2.7314	0.0000
MDRG_1118	179,335	-3.6405	-5.0179	-2.2630	0.0000
MDRG_1603	24,874	-2.2533	-4.2119	-0.2948	0.0241
MDRG_1604	157,721	-4.0841	-6.0220	-2.1461	0.0000
MDRG_1708	496	1.6368	-0.3159	3.5894	0.1004
MDRG_1801	163,460	0.4116	0.1953	0.6280	0.0002
MDRG_1802	29,817	-1.3852	-2.7632	-0.0072	0.0488
MDRG_1803	42,078	-2.4628	-4.4091	-0.5166	0.0131
MDRG_1805	19,784	-2.0438	-3.9947	-0.0930	0.0400
MDRG_1807	59,114	-3.6092	-5.5507	-1.6678	0.0003
MDRG_1808	1,382,082	-3.3880	-3.7795	-2.9964	0.0000
MDRG_2104	32,766	0.2242	-0.4166	0.8650	0.4928
MDRG_2108	53,099	-2.7924	-4.7330	-0.8518	0.0048
MDRG_2303	71,109	-3.4290	-5.3728	-1.4851	0.0005
TRNSFER	845,901	-0.0530	-0.2709	0.1649	0.6337

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16d. PSI 09 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-5.7001	-5.8302	-5.5701	0.0000
Age_50_54	132,043	-0.0540	-0.2483	0.1402	0.5857
Age_55_59	205,579	0.0235	-0.1430	0.1900	0.7817
Age_60_64	261,478	-0.0408	-0.2036	0.1219	0.6227
Age_65_69	1,107,700	-0.1009	-0.2400	0.0383	0.1554
Age_70_74	971,133	-0.0779	-0.2133	0.0574	0.2589
Age_75_79	756,172	-0.0438	-0.1838	0.0962	0.5401
Age_80_84	524,663	-0.0586	-0.2000	0.0828	0.4165
Age_85_89	331,813	-0.2246	-0.3909	-0.0583	0.0081
Age_90Plus	182,697	-0.5824	-0.7961	-0.3686	0.0000
MALE	2,199,218	0.1171	-0.0371	0.2712	0.1366
Age_50_54*MALE	66,460	0.0590	-0.1980	0.3159	0.6529
Age_55_59*MALE	104,955	-0.0279	-0.2554	0.1995	0.8097
Age_60_64*MALE	133,467	-0.0281	-0.2360	0.1798	0.7911
Age_65_69*MALE	548,206	-0.0556	-0.2359	0.1248	0.5460
Age_70_74*MALE	472,445	-0.0382	-0.2126	0.1363	0.6682
Age_75_79*MALE	355,364	0.0444	-0.1358	0.2247	0.6292
Age_80_84*MALE	233,482	-0.1042	-0.2966	0.0882	0.2885
Age_85_89*MALE	132,331	-0.0017	-0.2221	0.2186	0.9876
Age_90Plus*MALE	57,066	0.0205	-0.2903	0.3313	0.8972
DM	656,170	-0.2157	-0.2761	-0.1553	0.0000
CHF	415,582	0.3152	0.2449	0.3855	0.0000
DMCX	803,928	-0.1416	-0.1956	-0.0877	0.0000
DRUG	56,690	0.0712	-0.0947	0.2371	0.4004
METS	118,801	0.0747	-0.0415	0.1909	0.2074
PARA	154,170	0.1499	0.0583	0.2414	0.0013
HTN_C	3,195,033	-0.0699	-0.1124	-0.0275	0.0012
LYMPH	26,948	-0.1449	-0.3951	0.1054	0.2565
LYTES	679,690	-0.0652	-0.1218	-0.0086	0.0239
NEURO	366,578	-0.0646	-0.1426	0.0134	0.1048
OBESE	855,752	-0.0560	-0.1086	-0.0033	0.0372
PSYCH	117,907	-0.1181	-0.2432	0.0070	0.0642
TUMOR	99,941	0.0496	-0.0636	0.1627	0.3905
ULCER	35,576	0.3008	0.1237	0.4779	0.0009
VALVE	235,390	0.2537	0.1643	0.3431	0.0000
ANEMDEF	718,610	-0.3052	-0.3615	-0.2490	0.0000
DEPRESS	627,328	-0.0507	-0.1131	0.0118	0.1116
HYPOTHY	817,500	0.0312	-0.0216	0.0840	0.2467
CHRNLUNG	999,710	-0.0121	-0.0572	0.0329	0.5974
PERIVASC	450,953	0.2095	0.1581	0.2610	0.0000
PULMCIRC	70,955	0.2181	0.0799	0.3563	0.0020
RENLFAIL	834,723	0.1506	0.0983	0.2028	0.0000
WGHTLOSS	240,466	0.3812	0.3068	0.4556	0.0000
MDRG_103	45,590	0.4424	0.2929	0.5920	0.0000
MDRG_105	11,141	-3.8941	-6.6521	-1.1361	0.0057
MDRG_107	80,435	0.9519	0.8562	1.0477	0.0000
MDRG_301	9,265	0.9284	0.6549	1.2018	0.0000
MDRG_401	63,464	0.1504	0.0165	0.2842	0.0277
MDRG_402	44,522	-1.5922	-1.9517	-1.2327	0.0000
MDRG_501	6,975	1.2850	1.0431	1.5269	0.0000
MDRG_502	21,414	-0.7015	-1.0782	-0.3248	0.0003

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_503	61,885	1.4446	1.3453	1.5438	0.0000
MDRG_505	17,262	0.7908	0.5827	0.9989	0.0000
MDRG_507	91,869	0.8518	0.7486	0.9551	0.0000
MDRG_509	34,580	-0.2276	-0.4350	-0.0202	0.0315
MDRG_510	24,316	0.1807	-0.0465	0.4079	0.1190
MDRG_511	252,353	-1.0984	-1.2379	-0.9589	0.0000
MDRG_513	18,404	-0.8857	-1.3006	-0.4707	0.0000
MDRG_514	125,299	0.7732	0.6856	0.8608	0.0000
MDRG_540	56,828	0.1685	0.0043	0.3327	0.0442
MDRG_542	63,984	0.6442	0.5273	0.7610	0.0000
MDRG_543	32,802	-0.3320	-0.5806	-0.0834	0.0088
MDRG_601	56,630	0.1032	-0.0528	0.2591	0.1949
MDRG_602	219,897	-0.0484	-0.1413	0.0445	0.3074
MDRG_604	38,297	-0.3145	-0.5313	-0.0977	0.0045
MDRG_610	34,454	0.0339	-0.1578	0.2255	0.7292
MDRG_701	18,369	0.4187	0.1894	0.6479	0.0003
MDRG_704	12,545	-0.4109	-0.8327	0.0108	0.0562
MDRG_705	97,830	-0.9665	-1.1514	-0.7816	0.0000
MDRG_801	44,081	-0.5155	-0.7447	-0.2862	0.0000
MDRG_802	14,046	-0.9914	-1.4828	-0.5000	0.0001
MDRG_803	146,323	-1.0351	-1.2089	-0.8614	0.0000
MDRG_804	16,297	-2.7747	-3.9055	-1.6438	0.0000
MDRG_805	38,608	-0.2059	-0.4157	0.0039	0.0544
MDRG_806	83,157	-1.5602	-1.8394	-1.2810	0.0000
MDRG_807	992,352	-3.3353	-3.5482	-3.1224	0.0000
MDRG_810	21,840	-1.8783	-2.4990	-1.2577	0.0000
MDRG_811	236,468	-2.4300	-2.6890	-2.1711	0.0000
MDRG_812	118,288	-3.1610	-3.7192	-2.6028	0.0000
MDRG_816	73,581	-2.7292	-3.2609	-2.1976	0.0000
MDRG_820	10,752	-2.4814	-3.6619	-1.3009	0.0000
MDRG_824	13,271	-3.7106	-5.7797	-1.6415	0.0004
MDRG_826	59,004	-0.8242	-1.0654	-0.5830	0.0000
MDRG_903	24,764	1.4842	1.3664	1.6020	0.0000
MDRG_904	2,986	1.8351	1.5655	2.1048	0.0000
MDRG_1002	31,187	-0.6974	-0.9977	-0.3971	0.0000
MDRG_1003	35,940	-0.3958	-0.6437	-0.1479	0.0018
MDRG_1005	7,014	1.8789	1.6953	2.0626	0.0000
MDRG_1101	19,301	1.2712	1.0757	1.4668	0.0000
MDRG_1105	2,447	1.5812	1.2509	1.9115	0.0000
MDRG_1107	33,910	-2.2114	-2.7979	-1.6249	0.0000
MDRG_1109	16,766	0.4357	0.2136	0.6578	0.0001
MDRG_1201	35,643	-1.0904	-1.4311	-0.7498	0.0000
MDRG_1801	145,402	-0.4291	-0.5532	-0.3049	0.0000
MDRG_1802	27,225	0.1684	-0.0405	0.3773	0.1141
MDRG_7701	2,622	2.7287	2.4747	2.9828	0.0000
MDRG_7702	1,335	2.1648	1.7867	2.5430	0.0000
TRNSFER	281,730	0.0749	0.0055	0.1443	0.0343

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16e. PSI 10 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-6.9103	-7.2645	-6.5560	0.0000
Age_50_54	62,454	0.3542	-0.1545	0.8630	0.1724
Age_55_59	101,156	0.3975	-0.0441	0.8391	0.0777
Age_60_64	133,063	0.6368	0.2273	1.0463	0.0023
Age_65_69	725,867	0.6383	0.2756	1.0010	0.0006
Age_70_74	624,659	0.8406	0.4841	1.1971	0.0000
Age_75_79	453,477	1.0206	0.6511	1.3901	0.0000
Age_80_84	261,672	0.8675	0.4922	1.2429	0.0000
Age_85_89	119,509	0.6024	0.1812	1.0235	0.0051
Age_90Plus	35,578	0.0299	-0.5396	0.5995	0.9179
MALE	1,185,883	0.5064	0.0739	0.9389	0.0217
Age_50_54*MALE	26,185	-0.4082	-1.0765	0.2601	0.2312
Age_55_59*MALE	44,529	-0.5303	-1.1058	0.0452	0.0709
Age_60_64*MALE	60,035	-0.3757	-0.8946	0.1433	0.1560
Age_65_69*MALE	337,303	-0.2429	-0.6949	0.2092	0.2923
Age_70_74*MALE	290,624	-0.4330	-0.8788	0.0129	0.0570
Age_75_79*MALE	208,372	-0.5056	-0.9713	-0.0398	0.0334
Age_80_84*MALE	118,980	-0.4760	-0.9484	-0.0035	0.0483
Age_85_89*MALE	52,671	-0.5199	-1.0440	0.0043	0.0519
Age_90Plus*MALE	14,196	-0.5571	-1.3447	0.2305	0.1657
DM	397,850	-0.4746	-0.6215	-0.3277	0.0000
CHF	127,469	1.3133	1.1709	1.4557	0.0000
ARTH	122,091	-0.1824	-0.3741	0.0093	0.0622
COAG	58,876	0.4857	0.3605	0.6109	0.0000
DMCX	280,440	0.2285	0.1379	0.3192	0.0000
METS	51,976	-0.0380	-0.2278	0.1518	0.6945
PARA	41,030	-0.0008	-0.2363	0.2348	0.9949
HTN_C	1,759,849	-0.6938	-0.7666	-0.6209	0.0000
LIVER	57,161	0.3150	0.1455	0.4845	0.0003
LYTES	130,666	0.3594	0.2505	0.4684	0.0000
NEURO	142,699	-0.3609	-0.5337	-0.1882	0.0000
OBESE	518,186	0.3400	0.2546	0.4254	0.0000
PSYCH	50,708	-0.4897	-0.8486	-0.1309	0.0075
TUMOR	43,413	0.0661	-0.1436	0.2757	0.5368
ULCER	12,065	0.3725	0.0394	0.7055	0.0284
VALVE	106,753	0.1989	0.0300	0.3678	0.0210
ALCOHOL	28,141	-0.4563	-0.7983	-0.1143	0.0089
ANEMDEF	226,759	-0.1888	-0.3015	-0.0762	0.0010
DEPRESS	349,545	-0.5115	-0.6457	-0.3774	0.0000
HYPOTHY	451,108	-0.2301	-0.3307	-0.1296	0.0000
CHRNLUNG	516,913	0.0283	-0.0583	0.1149	0.5217
PERIVASC	199,026	0.2398	0.1508	0.3288	0.0000
PULMCIRC	21,231	0.7712	0.5557	0.9867	0.0000
RENLFAIL	296,259	1.0682	0.9742	1.1622	0.0000
WGHTLOSS	54,608	0.9248	0.8033	1.0464	0.0000
MDRG_103	34,825	-1.8497	-2.4919	-1.2076	0.0000
MDRG_106	9,266	-2.2346	-3.4271	-1.0421	0.0002
MDRG_107	66,297	-2.2649	-2.7432	-1.7867	0.0000
MDRG_401	49,590	-0.5351	-0.7931	-0.2771	0.0000
MDRG_402	13,627	-1.5657	-2.2331	-0.8983	0.0000
MDRG_501	1,773	1.7990	1.4529	2.1451	0.0000

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_503	64,645	1.3707	1.2490	1.4923	0.0000
MDRG_505	11,569	0.5835	0.3181	0.8489	0.0000
MDRG_507	56,925	0.6659	0.5067	0.8250	0.0000
MDRG_509	10,692	-1.0021	-1.4180	-0.5863	0.0000
MDRG_511	26,252	-0.5868	-0.9189	-0.2547	0.0005
MDRG_514	52,988	-1.0825	-1.3742	-0.7907	0.0000
MDRG_540	52,784	-0.5117	-0.7402	-0.2831	0.0000
MDRG_541	32,336	0.4521	0.2351	0.6690	0.0000
MDRG_542	23,228	0.6904	0.4860	0.8947	0.0000
MDRG_543	11,367	-1.4336	-2.2096	-0.6576	0.0003
MDRG_601	35,373	-0.0458	-0.2837	0.1921	0.7060
MDRG_610	19,580	-1.0857	-1.5816	-0.5897	0.0000
MDRG_701	16,416	0.6095	0.3828	0.8361	0.0000
MDRG_705	9,753	-1.2483	-1.9167	-0.5800	0.0003
MDRG_801	41,948	-1.4433	-1.9144	-0.9721	0.0000
MDRG_803	135,772	-1.8684	-2.2059	-1.5310	0.0000
MDRG_804	16,178	-2.1501	-3.1669	-1.1332	0.0000
MDRG_805	19,616	-1.1940	-1.7115	-0.6765	0.0000
MDRG_806	67,265	-2.0005	-2.4825	-1.5185	0.0000
MDRG_807	863,572	-3.0915	-3.3452	-2.8378	0.0000
MDRG_808	64,454	-2.3771	-2.9918	-1.7625	0.0000
MDRG_809	5,797	-2.5733	-4.0016	-1.1451	0.0004
MDRG_811	13,808	-2.4888	-3.6464	-1.3313	0.0000
MDRG_812	112,013	-4.1583	-5.2125	-3.1041	0.0000
MDRG_815	30,071	-2.2630	-3.0677	-1.4582	0.0000
MDRG_816	18,107	-2.3562	-3.3576	-1.3548	0.0000
MDRG_819	8,530	-1.4329	-2.3092	-0.5566	0.0014
MDRG_826	29,346	-1.8074	-2.4285	-1.1863	0.0000
MDRG_901	6,434	-2.2940	-3.6888	-0.8991	0.0013
MDRG_1002	5,692	-1.1448	-1.8371	-0.4524	0.0012
MDRG_1003	34,691	-1.1348	-1.6523	-0.6174	0.0000
MDRG_1102	11,593	0.3592	0.0762	0.6422	0.0129
MDRG_1201	31,912	-1.6599	-2.2507	-1.0691	0.0000
MDRG_1304	19,885	-1.8514	-2.7466	-0.9561	0.0001
MDRG_7701	1,215	1.7954	1.4159	2.1749	0.0000
MDRG_7702	726	1.6732	1.1782	2.1682	0.0000
TRNSFER	30,066	0.0079	-0.1888	0.2047	0.9370

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16f. PSI 11 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-5.8684	-5.9900	-5.7468	0.0000
Age_50_54	55,758	0.1962	0.0246	0.3678	0.0250
Age_55_59	86,870	0.2912	0.1386	0.4439	0.0002
Age_60_64	111,034	0.3415	0.1957	0.4873	0.0000

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Age_65_69	617,105	0.1657	0.0399	0.2915	0.0098
Age_70_74	511,862	0.2755	0.1485	0.4025	0.0000
Age_75_79	354,222	0.3937	0.2635	0.5239	0.0000
Age_80_84	190,940	0.4198	0.2799	0.5597	0.0000
Age_85_89	77,992	0.5330	0.3754	0.6906	0.0000
Age_90Plus	19,675	0.3012	0.0678	0.5347	0.0114
MALE	895,569	0.3786	0.2180	0.5391	0.0000
Age_50_54*MALE	22,421	-0.2775	-0.5269	-0.0281	0.0292
Age_55_59*MALE	35,995	-0.3618	-0.5823	-0.1413	0.0013
Age_60_64*MALE	46,778	-0.2916	-0.4987	-0.0846	0.0058
Age_65_69*MALE	269,277	-0.2560	-0.4322	-0.0798	0.0044
Age_70_74*MALE	221,921	-0.2528	-0.4299	-0.0758	0.0051
Age_75_79*MALE	150,130	-0.2269	-0.4077	-0.0461	0.0139
Age_80_84*MALE	79,407	-0.2372	-0.4315	-0.0429	0.0167
Age_85_89*MALE	31,174	-0.1188	-0.3374	0.0997	0.2865
Age_90Plus*MALE	7,118	0.0455	-0.2831	0.3740	0.7862
DM	324,522	-0.0498	-0.1028	0.0032	0.0653
CHF	110,312	0.8659	0.8172	0.9147	0.0000
COAG	35,658	0.6011	0.5229	0.6793	0.0000
DMCX	192,402	0.2686	0.2179	0.3192	0.0000
DRUG	15,762	0.3564	0.2109	0.5020	0.0000
METS	35,272	0.2103	0.1274	0.2931	0.0000
PARA	31,050	0.7019	0.6168	0.7870	0.0000
LIVER	47,014	0.0793	-0.0074	0.1659	0.0729
LYTES	96,041	0.7218	0.6718	0.7718	0.0000
NEURO	106,331	0.2735	0.2058	0.3412	0.0000
OBESE	433,819	0.4147	0.3731	0.4562	0.0000
PSYCH	42,832	0.3132	0.2111	0.4154	0.0000
TUMOR	29,801	0.1814	0.0859	0.2770	0.0002
ULCER	9,324	0.5620	0.4162	0.7078	0.0000
VALVE	96,255	0.1427	0.0755	0.2099	0.0000
ALCOHOL	19,795	0.5331	0.4140	0.6523	0.0000
DEPRESS	295,241	0.0458	-0.0043	0.0959	0.0732
CHRNLUNG	378,580	0.4074	0.3679	0.4469	0.0000
PERIVASC	93,563	0.2248	0.1617	0.2879	0.0000
PULMCIRC	17,953	0.6963	0.6021	0.7904	0.0000
RENLFAIL	201,181	0.3263	0.2776	0.3750	0.0000
WGHTLOSS	35,849	0.9324	0.8728	0.9920	0.0000
MDRG_103	33,336	0.7779	0.6790	0.8769	0.0000
MDRG_107	64,783	-0.3126	-0.4247	-0.2006	0.0000
MDRG_601	29,624	1.2245	1.1425	1.3065	0.0000
MDRG_602	124,216	0.8451	0.7880	0.9022	0.0000
MDRG_604	10,268	0.9452	0.8039	1.0866	0.0000
MDRG_610	19,250	0.6698	0.5480	0.7917	0.0000
MDRG_701	15,837	1.2246	1.1218	1.3274	0.0000
MDRG_704	4,541	0.7627	0.5476	0.9779	0.0000
MDRG_705	9,297	0.4482	0.2844	0.6121	0.0000
MDRG_801	40,831	0.5015	0.3986	0.6043	0.0000
MDRG_802	11,143	1.0480	0.9093	1.1868	0.0000
MDRG_803	133,033	-0.3740	-0.4646	-0.2834	0.0000
MDRG_804	16,009	-1.2407	-1.6278	-0.8536	0.0000
MDRG_805	18,700	-0.4538	-0.6418	-0.2657	0.0000
MDRG_806	65,441	-1.0466	-1.2036	-0.8895	0.0000

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_807	849,090	-1.4688	-1.5384	-1.3992	0.0000
MDRG_812	109,768	-1.0928	-1.2261	-0.9594	0.0000
MDRG_815	29,365	-0.6888	-0.8884	-0.4892	0.0000
MDRG_816	17,234	-0.6468	-0.8719	-0.4218	0.0000
MDRG_826	28,023	-0.6600	-0.8517	-0.4684	0.0000
MDRG_901	5,876	-0.7769	-1.1135	-0.4404	0.0000
MDRG_1002	5,702	-1.0633	-1.4104	-0.7163	0.0000
MDRG_1003	34,675	-0.2880	-0.4587	-0.1173	0.0009
MDRG_1102	11,239	1.0390	0.9114	1.1666	0.0000
MDRG_1103	28,337	0.6297	0.5241	0.7352	0.0000
MDRG_1201	34,024	-0.9860	-1.2325	-0.7394	0.0000
MDRG_1801	6,466	0.8874	0.7705	1.0043	0.0000
MDRG_2104	11,199	0.6049	0.4634	0.7464	0.0000
MDRG_7702	753	1.6956	1.4105	1.9807	0.0000
TRNSFER	15,233	0.3351	0.2281	0.4420	0.0000

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16g. PSI 12 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-5.8200	-5.9390	-5.7010	0.0000
Age_50_54	141,816	0.0594	-0.1181	0.2368	0.5121
Age_55_59	220,771	0.1307	-0.0161	0.2776	0.0811
Age_60_64	281,724	0.1795	0.0311	0.3279	0.0178
Age_65_69	1,170,762	0.2038	0.0878	0.3199	0.0006
Age_70_74	1,032,194	0.3027	0.1814	0.4240	0.0000
Age_75_79	809,941	0.4245	0.3026	0.5463	0.0000
Age_80_84	566,314	0.3582	0.2350	0.4815	0.0000
Age_85_89	359,615	0.3907	0.2607	0.5207	0.0000
Age_90Plus	198,129	0.2954	0.1529	0.4378	0.0000
MALE	2,387,789	0.1636	0.0049	0.3223	0.0434
Age_50_54*MALE	72,184	-0.0059	-0.2531	0.2412	0.9624
Age_55_59*MALE	113,661	-0.0787	-0.2945	0.1371	0.4747
Age_60_64*MALE	145,212	-0.0670	-0.2703	0.1363	0.5182
Age_65_69*MALE	586,421	-0.0512	-0.2202	0.1178	0.5528
Age_70_74*MALE	508,992	-0.0852	-0.2585	0.0881	0.3354
Age_75_79*MALE	387,902	-0.1876	-0.3621	-0.0131	0.0351
Age_80_84*MALE	258,219	-0.1432	-0.3208	0.0343	0.1139
Age_85_89*MALE	147,492	-0.1476	-0.3370	0.0418	0.1268
Age_90Plus*MALE	64,002	-0.2263	-0.4501	-0.0025	0.0475
DM	692,470	-0.2222	-0.2699	-0.1745	0.0000
CHF	467,077	0.3253	0.2778	0.3728	0.0000
ARTH	222,522	0.0349	-0.0388	0.1086	0.3533
COAG	205,037	0.3081	0.2515	0.3647	0.0000
DMCX	883,241	-0.1557	-0.2040	-0.1074	0.0000
METS	128,138	0.6548	0.5941	0.7156	0.0000

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
PARA	160,863	0.3498	0.2796	0.4199	0.0000
HTN_C	3,403,973	-0.1633	-0.1955	-0.1311	0.0000
LYMPH	33,438	0.2125	0.0713	0.3537	0.0032
LYTES	769,718	0.3678	0.3290	0.4067	0.0000
NEURO	394,040	0.0236	-0.0301	0.0773	0.3884
OBESE	912,025	0.3350	0.2950	0.3749	0.0000
TUMOR	111,352	0.3100	0.2339	0.3861	0.0000
ULCER	41,825	0.4716	0.3673	0.5758	0.0000
VALVE	261,065	-0.0145	-0.0787	0.0498	0.6587
ALCOHOL	92,776	-0.0811	-0.1860	0.0237	0.1294
ANEMDEF	805,335	-0.0780	-0.1227	-0.0334	0.0006
BLDLOSS	40,410	0.0941	-0.0358	0.2240	0.1556
DEPRESS	666,987	-0.1050	-0.1511	-0.0590	0.0000
HYPOTHY	874,959	-0.1167	-0.1589	-0.0745	0.0000
CHRNLUNG	1,075,812	0.0224	-0.0151	0.0600	0.2412
PERIVASC	497,263	-0.0069	-0.0573	0.0436	0.7900
PULMCIRC	71,035	0.6720	0.5858	0.7583	0.0000
RENLFAIL	943,353	-0.0213	-0.0663	0.0237	0.3534
WGHTLOSS	275,295	0.6483	0.6007	0.6959	0.0000
MDRG_102	16,445	0.0192	-0.1942	0.2326	0.8600
MDRG_103	48,881	0.0035	-0.1499	0.1568	0.9647
MDRG_104	12,504	0.1679	-0.0775	0.4133	0.1800
MDRG_106	12,577	-1.8884	-2.6855	-1.0913	0.0000
MDRG_107	81,830	-2.4594	-2.8424	-2.0765	0.0000
MDRG_301	9,413	-0.9661	-1.4285	-0.5036	0.0000
MDRG_401	66,092	-0.1365	-0.2637	-0.0093	0.0354
MDRG_402	46,520	-0.3490	-0.4894	-0.2086	0.0000
MDRG_502	22,640	-1.0929	-1.4316	-0.7542	0.0000
MDRG_504	32,135	-1.0362	-1.2983	-0.7741	0.0000
MDRG_506	3,485	0.7758	0.4234	1.1282	0.0000
MDRG_509	35,766	-0.4673	-0.6530	-0.2816	0.0000
MDRG_510	26,288	-0.7673	-1.0310	-0.5036	0.0000
MDRG_511	260,642	-1.6837	-1.8223	-1.5451	0.0000
MDRG_513	19,200	-1.5223	-2.0123	-1.0324	0.0000
MDRG_514	137,619	-0.6162	-0.7361	-0.4962	0.0000
MDRG_515	9,344	-2.1850	-3.1166	-1.2534	0.0000
MDRG_519	20,218	-0.8462	-1.1188	-0.5736	0.0000
MDRG_540	64,356	-1.2052	-1.4358	-0.9746	0.0000
MDRG_541	41,550	-0.7340	-0.9443	-0.5236	0.0000
MDRG_543	34,665	-0.9660	-1.2312	-0.7007	0.0000
MDRG_602	229,541	0.2417	0.1814	0.3019	0.0000
MDRG_604	39,868	0.1932	0.0611	0.3253	0.0041
MDRG_605	13,097	-0.7545	-1.1432	-0.3658	0.0001
MDRG_606	9,915	-1.1965	-1.7461	-0.6469	0.0000
MDRG_608	8,803	-1.3249	-1.9122	-0.7377	0.0000
MDRG_609	14,951	-0.5647	-0.8942	-0.2353	0.0008
MDRG_610	35,650	-0.1595	-0.3399	0.0209	0.0831
MDRG_611	23,657	0.0562	-0.1033	0.2157	0.4896
MDRG_701	20,576	0.3937	0.2361	0.5513	0.0000
MDRG_705	102,932	-1.1467	-1.3056	-0.9878	0.0000
MDRG_801	45,385	0.3883	0.2445	0.5322	0.0000
MDRG_802	14,972	1.1035	0.9469	1.2602	0.0000
MDRG_803	149,056	-0.1575	-0.2636	-0.0513	0.0036

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_804	16,662	0.4403	0.1880	0.6926	0.0006
MDRG_805	40,513	-0.0870	-0.2407	0.0667	0.2674
MDRG_807	1,017,197	-0.4856	-0.5462	-0.4249	0.0000
MDRG_808	75,356	-0.7074	-0.8816	-0.5332	0.0000
MDRG_811	256,236	0.0746	0.0098	0.1393	0.0239
MDRG_812	119,993	-1.3066	-1.4873	-1.1259	0.0000
MDRG_815	39,755	-0.2662	-0.4498	-0.0826	0.0045
MDRG_816	76,256	-0.4762	-0.6167	-0.3358	0.0000
MDRG_819	20,417	-0.6868	-0.9729	-0.4007	0.0000
MDRG_820	11,134	-1.2581	-1.7938	-0.7224	0.0000
MDRG_824	13,504	-1.0564	-1.5011	-0.6118	0.0000
MDRG_826	60,949	-0.3511	-0.5066	-0.1956	0.0000
MDRG_901	27,317	-1.2260	-1.5312	-0.9208	0.0000
MDRG_903	27,002	-1.2558	-1.5827	-0.9289	0.0000
MDRG_1002	32,160	-1.4083	-1.7491	-1.0676	0.0000
MDRG_1003	36,235	-1.7736	-2.2451	-1.3022	0.0000
MDRG_1004	14,172	-1.1430	-1.5871	-0.6989	0.0000
MDRG_1006	16,888	-0.6075	-0.9068	-0.3083	0.0001
MDRG_1101	21,899	-1.0514	-1.4446	-0.6581	0.0000
MDRG_1102	13,974	0.7390	0.5677	0.9103	0.0000
MDRG_1103	31,780	-0.0122	-0.1988	0.1743	0.8977
MDRG_1104	23,670	-0.5745	-0.8232	-0.3259	0.0000
MDRG_1107	35,852	-0.7428	-0.9566	-0.5290	0.0000
MDRG_1201	35,998	-1.1008	-1.4013	-0.8003	0.0000
MDRG_1302	7,258	0.8639	0.6260	1.1019	0.0000
MDRG_1303	9,576	0.4000	0.1406	0.6594	0.0025
MDRG_1708	9,543	0.1534	-0.0786	0.3853	0.1950
MDRG_1709	5,539	0.6799	0.3961	0.9636	0.0000
MDRG_1801	171,552	0.1063	0.0433	0.1692	0.0009
MDRG_2104	35,990	-0.0201	-0.1711	0.1310	0.7946
MDRG_2407	8,795	0.9920	0.8021	1.1819	0.0000
MDRG_2408	5,138	1.5590	1.3682	1.7498	0.0000
MDRG_7701	4,361	0.8303	0.5532	1.1075	0.0000
TRNSFER	310,178	0.4990	0.4387	0.5592	0.0000

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16h. PSI 13 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-4.9162	-5.0463	-4.7862	0.0000
Age_50_54	60,461	0.1570	-0.0357	0.3496	0.1102
Age_55_59	97,871	0.1133	-0.0617	0.2884	0.2044
Age_60_64	129,109	0.1475	-0.0175	0.3124	0.0797
Age_65_69	707,614	-0.0523	-0.1895	0.0848	0.4547
Age_70_74	610,025	0.0524	-0.0856	0.1904	0.4566
Age_75_79	443,551	0.1481	0.0074	0.2887	0.0390
Age_80_84	255,658	0.1230	-0.0271	0.2731	0.1081

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Age_85_89	116,054	0.1483	-0.0193	0.3159	0.0830
Age_90Plus	33,856	-0.2041	-0.4509	0.0427	0.1051
MALE	1,161,911	0.1706	-0.0040	0.3453	0.0555
Age_50_54*MALE	25,280	-0.1552	-0.4293	0.1190	0.2673
Age_55_59*MALE	42,960	-0.0119	-0.2545	0.2306	0.9232
Age_60_64*MALE	58,206	-0.0431	-0.2716	0.1853	0.7113
Age_65_69*MALE	330,260	0.0806	-0.1103	0.2716	0.4078
Age_70_74*MALE	285,286	0.1069	-0.0842	0.2979	0.2730
Age_75_79*MALE	204,889	0.1161	-0.0778	0.3099	0.2407
Age_80_84*MALE	116,948	0.1689	-0.0360	0.3737	0.1061
Age_85_89*MALE	51,633	0.1169	-0.1124	0.3461	0.3178
Age_90Plus*MALE	13,754	0.2165	-0.1218	0.5548	0.2097
DM	387,946	-0.3613	-0.4247	-0.2978	0.0000
CHF	119,967	1.0089	0.9548	1.0631	0.0000
COAG	57,748	0.4594	0.3849	0.5338	0.0000
DMCX	271,439	0.0516	-0.0009	0.1041	0.0542
DRUG	17,447	0.4957	0.3394	0.6520	0.0000
METS	50,828	0.3544	0.2834	0.4254	0.0000
PARA	34,907	0.6196	0.5235	0.7157	0.0000
HTN_C	1,707,054	-0.2862	-0.3252	-0.2472	0.0000
LIVER	55,093	0.1940	0.1065	0.2816	0.0000
LYTES	124,927	0.5702	0.5180	0.6223	0.0000
NEURO	136,400	0.1549	0.0833	0.2265	0.0000
OBESE	503,981	0.1421	0.0945	0.1897	0.0000
ULCER	11,507	0.4589	0.3027	0.6151	0.0000
ALCOHOL	26,674	0.2244	0.0944	0.3543	0.0007
ANEMDEF	224,542	-0.1145	-0.1663	-0.0627	0.0000
DEPRESS	337,028	-0.1121	-0.1710	-0.0531	0.0002
HYPOTHY	438,186	-0.0946	-0.1467	-0.0426	0.0004
CHRNLUNG	500,912	-0.0923	-0.1371	-0.0474	0.0001
PERIVASC	191,456	0.0847	0.0254	0.1440	0.0051
PULMCIRC	20,274	0.5275	0.4180	0.6370	0.0000
RENLFAIL	298,010	0.4421	0.3938	0.4903	0.0000
WGHTLOSS	47,488	1.2715	1.2185	1.3244	0.0000
MDC_1	132,947	-0.9813	-1.1104	-0.8522	0.0000
MDC_4	66,476	0.8676	0.7251	1.0101	0.0000
MDC_7	29,144	0.6582	0.5680	0.7485	0.0000
MDC_9	21,160	-1.3550	-1.6159	-1.0940	0.0000
MDC_10	60,316	-1.3922	-1.5562	-1.2282	0.0000
MDC_12	46,050	-0.4581	-0.6737	-0.2425	0.0000
MDC_17	12,342	0.2268	0.0754	0.3783	0.0033
MDRG_107	66,480	-1.6041	-1.9162	-1.2919	0.0000
MDRG_401	48,645	-1.0610	-1.2301	-0.8918	0.0000
MDRG_402	12,925	-1.9952	-2.2923	-1.6980	0.0000
MDRG_503	65,051	-0.0230	-0.1090	0.0631	0.6009
MDRG_504	6,463	-1.1277	-1.5088	-0.7467	0.0000
MDRG_507	57,782	-0.4714	-0.5823	-0.3606	0.0000
MDRG_511	27,074	-1.3106	-1.5372	-1.0839	0.0000
MDRG_514	53,884	-0.9727	-1.1118	-0.8336	0.0000
MDRG_519	3,054	-1.1835	-1.5747	-0.7922	0.0000
MDRG_540	53,447	-1.4720	-1.6452	-1.2989	0.0000
MDRG_541	32,414	-0.5878	-0.7429	-0.4327	0.0000
MDRG_601	35,064	0.6161	0.5332	0.6989	0.0000

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_602	110,274	0.3799	0.3206	0.4392	0.0000
MDRG_610	19,666	-0.6432	-0.8410	-0.4454	0.0000
MDRG_705	5,157	-0.9027	-1.1902	-0.6152	0.0000
MDRG_801	41,852	-0.8372	-1.0009	-0.6736	0.0000
MDRG_803	135,703	-1.2590	-1.3758	-1.1422	0.0000
MDRG_804	16,136	-1.4719	-1.8386	-1.1051	0.0000
MDRG_805	8,530	-1.3070	-1.7019	-0.9120	0.0000
MDRG_806	60,664	-1.9341	-2.1596	-1.7086	0.0000
MDRG_807	863,168	-2.4880	-2.5786	-2.3975	0.0000
MDRG_808	64,514	-1.7328	-1.9378	-1.5277	0.0000
MDRG_811	13,232	-1.3518	-1.6517	-1.0519	0.0000
MDRG_812	111,390	-2.6109	-2.8519	-2.3698	0.0000
MDRG_815	30,042	-1.5968	-1.8643	-1.3294	0.0000
MDRG_816	16,565	-1.7021	-2.0568	-1.3475	0.0000
MDRG_826	28,829	-1.4926	-1.7429	-1.2423	0.0000
MDRG_1102	11,611	0.8598	0.7447	0.9749	0.0000
MDRG_1104	14,387	-0.0469	-0.2064	0.1127	0.5649
MDRG_1201	34,271	-1.1239	-1.4518	-0.7959	0.0000
MDRG_1302	6,320	0.3694	0.1434	0.5954	0.0014
MDRG_1304	19,759	-0.9406	-1.2034	-0.6779	0.0000
MDRG_7701	1,259	1.3342	1.1031	1.5653	0.0000
TRNSFER	27,069	0.2566	0.1458	0.3673	0.0000

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16i. PSI 14A Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-7.3300	-7.9249	-6.7351	0.0000
Age_50_54	14,659	0.4037	-0.4287	1.2361	0.3418
Age_55_59	22,125	0.0063	-0.8681	0.8807	0.9887
Age_60_64	28,083	0.5928	-0.1196	1.3052	0.1029
Age_65_69	130,592	0.1997	-0.4350	0.8345	0.5374
Age_70_74	115,302	0.2199	-0.4069	0.8467	0.4918
Age_75_79	88,568	0.3840	-0.2681	1.0361	0.2484
Age_80_84	57,774	0.1797	-0.4883	0.8478	0.5980
Age_85_89	33,428	-0.1367	-0.8946	0.6213	0.7238
Age_90Plus	15,215	-0.2855	-1.1895	0.6184	0.5358
MALE	239,954	0.8414	0.1245	1.5582	0.0214
Age_50_54*MALE	6,217	0.3100	-0.6979	1.3178	0.5466
Age_55_59*MALE	10,004	0.3902	-0.6428	1.4233	0.4591
Age_60_64*MALE	13,148	0.0513	-0.8220	0.9246	0.9083
Age_65_69*MALE	60,913	0.0499	-0.7177	0.8174	0.8986
Age_70_74*MALE	54,061	0.1403	-0.6160	0.8965	0.7162
Age_75_79*MALE	41,217	-0.2598	-1.0597	0.5401	0.5244
Age_80_84*MALE	26,007	-0.0304	-0.8621	0.8013	0.9429
Age_85_89*MALE	13,773	0.2497	-0.6709	1.1704	0.5950

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Age_90Plus*MALE	5,420	0.5057	-0.5966	1.6080	0.3685
DM	77,502	-0.6523	-0.8949	-0.4098	0.0000
CHF	48,342	0.2508	0.0435	0.4580	0.0177
COAG	18,888	-0.1632	-0.4772	0.1507	0.3083
DMCX	59,267	-0.5148	-0.7583	-0.2713	0.0000
DRUG	6,696	0.4753	0.0866	0.8639	0.0166
METS	31,725	-0.1186	-0.3625	0.1252	0.3404
PARA	15,583	0.1785	-0.1456	0.5026	0.2805
LIVER	19,335	0.0636	-0.2383	0.3656	0.6796
LYTES	97,230	0.2290	0.0739	0.3841	0.0038
NEURO	39,052	0.1565	-0.0642	0.3772	0.1645
OBESE	98,497	0.1969	0.0238	0.3699	0.0258
PSYCH	15,219	0.3442	0.0423	0.6462	0.0255
TUMOR	18,426	0.0736	-0.2123	0.3596	0.6137
ULCER	7,867	0.1224	-0.2944	0.5392	0.5649
VALVE	28,367	-0.1556	-0.4653	0.1542	0.3250
ALCOHOL	9,904	0.5161	0.2263	0.8059	0.0005
ANEMDEF	69,969	0.0673	-0.1125	0.2471	0.4631
BLDLOSS	7,065	-0.2771	-0.8259	0.2718	0.3225
HYPOTHY	88,773	-0.0377	-0.2283	0.1528	0.6981
CHRNLUNG	113,035	0.5593	0.4120	0.7066	0.0000
PULMCIRC	8,836	0.2053	-0.2094	0.6201	0.3319
RENLFAIL	61,473	0.0894	-0.1097	0.2886	0.3789
WGHTLOSS	35,385	0.5306	0.3483	0.7129	0.0000
MDC_1	3,681	-1.3724	-2.7958	0.0511	0.0588
MDRG_519	823	0.9363	0.0423	1.8302	0.0401
MDRG_541	4,608	0.6130	0.1280	1.0980	0.0132
MDRG_542	11,411	-0.2435	-0.7153	0.2283	0.3117
MDRG_601	22,310	0.0417	-0.2891	0.3725	0.8049
MDRG_602	146,372	0.5108	0.3365	0.6852	0.0000
MDRG_604	24,762	-0.3347	-0.7188	0.0494	0.0877
MDRG_609	10,051	-1.9445	-3.0750	-0.8141	0.0007
MDRG_610	23,528	-2.2589	-3.2412	-1.2765	0.0000
MDRG_611	5,211	-0.8844	-1.8830	0.1142	0.0826
MDRG_704	11,872	-1.0320	-1.7388	-0.3253	0.0042
MDRG_705	2,688	0.7020	0.0727	1.3312	0.0288
MDRG_801	26,031	-2.6055	-3.7334	-1.4775	0.0000
MDRG_803	41,254	-2.7369	-3.7197	-1.7540	0.0000
MDRG_819	559	1.4520	0.4676	2.4365	0.0038
MDRG_826	17,538	-3.3333	-5.2487	-1.4179	0.0006
MDRG_1102	9,018	1.0728	0.7655	1.3800	0.0000
MDRG_1103	11,486	0.1062	-0.3377	0.5502	0.6390
MDRG_1201	5,613	-1.3758	-2.5149	-0.2367	0.0179
MDRG_1304	12,661	-0.8589	-1.7622	0.0444	0.0624
MDRG_1707	1,060	0.9775	0.0781	1.8768	0.0332
MDRG_1801	26,793	0.4013	0.1310	0.6717	0.0036
MDRG_2408	1,376	0.5975	-0.3477	1.5428	0.2154
TRNSFER	23,657	0.3697	0.1209	0.6184	0.0036

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16j. PSI 14B Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-12.9040	-14.7276	-11.0804	0.0000
Age_55_59	27,974	0.4239	-1.7597	2.6074	0.7036
Age_65_69	134,628	0.0361	-1.3780	1.4503	0.9600
Age_70_74	110,689	-0.1561	-1.6354	1.3233	0.8362
Age_75_79	85,717	-1.1120	-3.2467	1.0227	0.3072
Age_80_84	62,563	-0.1379	-1.8015	1.5258	0.8710
Age_85_89	40,061	0.2156	-1.5982	2.0295	0.8157
Age_90Plus	19,697	0.2995	-2.1014	2.7004	0.8068
MALE	278,644	-0.5786	-2.2348	1.0776	0.4935
Age_55_59*MALE	13,752	1.2838	-1.7101	4.2776	0.4007
Age_65_69*MALE	69,054	0.7704	-1.3304	2.8711	0.4723
Age_70_74*MALE	57,031	0.4171	-1.8590	2.6932	0.7195
Age_75_79*MALE	43,118	1.6990	-1.1029	4.5008	0.2347
Age_85_89*MALE	18,633	0.8598	-1.7318	3.4514	0.5156
Age_90Plus*MALE	8,134	1.6669	-1.3017	4.6355	0.2711
DM	93,117	-0.2218	-1.2097	0.7661	0.6599
CHF	71,794	0.3645	-0.7769	1.5058	0.5314
COAG	52,255	-0.5970	-2.8291	1.6352	0.6002
DMCX	86,926	-0.2323	-1.3695	0.9049	0.6889
METS	35,250	0.2550	-0.9103	1.4203	0.6680
PARA	21,165	1.3258	-0.0562	2.7078	0.0601
HTN_C	373,037	0.4593	-0.3547	1.2733	0.2688
LIVER	65,246	-0.1473	-1.5971	1.3024	0.8421
LYTES	165,693	-0.1871	-1.1168	0.7426	0.6932
NEURO	47,023	-1.1996	-3.2792	0.8800	0.2582
OBESE	94,264	-0.1801	-1.1292	0.7691	0.7100
PSYCH	17,770	0.1368	-1.9578	2.2315	0.8981
VALVE	34,739	0.4040	-0.7279	1.5359	0.4842
ALCOHOL	25,946	0.2624	-1.6639	2.1886	0.7895
ANEMDEF	118,937	0.8715	0.1289	1.6141	0.0214
BLDLOSS	8,218	0.3619	-1.7376	2.4615	0.7355
DEPRESS	73,301	0.1647	-0.8117	1.1411	0.7409
HYPOTHY	96,809	0.0568	-0.8736	0.9871	0.9048
CHRNLUNG	117,751	0.5853	-0.2140	1.3846	0.1512
PERIVASC	33,048	-0.1935	-1.6668	1.2798	0.7969
PULMCIRC	14,250	0.0191	-1.9727	2.0109	0.9850
RENLFAIL	94,180	0.1109	-0.9085	1.1302	0.8312
WGHTLOSS	43,981	-0.1061	-1.4893	1.2771	0.8805
MDRG_601	21,935	3.3134	1.3332	5.2935	0.0010
MDRG_602	55,583	3.9737	2.3542	5.5933	0.0000
MDRG_609	2,102	4.3858	1.8458	6.9259	0.0007
MDRG_701	4,481	3.8224	1.1653	6.4796	0.0048
MDRG_705	88,776	2.1566	0.3711	3.9422	0.0179
MDRG_1003	20,875	2.2525	-0.1988	4.7038	0.0717
MDRG_1102	2,953	4.9506	2.9859	6.9152	0.0000
MDRG_1103	14,850	2.7811	0.2959	5.2663	0.0283
MDRG_1201	12,521	4.0888	2.2275	5.9502	0.0000
MDRG_1801	25,787	2.8546	1.0116	4.6976	0.0024
TRNSFER	41,555	0.8208	-0.4622	2.1039	0.2099

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

Table 16k. PSI 15 Risk-Adjustment Model: All Risk Factors with Prevalences, v10.0 Parameter Estimates with 95% Confidence Intervals, and Associated p Values

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
Intercept	- -	-7.0785	-7.2865	-6.8704	0.0000
Age_50_54	101,471	0.2598	-0.0389	0.5584	0.0882
Age_55_59	150,379	0.1962	-0.0824	0.4748	0.1676
Age_60_64	186,430	0.2545	-0.0056	0.5147	0.0551
Age_65_69	620,346	0.1251	-0.0843	0.3346	0.2415
Age_70_74	561,394	0.2123	0.0027	0.4219	0.0472
Age_75_79	479,393	0.2495	0.0355	0.4636	0.0223
Age_80_84	381,800	0.2328	0.0152	0.4503	0.0360
Age_85_89	277,173	0.1263	-0.1116	0.3641	0.2981
Age_90Plus	158,154	-0.1133	-0.4071	0.1805	0.4497
MALE	1,543,470	-0.1873	-0.4867	0.1121	0.2201
Age_50_54*MALE	50,712	0.1800	-0.2746	0.6345	0.4378
Age_55_59*MALE	77,861	-0.1532	-0.5805	0.2742	0.4823
Age_60_64*MALE	97,854	-0.0162	-0.4267	0.3942	0.9382
Age_65_69*MALE	323,590	0.0340	-0.2967	0.3648	0.8403
Age_70_74*MALE	290,310	-0.1465	-0.4782	0.1851	0.3865
Age_75_79*MALE	242,436	-0.0153	-0.3480	0.3174	0.9281
Age_80_84*MALE	186,845	0.0950	-0.2408	0.4308	0.5792
Age_85_89*MALE	127,549	0.0932	-0.2759	0.4623	0.6207
Age_90Plus*MALE	64,428	0.1014	-0.3844	0.5872	0.6825
DM	435,277	-0.4165	-0.5271	-0.3059	0.0000
CHF	512,891	0.2427	0.1394	0.3460	0.0000
ARTH	126,321	0.0978	-0.0613	0.2568	0.2285
COAG	259,168	0.1822	0.0537	0.3107	0.0055
DMCX	585,231	-0.2447	-0.3516	-0.1378	0.0000
METS	148,445	-0.0318	-0.1628	0.0992	0.6344
HTN_C	2,108,055	-0.1522	-0.2209	-0.0836	0.0000
LIVER	236,700	-0.1188	-0.2642	0.0266	0.1094
LYMPH	30,902	-0.3778	-0.7629	0.0072	0.0545
LYTES	990,948	0.1025	0.0223	0.1828	0.0123
NEURO	326,501	-0.0956	-0.2135	0.0223	0.1121
OBESE	479,188	0.1364	0.0477	0.2252	0.0026
PSYCH	104,002	0.1353	-0.0414	0.3120	0.1335
TUMOR	128,818	0.0030	-0.1518	0.1578	0.9698
VALVE	222,771	-0.0825	-0.2240	0.0589	0.2529
ALCOHOL	109,274	-0.0257	-0.2372	0.1857	0.8114
ANEMDEF	752,589	-0.2696	-0.3584	-0.1809	0.0000
DEPRESS	418,865	-0.1099	-0.2099	-0.0100	0.0311
HYPOTHY	539,355	-0.1045	-0.1961	-0.0129	0.0253
PERIVASC	289,675	0.0979	-0.0072	0.2030	0.0679
PULMCIRC	92,673	0.0034	-0.1969	0.2037	0.9736
RENLFAIL	720,449	0.1288	0.0326	0.2249	0.0087
WGHTLOSS	408,572	0.6653	0.5760	0.7546	0.0000

Risk factor	Discharges (N)	CMS v10.0 Coefficients	Lower Confidence Interval	Upper Confidence Interval	p-value
MDRG_114	24,581	-1.2382	-2.0050	-0.4713	0.0016
MDRG_401	5,903	1.0461	0.5432	1.5490	0.0000
MDRG_404	20,462	-1.2749	-2.0722	-0.4776	0.0017
MDRG_410	13,335	-2.5312	-4.5205	-0.5419	0.0126
MDRG_416	18,194	-2.2935	-3.7363	-0.8507	0.0018
MDRG_503	7,659	1.6362	1.2983	1.9741	0.0000
MDRG_507	4,896	1.4199	0.9506	1.8893	0.0000
MDRG_514	33,813	0.1008	-0.2560	0.4577	0.5797
MDRG_524	38,082	-2.8307	-4.2462	-1.4153	0.0001
MDRG_540	3,045	2.0402	1.5977	2.4827	0.0000
MDRG_541	41,385	0.3840	0.0897	0.6782	0.0105
MDRG_542	30,023	0.8745	0.6098	1.1392	0.0000
MDRG_601	65,150	1.6232	1.4899	1.7566	0.0000
MDRG_602	231,316	1.7317	1.6360	1.8273	0.0000
MDRG_605	13,157	-2.4084	-4.3978	-0.4190	0.0177
MDRG_607	10,897	0.4344	-0.0831	0.9520	0.1000
MDRG_610	35,776	-1.4627	-2.2129	-0.7125	0.0001
MDRG_613	18,640	-1.5312	-2.5376	-0.5248	0.0029
MDRG_614	30,969	-1.5540	-2.3313	-0.7767	0.0001
MDRG_615	311,273	-2.2181	-2.5852	-1.8509	0.0000
MDRG_616	23,842	-1.3319	-2.1383	-0.5256	0.0012
MDRG_619	41,642	-2.6016	-3.8071	-1.3962	0.0000
MDRG_620	93,093	-1.8297	-2.3596	-1.2997	0.0000
MDRG_621	70,401	-1.8789	-2.5111	-1.2467	0.0000
MDRG_701	20,891	1.2081	0.9443	1.4720	0.0000
MDRG_702	6,078	2.4417	2.1737	2.7098	0.0000
MDRG_704	13,320	1.8637	1.6187	2.1088	0.0000
MDRG_705	103,193	0.5204	0.3433	0.6974	0.0000
MDRG_708	34,566	-2.8686	-4.3127	-1.4246	0.0001
MDRG_710	19,913	-2.9964	-5.1517	-0.8412	0.0064
MDRG_711	26,654	-2.2493	-3.4562	-1.0424	0.0003
MDRG_801	30,343	0.4185	0.0912	0.7458	0.0122
MDRG_803	55,551	-0.6756	-1.0999	-0.2513	0.0018
MDRG_815	17,712	-2.7427	-4.9479	-0.5375	0.0148
MDRG_1102	14,466	2.0538	1.8100	2.2976	0.0000
MDRG_1103	32,060	0.7625	0.4822	1.0429	0.0000
MDRG_1104	26,574	1.0276	0.7680	1.2871	0.0000
MDRG_1110	50,272	-3.9037	-5.8841	-1.9233	0.0001
MDRG_1113	30,304	-1.6999	-2.5830	-0.8169	0.0002
MDRG_1118	43,452	-2.6562	-3.8992	-1.4132	0.0000
MDRG_1303	9,671	1.0518	0.5937	1.5099	0.0000
MDRG_1304	23,044	0.9532	0.6453	1.2612	0.0000
MDRG_1604	46,069	-2.3724	-3.3485	-1.3963	0.0000
MDRG_1709	3,673	1.6246	1.1338	2.1154	0.0000
MDRG_1801	101,219	1.0191	0.8799	1.1583	0.0000
MDRG_1802	8,316	1.1147	0.7157	1.5136	0.0000
MDRG_1807	15,826	-0.9865	-1.7563	-0.2167	0.0120
MDRG_1808	171,265	-2.0320	-2.4483	-1.6157	0.0000
MDRG_2104	15,677	1.2842	0.9942	1.5741	0.0000
TRNSFER	224,269	0.2818	0.1678	0.3957	0.0000

Source: CVP/Mathematica Risk-Adjustment Report, October 2019. The final model is analysis of Medicare FFS discharges from IPPS hospitals July 1, 2016, through June 30, 2018, processed through the CMS v10.0 PSI software.

Abbreviations: MDC – Major Diagnostic Category; MDRG – Modified Medicare Severity-Diagnosis Related Group

- - cell intentionally left blank

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not applicable

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*) **Also discuss any "ordering" of risk factor inclusion;** for example, are social risk factors added after all clinical factors?

The conceptual approach to PSI risk-adjustment includes age and gender as key demographic characteristics, the reason(s) why the patient was admitted to the hospital, the type of operating room procedure(s) that the patient received (for postprocedural events), comorbid conditions that are associated with clinically significant disabilities or increase the risk of adverse events, and transfer-in as an indicator of recent health service use at a different facility. Because the PSIs focus on adverse events occurring within acute care hospitals, often after a major operating room procedure, social risk factors are not included in the conceptual approach.

Through testing and refinement over the past decade, these risk factor concepts are now operationalized using open-source tools, as follows:

- Age categories, generally 5 years in width, using the youngest category as the omitted referent. (Five year categories have been shown to work better, in general, than either wider categories or linear/quadratic specifications.)
- Sex categories, using female as the omitted referent.
- Two-way age-sex interactions, which allow for different age-outcome relationships among men versus women. (Older men often show lower risk than older women, presumably due to "healthy survivor" or "surgical selection" effects.)
- AHRQ (Elixhauser) comorbidities, which are coded using publicly available HCUP software, annually updated, and extensively validated.[13] (Comorbidities are identified using only ICD-10-CM diagnoses reported as present on admission, including "clinically undetermined" diagnoses and codes classified as POA-exempt.)
- Major Diagnostic Categories (MDCs) based on the body system of the principal diagnosis[14]
- Modified Diagnosis Related Groups (MDRGs) based on aggregation of adjacent Medicare Severity (MS) DRGs with or without comorbidities and complications, which capture both the reason for admission and major operating procedures, without adjusting for hospital-associated complications.[15] (Certain MS-DRGs are omitted from feature selection because they capture complications of hospital care – 003, 004, 011, 012, or 013 for ECMO or tracheostomy – or because they are clinically uninterpretable –

[13] https://www.hcup-us.ahrq.gov/toolssoftware/comorbidityicd10/comorbidity_icd10.jsp

[14] MDCs fully described in Table A.3 of the PSI Parameter Estimates
https://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V2019/Parameter_Estimates_PSI_v2019.pdf

[15] MDRGs fully described in Table A.2 of the PSI Parameter Estimates
https://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V2019/Parameter_Estimates_PSI_v2019.pdf

981-989 for OR procedures unrelated to the principal diagnosis, 998 for invalid principal diagnosis, and 999 for ungroupable.)

- Point of origin indicating transfer in from another hospital.

CMS starts de novo with the full set of available risk factors and performs risk factor selection methods to develop risk-adjustment models. In particular, after filtering out risk factors that have small denominators (e.g., fewer than 30 records), and that have quasi completeness or high collinearity, (e.g., Variance Inflation Factor of 1000 or higher), the least absolute shrinkage and selection operator (LASSO) feature selection method, a penalized regression approach, is used to select risk factors that provide good balance between model performance and model complexity. The LASSO method is used because the traditional p-value or stepwise based selection methods use sequential fitting, which could lead to biased coefficient estimates and less optimal models. Generalized estimating equations (GEE) with the logit-binomial link function are used to address the clustering of patients within hospitals.

The exceptions to this method are PSI 11 (Postoperative Respiratory Failure Rate) and PSI 13 (Postoperative Sepsis Rate). After fitting the PSI 11 and PSI 13 risk-adjustment models using GEE, the resulting model fit was poor particularly for discharges at the lowest risk of a PSI event. In an effort to improve model fit, the PSI 11 and PSI 13 risk-adjustment models were refit with ordinary logistic regression on the same risk factors. The C statistics for the overall model are similar using the logistic regression model compared with the GEE model with logit-binomial link function; however, model calibration is substantially improved for discharges at the lowest risk of a PSI event.

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☒ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

In general, any observed relationship between social or sociodemographic status (SDS) factors and PSI events could work through three mechanisms:

1. Patients with social risk factors may have worse health at the time of hospital admission (in ways that cannot be captured through other measured variables, such as comorbidities, diagnoses, and procedures or services used).
2. Patients with social risk factors may receive care at lower-quality hospitals.
3. Patients with social risk factors may receive poorer care within hospitals, even accounting for their severity of illness and their distribution across hospitals.

The second and third mechanisms have been extensively studied and validated in the peer-reviewed literature as well as research reported by the Office of the Assistant Secretary for Planning and Evaluation (ASPE) pursuant to the IMPACT Act of 2014.

Therefore, CMS follows ASPE guidance, as summarized in its March 2020 *Report to Congress on Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program*: "Recommendation 1.6: Composite scores should not be adjusted for social risk factors for public reporting. Composite measures used for public reporting should NOT use measures that are adjusted for public reporting. They should also not use other

methods to account for social risk, such as peer grouping." In the case of PSI 90, these recommendations are generally consistent with guidance contained in the *NQF Technical Report on Risk-Adjustment for Socioeconomic Status or Other Sociodemographic Factors* (August 15, 2014): "For example, the outcome of central line infection occurring during a hospital stay would not have a conceptual basis for SDS adjustment, as there is no logical reason why these measures should be affected by variables such as poverty, illiteracy, or limited English proficiency. Important considerations include whether the key processes leading to an outcome are directly under the control of the healthcare unit and do not depend on active patient participation as in the examples noted above." Specifically, based on NQF's suggested questions for identifying a conceptual basis for adjusting for sociodemographic factors (p. 36):

- Prior research does not indicate a consistent relationship between SDS and PSI outcomes.
- There is no clear theory supporting a relationship between SDS and PSI outcomes.
- There is no passage of time between the hospital's treatment and PSI outcomes, during which other factors may have an effect.
- Patient actions or decisions affected by SDS do not consistently influence PSI outcomes (e.g., ability to purchase medications).
- The patient community outside the hospital has no clear influence on PSI outcomes (e.g., distance to pharmacies, groceries, other resources).

CMS continues to explore ways to improve the PSI risk-adjustment models to include additional measures of risk that are correlated with social risk factors, consistent with ASPE recommendations, including functional risk and prior health service use. For example, a variable indicating transfer-in from another hospital has been added to all risk-adjustment models. Variables indicating transfer-in from skilled nursing care, assisted living facilities, and hospice care (as measures of functional risk) are currently being tested, and will be added to the next version of the CMS Medicare PSI software if appropriate.

2b3.4a. What were the statistical results of the analyses used to select risk factors?

Please refer to detailed model results above; the table below summarizes the results of feature selection for PSI risk-adjustment models. HCUP results are shown with Medicare FFS results to demonstrate the robustness of the approach across data sources, given that previous NQF endorsement was based on HCUP all-payer data.

Table 17. Number of Covariates Selected for PSI Risk-Adjustment Models in CMS v10.0 (Current), CMS v9.0, and AHRQ v2019

PSIs	CMS Medicare FFS Data (CMS v10.0)	CMS Medicare FFS Data (CMS v9.0)	HCUP All-Payer Data (v2019)
PSI 03	138	135	147
PSI 06	127	94	125
PSI 08	131	130	116
PSI 09	95	98	113
PSI 10	77	88	91

PSIs	CMS Medicare FFS Data (CMS v10.0)	CMS Medicare FFS Data (CMS v9.0)	HCUP All-Payer Data (v2019)
PSI 11	70	72	86
PSI 12	125	113	124
PSI 13	82	83	92
PSI 14A	*	67	79
PSI 14B	*	49	23
PSI 15	132	93	103

Source: CVP/Mathematica Scientific Acceptability Report, December 2019; AHRQ v2019 PSI Parameter Estimates

Note: PSI 14A reflects postoperative wound dehiscence with an open approach and PSI 14B reflects non-open approach.

*cell intentionally left blank

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.) Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.

Not applicable; see 2b3.3b above.

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

For each PSI, we summarize model fit and adequacy using the following measures:

- Overall model discrimination as assessed by C-statistic. We calculated C-statistics using Medicare FFS data to determine the predicted probability from the CMS v10.0 model developed on the July 2016–June 2018 Medicare FFS reference population. The C-statistic is the area under the receiver-operator curve that measures the discriminative ability of a regression model. It also describes the probability that a randomly selected patient who experienced a PSI event had a higher expected value than a randomly selected patient who did not experience that event.
- Model fit by deciles of patient risk using Hosmer-Lemeshow plots. The Hosmer-Lemeshow plots show the observed-to-predicted ratio for deciles of risk with July 2016–June 2018 Medicare FFS data processed through CMS v10.0 models. For each PSI, the deciles of risk are ten mutually exclusive groups containing an equal number 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.

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

HCUP results are shown with Medicare FFS results to demonstrate the robustness of the approach across data sources, given that previous NQF endorsement was based on HCUP all-payer data.

Table 18. Discrimination of PSI Risk-Adjustment Models in CMS v10.0 (Current), CMS v9.0, and AHRQ v2019

PSIs	CMS Medicare FFS Data (v10.0) C-statistic	CMS Medicare FFS Data (v9.0) C-statistic	HCUP All-Payer Data (v2019) C-statistic
PSI 03	0.814	0.812	0.809
PSI 06	0.852	0.852	0.847
PSI 08	0.871	0.852	0.861
PSI 09	0.791	0.799	0.771
PSI 10	0.902	0.907	0.906
PSI 11	0.828	0.825	0.825
PSI 12	0.711	0.712	0.740
PSI 13	0.847	0.843	0.847
PSI 14A	0.777	*	0.805
PSI 14B	0.889	*	0.774
PSI 15	0.902	0.907	0.779

Source: CVP/Mathematica Scientific Acceptability Report, December 2019; AHRQ v2019 PSI Parameter Estimates

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

C-statistics indicate the area under the receiver-operator curve that measures the discriminative capacity of a regression model.

Risk-adjustment for PSI 14 was done at the individual component level for CMS v10.0 and the overall level for CMS v9.0, resulting in NA values for PSI 14A and PSI 14B.

*cell intentionally left blank

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Due to the very large sample sizes used to test these measures, all Hosmer-Lemeshow chi square statistics are significant ($p<0.05$) and uninformative. Therefore, we rely on risk decile plots as shown below (Figures 2-12).

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Table 19. Number and Rate (per 1000) of PSI 90 Component Numerator Events in Each Predicted Risk Decile

Predicted Risk Decile	1	2	3	4	5	6	7	8	9	10
PSI 03 N	53	130	191	241	299	391	536	765	1,374	4,146

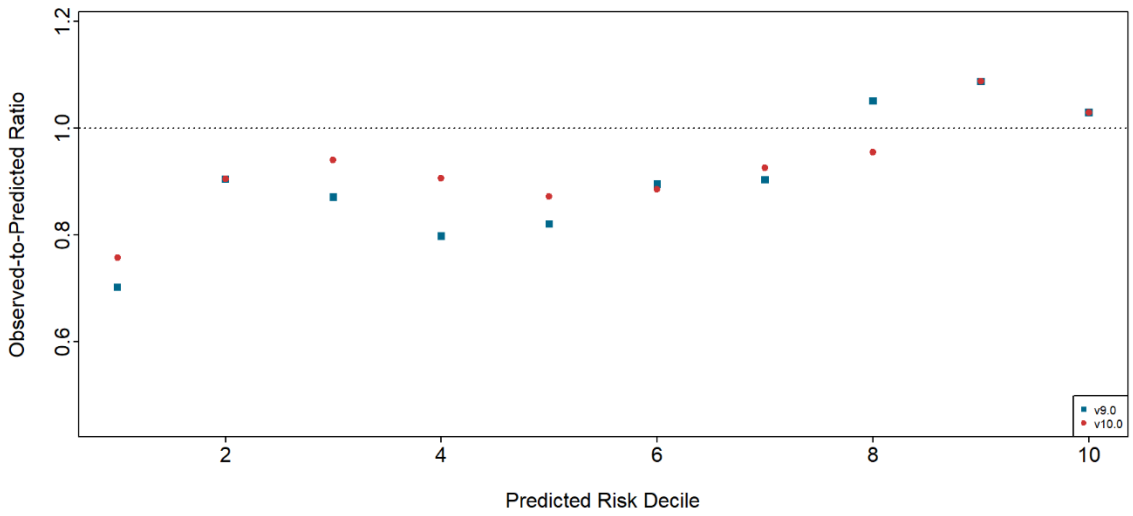
Predicted Risk Decile	1	2	3	4	5	6	7	8	9	10
PSI 03 Rate	0.039	0.096	0.142	0.179	0.222	0.290	0.398	0.568	1.019	3.076
PSI 06 N	21	45	86	115	156	196	260	341	420	2,781
PSI 06 Rate	0.012	0.026	0.049	0.066	0.089	0.112	0.149	0.195	0.241	1.594
PSI 08 N	5	10	20	29	39	49	80	96	276	1,056
PSI 08 Rate	0.003	0.007	0.013	0.019	0.025	0.032	0.052	0.063	0.180	0.689
PSI 09 N	41	45	92	346	532	996	1,270	1,535	2,330	4,457
PSI 09 Rate	0.088	0.096	0.197	0.740	1.138	2.131	2.718	3.285	4.986	9.538
PSI 10 N	3	4	11	14	28	70	163	281	535	2,420
PSI 10 Rate	0.012	0.015	0.042	0.054	0.108	0.269	0.627	1.080	2.057	9.303
PSI 11 N	72	135	181	295	451	613	847	1,431	2,310	6,619
PSI 11 Rate	0.342	0.641	0.860	1.401	2.142	2.912	4.023	6.797	10.972	31.439
PSI 12 N	333	608	955	961	1,186	1,544	1,811	2,331	3,221	5,838
PSI 12 Rate	0.666	1.217	1.911	1.923	2.374	3.090	3.625	4.665	6.447	11.684
PSI 13 N	76	78	119	176	336	480	748	1,299	2,263	6,567
PSI 13 Rate	0.300	0.307	0.469	0.694	1.324	1.892	2.948	5.119	8.918	25.880
PSI 14A N	4	5	23	31	51	67	99	134	211	339
PSI 14A Rate	0.076	0.094	0.434	0.585	0.963	1.265	1.870	2.531	3.985	6.402
PSI 14B N	0	0	1	0	0	1	1	4	5	24
PSI 14B Rate	0.000	0.000	0.018	0.000	0.000	0.018	0.018	0.071	0.089	0.425
PSI 15 N	16	23	45	74	153	235	329	443	830	1,762
PSI 15 Rate	0.052	0.074	0.145	0.239	0.494	0.759	1.062	1.430	2.679	5.687

Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

PSI 14A reflects postoperative wound dehiscence with an open approach and PSI 14B reflects non-open approach.

Figure 2. PSI 03 Hosmer-Lemeshow plot

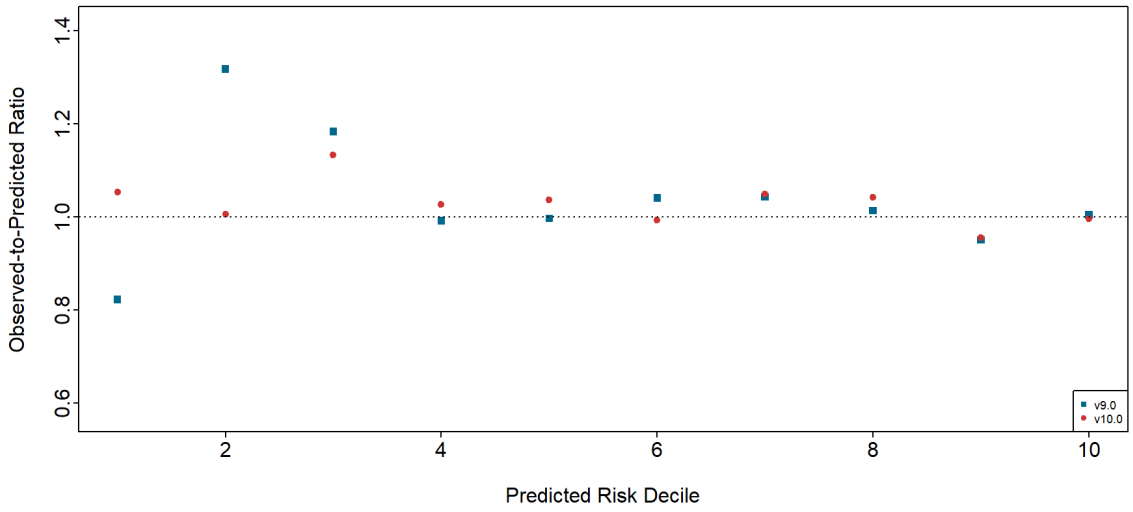


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 3. PSI 06 Hosmer-Lemeshow plot

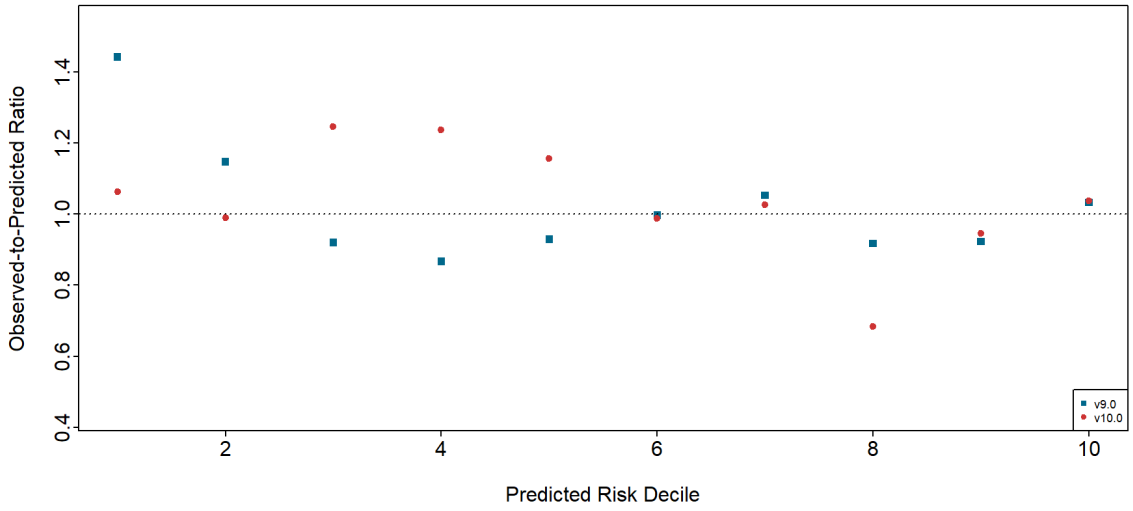


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 4. PSI 08 Hosmer-Lemeshow plot

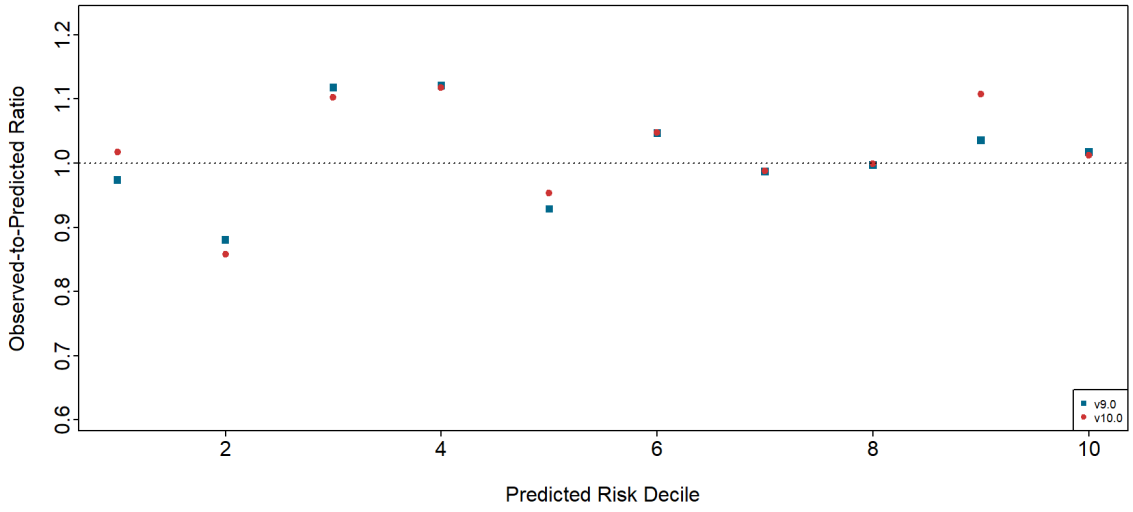


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 5. PSI 09 Hosmer-Lemeshow plot

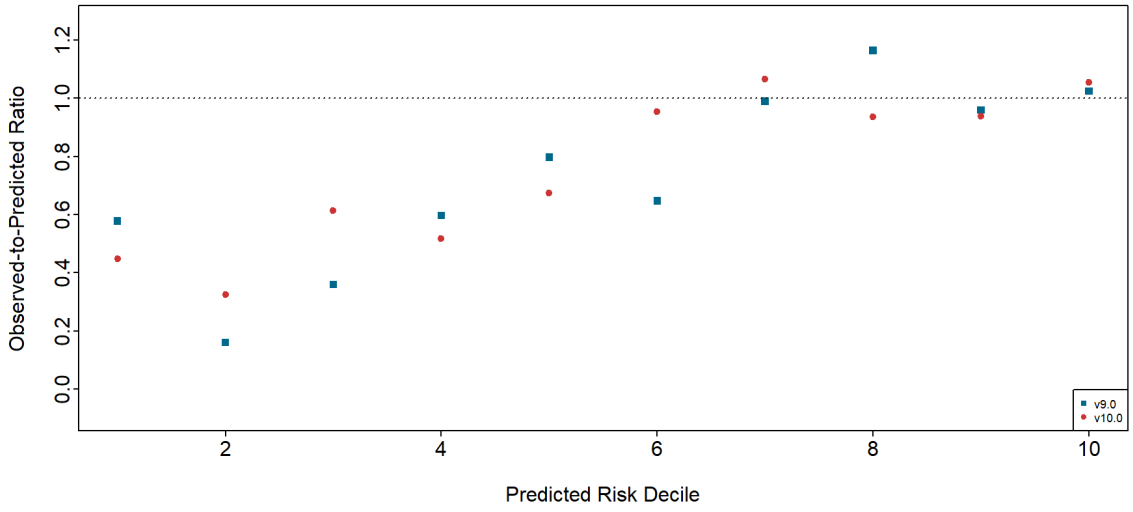


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software..

Figure 6. PSI 10 Hosmer-Lemeshow plot

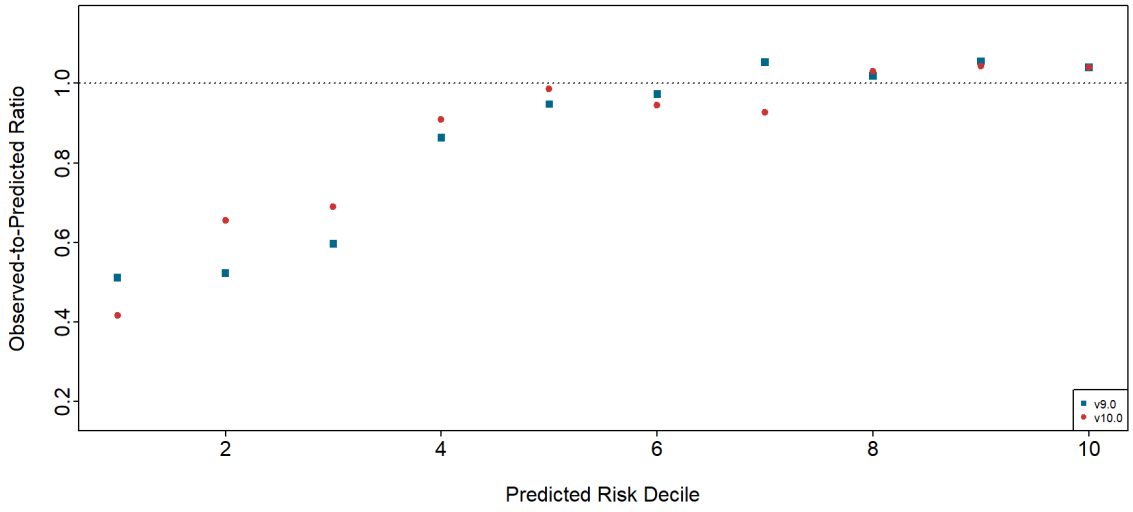


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 7. PSI 11 Hosmer-Lemeshow plot

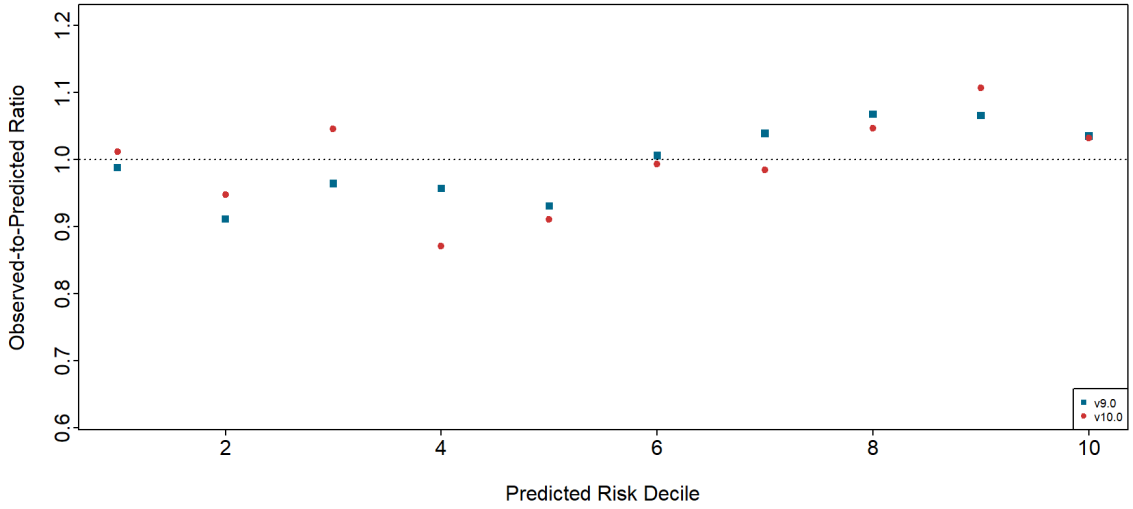


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 8. PSI 12 Hosmer-Lemeshow plot

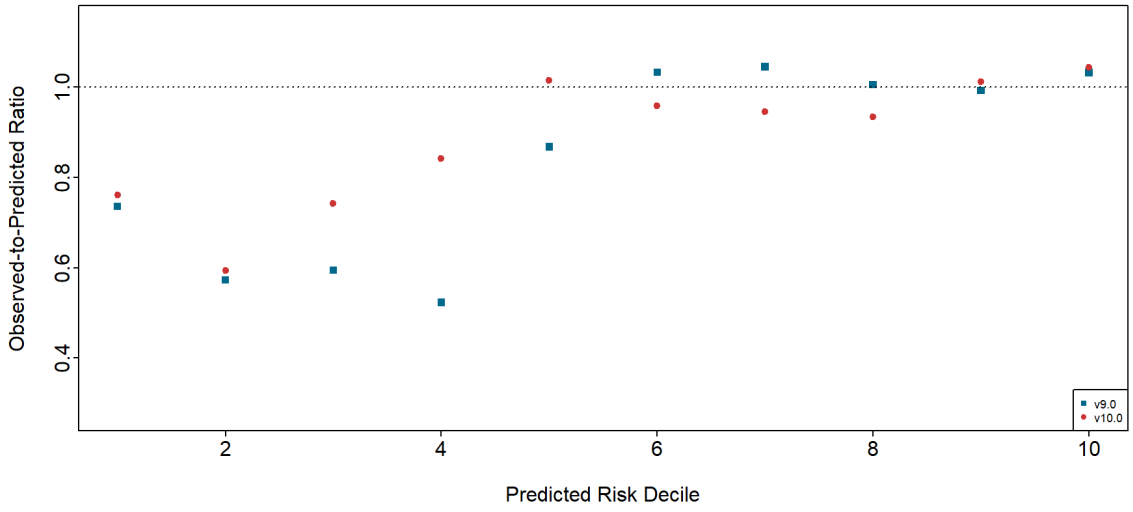


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 9. PSI 13 Hosmer-Lemeshow plot

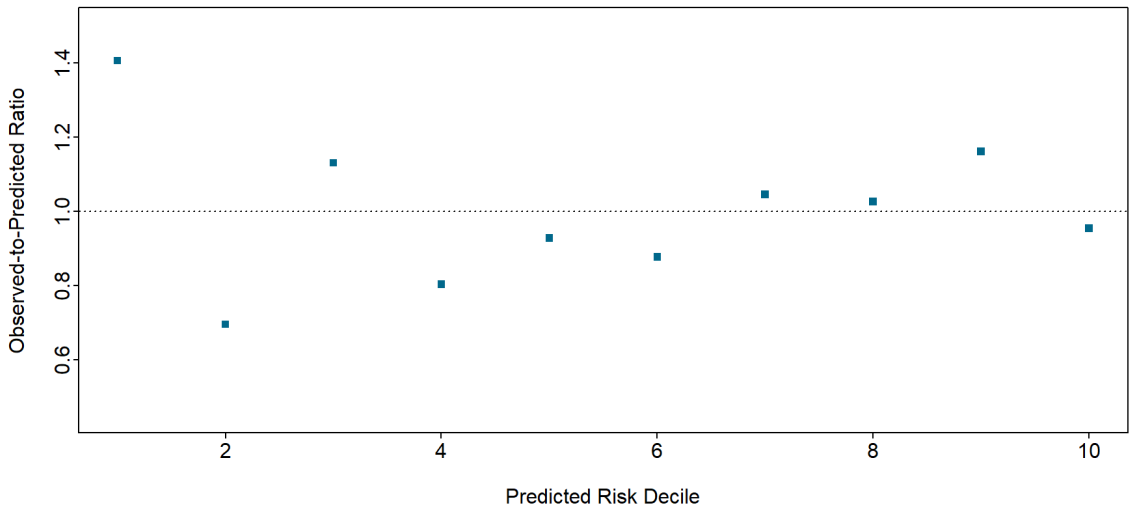


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 10. PSI 14A Hosmer-Lemeshow plot

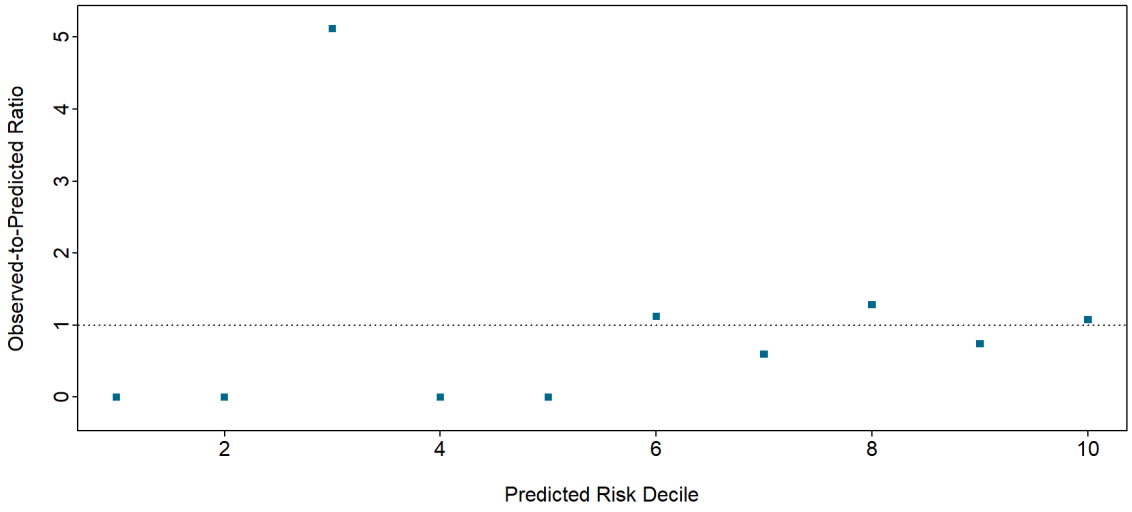


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 11. PSI 14B Hosmer-Lemeshow plot

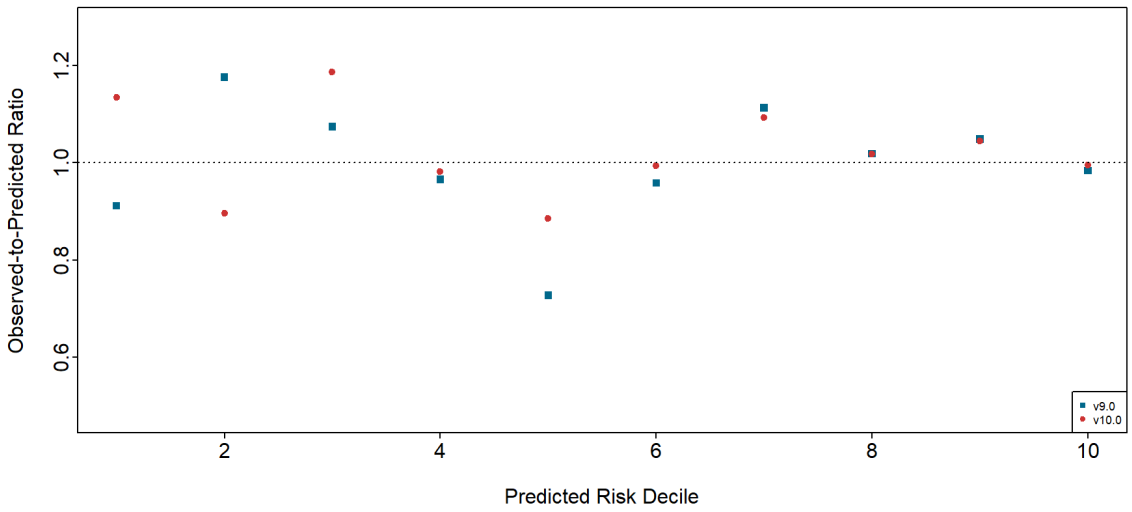


Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

Figure 12. PSI 15 Hosmer-Lemeshow plot



Source: CVP/Mathematica Risk-Adjustment Report, October 2019

Note: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with v10.0 CMS software.

CMS v9.0 represents analysis of the same discharges processed through CMS v9.0 software.

2b3.9. Results of Risk Stratification Analysis:

Not applicable

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

The c statistic is a measure of the extent to which a statistical model is able to discriminate between patients with and without the outcome, equivalent to the area under a receiver operating characteristic (ROC) curve. The c statistic ranges between 0.5 for a model that is no better than random prediction to 1.0 for a model with perfect prediction, in which outcomes are fully explained by patient characteristics and quality-of-care plays no role. As shown in **Table 18**, PSI risk-adjustment models are very strong, with discrimination (c) statistics over 0.75 except for PSI 12 (c=0.71). In general, c-statistics >0.75 are considered excellent for these types of risk-adjustment models. As shown in **Table 19**, these models also sort patients very well based on their risk, with 18-fold (11.684/0.666, PSI 12) to 807-fold (9.303/0.012, PSI 10) differences in risk between the decile of lowest risk patients and the decile of highest risk patients. Finally, **Figures 2-12** show excellent model calibration, with observed-to-expected ratios close to 1.0 across nearly all deciles, especially for the current version 10.0 (in comparison with the previous version 9.0). Where significant deviations from 1.0 exist, as in **Figures 2, 3, 6, 7, and 9**, they are generally in the lowest risk deciles, which include less than 1% of all events. In other words, calibration error is limited to the portion of the risk distribution in which events are extremely rare.

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

All risk models have been tested and recalibrated using both Medicare FFS data and all-payer HCUP data; see results in **Tables 17-18** above. The results of feature selection, model discrimination, and model calibration are generally very consistent between the two data sources.

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

Note: *Applies to the composite performance measure.*

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

First, the distribution of CMS Medicare PSI 90 scores was tabulated across hospitals to examine the magnitude of relative differences in performance. The mean, standard deviation, 5th, 25th, 50th, 75th, and 95th percentiles are shown in **Table 20**. The interquartile range represents the difference between the 25th and 75th percentiles.

Second, 95% confidence intervals are computed around each hospital's estimated CMS Medicare PSI 90 score, based on the square root of its estimated variance. The estimated variance is computed based on the signal variance-covariance matrix in the reference population and the hospital's own reliability weights. This calculation is based on the assumption of independence among the component PSIs – that is, component PSI rates are uncorrelated within hospitals. Hospitals for which the 95% confidence interval does not include 1 (the value based on the national reference population) are classified as outliers. The CMS Medicare PSI 90 score values for these outliers are shown in **Table 21** and graphically displayed in **Figure 13**.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

Table 20. Distribution of Hospital Performance on PSI 90 across Three Time Periods Used for Reporting

Years	N	Mean score	SD	5th percentile	25th percentile	Median score	75th percentile	95th percentile
2016-17	3305	0.995	0.173	0.805	0.908	0.972	1.035	1.288
2017-18	3287	0.995	0.165	0.808	0.908	0.971	1.028	1.287
2018-19	3249	0.996	0.16	0.804	0.913	0.972	1.031	1.276

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

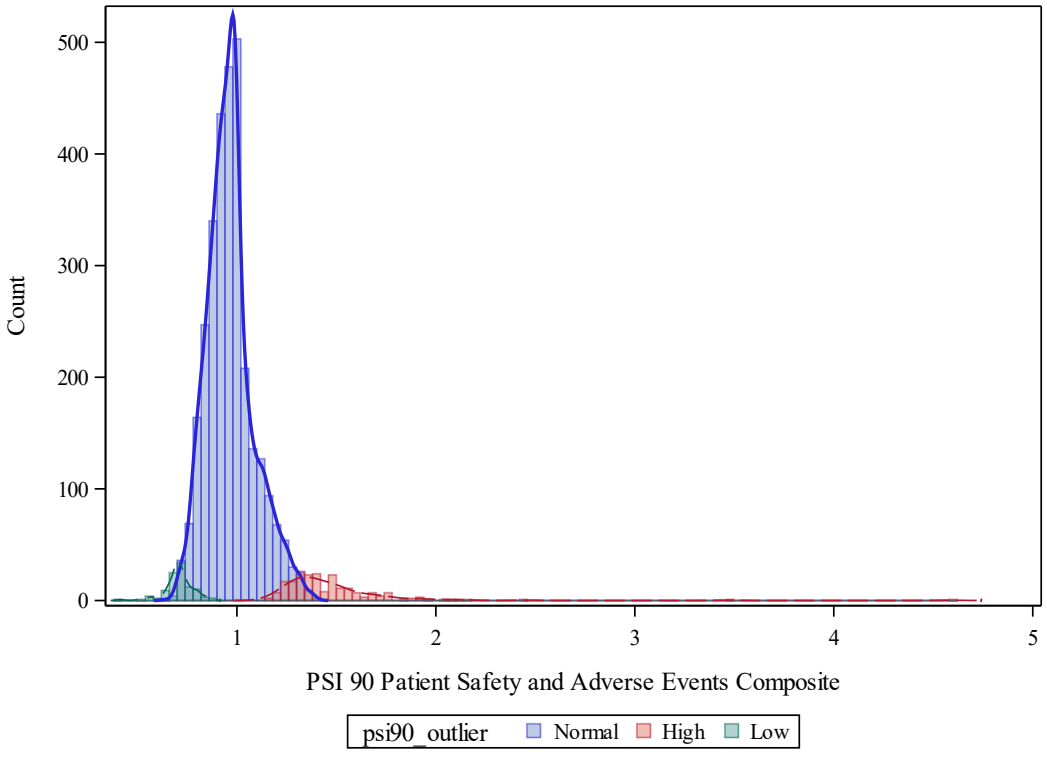
Abbreviations: SD=standard deviation; p=percentile

Table 21. Descriptive Statistics for PSI 90 Scores across Statistically Determined Hospital Outlier Groups

Outlier Group	Number of Hospitals (%)	Mean	Minimum	25th percentile	Median	75th percentile	Maximum
High performing	103 (3.1%)	0.71	0.41	0.68	0.71	0.75	0.87
Neither	3031 (90.6%)	0.97	0.64	0.89	0.96	1.02	1.40
Low performing	211 (6.3%)	1.48	1.14	1.32	1.40	1.54	4.58

Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2019) processed with CMS v10.0 software.

Figure 13. Distribution of PSI 90 Scores across Statistically Determined Hospital Outlier Groups



Source: Medicare FFS discharges from IPPS hospitals (7/1/2016-6/30/2018) processed with CMS v10.0 software.

Note: hospital counts along the Y axis

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

Statistically significant and practically/clinically meaningful differences in performance can be identified. Although the interquartile range in **Table 20** indicates only 13-14% relative difference in PSI 90 scores between the 25th and 75th percentiles (e.g., 1.035/0.908), this finding reflects removal of confounding effects through risk-adjustment and removal of noise through reliability-adjustment. The relative difference between the 5th and 95th percentiles in **Table 20** (e.g., 1.288/0.805) is 58-60%, which represents a substantial difference in the incidence of clinically important complications (after both risk-adjustment and reliability-adjustment). As shown in **Table 21**, PSI 90 identifies about 10% of hospitals (e.g., 314/3145) as performance outliers, using 24 months of Medicare FFS claims data. High-performing outliers have PSI 90 scores of 0.41 to 0.87 (median 0.71, IQR 0.68-0.75), indicating substantially fewer complications than expected from the reference population. Low-performing outliers have PSI 90 scores of 1.14 to 4.58 (median 1.40, IQR 1.32-1.54), indicating substantially more complications than expected from the reference population. **Figure 13** shows that when the PSI 90 score is less than about 0.5 or more than about 1.3, the majority of hospitals can be identified as statistical outliers (because the histogram for the outlier group surpasses the histogram for the nonoutlier group).

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

Note: *Applies to all component measures, unless already endorsed or are being submitted for individual endorsement.*

If only one set of specifications, this section can be skipped.

Note*: This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

Not applicable

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

Not applicable

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Not applicable

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

Note: *Applies to the overall composite measure.*

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

For a claims-based measure such as PSI 90, a discharge may be missing a key data element (for example, present-on-admission codes); claims may be missing from the analytic file; or measure results may be missing from certain hospitals that do not have sufficient numbers of denominator-eligible cases. In general, Medicare claims are essentially 100% complete on all of the necessary data elements because payment is contingent on submission of a complete claim. Therefore, we focused on hospitals that do not have PSI 90 component values. When a hospital has fewer than three denominator cases, the CMS Medicare PSI software substitutes the component indicator value with the observed-to-expected ratio in the reference population (1.0) to construct the PSI 90 composite. Although it happens infrequently, a hospital can receive a CMS Medicare PSI 90 composite value if one to ten of the 10 components are imputed using the observed-to-expected ratio in the reference population (i.e., 1.0). We examined the number of components missing from the PSI 90 composite for each hospital and the proportion of hospitals missing combinations of component indicators.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

Table 22. Number and Percentage of Hospitals with Missing PSI 90 Component Indicators, by Number of Missing Components, with Mean Imputed PSI 90 Scores

Number of missing PSI components	Number of hospitals	Percentage of hospitals	Mean PSI 90 composite value	Standard Deviation
0	2,947	89.0%	0.992	0.202
1	31	0.9%	0.976	0.076
2	24	0.7%	0.970	0.064
3	101	3.0%	0.972	0.113

Number of missing PSI components	Number of hospitals	Percentage of hospitals	Mean PSI 90 composite value	Standard Deviation
4	24	0.7%	0.992	0.059
5	6	0.2%	0.992	0.005
6	45	1.4%	1.018	0.224
7	119	3.6%	0.990	0.027
8	8	0.2%	1.000	0.000
10	8	0.2%	1.000	0.000

Source: CVP/Mathematica Scientific Acceptability Report, December 2019

Medicare FFS discharges from 3,313 IPPS hospitals, July 1, 2016, through June 30, 2018 processed with CMS v10.0 PSI software.

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

Missing data are extremely uncommon (<0.01%) at the component measure level, so they cannot have any meaningful impact on PSI 90 scores. Missing component measures are more common, affecting about 11% of eligible hospitals, as shown in **Table 23**. Eight hospitals are missing all 10 components, and another 8 are missing 8 or 9 components. The majority (89%) of hospitals, as shown in **Table 23**, have all 10 PSI components contributing to PSI 90. These hospitals have a composite value of 0.992, slightly better than the national average, 1.000. Three percent of hospitals have the national observed-to-expected ratio substituted for three PSIs in the composite calculation. These hospitals have a slightly better-than-average composite value, 0.972. For accountability applications, users are encouraged to set a minimum threshold for the allowable number of PSI 90 component measures. CMS is considering suppressing public reporting for hospitals with 4 or more missing component measures, as 7 non-missing component measures are sufficient to estimate about half of the total weight of PSI 90.

2c. EMPIRICAL ANALYSIS TO SUPPORT COMPOSITE CONSTRUCTION APPROACH

Note: *If empirical analyses do not provide adequate results—or are not conducted—justification must be provided and accepted in order to meet the must-pass criterion of Scientific Acceptability of Measure Properties. Each of the following questions has instructions if there is no empirical analysis.*

2d1. Empirical analysis demonstrating that the component measures fit the quality construct, add value to the overall composite, and achieve the object of parsimony to the extent possible.

2d1.1 Describe the method used (*describe the steps—do not just name a method; what statistical*

analysis was used; if no empirical analysis, provide justification)

We computed weighted Pearson and Spearman (rank) correlations between hospitals' PSI 90 scores and each of their component smoothed risk-adjusted rates (RAR). These correlations are equal to the correlations with the corresponding smoothed observed/expected ratios because the RAR is a constant multiple of the observed/expected ratio within each PSI. The weighted Pearson correlation uses the PSI denominator for a weight.

We also computed Spearman rank correlations among all of the component indicators, after risk-adjustment and smoothing. One tenet of this composite is that each component measure is correlated with an aspect of each hospital's underlying quality of care. Therefore, we expect to observe positive hospital-level correlations among the individual measures within the composite.

2d1.2. What were the statistical results obtained from the analysis of the components? (e.g., *correlations, contribution of each component to the composite score, etc.*; *if no empirical analysis, identify the components that were considered and the pros and cons of each*)

Table 23. PSI 90 Component Weights, by Version, and Item-Total Spearman Rank Correlations

PSIs	CMS v9.0 weight	CMS v10.0 weight	CMS v10.0 Item-Total Correlation (2016-2018)
PSI 03	0.134	0.161	0.659
PSI 06	0.041	0.039	0.172
PSI 08	0.015	0.015	0.046
PSI 09	0.042	0.043	0.160
PSI 10	0.078	0.081	0.232
PSI 11	0.212	0.185	0.522
PSI 12	0.185	0.188	0.494
PSI 13	0.247	0.242	0.472
PSI 14	0.009	0.009	0.038
PSI 15	0.037	0.037	0.216

Source: First two columns are from the CVP/Mathematica Scientific Acceptability Report, December 2019.

Note: Results based on CMS v10.0 PSI software with parameters derived from Medicare FFS discharges from 3,313 IPPS hospitals, 7/1/2016-6/30/2018. CMS v9.0 weights were derived using data from 7/1/2015-6/30/2017.

Table 24. Spearman Rank Correlation Between PSI 90 Composite Score and Each of the PSI Components' Smoothed Risk-Adjusted Rate

Correlation	PSI 03	PSI 06	PSI 08	PSI 09	PSI 10	PSI 11	PSI 12	PSI 13	PSI 14	PSI 15
Spearman Correlation	0.6559	0.168	0.080	0.160	0.223	0.538	0.484	0.440	0.051	0.200
N	3,289	3,298	3,294	3,117	3,001	2,995	3,117	2,984	3,042	3,129
Weighted Pearson Correlation	0.7835	0.176	0.023	0.242	0.284	0.465	0.471	0.438	0.065	0.252

Source: Medicare FFS discharges from IPPS hospitals (7/1/2017-6/30/2019) processed with v10.0 CMS software

Note: all correlations were significant (<0.0001) with the exception of the weighted Pearson correlation for PSI 08 (p=0.1947)

Table 25. Spearman Rank Correlations Among Risk-Adjusted PSI 90 Component Indicators

PSI	Corr.	PSI03	PSI06	PSI08	PSI09	PSI10	PSI11	PSI12	PSI13	PSI14	PSI15
PSI 03	Corr.	1.000	0.302	0.185	0.245	0.258	0.232	0.314	0.256	0.188	0.279
PSI 03	N	3289	3289	3285	3113	2997	2991	3113	2980	3041	3127
PSI 06	Corr.	0.302	1.000	0.217	0.267	0.301	0.272	0.273	0.285	0.193	0.273
PSI 06	N	3289	3298	3294	3117	3001	2995	3117	2984	3042	3129
PSI 08	Corr.	0.184	0.217	1	0.139	0.197	0.109	0.182	0.162	0.119	0.177
PSI 08	N	3285	3294	3294	3117	3001	2995	3117	2984	3042	3129
PSI 09	Corr.	0.245	0.267	0.139	1.000	0.309	0.224	0.280	0.245	0.175	0.265
PSI 09	N	3113	3117	3117	3117	3001	2995	3117	2984	3041	3089
PSI 10	Corr.	0.258	0.301	0.197	0.309	1.000	0.285	0.239	0.371	0.262	0.289
PSI 10	N	2997	3001	3001	3001	3001	2994	3001	2983	2947	2978
PSI 11	Corr.	0.232	0.272	0.109	0.224	0.285	1.000	0.239	0.413	0.139	0.232
PSI 11	N	2991	2995	2995	2995	2994	2995	2995	2981	2943	2972
PSI 12	Corr.	0.314	0.273	0.182	0.280	0.239	0.239	1.000	0.277	0.180	0.245
PSI 12	N	3113	3117	3117	3117	3001	2995	3117	2984	3041	3089
PSI 13	Corr.	0.256	0.285	0.162	0.245	0.371	0.413	0.277	1.000	0.193	0.286
PSI 13	N	2980	2984	2984	2984	2983	2981	2984	2984	2936	2963
PSI 14	Corr.	0.188	0.193	0.119	0.175	0.262	0.139	0.180	0.193	1.000	0.224
PSI 14	N	3041	3042	3042	3041	2947	2943	3041	2936	3042	3042
PSI 15	Corr.	0.279	0.273	0.177	0.265	0.289	0.232	0.245	0.286	0.224	1.000

PSI	Corr.	PSI03	PSI06	PSI08	PSI09	PSI10	PSI11	PSI12	PSI13	PSI14	PSI15
PSI 15	N	3127	3129	3129	3089	2978	2972	3089	2963	3042	3129

Source: Medicare FFS discharges from IPPS hospitals (7/1/2017-6/30/2019) processed with v10.0 CMS software

Abbreviations: Corr=correlation, N: number of hospitals included in that correlation analysis (based on having at least 3 denominator-eligible records)

Note: all correlations were statistically significant (<0.0001)

2d1.3. What is your interpretation of the results in terms of demonstrating that the components included in the composite are consistent with the described quality construct and add value to the overall composite? (i.e., *what do the results mean in terms of supporting inclusion of the components; if no empirical analysis, provide rationale for the components that were selected)*

Table 23 shows the empirically derived PSI component weights from the current software (v10.0, based on data from 7/1/2016-6/30/2018) and the previous version (v9.0, based on data from 7/1/2015-6/30/2017). The first two columns demonstrate substantial consistency in these weights over time; the increased weight on PSI 03 (Pressure Ulcer) was attributable to dropping several undesirable exclusion criteria (e.g., exclusion of patients transferred from other hospitals or long-term care facilities that developed new pressure injuries after admission). The third column shows that item-total correlations are much higher than component weights, suggesting that the composite is leveraging shared variation that exceeds what would be expected simply from the construction of the composite.

Table 24 updates these item-total correlations with a more recent year of data (7/1/2017-6/30/2019), adding p values and weighted Pearson correlations. These hospital-level correlations vary from low (<0.1) for PSIs 08 and 14 to high (>0.4) for PSIs 03, 11, 12, and 13, but all are consistently positive. Finally, **Table 25** shows that all of the inter-item correlations among risk-adjusted PSI component measures are positive and highly significant. The highest correlation of 0.413 was between PSI 11 (Postoperative Respiratory Failure) and PSI 13 (Postoperative Sepsis), whereas the lowest correlation of 0.109 was between PSI 11 and PSI 08 (In-Hospital Hip Fracture). These findings support the design of PSI 90 as a single composite summarizing various hospital harms.

2d2. Empirical analysis demonstrating that the aggregations and weighting rules are consistent with the quality construct and achieve the objective of simplicity to the extent possible

2d2.1 Describe the method used (*describe the steps—do not just name a method; what statistical analysis was used; if no empirical analysis, provide justification)*

Each component PSI indicator, q, that is part of PSI 90 receives a weight defined by:

$$weight_q = \frac{volume_q \sum_{h=1}^{H} harm_{qh} disutility_{qh}}{\sum_{q=1}^{Q} volume_q \sum_{h=1}^{H} harm_{qh}\, disutility_{qh}}$$

Where:
Q is the total number of component quality indicators, q, in PSI 90.

H is the total number of outcome types (harms), h, related to each component indicator.

volume is the numerator count, or the number of total QI events within the component indicator in the reference population.

harm is the excess risk (risk difference) of each type of outcome (i.e. harm) within each component indicator estimated from a model comparing people with PSI events to those without PSI events in an "at risk" cohort.

disutility is the complement of a utility weight (1-utility_wt) assigned to each excess occurrence of each type of outcome within each component indicator.

For each component indicator in the modified version of PSI 90 composite, two sets of values need to be computed or estimated. The first is the excess risk of each harm outcome (risk difference) that may occur in association with the component PSI event. These harm risks are multiplied by harm-specific disutility scores, which reflect the relative valuation of various outcome states by patients and clinicians, and then summed across all of the harms relevant to a component PSI, to obtain the summed harm weight for each PSI. Next numerator weights are calculated from the volume (count) of each PSI component event in the CMS FFS reference population. Finally, the volume weight for each PSI is multiplied by its summed harm weight, and the resulting product is rescaled across all 10 components so that the sum of the final weights is 1.

2d2.2. What were the statistical results obtained from the analysis of the aggregation and weighting rules? (e.g., *results of sensitivity analysis of effect of different aggregations and/or weighting rules; if no empirical analysis, identify the aggregation and weighting rules that were considered and the pros and cons of each*)

HCUP results are shown with Medicare FFS results to demonstrate the robustness of the approach across data sources, given that previous NQF endorsement was based on HCUP all-payer data.

Table 26. Final PSI Component Weights Reflect Both Indicator-Specific Harm Weights and Population-Specific Volume Weights

Component	PSI 03	PSI 06	PSI 08	PSI 09	PSI 10	PSI 11	PSI 12	PSI 13	PSI 14	PSI 15
Harm Weight	0.3080	0.1381	0.1440	0.0570	0.3584	0.2219	0.1557	0.3102	0.1441	0.1474
CMS Medicare FFS Volume Weight	0.1039	0.0566	0.0212	0.1489	0.0451	0.1657	0.2403	0.1553	0.0128	0.0500
Final CMS Medicare FFS Weight	0.1608	0.0392	0.0154	0.0426	0.0812	0.1846	0.1879	0.2419	0.0093	0.0370
AHRQ All-Payer Volume Weight	0.0860	0.0538	0.0172	0.1598	0.0280	0.1821	0.2543	0.1550	0.0138	0.0500
Final AHRQ All-Payer Weight	0.1373	0.0385	0.0128	0.0472	0.0520	0.2094	0.2052	0.2491	0.0103	0.0382

2d2.3. What is your interpretation of the results in terms of demonstrating the aggregation and weighting rules are consistent with the described quality construct? (i.*e., what do the results mean in terms of supporting the selected rules for aggregation and weighting; if no empirical analysis, provide rationale for the selected rules for aggregation and weighting*)

In the first row, **Table 26** shows the harm weight (excess harm and disutility) for each PSI 90 component; these harm weights are updated every 3-5 years and are intended for use with any application of PSI 90. These harm weights demonstrated the expected patterns, with indicators such as postoperative sepsis, which are associated with higher mortality rates, having higher harm weights than less serious events.

The second row shows volume weights from the Medicare FFS population, while the fourth row shows volume weights from the AHRQ all-payer population. The final weight on each component measure is proportional to the relative incidence of that event in the appropriate reference population, reflecting the overall level of harm associated with each PSI in that reference population. For example, PSI 03 carries a higher weight in CMS Medicare PSI 90 than in AHRQ's implementation simply because pressure injuries have higher incidence in the Medicare FFS population than in the all-payer population.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in electronic claims

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

Because CMS PSI 90 is based on readily available administrative claims data, feasibility is not an issue. This version of the indicator requires present-on-admission (POA) data. Present-on-Admission was added as a data element to the uniform bill form (UB-04) effective October 1, 2007, and hospitals incurred a payment penalty for not including POA on Medicare records beginning October 1, 2008. Each of the several diagnoses in a discharge record can be flagged as "present at the time the order for inpatient admission occurs" or not (see http://www.cdc.gov/nchs/icd/icd9cm_addenda_guidelines.htm). 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.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

No fees. CMS v10.0 is available by request through the CMS Quality Net Help Desk (https://www.qualitynet.org/inpatient/measures/psi/resources).

Article II. 4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
*	Public Reporting CMS Medicare Hospital Compare Program https://www.medicare.gov/HospitalCompare/Data/Serious-Complications.html CMS Medicare Hospital Compare Program https://www.medicare.gov/HospitalCompare/Data/Serious-Complications.html Payment Program https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/HAC-Reduction-Program CMS Hospital-Acquired Condition (HAC) Reduction Program (HACRP) CMS Hospital Value-Based Purchasing Program (HVBP) https://www.qualitynet.org/inpatient/hvbp/measures Regulatory and Accreditation Programs Statewide Quality Advisory Committee (Massachusetts) http://chiamass.gov/sqms/

*cell intentionally left blank

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

Public Reporting

CMS publicly reports these measures to increase the transparency of hospital care, provide useful information for consumers, and assist hospitals in their quality improvement efforts.

CMS Medicare Hospital Compare Program

Publicly available database containing information about the quality of care at over 4,000 Medicare-certified hospitals across the U.S. PSI data are only calculated for hospitals that are paid through the Inpatient Prospective Payment System (IPPS), which excludes critical access hospitals (CAHs), long-term care hospitals (LTCHs),cancer hospitals, children´s inpatient facilities, rural health clinics, federally qualified health centers, inpatient psychiatric hospitals, inpatient rehabilitation facilities, Veterans Administration/ Department of Defense hospitals, and religious, non-medical health care institutions.

https://www.medicare.gov/HospitalCompare/Data/Serious-Complications.html

We report the number of Medicare FFS patients who fall into the denominator and experience each of the component Patient Safety Indicator events in Table 13 below.

Table 13. Medicare FFS Beneficiaries Reported for Component Indicators for Medicare FFS IPPS Hospitals

Component	Numerator	Denominator
PSI 03	8,126	13,477,287
PSI 06	4,421	17,444,847
PSI 08	1,661	15,370,433
PSI 09	11,657	4,711,559
PSI 10	3,551	2,603,987
PSI 11	12,995	2,106,016

PSI 12	19,064	5,035,140
PSI 13	12,150	2,539,548
PSI 14	1,000	1,092,647
PSI 15	3,910	3,096,764

Source: National CMS PSI Results for the 2016-2018 Medicare Population, Supplementary Information July 2019 Public Reporting
https://www.qualitynet.org/files/5d0d3919764be766b01030f4?filename=July2019_Ntl_CMS_PSI_Results_2.pdf
Numerator = Actual number of outcomes that occurred in the July 2016 – June 2018 Medicare FFS IPPS hospital population. An outcome will not count if its associated discharge is not part of the denominator.
Denominator=Number of discharges in the July 2016 – June 2018 Medicare FFS IPPS hospital population that meet the inclusion criteria for each CMS PSI.
Payment Programs:
CMS Hospital-Acquired Condition Reduction Program (HACRP):
Section 3008 of the Affordable Care Act requires CMS to establish a program for IPPS hospitals to improve patient safety, by imposing financial penalties on hospitals that perform poorly with regard to hospital-acquired conditions.
https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/HAC-Reduction-Program
The HACRP includes Medicare-participating acute care hospitals from all states (except Maryland) and the District of Columbia. The most recently reported analysis from 7/1/2016 through 6/30/2108 included 3,177 hospitals, of which 3,134 had valid PSI 90 Z scores (98.65%).
CMS Hospital Value-Based Purchasing Program (HVBP):
Congress authorized the Inpatient Hospital VBP in Section 3001(a) of the Affordable Care Act. The Hospital VBP Program rewards acute care hospitals with incentive payments for the quality of care provided in the inpatient hospital setting. This program encourages hospitals to improve the quality, efficiency, patient experience and safety of care that Medicare beneficiaries receive during acute care inpatient stays by:

- Eliminating or reducing adverse events (healthcare errors resulting in patient harm).
- Adopting evidence-based care standards and protocols in order to obtain the best outcomes for Medicare patients.
- Incentivizing hospitals to develop processes that improve patient experience.
- Increasing the transparency of care quality for consumers, clinicians, and others.
- Recognizing hospitals that provide high-quality care at a lower cost to Medicare.

CMS removed the CMS PSI 90 measure from the Hospital Inpatient Quality Reporting Program in FY 2020 due to substantive changes in the design of the composite that interfered with measuring improvement over time. The CMS PSI 90 will be added to the Hospital VBP Program beginning with FY2023 payment determination. The HVBP includes Medicare-participating acute care hospitals from all states (except Maryland) and the District of Columbia. The most recently reported performance period included 2,731 hospitals, but PSI 90 was not computed.
https://www.qualitynet.org/inpatient/measures/psi/resources
https://www.qualitynet.org/inpatient/hvbp/measures
Regulatory and Accreditation Programs:
Statewide Quality Advisory Committee (Massachusetts):
The committee annually recommends a standard set of health metrics to use throughout statewide health quality efforts.
http://chiamass.gov/sqms/

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

Not applicable

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

Not applicable

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

CMS provides free software, in both SAS and Windows format, to calculate the CMS PSIs. Users may use their own ICD-10-CM/PCS coded hospital administrative data to calculate the PSIs using this software.

In addition, CMS provides technical assistance to users through an online Q&A form (https://cmsqualitysupport.servicenowservices.com/qnet_qa?id=ask_a_question). CMS triages, troubleshoots, and responds to technical inquiries related to methodology and rationale behind the indicators and general questions related to the use of the software. During a calendar year, CMS typically provides technical support to over 1,000 queries.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

The CMS PSI software is updated annually. Technical support is available on an on-going basis. No data updates are necessary; users apply the CMS PSI software to their own hospital administrative data.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

Feedback is obtained from users through a variety of channels, particularly through a technical assistance mechanism described above. In addition, CMS incorporates input on PSI implementation from technical expert panels convened to support PSI development and maintenance, stakeholder committees such as the NQF standing committees, and peer-reviewed or other research publications.

4a2.2.2. Summarize the feedback obtained from those being measured.

CMS' PSI support team routinely receives user inquires via the technical assistance mechanism described above. These inquiries commonly involve clarification regarding the technical specifications of the component indicators (most commonly, PSI 03, PSI 08, and PSI 12), as well as clarification about the population subject to inclusion in the PSI 90 composite, eligible admission types, the number of diagnosis fields used to calculate the component measures, and the Medicare fee-for-service date ranges used to calculate PSI rates.

Specific suggestions for refining or enhancing the PSI specifications are addressed by CMS in consultation with AHRQ as needed, in its capacity as the original developer of PSI 90.

4a2.2.3. Summarize the feedback obtained from other users

Not separately evaluated.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

The CMS PSIs are updated annually, including updating indicator technical specifications in accordance with the latest coding guidance; suggestions from users and other stakeholders obtained through Technical Assistance, committees, or workgroups; and the latest clinical and scientific research. CMS regularly reviews these sources,

identifies possible indicator updates, and prioritizes updates for each indicator and software update based on expected impact on users.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

Over this three-year period from July 2016 through June 2018, based on national Medicare fee-for-service claims data as described in the Testing attachment, PSI 90 has shown minimal change in mean and median values. However, the 75th, 90th, 95th, and 100th percentile values have decreased, suggesting that the hospitals with the highest PSI event rates have been able to reduce their rates. Data from before October 1, 2015 cannot be compared with later data due to the code set conversion from ICD-9-CM to ICD-10-CM/PCS.

However, these results for PSI 90 do not tell the full story, because each component indicator is separately risk-adjusted and reliability-adjusted at the hospital level before it is put into PSI 90. The observed rates of the component indicators are also shown in 1b above and eight of the ten components demonstrate consistent improvement over time between the 7/1/2016-6/30/2017 year and the 7/1/2018-6/30/2019 year. Specifically, overall national observed rates of PSI 03, 06, 08, 09, 10, 11, 12, 13, 14, and 15 have decreased by 2.2%, 14.0%, 7.6%, -3.0%, -1.0%, 28.3%, 5.6%, 4.5%, 10.0%, and 7.3%, respectively. For all components except PSI 14, the overall national observed rate in 2018-19 was lower than the corresponding rate in 2017-18.

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

One consequence of all quality measurement programs that are used for accountability applications is that health care providers focus their attention on the accuracy of the data and try to minimize both inadvertent underreporting of desired processes of care and inadvertent overreporting of undesired outcomes of care. This is not the intended consequence of quality measurement, but it is certainly an expected consequence. In the case of the NQF-endorsed Patient Safety Indicators, there is anecdotal published evidence of efforts to clarify clinical documentation such that clinically inconsequential events and "incidental" injuries "inherent" to a surgical procedure are no longer coded (and thus no longer reported to payers and state health data organizations.(1) Several large hospitals, such as New York University Langone Medical Center and the University of Washington Medical Center have established "prebilling review processes" with "prompt review of documentation and coding to confirm accuracy [of potential PSI diagnoses] and to identify opportunities to improve care quality and safety."(2)

The AHRQ QI Toolkit offers specific guidance to hospitals and quality improvement leaders about "how to establish an effective coding communication and review process."(3) The implication of these efforts is that some of the observed decrease in the incidence of this event over the last decade may be due to more accurate clinical documentation and coding, rather than to true improvements in patient outcomes and quality of care. Therefore, users should be cautious about interpreting recently observed changes in the incidence of component events. There is no evidence that more accurate clinical documentation and coding have had any

negative consequences for individuals or populations. Any harm from increasing providers' attention to documentation is likely to be counterbalanced by the benefits of more accurate data and more careful reflection on adverse events. In addition, these efforts appear to lead to "one-time corrections" in PSI rates, as hospitals implement processes to prevent overreporting, but do not affect the prior or subsequent trend lines. For example, both the University of Washington Medical Center and Cedars Sinai Medical Center (CSMC) reported that concurrent review of clinical documentation was only the first step toward improving PSI performance.(4) CSMC noted that "task forces that include staff from many different departments and disciplines are assigned to carry out a "leave-no-stone-unturned" search for opportunities to prevent harm across the board… all ideas are important…"

Finally, some users have raised a specific concern about unintended consequences of PSI 12, Perioperative Pulmonary Embolism or Deep Vein Thrombosis Rate. Specifically, the concern is that higher rates are a result of "increased vigilance in detection" at some hospitals.(5) Following this argument, high rates may be nonpreventable – even desirable – because perioperative PEs and DVTs are being diagnosed early (i.e., before symptoms develop) and treated aggressively at these "high surveillance" hospitals. Proponents of this argument cite Medicare claims data showing that "postoperative VTE imaging rates ranged from 85.26 per 1000 discharges in the lowest quartile of hospitals… to 168.86 in the highest quartile… drivers of high imaging rates at the 90th quantile were high resident-to-bed ratio (coefficient=51.35, $p<0.01$), Joint Commission accreditation (coefficient=19.05, $p<0.01$), presence of other hospitals in the same market with high imaging rates (coefficient=15.29, $p<0.01$), case severity (coefficient=11.97, $p<0.01$)…" (suggesting that more imaging is associated with higher quality hospitals).(6) Bilimoria et al. examined 2010 data from Hospital Compare and the American Hospital Association and 2009-2010 Medicare claims data; they reported that greater hospital adherence to VTE prophylaxis was very weakly associated with higher risk-adjusted VTE rates ($r2=4.2\%$, $p=0.03$) but risk-adjusted VTE rates increased concordantly with VTE imaging use rates ($p<0.001$). Ju et al. similarly used National Surgical Quality Improvement Program data to identify VTE events and Medicare claims data to obtain information about VTE imaging;(7) mean risk-adjusted VTE rates (within 30 days after surgery) were significantly lower in hospitals in the lowest quartile of VTE imaging use (1.13%) than in hospitals in the highest quartile (1.92%, $p<0.001$). Similarly, Pierce et al. showed in the National Trauma Data Bank, with 147 hospitals from 2001-2005, that "hospitals with an ultrasound rate of 2% or greater had a 1.07% (95% CI: 1.05-1.09%) increase in reported DVT rate for every 1% increase in ultrasound rate."(8) Admission to a "screening trauma center" that performed vascular ultrasound on at least 2% of admitted trauma patients was independently associated with 2.2 (95% CI 1.1-4.3) times higher odds of DVT, after adjusting for age, injury type, injury severity, need for major surgery, and ventilator days.(9)

The critical question, however, is whether more venous imaging, and hence more diagnosis of VTE, is actually better for patients. Over diagnosis of VTE among asymptomatic or minimally symptomatic patients may lead to overtreatment, with the known adverse effects of anticoagulation and/or IVC device placement. Evidence-based guidelines note that "although distal DVT may be present in patients with a normal proximal ultrasound, it is seldom if ever associated with important clinical sequelae."(10) With respect to treatment, the American College of Chest Physicians also states, "in patients with acute isolated distal DVT of the leg and without severe symptoms or risk factors for extension… we suggest serial imaging of the deep veins for 2 weeks over initial anticoagulation (Grade 2C)."(11) To explore this problem, White et al. (personal communication) undertook a local root-cause analysis of all hospital-acquired VTEs at one academic center that had a relatively high PSI 12 rate. They found that some surgical house staff routinely order venous imaging in all febrile patients because they believe that DVT causes postoperative fever. The hospital's vascular laboratory then routinely scans calf veins and reports the presence of DVT in soleal or gastrocnemius muscular branches, despite evidence that sonography limited to proximal veins is equally safe.(12) Indeed, the American College of Radiology's Appropriateness Criteria for Suspected Lower-Extremity Deep Vein Thrombosis specifically advise radiologists (with the maximum rating of 9) that "the use of this procedure [ultrasound with Doppler] is limited to between the inguinal ligament and knee."(13) The American Academy of Orthopedic Surgeons also recommends

"against routine post-operative duplex ultrasonography screening of patients who undergo elective hip or knee arthroplasty."

In a similar way, pulmonary embolism is now being over-diagnosed because small sub-segmental filling defects are being read as pulmonary emboli (rather than as "small sub-segmental filling defects of undetermined significance", which is a more appropriate term).(14) This problem of overdiagnosis and overtreatment (labeled as "surveillance bias" by some authors) has received increasing attention in the clinical and epidemiologic literature.(15-17) The three key hallmarks of overdiagnosis are: (1) increasing incidence over time; (2) decreasing case fatality over time; and (3) no change in overall attributable mortality over time. All of these hallmarks have been supported with respect to pulmonary emboli; therefore, it seems more accurate to describe this concern as "overdiagnosis bias" rather "surveillance bias."

To address these concerns, CMS has made two important changes to PSI 12 to make it less sensitive to overdiagnosis bias: (1) PSI 12 now captures only proximal (groin/thigh), not distal (calf) vein thromboses; and (2) PSI 12 no longer captures solitary subsegmental pulmonary emboli. With these changes, CMS is now seeing a decreasing temporal trend in PSI 12 rates (down 10.2% from 7/1/2016-6/30/2017 to 7/1/2017-6/30/2018) and no change in case fatality over time. These results provide reassurance that the current specification of PSI 12 is not sensitive to overdiagnosis bias, because it focuses on clinically important events that are consistently diagnosed and treated across all hospitals.

1. Utter GH, Vermoch KL, Rogers S. Clinical documentation improvement and the agency for healthcare research and quality accidental puncture or laceration patient safety indicator. JAMA Surg. 2015;150(5):388-389.
2. http://www.hcpro.com/content/280170.ppt
3. https://www.ahrq.gov/patient-safety/settings/hospital/resource/qitool/index.html
4. http://www.ahrq.gov/professionals/systems/hospital/qitoolkit/qicasestudy.html
5. Rajaram R, Barnard C, Bilimoria KY. Concerns about using the patient safety indicator-90 composite in pay-for-performance programs. JAMA. 2015 Mar 3;313(9):897-8.
6. Chung JW, Ju MH, Kinnier CV, Haut ER, Baker DW, Bilimoria KY. Evaluation of hospital factors associated with hospital postoperative venous thromboembolism imaging utilisation practices. BMJ Qual Saf. 2014;23(11):947-56.
7. Ju MH, Chung JW, Kinnier CV, et al. Association between hospital imaging use and venous thromboembolism events rates based on clinical data. Ann Surg 2014;260(3):558-564.
8. Pierce CA, Haut ER, Kardooni S, et al. Surveillance bias and deep vein thrombosis in the national trauma data bank: the more we look, the more we find. J Trauma 2008;64(4): 932-6.
9. Haut ER, Chang DC, Pierce CA, et al. Predictors of posttraumatic deep vein thrombosis (DVT): hospital practice versus patient factors-an analysis of the National Trauma Data Bank (NTDB). J Trauma 2009;66(4):994-9.
10. http://journal.publications.chestnet.org/data/Journals/CHEST/23443/112299.pdf
11. http://journal.publications.chestnet.org/article.aspx?articleID=1159520
12. Bernardi E, Camporese G, Büller HR, et al. Serial 2-point ultrasonography plus D-dimer vs whole-leg color-coded Doppler ultrasonography for diagnosing suspected symptomatic deep vein thrombosis: a randomized controlled trial. JAMA 2008;300(14): 1653-9.
13. https://acsearch.acr.org/docs/69416/Narrative
14. Suh TM, Cronan JJ, Healey TT. Dots are not clots: the over-diagnosis and over-treatment of PE. Emerg Radiol (2010) 17:347–52.
15. Righini M, Le Gal G, Bounameaux H. Venous thromboembolism diagnosis: unresolved issues. Thromb Haemost 2015;113(6):1184-92.

16. Wiener RS, Schwartz LM, Woloshin S. Time trends in pulmonary embolism in the United States: evidence of overdiagnosis. Arch Intern Med 2011;171(9):831-7.
17. D´Apuzzo MR, Keller TC, Novicoff WM, et al. CT pulmonary angiography after total joint arthroplasty: overdiagnosis and iatrogenic harm? Clin Orthop Relat Res 2013;471(9):2737-42.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

None.

Article III. 5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

No

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):

Are the measure specifications harmonized to the extent possible?

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

Not applicable.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);

OR

Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Not applicable.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested

information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment **Attachment:** Appendix_H-PSI_90_NQF_0531_Conceptual_Framework.pdf

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare and Medicaid Services

Co.2 Point of Contact: Yuling, Li, Yuling.Li@cms.hhs.gov, 410-786-8421-

Co.3 Measure Developer if different from Measure Steward: IMPAQ International

Co.4 Point of Contact: Stacie, Schilling, nqf@impaqint.com, 443-259-5133-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

Technical Expert Panel (TEP):

D'Anna Holmes, MSHA, CPXP, Caregiver Representative

Patient Representative, Florida

Ann Borzecki, MD, MPH, Veterans Health Administration

John Bott, MBA, MSW, Consultant – Healthcare Performance Measurement

Chad Craig, MD, FACP, Weill Medical College of Cornell University

Irene Fraser, PhD, NORC at the University of Chicago

Kathryn Hallock, RHIA, CDIP, Vanderbilt University Medical Center

Sharon Hibay, RN, DNP, Advanced Health Outcomes LLC

Stefanie Ledbetter, RN, BSN, MHI, East Alabama Medical Center and EAMC

Michelle Martin, MBA, VP, HR, CBS Corporation

Amy Rosen, PhD, BA, Boston University School of Medicine

Ilan Rubinfeld, MD, MBA, FACS, FCCP, FCCM, Henry Ford Hospital, Detroit, MI

Bruce Spurlock, MD, Cynosure Health

Eleni Theodoropoulos, BS, CPHIMS, URAC

Julie Wall, RN, MBA, FACMPE, Benefis Health System

Patricia Zrelak, PhD, RN, NEA-bc, SCRN, CNRN, Kaiser Foundation Hospitals

TEP members responded to the posted Call for TEP and provide feedback on clinical acceptability of measure specifications and feasibility of the measure.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2009

Ad.3 Month and Year of most recent revision: 07, 2019

Ad.4 What is your frequency for review/update of this measure? Annually

Ad.5 When is the next scheduled review/update for this measure? 07, 2020

Ad.6 Copyright statement: Limited proprietary coding is contained in the Measure specifications for user convenience. Users of proprietary code sets should obtain all necessary licenses from the owners of the code sets. The ICD-10 is copyrighted by the World Health Organization (WHO) , which owns and publishes the classification. WHO has authorized the development of an adaptation of ICD-10 for use in the United States for U.S. government purposes. As agreed, all modifications to the ICD-10 must conform to WHO conventions for the ICD. All Rights Reserved.

Ad.7 Disclaimers: This measure and specifications are subject to further revisions. This performance measure is not a clinical guideline and does not establish a standard of medical care, and has not been tested for all potential applications. THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Ad.8 Additional Information/Comments: This measure was originally developed, specified, and tested by the Agency for Healthcare Research and Quality (AHRQ) and the responsibility for stewardship of this measure was assumed by the Centers for Medicare & Medicaid Services (CMS) in 2020. IMPAQ International LLC also wishes to recognize our colleagues at University of California at Davis, led by Patrick S. Romano, MD MPH FAAP FACP, who have developed and maintained PSI 90 via subcontract under our Measure & Instrument Development and Support (MIDS) Patient Safety Measure Development and Maintenance contract with CMS.