

UC Davis

UC Davis Previously Published Works

Title

Postoperative respiratory failure: An update on the validity of the Agency for Healthcare Research and Quality Patient Safety Indicator 11 in an era of clinical documentation improvement programs.

Permalink

<https://escholarship.org/uc/item/7zv0v36t>

Journal

American journal of surgery, 220(1)

ISSN

1879-1883

Authors

Stocking, Jacqueline C
Utter, Garth H
Drake, Christiana
et al.

Publication Date

2020-07-14

Peer reviewed



Postoperative respiratory failure: An update on the validity of the Agency for Healthcare Research and Quality Patient Safety Indicator 11 in an era of clinical documentation improvement programs

Jacqueline C. Stocking^{a,*}, Garth H. Utter^{b,c}, Christiana Drake^d, J. Matthew Aldrich^e, Michael K. Ong^{f,g}, Alpesh Amin^h, Rebecca A. Marmorⁱ, Laura Godatⁱ, Maxime Cannesson^j, Michael A. Gropper^e, Patrick S. Romano^{a,c}

^a Department of Internal Medicine, University of California Davis, Sacramento, CA, USA

^b Department of Surgery, Outcomes Research Group, University of California Davis, Sacramento, CA, USA

^c Center for Healthcare Policy and Research, University of California Davis, Sacramento, CA, USA

^d Department of Statistics, University of California Davis, Davis, CA, USA

^e Department of Anesthesia and Perioperative Care, University of California San Francisco, San Francisco, CA, USA

^f Department of Medicine, University of California Los Angeles, Los Angeles, CA, USA

^g VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA

^h Department of Hospital Medicine, University of California Irvine, Irvine, CA, USA

ⁱ Department of Surgery, University of California San Diego, San Diego, CA, USA

^j Department of Anesthesiology and Perioperative Medicine, University of California Los Angeles, Los Angeles, CA, USA

ARTICLE INFO

Article history:

Received 29 July 2019

Received in revised form

28 October 2019

Accepted 10 November 2019

Keywords:

postoperative respiratory failure

positive predictive value

validity

Agency for Healthcare Research and Quality (AHRQ)

Patient Safety Indicator 11 (PSI 11)

ABSTRACT

Background: Administrative data can be used to identify cases of postoperative respiratory failure (PRF). We aimed to determine if recent changes to the Agency for Healthcare Research and Quality Patient Safety Indicator 11 (PSI 11) and adoption of clinical documentation improvement programs have improved the validity of PSI 11. We also analyzed reasons why PSI 11 was falsely triggered.

Study design: Cross-sectional study of all eligible discharges using health record data from five academic medical centers between October 1, 2012 and September 30, 2015.

Results: Of 437 flagged records, 434 (99.3%) were accurately coded and 414 (94.7%) represented true clinical PRF. None of the false positive records involved respiratory failure present on admission. Most (78.3%) false positive records required airway protection but did not have respiratory failure.

Conclusion: The validity of PSI 11 has improved with recent changes to the code criterion and adoption of clinical documentation improvement programs.

© 2019 Elsevier Inc. All rights reserved.

Introduction

Postoperative respiratory failure (PRF)—defined variously as unplanned reintubation, mechanical ventilation beyond 48 (or 96)

hours after surgery, or inadequate oxygenation or ventilation—is the most common serious pulmonary complication, with an incidence of 0.2–7.5%^{1–11} and attributable in-hospital mortality of 25–40%.^{1,12} Each case is associated with approximately \$53,000 in excess charges and nine excess days of hospitalization after adjusting for preoperative risk factors,¹³ making it among the most expensive and deadly of postoperative complications.

One method to identify cases of PRF is through the Agency for Healthcare Research and Quality (AHRQ) Patient Safety Indicator (PSI) 11. PSI 11 is efficient because it utilizes administrative data to identify potential instances of PRF for the purpose of informing targeted institution-level quality improvement efforts.¹⁴ However, its validity has been constrained by the accuracy of clinical

Abbreviations: Agency for Healthcare Research and Quality, (AHRQ); Institutional Review Board, (IRB); International Classification of Diseases 9th Revision Clinical Modification, (ICD-9-CM); Interquartile Range, (IQR); Major Diagnostic Category, (MDC); Patient Safety Indicator, (PSI); Positive Predictive Value, (PPV); Postoperative Respiratory Failure, (PRF); Present-on-Admission, (POA); Veteran's Health Administration, (VHA).

* Corresponding author. Department of Internal Medicine, University of California Davis Health, 4150 V Street, Suite 3400 Sacramento, CA, 95817, USA.

E-mail address: jcstocking@ucdavis.edu (J.C. Stocking).

documentation by physicians, chart review by coders (and, increasingly, clinical documentation specialists),¹⁵ and available International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnosis and procedure codes.¹⁶

Researchers have validated PSI 11 under ICD-9-CM specifications in academic centers and the Veteran's Health Administration using chart review as the criterion standard and found the positive predictive value (PPV) was 67–83%.^{15,17–20} Subsequently, the AHRQ took advantage of present-on-admission (POA) flags starting in 2008 to exclude preoperative respiratory failure and revised the diagnosis code criteria in 2011 to better focus PSI 11 on cases following trauma or surgery. Contemporaneously, hospitals began verifying PSI-flagged events through clinical documentation improvement programs, which involve concurrent communication between physicians, clinical documentation specialists, and coders.^{21–24} Our primary objective was to study the effect of the new ICD-9-CM diagnosis codes and the POA flag—in the context of clinical documentation improvement programs—on the ability of PSI 11 to accurately identify PRF. Our secondary objective was to analyze records that triggered PSI 11 but that did not represent PRF to understand how PSI 11 might be further refined.

Methods

This study was approved by the Institutional Review Board (IRB) of the University of California, Davis as the lead site and endorsed through the IRB Reliance process by the IRBs at the University of California Irvine, Los Angeles, San Diego, and San Francisco.

Study design and population

This was a multisite cross-sectional health record-based study of all eligible discharges from October 1, 2012 through September 30, 2015 that met criteria for AHRQ PSI 11 at the five University of California academic medical centers (San Francisco, Davis, Irvine, Los Angeles, and San Diego campuses). Participation by each center was voluntary. Eligible discharges are defined by AHRQ PSI 11 technical specifications²⁵ (Table 1). We chose the start date based on: 1) the availability of data from all sites (as well as allowance of a

phase-in period for coders to adjust to the updated ICD-9-CM diagnosis codes relevant to PSI 11, first in use in October 2011), and 2) the presence of a clinical documentation improvement program at each site. The end date was chosen based on the October 1, 2015 implementation of International Classification of Diseases, 10th Revision, Clinical Modification/Procedure Classification System in the United States.

Instrument development

We modified an abstraction instrument from a prior study¹⁵ for use via REDCap™. Modifications occurred iteratively through review and pilot testing by the research team. The final instrument included information on demographic characteristics, diagnoses and procedures, length of stay, discharge disposition, preexisting comorbid conditions, preoperative laboratory and radiographic test results, operating room treatment, postoperative care, and open-ended text fields to allow abstractors to provide additional clarification or to pose questions for further investigation (see supplementary material/electronic appendix).

Data collection

We identified records at participating hospitals through application of the AHRQ PSI software to the Vizient™ (formerly University Healthsystem Consortium) Clinical Data Base for the timeframe of interest. The two data sources utilized were administrative data from the Vizient™ Clinical Data Base and the paper and electronic health records of each participating hospital.

Five data abstractors (AA, EW, BM, RK, SB) were trained through a combination of written training materials, teleconferences, and regular in-person meetings. The principal investigator (JS), a critical care registered nurse with over twenty years of experience, supervised the training and reviewed all abstracted data for accuracy. We did not formally assess inter-rater reliability.

Analysis

The primary clinical team (PS, GU, JS) used the AHRQ PSI 11

Table 1
Postoperative respiratory failure PSI 11 ICD-9-CM technical specifications.²⁵

Definition	Postoperative respiratory failure (secondary diagnosis), prolonged mechanical ventilation, or reintubation cases per 1000 elective surgical discharges for patients ages 18 years and older.
Numerator	Discharges, among cases meeting the inclusion and exclusion rules for the denominator, with either: <ul style="list-style-type: none"> • any secondary ICD-9-CM diagnosis code for acute (518.51) or acute and chronic (518.53) respiratory failure following trauma and surgery; or • any secondary ICD-9-CM procedure codes for mechanical ventilation for 96 consecutive hours or more (96.72) that occurs zero or more days after the first major operating room procedure code; or • any secondary ICD-9-CM procedure codes for mechanical ventilation for less than 96 consecutive hours (96.71) or undetermined duration (96.70) that occurs two or more days after the first major operating room procedure code; or • any secondary ICD-9-CM procedure codes for insertion of an endotracheal tube (96.04) that occurs one or more days after the first major operating room procedure code.
Denominator	Elective surgical discharges, for patients ages 18 years and older, with any-listed ICD-9-CM procedure codes for a major operating room procedure. Exclude cases: <ul style="list-style-type: none"> • with a principal ICD-9-CM diagnosis code (or secondary diagnosis present on admission) for acute respiratory failure • where the only operating room procedure is tracheostomy • where a procedure for tracheostomy occurs before the first operating room procedure • with any-listed ICD-9-CM diagnosis codes for neuromuscular disorder • with any-listed ICD-9-CM procedure codes for laryngeal or pharyngeal, nose, mouth, pharynx or facial surgery • with any-listed ICD-9-CM procedure codes for esophageal resection • with any-listed ICD-9-CM procedure codes for lung cancer • with any-listed ICD-9-CM diagnosis codes for degenerative neurological disorder • with any-listed ICD-9-CM procedure codes for lung transplant • with MDC 4 (diseases/disorders of respiratory system) • with MDC 5 (diseases/disorders of circulatory system) • with MDC 14 (pregnancy, childbirth, and puerperium) • with missing gender, age, quarter, year, or principal diagnosis

Full AHRQ PSI 11 technical specifications for ICD-9-CM available at: <https://www.qualityindicators.ahrq.gov/Archive/default.aspx>.

criteria, the ICD-9-CM coding data, and the clinical records to categorize each hospitalization as either a true or false positive from both a coding and a clinical perspective. This was a two-step process. The principal investigator first validated the data abstraction of each chart and made an initial determination of each hospitalization as a true or false positive. The three primary clinical investigators then met in person to review each case and determine, by consensus, the final assignment of true or false positive status. This allowed us to generate two estimates of PPV ($[\text{true positive}/\text{true positive} + \text{false positive}] \times 100$) and the associated 95% confidence interval (CI): 1) “coding PPV,” i.e., did coders apply the appropriate ICD-9-CM code(s) to each flagged record, and 2) “clinical PPV,” i.e., did the record represent true PRF from a clinical perspective. From the clinical perspective, we also assessed the “marginal PPV”²⁶ of specific PSI 11 numerator criteria (single diagnosis or procedure codes) among those records flagged only because of that particular criterion of the numerator.

We classified records as false positive by coding criteria if they had been assigned an incorrect ICD-9-CM code, such as a 518.51 code for acute respiratory failure following trauma or surgery in the absence of any physician documentation to support assignment of this code. We classified records as false positive by clinical criteria if they failed to represent PRF from a purely clinical perspective, such as reintubation and/or prolonged ventilator management with a documented rationale of airway protection, not respiratory failure.

We compared characteristics of true and false positive records and evaluated the reasons behind false positive records. Among records flagged by PSI 11 based on only a single diagnosis or procedure code criterion, we calculated the marginal PPV of each criterion. All analyses were done using Stata SE® version 14.2 (Stata Corp). We compared the true and false positive groups using Student's *t*-test or the Wilcoxon Mann-Whitney test for continuous variables, and Pearson's chi square test or Fisher's exact test for categorical variables. Significance was set a priori at $p < 0.05$.

Results

There were a total of 437 PSI 11-flagged records among 59,073 eligible discharges from the five sites. The overall rate of PRF was 7.4 records per 1000 eligible discharges, with a range of 5.5–11.4. Of the 437 records, 126 (28.8%) met diagnosis and procedure code criteria, 217 (49.7%) met only procedure code criteria, and 94 (21.5%) met only diagnosis code criteria (Table 2). Among all flagged records, 434 (99.3%; 95% CI, 98.0–99.9%) met the ICD-9-CM coding criteria (“coding validity”) and 414 (94.7%; 95% CI, 92.2–96.6%) met clinical criteria (“clinical validity”) for PRF.

A total of 834 PSI 11 numerator diagnosis and procedure codes were assigned to the 437 records (mean 1.91, range 1–4), with

diagnosis code 518.51 and procedure code 96.04 being the most frequently assigned (Table 3).

Characteristics of flagged records

Patients whose records were flagged had a median age of 62 years and were more likely to be female (51.5%) and white (70.7%) (Table 4). The median hospital length of stay was 15 days (interquartile range [IQR] 8, 29 days) and the median intensive care unit length of stay was 4 days (IQR 1, 13 days). The diagnosis or procedure triggering PSI 11 was documented a median of 2 days (IQR 0, 4 days) postoperatively. The highest percentage of flagged records ($n = 152$, 34.8%) were identified on the day of surgery, with most records identified during post-operative days 0–5 ($n = 340$, 77.8%) and almost all by post-operative day 15 ($n = 413$, 94.5%). Seventy-five patients (17.2%) died during their hospitalization.

Flagged records had a median of 2 comorbidities (IQR 1, 4 and range 0–8). Only 37 (8.5%) flagged records did not have any comorbid conditions. A majority (72.3%) of flagged records had an American Society of Anesthesiologists class of III (moderate to severe systemic disease) (Table 4). The most common comorbid conditions were hypertension (54.7%), gastroesophageal reflux disease (30%), and cardiovascular disease (24%) (Table 4). Most patients (92.9%) were functionally independent with activities of daily living prior to admission.

Among all flagged records, the median duration of surgery (initial incision to closure) was 284 min (IQR 176, 443 min) (Table 4). The median time under anesthesia care was 411 min (IQR 271, 561 min). General anesthesia was used for 428 patients (98%), with 399 (91.3%) receiving neuromuscular blockade. The most common neuromuscular blocking agent used was rocuronium ($n = 270$, 67.7%), followed by succinylcholine ($n = 65$, 16.3%). Fifty-eight patients (13.4%) received more than one neuromuscular blocking agent, with succinylcholine followed by rocuronium being the most prevalent combination ($n = 23$). Of the patients who received general anesthesia, 29 (7.3%) also received local or regional analgesia. Surgical procedures involving the abdomen or pelvis were most common (46%), followed by surgical procedures on the head or soft tissue of the neck (16.7%). Most patients received intravenous opioid analgesics (95.7%) and benzodiazepines (67.7%) intraoperatively.

Coding validity

The three records that were false positives from a coding perspective were all flagged solely due to diagnosis code 518.51 (acute respiratory failure following trauma and surgery). All three records lacked objective clinical criteria or explicit physician

Table 2
PSI 11 diagnosis and procedure code criteria for the 437 total records flagged positive by the indicator.

PSI 11 Numerator Criteria, n (%)	All Flagged Records (n = 437)	True Positives Coding (n = 434)	False Positives Coding (n = 3)	True Positives Clinical (n = 414)	False Positives Clinical (n = 23)
Diagnosis code only (518.51, 518.53)	94 (21.5)	91 (21.0)	3 (100)	88 (21.3)	6 (26.1)
Procedure code only (96.72, 96.71, 96.70, 96.04)	217 (49.7)	217 (50.0)	0 (0)	201 (48.6)	16 (69.6)
Diagnosis code (518.51, 518.53) and Procedure code (96.72, 96.71, 96.70, 96.04)	126 (28.8)	126 (29.0)	0 (0)	125 (30.2)	1 (4.3)
Total	437	434	3	414	23

Key.

518.51: Acute Respiratory Failure Following Trauma and Surgery.

518.53: Acute and Chronic Respiratory Failure Following Trauma and Surgery.

96.70: Continuous Mechanical Ventilation of Unspecified Duration.

97.71: Continuous Mechanical Ventilation for Less Than 96 Consecutive Hours.

96.72: Continuous Mechanical Ventilation for 96 Consecutive Hours or More.

96.04: Insertion of Endotracheal Tube.

Table 3
Contributions of different combinations of procedure and diagnosis code criteria for the 437 flagged records*.

Total n (True Positive, False Positive)	Any Diagnosis Code Present (518.51 or 518.53)	518.51	518.53	518.51 and 518.53	No Diagnosis Code	Total
Any Procedure Code Present (96.04, 96.72, 96.71, or 96.70)	126 (125,1)	125 (124,1)	1 (1,0)	0 (0,0)	217 (201,16)	343 (326,17)
96.04	5 (5,0)	5 (5,0)	0 (0,0)	0 (0,0)	5 (4,1)	10 (9,1)
96.70	0 (0,0)	—	—	—	0 (0,0)	0 (0,0)
96.71	23 (22,1)	23 (22,1)	0 (0,0)	0 (0,0)	13 (10,3)	36 (32,4)
96.72	7 (7,0)	7 (7,0)	0 (0,0)	0 (0,0)	27 (23,4)	34 (30,4)
96.04 and 96.70	0 (0,0)	—	—	—	0 (0,0)	0 (0,0)
96.04 and 96.71	50 (50,0)	50 (50,0)	0 (0,0)	0 (0,0)	81 (75,6)	131 (125,6)
96.04 and 96.72	38 (38,0)	37 (37,0)	1 (1,0)	0 (0,0)	85 (83,2)	123 (121,2)
96.70 and 96.71	0 (0,0)	—	—	—	0 (0,0)	0 (0,0)
96.70 and 96.72	0 (0,0)	—	—	—	0 (0,0)	0 (0,0)
96.71 and 96.72	0 (0,0)	—	—	—	1 (1,0)	1 (1,0)
96.04, 96.70, and 96.71	0 (0,0)	—	—	—	0 (0,0)	0 (0,0)
96.04, 96.70, and 96.72	0 (0,0)	—	—	—	1 (1,0)	1 (1,0)
96.04, 96.71, and 96.72	3 (3,0)	3 (3,0)	0 (0,0)	0 (0,0)	4 (4,0)	7 (7,0)
96.70, 96.71, and 96.72	0 (0,0)	—	—	—	0 (0,0)	0 (0,0)
96.04, 96.70, 96.71, and 96.72	0 (0,0)	—	—	—	0 (0,0)	0 (0,0)
No Procedure Code Present	94 (88,6)	91 (87,4)	3 (1,2)	0 (0,0)	0 (0,0)	94 (88,6)
Total	220 (213,7)	216 (211,5)	4 (2,2)	0 (0,0)	217 (201,16)	437 (414, 23)

Records were considered false positive if they failed to meet either coding or clinical criteria.

The total n is provided, with the number of true positives and false positives provided in parenthesis: [e.g. "125 (124,1)" where n = 125, true positives = 124, false positives = 1].

Key.

518.51: Acute Respiratory Failure Following Trauma and Surgery.

518.53: Acute and Chronic Respiratory Failure Following Trauma and Surgery.

96.70: Continuous Mechanical Ventilation of Unspecified Duration.

97.71: Continuous Mechanical Ventilation for Less Than 96 Consecutive Hours.

96.72: Continuous Mechanical Ventilation for 96 Consecutive Hours or More.

96.04: Insertion of Endotracheal Tube.

documentation to support the diagnosis.

Clinical validity

The 23 records (5.3%) that were false positives from a clinical perspective involved: intubation and/or mechanical ventilation for airway protection, not respiratory failure (e.g., high cervical spine laminectomy with concern for airway compromise due to swelling) (n = 18, 78.3%); and documentation and/or coding errors where there was no explicit clinical documentation to support a diagnosis and subsequent coding of respiratory failure (e.g., diagnosis of acute respiratory failure in a consult service note with no other supporting information) (n = 5, 21.7%).

Differences between true positive and false positive clinical records

Records with evidence of true clinical PRF had longer hospital and intensive care unit lengths of stay and higher documented operative peak inspiratory pressures (Table 4). True positive records were more likely to involve an abdominal/pelvic operation, while false positive records were more likely to involve a head/soft tissue neck procedure (Table 4). Of 340 true positive patients who survived to discharge, only 22% were functionally independent. Discharge dispositions for survivors included: home with home health care (n = 95, 27.9%), skilled nursing facility (n = 71, 20.9%), rehabilitation facility (n = 54, 15.9%), long term care facility (n = 13, 3.8%), another hospital (n = 9, 2.6%), hospice (n = 2, 0.6%), and other facilities (n = 21, 6.2%). Only 22% of survivors were discharged home under self-care.

Marginal positive predictive value

Of the 437 flagged records, 139 (31.8%) met criteria for only a

single diagnosis or procedure code criterion of the PSI 11 numerator. Of these 139 records, the code was properly assigned in 127 (91.4%) and incorrectly assigned in 12 (8.6%) (Table 5). The most frequently utilized sole criteria were 518.51 (acute respiratory failure following trauma and surgery) (n = 91, marginal PPV = 92.3%), followed by 96.72 (continuous mechanical ventilation for 96 consecutive hours or more) (n = 27, marginal PPV = 92.5%). Ignoring the infrequently used 518.53 (n = 3) and 96.70 (n = 0) codes, the marginal PPV of all remaining criteria was at least 84%.

Discussion

We found the addition of a POA flag in 2008 and revisions to the diagnosis code criteria in 2011 increased the clinical and coding PPV of AHRQ PSI 11 when compared to available data from two previous studies that used chart review as the criterion standard.^{15,19} We also found increased marginal PPV of PSI 11 when compared to a previous study.¹⁵ None of the flagged records in our dataset had respiratory failure present on admission, which indicates the POA flag functions well. These changes occurred during a timeframe when hospitals were implementing clinical documentation improvement programs.

Many clinical false positive records involved mechanical ventilation for airway protection, consistent with prior studies.^{15,19} While these records did not represent PRF, they might still represent opportunities to improve or standardize care. Many patients had multiple non-modifiable risk factors such as advanced age, preexisting cardiopulmonary disease, and complex, lengthy, open torso procedures. In this population, there may be little surgeons and anesthesiologists can do to mitigate risk, other than modify patient selection,¹⁵ procedure type, and anesthesia techniques – neuraxial anesthesia and lung protective ventilation – that may

Table 4
Characteristics of records flagged positive by AHRQ PSI 11.

Characteristic	All Flagged Records 437 (100%)	True Positives Clinical 414 (94.7%)	False Positives Clinical 23 (5.3%)	p Value
Age (y), median (IQR)	62.0 (52.0–70.0)	62.5 (52.0–70.0)	56.0 (51.0–70.0)	0.36
Male sex, n (%)	212 (48.5)	203 (49)	9 (39.1)	0.36
Race, n (%)				0.83
White	309 (70.7)	294 (71)	15 (65.2)	
Black	21 (4.8)	20 (4.8)	1 (4.4)	
Asian	42 (9.6)	39 (9.4)	3 (13)	
Other	65 (14.9)	61 (14.7)	4 (17.4)	
Hospital, n (%)				0.07
Site 1	64 (14.7)	57 (89.1)	7 (10.9)	
Site 2	125 (28.6)	117 (93.6)	8 (6.4)	
Site 3	86 (19.7)	85 (98.8)	1 (1.2)	
Site 4	84 (19.2)	79 (94.0)	5 (6.0)	
Site 5	78 (17.9)	76 (97.4)	2 (2.6)	
Hospital length of stay, days, median (IQR)	15.0 (8.0–29.0)	15.0 (8.0–30.0)	10.0 (6.0–19.0)	0.01
ICU length of stay, days, median (IQR)	4.0 (1.0–13.0)	4.0 (1.0–14.0)	3.0 (0.0–6.0)	0.01
Postoperative day PRF criterion occurred, median (IQR)	1.0 (0.0–4.0)	1.0 (0.0–5.0)	1.0 (0.0–3.0)	0.1
Body mass index, median (IQR)	27.4 (23.4–32.3)	27.4 (23.5–32.4)	26.9 (23.0–30.4)	0.71
Duration of anesthesia care, minutes, median (IQR)	411.0 (271.0–561.0)	405.5 (268.0–554.5)	464.0 (334.0–647.0)	0.13
Duration of surgical procedure, minutes, median (IQR)	284.0 (176.0–443.0)	278.0 (171.5–441.5)	362.0 (207.0–520.0)	0.19
Intraoperative				
Maximum positive end expiratory pressure (PEEP), cm H ₂ O, median (IQR)	6.0 (5.0–8.0)	6.0 (5.0–8.0)	5.0 (5.0–7.0)	0.39
Maximum peak inspiratory pressure (PIP), cm H ₂ O, median (IQR)	26.0 (22.0–32.0)	26.0 (22.0–32.0)	23.0 (20.0–28.0)	0.02
Maximum tidal volume, mL/kg ideal body weight, median (IQR)	9.33 (8.01–10.70)	9.30 (7.97–10.63)	9.54 (8.97–11.29)	0.12
Morphine equivalent dose, mg/kg ideal body weight, median (IQR)	0.36 (0.23–0.66)	0.36 (0.23–0.66)	0.41 (0.31–0.67)	0.45
Benzodiazepine equivalent dose, mg/kg ideal body weight, median (IQR)	0.11 (0.00–0.15)	0.11 (0.00–0.15)	0.12 (0.05–0.17)	0.19
Net intraoperative fluid, mL/kg ideal body weight, median (IQR)	34.73 (16.14–64.81)	33.73 (16.42–65.37)	38.17 (14.61–49.66)	0.57
Other				
Net fluid during the first 24-postoperative hours, mL/kg ideal body weight, median (IQR)	32.88 (11.46–64.55)	33.03 (11.23–67.15)	25.67 (12.88–56.28)	0.79
Net fluid given intraoperatively plus the first 24-postoperative hours, mL/kg ideal body weight, median (IQR)	67.10 (34.04–121.59)	67.51 (34.93–121.90)	59.29 (30.06–102.09)	0.58
Comorbid Conditions				
American Society of Anesthesiologists (ASA) Class, n (%)				0.56
I	0 (0)	0 (0)	0 (0)	
II	74 (16.9)	68 (16.4)	6 (26.1)	
III	316 (72.3)	301 (72.7)	15 (65.2)	
IV	44 (10.1)	42 (10.1)	2 (8.7)	
V	1 (0.2)	1 (0.2)	0 (0)	
Not assessed	2 (0.5)	2 (0.5)	0 (0)	
Cardiovascular disease	105 (24.0)	99 (23.9)	6 (26.1)	0.81
Chronic kidney disease	64 (14.7)	62 (15.0)	2 (8.7)	0.55
Chronic obstructive pulmonary disease, n (%)	41 (9.4)	38 (9.2)	3 (13)	0.37
Current smoker, n (%)	41 (9.4)	37 (8.9)	4 (17.4)	0.16
Diabetes mellitus, n (%)	97 (22.2)	90 (21.7)	7 (30.4)	0.33
Functional status, n (%)				0.09
Independent	406 (92.9)	386 (93.2)	20 (87)	
Partially dependent	23 (5.3)	22 (5.3)	1 (4.4)	
Totally dependent	8 (1.8)	6 (1.5)	2 (8.7)	
Gastroesophageal reflux disease (GERD), n (%)	131 (30)	127 (30.7)	4 (17.4)	0.13
Heart failure, n (%)	15 (3.4)	14 (3.4)	1 (4.4)	0.54
Hypertension, n (%)	239 (54.7)	228 (55.1)	11 (47.8)	0.50
Obstructive sleep apnea, n (%)	56 (12.8)	55 (13.3)	1 (4.4)	0.18
Anatomic region of surgical procedure ^a , n (%)				0.003
Head/Soft Tissue Neck	73 (16.7)	65 (15.7)	8 (34.8)	
Cervical Spine	15 (3.4)	14 (3.4)	1 (4.4)	
Chest	5 (1.1)	4 (1)	1 (4.4)	
Abdomen/Pelvis	201 (46)	198 (47.8)	3 (13)	
Thoracolumbar Spine	35 (8)	32 (7.7)	3 (13)	
Extremity	19 (4.4)	17 (4.1)	2 (8.7)	
Percutaneous/Endoscopic	45 (10.3)	44 (10.6)	1 (4.4)	
Unable to Determine ^b	44 (10.1)	40 (9.7)	4 (17.4)	
Death (in-hospital), n (%)	75 (17.2)	74 (17.9)	1 (4.4)	0.07

^a All are open approach surgical procedures categorized by anatomical region with the exception of the “Percutaneous/Endoscopic” category, which includes all non- and minimally-invasive surgical approaches regardless of anatomical region.

^b This categorization is based solely on ICD-9-CM procedure codes and, as such, the anatomical location could not be determined in a small percentage of records.

reduce harm.

Many clinical true positive records involved procedures on the abdominal and pelvic cavities. This is not surprising due to the anatomic proximity of the muscles of respiration, the likelihood of increased intra-abdominal pressure from postoperative ileus, and the potentially long duration of complex abdominal procedures.

Just as the factors contributing to acute respiratory failure may differ depending on the anatomic region of the index operation, the criterion validity of PSI 11 may also vary depending on the anatomic region.

Although all five sites had a clinical documentation improvement program in place during the study period, some false positive

Table 5

Marginal positive predictive values among the 139 records flagged on the basis of a single diagnosis or procedure code criterion.

Code Criterion	True Positive (n)	False Positive (n)	Marginal PPV ^a (%)
518.51	84	7	92.3
518.53	2	1	66.7
96.04	5	0	100
96.70	0	0	—
96.71	11	2	84.6
96.72	25	2	92.5

Key.

518.51: Acute Respiratory Failure Following Trauma and Surgery.

518.53: Acute and Chronic Respiratory Failure Following Trauma and Surgery.

96.70: Continuous Mechanical Ventilation of Unspecified Duration.

97.71: Continuous Mechanical Ventilation for Less Than 96 Consecutive Hours.

96.72: Continuous Mechanical Ventilation for 96 Consecutive Hours or More.

96.04: Insertion of Endotracheal Tube.

^a The PPV among records flagged solely on the basis of that code criterion.

records were due to documentation and coding errors. The AHRQ provides a toolkit²⁷ with a section on clinical documentation improvement programs, to include physician documentation, concurrent review by clinical documentation specialists, a standardized query process when documentation is unclear, and subsequent coding. Implementation of such programs in departments of surgery has coincided with improved accuracy of documentation,²² better compliance with surgical quality measures,²³ and a reduction in the false positive rate for PSI 11 specifically.²⁴ However, these programs require resources and can burden physicians with additional queries. Because such programs primarily focus on reviewing PSI-flagged records and narrowing the definitions of PSI numerator-relevant diagnoses, they may also increase false negative rates. With increased emphasis on pay for performance, clinical documentation improvement programs are poised to have an ongoing impact on PSI 11 rates.

Analysis of the marginal PPV provides additional insight into the validity of PSI 11. About one-fifth of records in our study were flagged solely based on a diagnosis code, and these diagnoses seemed particularly sensitive to differences in documentation and coding practices. Contributing factors include the lack of a consensus definition of acute respiratory failure,²⁴ disagreement on whether the PRF is a result of surgery, and inconsistent documentation, querying, and coding practices. While it is possible to eliminate the diagnosis codes from the PSI 11 numerator, doing so would only minimally impact the overall PPV because the marginal PPV of the diagnosis codes (86/94 = 91.4%) was only slightly less than the PPV of the remaining records flagged by at least one procedure code (326/343 = 95.0%).

Our data suggest the PPV of PSI 11 could be improved. Additional surgical procedures involving the head or soft tissue of the neck could be excluded as, in our sample, these patients were more likely to require endotracheal intubation and prolonged mechanical ventilation for airway protection rather than for respiratory failure. Such exclusions would need to be carefully selected as they would categorically eliminate from the denominator patients undergoing certain procedures who might otherwise be at high risk for postoperative respiratory failure. The additional available codes and new classification approach within the International Classification of Diseases, 10th Revision, Clinical Modification and Procedure Coding System might allow for more precise exclusions.²⁸ While we did not identify any records that had PRF present on admission, we did find two flagged records in which the patients had a tracheostomy in place and were chronically mechanically ventilated and were admitted for elective surgical revisions of prior traumatic neurologic insults. Excluding such patients might also improve the PPV of PSI 11.

Our study has limitations. All five sites are tertiary academic medical centers located within a single healthcare system in California. As such, our findings may not be generalizable. Differences in electronic health record platforms across the five centers may have led to some misclassification of collected data; however, we verified all records by a second abstraction process and had three members of our team adjudicate the findings. We also lacked the resources to collect data on an appropriate sample of the over 59,000 encounters during the study period that were not flagged by PSI 11; thus, we were unable to estimate the sensitivity of PSI 11.

Conclusions

The positive predictive value of PSI 11 has improved since the 2008 addition of a present-on-admission flag and 2011 changes to the diagnosis codes used in its numerator. Clinical documentation improvement programs may have contributed to this improvement, but the impact on the sensitivity of PSI 11 remains unknown. Further refinements to the PSI 11 specifications might focus on exclusion criteria applicable to patients who experience prolonged mechanical ventilation primarily for airway protection rather than for respiratory failure.

Funding

This work was supported by: 1) The Betty Irene Moore School of Nursing (doctoral scholarship - PI); 2) The Board of Certification for Emergency Nursing, Flight and Transport Nurse Scholarship for Air & Surface Transport Nurse (ASTNA) Members (doctoral scholarship - PI); 3) The American Association of Critical-Care Nurses (AACN) (Impact Research Grant, # 20297); 4) The University of California Davis Clinical and Translational Science Center (CTSC) support for the Research Electronic Data Capture (REDCapTM) database (National Center for Advancing Translational Sciences [NCATS, National Institutes of Health [NIH], (grant UL1 TR000002). The funding sources had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Acknowledgements

The author(s) wish to thank Anna Aledia, Esther Wang, Benjamin Mooso, Rebecca Kim, and Sabrina Berci for their assistance with data abstraction. We would also like to acknowledge Holly Thurston, PhD, MSW for her consultation with data cleaning, coding, and preparation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.amjsurg.2019.11.019>.

References

1. Arozullah AM, et al. Multifactorial risk index for predicting postoperative respiratory failure in men after major noncardiac surgery. The National Veterans Administration Surgical Quality Improvement Program. *Ann Surg*. 2000;232(2): 242–253.
2. Blum JM, et al. Preoperative and intraoperative predictors of postoperative acute respiratory distress syndrome in a general surgical population. *Anesthesiology*. 2013;118(1), 10.1097/ALN.0b013e3182794975.
3. Brueckmann B, et al. Development and validation of a score for prediction of postoperative respiratory complications. *Anesthesiology*. 2013;118(6): 1276–1285.
4. Canet J, et al. Development and validation of a score to predict postoperative respiratory failure in a multicentre European cohort: a prospective, observational study. *Eur J Anaesthesiol*. 2015;32(7):458–470.
5. Gupta H, et al. Development and validation of a risk calculator predicting

- postoperative respiratory failure. *Chest*. 2011;140(5):1207–1215.
6. Hua M, Brady JE, Li G. A scoring system to predict unplanned intubation in patients having undergone major surgical procedures. *Anesth Analg*. 2012;115(1):88–94.
 7. Johnson RG, et al. Multivariable predictors of postoperative respiratory failure after general and vascular surgery: results from the patient safety in surgery study. *J Am Coll Surg*. 2007;204(6):1188–1198.
 8. Johnson AP, et al. Predicting the risk of postoperative respiratory failure in elective abdominal and vascular operations using the national surgical quality improvement program (NSQIP) participant use data file. *Ann Surg*. 2016;266(6):968–974.
 9. Kor DJ, et al. Derivation and diagnostic accuracy of the surgical lung injury prediction model. *Anesthesiology*. 2011;115(1):117–128.
 10. Ramachandran SK, et al. Independent predictors and outcomes of unanticipated early postoperative tracheal intubation after nonemergent, noncardiac surgery. *Anesthesiology*. 2011;115(1):44–53.
 11. Kor DJ, et al. Predicting risk of postoperative lung injury in high-risk surgical patients: a multicenter cohort study. *Anesthesiology*. 2014;120(5):1168–1181.
 12. Canet J, Gallart L. Postoperative respiratory failure: pathogenesis, prediction, and prevention. *Curr Opin Crit Care*. 2014;20(1):56–62.
 13. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *Jama*. 2003;290(14):1868–1874.
 14. Miller MR, et al. Patient Safety Indicators: using administrative data to identify potential patient safety concerns. *Health Serv Res*. 2001;36(6 Pt 2):110–132.
 15. Utter GH, et al. Detection of postoperative respiratory failure: how predictive is the agency for healthcare research and quality's patient safety indicator? *J Am Coll Surg*. 2010;211(3):347–354. e1–29.
 16. Nguyen MC, et al. Agency for healthcare research and quality (AHRQ) patient safety indicator for postoperative respiratory failure (PSI 11) does not identify accurately patients who received unsafe care. *Surgery*. 2016;160(4):858–868.
 17. Romano PS, et al. Validity of selected AHRQ patient safety indicators based on VA National Surgical Quality Improvement Program data. *Health Serv Res*. 2009;44(1):182–204.
 18. Kaafarani HMA, Rosen AK. Using administrative data to identify surgical adverse events: an introduction to the Patient Safety Indicators. *Am J Surg*. 2009;198(5, Supplement):S63–S68.
 19. Borzecki AM, et al. How valid is the AHRQ Patient Safety Indicator “postoperative respiratory failure”? *J Am Coll Surg*. 2011;212(6):935–945.
 20. Rosen AK, et al. Validating the patient safety indicators in the Veterans Health Administration: do they accurately identify true safety events? *Med Care*. 2012;50(1):74–85.
 21. Fox N, et al. A critical review of patient safety indicators attributed to trauma surgeons. *Injury*. 2017;48(9):1994–1998.
 22. Fox N, et al. Lost in translation: focused documentation improvement benefits trauma surgeons. *Injury*. 2016;47(9):1919–1923.
 23. Reyes C, et al. Implementation of a clinical documentation improvement curriculum improves quality metrics and hospital charges in an academic surgery department. *J Am Coll Surg*. 2017;224(3):301–309.
 24. Najjar P, et al. A multidisciplinary three-phase approach to improve the clinical utility of patient safety indicators. *Qual Manag Health Care*. 2015;24(2):62–68.
 25. Agency for Healthcare Research and Quality. *AHRQ QIP™ Research Version 5.0, Patient Safety Indicators 11, Technical Specifications, Postoperative Respiratory Failure Rate*. Rockville, MD: Agency for Healthcare Research and Quality; 2015.
 26. Pepe MS. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. , New York: Oxford University Press; 2003.
 27. Agency for Healthcare Research and Quality. *Toolkit for using the AHRQ quality indicators*. Available from <http://www.ahrq.gov/professionals/systems/hospital/qitoolkit/index.html>; 2017.
 28. Utter GH, et al. Challenges and opportunities with ICD-10-CM/PCS: implications for surgical research involving administrative data. *J Am Coll Surg*. 2013;217(3):516–526.