Detection of Postoperative Respiratory Failure: How Predictive Is the Agency for Healthcare Research and Quality's Patient Safety Indicator?

Garth H Utter, MD, MSC, FACS, Joanne Cuny, RN, MBA, Pradeep Sama, MBA, Michael R Silver, MD, Patricia A Zrelak, PhD, RN, Ruth Baron, RN, Saskia E Drösler, MD, Patrick S Romano, MD, MPH

BACKGROUND: Patient Safety Indicator (PSI) 11, or postoperative respiratory failure, was developed by the US Agency for Healthcare Research and Quality to detect incident cases of respiratory failure after elective operations through use of ICD-9-CM diagnosis and procedure codes. We sought to determine the positive predictive value (PPV) of this indicator.

STUDY DESIGN: We conducted a retrospective cross-sectional study, sampling consecutive cases that met PSI 11 criteria from 18 geographically diverse academic medical centers on or before June 30, 2007. Trained abstractors from each center reviewed medical records using a standard instrument. We assessed the PPV of the indicator (with 95% CI adjusted for clustering within centers) and conducted descriptive analyses of the cases.

RESULTS: Of 609 cases that met PSI 11 criteria, 551 (90.5%; 95% CI, 86.5–94.4%) satisfied the technical criteria of the indicator and 507 (83.2%; 95% CI, 77.2–89.3%) represented true cases of postoperative respiratory failure from a clinical standpoint. The most frequent reasons for being falsely positive were nonelective hospitalization, prolonged intubation for airway protection, and insufficient evidence to support a diagnosis of acute respiratory failure. Fifty percent of true-positive cases involved substantial baseline comorbidities, and 23% resulted in death.

CONCLUSIONS: Although PSI 11 predicts true postoperative respiratory failure with relatively high frequency, the indicator does not limit detection to preventable cases. The PPV of PSI 11 might be increased by excluding cases with a principal diagnosis suggestive of a nonelective hospitalization and those with head or neck procedures. Removing the diagnosis code criterion from the indicator might also increase PPV, but would decrease the number of true positive cases detected by 20%. (J Am Coll Surg 2010;211:347–354. © 2010 by the American College of Surgeons)

Respiratory failure—frequently defined by the need for intubation or prolonged mechanical ventilation—is a relatively common postoperative complication, affecting an estimated 0.8% to 1.2% of all patients undergoing elective nonthoracic surgical procedures.[1,2] Postoperative respiratory failure (PRF) has become a focus of several organizations that measure quality of care[3] because it is associated with increased use of resources (hospitalization costs[4] and length of stay) and a 7-fold increase in mortality.[5]

There are currently 2 main strategies to retrospectively detect cases of PRF: detailed review of medical records (eg, the American College of Surgeons' National Surgical Quality Improvement Program [NSQIP][6]) and application of an algorithm to routinely collected administrative data (eg, Patient Safety Indicator [PSI] 11 (Postoperative Respiratory Failure) of the Agency for Healthcare Research and Quality[7]). Both strategies show that PRF is associated with excess inpatient mortality (21% to 34%), hospital stay (4.5 to 9.1 days), and hospital costs ($9,641 to $53,502) relative

Disclosure Information: Nothing to disclose.

This project was supported by contract #290-04-0020 from the US Agency for Healthcare Research and Quality. University HealthSystem Consortium and the 18 University HealthSystem Consortium member institutions that voluntarily participated also provided in kind support of this study.

Preliminary findings presented at the University HealthSystem Consortium Postoperative Respiratory Failure Knowledge Transfer Meeting, Oak Brook, IL, May 2008; AcademyHealth 2009 Annual Research Meeting, Washington, DC, June 2009; and the Organization for Economic Cooperation and Development Health Care Quality Indicators Patient Safety Subgroup Meeting, Paris, France, October 2009.

Received February 1, 2010; Revised April 26, 2010; Accepted April 27, 2010.
From the Departments of Surgery (Utter) , and Internal Medicine (Romano), and the Center for Healthcare Policy and Research (Utter, Zrelak, Baron, Romano), University of California, Davis, Medical Center, Sacramento, CA,University HealthSystem Consortium, Oak Brook, IL (Cuny, Sama), Rush University Medical Center, Chicago, IL (Silver), and Niederrhein University of Applied Sciences, Krefeld, Germany (Drösler).

Correspondence address: Garth H Utter, MD, MSc, Department of Surgery, University of California, Davis, Medical Center, 2315 Stockton Blvd, Rm 4206 MH, Sacramento, CA 95817. email: garth.utter@ucdmc.ucdavis.edu

ISSN 1072-7515/10/$36.00
doi:10.1016/j.jamcollsurg.2010.04.022

Abbreviations and Acronyms

MDC = major diagnostic category
NSQIP = National Surgical Quality Improvement Program
PPV = positive predictive value
PRF = postoperative respiratory failure
PSI = patient safety indicator
UHC = University HealthSystem Consortium

to carefully matched control patients without PRF.[8-10] Although NSQIP-based detection of PRF is reliable, it is resource-intensive and only encompasses a subset of at-risk patients at participating hospitals. PSI 11 offers an efficient, comprehensive alternative, but it can be constrained by available diagnosis and procedure codes and the limited accuracy of documentation and coding.

Early reports of the accuracy, or criterion validity, of PSI 11 suggest that the indicator holds promise. A forerunner of PSI 11, with a somewhat broader definition, demonstrated a positive predictive value (PPV) of 72% to 75% in Medicare claims data.[11,12] Also, <7% of PSI 11 cases were reported to be present on admission (ie, before an operation),[13,14] a common weakness of other PSIs. Although the Agency for Healthcare Research and Quality and NSQIP definitions of PRF differ slightly, PSI 11 had a sensitivity of 63% and PPV of 68% in a previous study that linked administrative and NSQIP data (with NSQIP-defined PRF considered the standard).[15]

As part of the University HealthSystem Consortium (UHC) PRF Benchmarking Project, we set out to describe the PPV of PSI 11 and to assess whether it is sufficiently predictive of true cases to warrant use as a tool to monitor quality. We also sought a better understanding of the types of events this indicator detects and whether cases might have been preventable.

METHODS

Our analysis of deidentified patient-level information was approved by the Institutional Review Board at the University of California, Davis.

Study design

We conducted a retrospective cross-sectional study of hospitalization records that met criteria for PSI 11. The denominator for this indicator includes all hospitalizations of adults aged 18 years or older with an "elective" admission type involving a surgical diagnosis related group and an ICD-9-CM procedure code for a "major operating room procedure," as defined by the Center for Medicare and Medicaid Services.[7] Excluded from consideration are patients with a principal diagnosis of acute respiratory failure (or a secondary such diagnosis present on admission); a diagnosis of a neuromuscular disorder; tracheostomy as the only procedure or occurring before the first operating room procedure; or a principal diagnosis involving major diagnostic category (MDC) 4 (ie, diseases/disorders of respiratory system), 5 (ie, diseases/disorders of circulatory system), or 14 (ie, pregnancy, childbirth, and puerperium).

Among hospitalizations meeting the denominator criteria, PRF is defined as a diagnosis of "acute respiratory failure" (518.81) or "acute and chronic respiratory failure" (518.84); or "insertion of endotracheal tube" (96.04) 1 or more days after the earliest "major operating room procedure"; continuous mechanical ventilation "of unspecified duration" (96.70) or "for less than 96 consecutive hours" (96.71) starting 2 or more days after the earliest procedure; or "continuous mechanical ventilation for 96 consecutive hours or more" (96.72) 0 or more days after the earliest procedure.

Study population

The UHC PRF Benchmarking Project involved 18 academic medical centers from 15 states. Participation was voluntary and without compensation; no hospitals that volunteered were denied participation. The rate of PRF (ie, PSI 11) in 2007 was 14.7/1,000 (range 2.3 to 29.2/1,000) among the 18 centers participating in this study and 15.0/1,000 among all 87 UHC member institutions. Average for all US hospitals in 2006 was 10.4/1,000.[1]

Instrument development

UHC project leaders developed a data abstraction instrument and guidelines in collaboration with the University of California Davis investigators and the PRF Benchmarking Project Steering Committee. Members of the multidisciplinary Steering Committee pilot-tested the instrument at their respective institutions.

The final instrument (see Appendix, available online) included questions about demographic characteristics, length of stay, disposition, potentially exclusionary conditions or procedures, the nature of the procedure and anesthesia associated with PRF, features and severity of PRF, and outcomes. Because the complication could be associated with any of a series of operations or periods of endotracheal intubation, we structured the instrument with discrete sets of questions corresponding to each qualifying operation and each period of intubation before diagnosis of PRF. To ensure standardization of responses, we presented most questions with either a menu of response options or specific numeric fields. Open-ended text fields allowed abstractors to explain why, if applicable, a diagnosis of acute respiratory failure was not made during the hospitalization