A Scoring System to Predict Unplanned Intubation in Patients Having Undergone Major Surgical Procedures

May Hua, MD,* Joanne E. Brady, SM,*† and Guohua Li, MD, DrPH†‡

BACKGROUND: Unplanned tracheal intubation after surgery has been associated with high mortality. Few studies have examined the risk factors for this complication.

METHODS: The American College of Surgeons National Surgical Quality Improvement Program (NSQIP) is a multicenter, prospective, outcome-oriented database for patients having undergone major surgical procedures. Using the NSQIP data for the years 2005 to 2007 (n = 231,548) and Cox proportional hazards modeling, we identified risk factors and used them to derive a scoring system to stratify patients’ risk of having an unplanned intubation outcome. NSQIP data for the year 2008 (n = 176,031) were then used to validate the scoring system.

RESULTS: The variables most predictive of unplanned intubation were patient age (0–4 points), ASA physical status (0–7 points), the presence of preoperative sepsis (3 points), and total operative time (0–4 points). The Unplanned Intubation Risk Index based on the adjusted hazard ratios for these variables, ranging from 0 (lowest risk) to 18 (highest risk), had a 79% accuracy in distinguishing patients requiring unplanned intubation from those not requiring it (area under the receiver operating characteristic curve 0.79, 95% confidence interval 0.79–0.80). When the scoring system was applied to the validation cohort data, its discriminative performance remained virtually unchanged (area under the receiver operating characteristic curve 0.79, 95% confidence interval 0.79–0.80).

CONCLUSIONS: A scoring system based on clinical risk factors was able to accurately predict unplanned intubation after surgery. Further investigation is needed to assess the utility of the Unplanned Intubation Risk Index in reducing the incidence of unplanned intubation through improved risk stratification and management in perioperative care. (Anesth Analg 2012;115:88–94)
likelihood of requiring unplanned intubation in patients undergoing major surgical procedures.

**METHODS**

**Patients and Data Collection**

The study protocol was reviewed and approved by the IRB of Columbia University Medical Center (New York, NY). Written informed consent was waived. Data for this study came from the multicenter, prospective, outcome-oriented database of the American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP) for the years 2005 to 2006. These data were collected from 251 participating hospitals for patients who underwent major surgical procedures. Major surgical procedures included any case performed under general, spinal, or epidural anesthesia, as well as the following procedures regardless of anesthetic technique: carotid endarterectomy, inguinal herniorrhaphy, parathyroidectomy, thyroidectomy, breast lumpectomy, and endovascular abdominal aortic aneurysm repair. To participate, hospitals must submit a minimum of 900 cases annually. Surgical cases are sampled in 8-day cycles, with the first 40 consecutive general or vascular cases performed under general, spinal, and epidural anesthesia included. Excluded from the ACS NSQIP were cases performed under monitored anesthesia care, peripheral nerve block, or local anesthesia, patients younger than 16 years, trauma cases, and transplant cases. No more than 3 breast lumpectomies, inguinal herniorrhaphies, laparoscopic cholecystectomies, transurethral resections of the prostate, or transurethral resections of the bladder are included in any 8-day sampling period, because these are considered to be low-risk but high-volume cases. For subspecialty operations, NSQIP samples using both a high-volume model, where hospitals must submit 20% of their cases, and a low-volume model, where a minimum of 900 cases are submitted annually. Gynecologic, neurologic, orthopedic, otolaryngologic, plastic, cardiac, thoracic, urologic, and vascular surgical procedures are included. In each participating hospital, a trained surgical nurse abstracted information for 135 variables, including demographic characteristics, preoperative and intraoperative variables, and 30-day postoperative morbidity and mortality outcomes from medical records using standard protocols. Case selection and case mix are monitored weekly to ensure proper sampling. Further detailed information regarding the database and its methods has been published.

There are several quality assurance measures to ensure that only data of the highest quality are recorded in the participant use data file. Hospitals with a 30-day follow-up rate <80% and whose surgical volume does not meet eligibility criteria are excluded. Furthermore, the consistency in data recording and reporting is checked with the Inter-Rater Reliability Audit, which is a process involving the review of 20 charts, with some cases selected randomly and some cases selected based on predetermined criteria; an interrater agreement rate of ≥95% is deemed acceptable. Combined results of the audits for the 2005 to 2008 data revealed an interrater agreement rate of 98%.

These data comprise the ACS NSQIP participant data use file. After exclusion of patients with preoperative ventilator dependence (n = 3901) and outpatients (n = 128,488), the derivation cohort consisted of 231,548 patients in the NSQIP database for the years 2005 to 2007. After applying the same exclusion criteria, data for the year 2008 were used as the validation cohort (n = 176,031).

Unplanned intubation was the primary outcome measure, which is operationally defined in the NSQIP database as requiring placement of an endotracheal tube secondary to the onset of respiratory or cardiac failure as evidenced by severe respiratory distress, hypoxia, hypercarbia, or respiratory acidosis within 30 days of the operation. For patients who were intubated for surgery, any intubation after extubation was considered an unplanned intubation event; in patients who were not intubated during surgery, any postoperative intubation was considered to be unplanned.

**Variable Selection**

Variables thought to be predictive of the primary outcome were broadly selected based on the methods of prior studies of postoperative respiratory complications. Demographic variables included age, race, and gender. Lifestyle variables included alcohol use (defined as >2 drinks per day in the 2 weeks before admission) and smoking (current smoking within 1 year of surgery). General factors included ASA physical status (ASA I, normal healthy patient; ASA II, patient with mild systemic disease; ASA III, patient with severe systemic disease; ASA IV, patient with severe systemic disease that is a constant threat to life; and ASA V, moribund patient who is not expected to survive without the operation), transfer status (admitted from home, acute care facility, or chronic care facility), functional status (independent, partially dependent, or totally dependent), emergency status, and body mass index. Laboratory values included preoperative hematocrit, white blood cell count, platelet count, serum sodium blood urea nitrogen, creatinine, albumin, bilirubin, serum glutamic oxaloacetic transaminase, prothrombin time, and partial thromboplastin time.

Preoperative comorbidities included in the NSQIP database were recoded into a comorbidity index modified from the Charlson comorbidity index. The following comorbidities were assigned a score of 1: history of chronic obstructive pulmonary disease, history of chronic heart failure, history of myocardial infarction, peripheral vascular disease, any diabetes, and cerebrovascular disease. Dialysis, and patients with radiation and chemotherapy without disseminated cancer, and hemiplegia were coded as a 2. Patients with ascites received a score of 3, and patients with disseminated cancer received a 6. A comorbidity score was tabulated for each patient. Other preoperative comorbidities included in the model that were not part of the Charlson index were sepsis (which includes the systemic

---

*a* The American College of Surgeons National Surgical Quality Improvement Program and the hospitals participating in the ACS NSQIP are the source of data used herein; they have not verified and are not responsible for the statistical validity of the data analysis or the conclusions derived by the authors.

Unplanned Intubation Risk Index

inflammatory response syndrome, sepsis, severe sepsis, and septic shock), dyspnea (at rest, moderate, or with exertion), and weight loss (defined as unintentional loss of 10% of body weight in the 6 months before surgery), chemotherapy, and transfusion requirement (requiring >4 U of packed red blood cells in the 72 hours before surgery).

Regarding surgical variables, the Current Procedural Terminology codes were identified and grouped by surgical specialty. Because there is an increased incidence of respiratory complications with incisions in closer proximity to the diaphragm, general surgery was separated into abdominal and nonabdominal categories. Vascular surgery was separated into abdominal and nonabdominal cases as well as an endovascular category. Total operative time was defined as surgical start to surgical stop. This variable was highly correlated with all other times reported in the NSQIP database (duration from anesthesia start to surgery start, duration from surgery stop to anesthesia stop, duration patient was in room, and duration of anesthesia). To determine categories for total operative time, groups were broken down into 60-minute intervals and groups with like odds ratios were grouped together.

Statistical Analysis

To develop a scoring system, predictors of unplanned intubation were identified as risk factors for unplanned intubation using χ² tests. A Cox proportional hazards model was used to determine association and strength of independent predictors of time until unplanned intubation. Risk factors statistically significant at an independent predictors of time until unplanned intubation. Model was used to determine association and strength of predictors. To develop a scoring system, variables with the highest P values were sequentially eliminated, and at each step, the receiver operating characteristic (ROC) curve was assessed to maintain the model’s discriminative power. A scoring system was created based on hazard ratios (HRs) from the final Cox model. Points for each category were generated as follows: an HR between 1.00 and 1.20 was dropped, whereas an HR between 1.21 and 1.49 for a given variable was given 1 point, and HR of 1.50 to 2.49 would yield 2 points, and so forth. Points for each variable were summed to create a total score. A logistic regression model of unplanned intubation status as predicted by the total score was given 1 point, and HR of 1.50 to 2.49 would yield 2 points, whereas an HR between 1.21 and 1.49 for a given variable was given 1 point, and HR of 1.50 to 2.49 would yield 2 points, and so forth. Points for each variable were summed to create a total score. A logistic regression model of unplanned intubation status as predicted by the total score was modeled as a continuous variable. The performance of the scoring system was then validated using the NSQIP data for the year 2008. Discrimination of the model was assessed using the C statistic and ROC curve. Calibration of the model was assessed using the Hosmer-Lemeshow goodness-of-fit test. The overall performance of the model was assessed using the Brier score. Statistical analysis was performed using SAS version 9.2 (SAS Institute, Cary, NC) and R version 2.14.0 (2011-10-31, Copyright © 2011 The R Foundation for Statistical Computing).

RESULTS

In the derivation cohort, 5028 patients (2.2%) had an unplanned intubation event. Unadjusted 30-day mortality was 28.1% for patients who experienced unplanned intubation versus 1.5% for those who did not (P < 0.0001). Approximately 50% of the unplanned intubations occurred within the first 3 days, and 70% of the unplanned intubation events occurred within the first 7 days after surgery. In the validation cohort, there were 3327 unplanned intubations of 176,031 patients (1.9%). The incidence rate of unplanned intubation in the validation cohort was significantly lower than in the derivation cohort (P < 0.001). Thirty-day mortality was similar to that of the derivation cohort (28.0% for patients who experienced unplanned intubation versus 1.5% for those who did not, P < 0.0001). Patient characteristics in the derivation and validation cohorts were similar in demographic makeup, ASA physical status, comorbidity, transfusion status, and surgical variables (Table 1).

Development and Validation of the Unplanned Intubation Risk Index

Variables that were significantly associated with unplanned intubation included age in years (grouped 16–29 and in 10-year increments), general anesthesia, surgical specialty, emergency status, transfer status, modified Charlson comorbidity index, total operative time (<120 minutes, 120–299 minutes, 300–359 minutes, ≥360 minutes), surgical procedure type (cardiothoracic, vascular abdominal, and other), ASA physical status, preoperative weight loss, any preoperative sepsis, and any preoperative dyspnea (Table 2). The 4 variables most predictive of unplanned intubation (age, ASA physical status, preoperative sepsis, and total operative time) were retained in the parsimonious model to create the Unplanned Intubation Risk Index (UIRI) (Table 3). Starting at age 40 years, each age decile was associated with increased hazard for unplanned intubation, with patients aged >80 years having the greatest hazard (adjusted HR 3.93, 95% confidence interval [CI] 3.38–4.58). Regarding the presence of preoperative comorbidities, the hazard of unplanned intubation was increased for ASA physical status III patients (adjusted HR 3.46, 95% CI 3.15–3.79), ASA physical status IV to V patients (adjusted HR 7.74, 95% CI 6.63–8.14), and patients with sepsis (adjusted HR 2.81, 95% CI 2.64–3.00). Finally, total operative time significantly increased the hazard for unplanned intubation, with the greatest hazard occurring for patients with operative times >360 minutes (adjusted HR 4.00, 95% CI 3.63–4.41). Based on these adjusted HRs, the UIRI was created to measure the risk of unplanned intubation within the first 30 postoperative days (Table 3). Ranging from 0 (lowest risk) to 18 (highest risk), the UIRI had a 79% accuracy in distinguishing patients who did and did not require unplanned intubation (area under the ROC curve [AUC] 0.79, 95% CI 0.79–0.80) (Fig. 1). Adding other risk factors to the scoring system did not improve the performance of the scoring system to any meaningful degree. When the risk index was applied to the validation cohort data, its diagnostic performance remained virtually unchanged (AUC 0.79, 95% CI 0.79–0.80) (Fig. 1). The incidence of unplanned intubation increased progressively with the UIRI score for both the derivation cohort and the validation cohort (Fig. 2). Regarding calibration of the model, the Hosmer-Lemeshow goodness-of-fit test was significant (P < 0.0001); however, this test has been shown...
to perform poorly with large sample sizes.11–13 We mea-

tured the overall model performance, based on the overall 
model discrimination and calibration, using the Brier score, 
where a score of 0 indicates a perfect model and a score of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adjusted hazard ratio</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16–29</td>
<td>0.80</td>
<td>0.60–1.06</td>
</tr>
<tr>
<td>40–49</td>
<td>1.27</td>
<td>1.05–1.54</td>
</tr>
<tr>
<td>50–59</td>
<td>1.66</td>
<td>1.39–1.99</td>
</tr>
<tr>
<td>60–69</td>
<td>2.15</td>
<td>1.81–2.57</td>
</tr>
<tr>
<td>70–79</td>
<td>2.67</td>
<td>2.26–3.21</td>
</tr>
<tr>
<td>≥80</td>
<td>3.30</td>
<td>2.76–3.94</td>
</tr>
<tr>
<td>General anesthesia</td>
<td>2.12</td>
<td>1.77–2.55</td>
</tr>
<tr>
<td>Emergency status</td>
<td>1.64</td>
<td>1.52–1.76</td>
</tr>
<tr>
<td>Transfer status</td>
<td>1.43</td>
<td>1.31–1.55</td>
</tr>
<tr>
<td>Comorbidity score I</td>
<td>1.13</td>
<td>1.05–1.22</td>
</tr>
<tr>
<td>2–3</td>
<td>1.36</td>
<td>1.25–1.47</td>
</tr>
<tr>
<td>4–5</td>
<td>1.83</td>
<td>1.62–2.06</td>
</tr>
<tr>
<td>≥6</td>
<td>1.74</td>
<td>1.56–1.95</td>
</tr>
<tr>
<td>ASA physical status II</td>
<td>2.19</td>
<td>1.51–3.19</td>
</tr>
<tr>
<td>III</td>
<td>6.16</td>
<td>4.25–8.91</td>
</tr>
<tr>
<td>IV–V</td>
<td>9.75</td>
<td>6.70–14.17</td>
</tr>
<tr>
<td>Preoperative weight loss</td>
<td>1.63</td>
<td>1.48–1.80</td>
</tr>
<tr>
<td>Any sepsis</td>
<td>2.09</td>
<td>1.95–2.25</td>
</tr>
<tr>
<td>Any dyspnea</td>
<td>1.46</td>
<td>1.37–1.56</td>
</tr>
<tr>
<td>Cardiothoracic surgery</td>
<td>1.84</td>
<td>1.64–2.06</td>
</tr>
<tr>
<td>Vascular abdominal surgery</td>
<td>1.75</td>
<td>1.57–1.94</td>
</tr>
<tr>
<td>Total operative time (min) ≥360</td>
<td>4.00</td>
<td>3.63–4.41</td>
</tr>
</tbody>
</table>

*P < 0.0001.
Figure 1. The Unplanned Intubation Risk Index. The performance of the Unplanned Intubation Risk Index was assessed for both the derivation and the validation cohorts using a receiver operating characteristic curve and quantified as the area under the curve.

Figure 2. Percentage of patients with unplanned intubation by the Unplanned Intubation Risk Index (UIRI) score. The percentage of patients who experienced unplanned intubation for each scoring category in the UIRI is shown for both the derivation and validation cohorts. T bars indicate standard error. *Scores 17 and 18 were combined because of the small numbers of unplanned intubation cases in the UIRI 18 group (2 in the derivation cohort and 1 in the validation cohort).
0.25 indicates a noninformative model. The Brier score was 0.021, indicating good model performance.

DISCUSSION

In this study of patients having undergone major surgical procedures, we confirm that unplanned intubation occurs at a rate of approximately 2% and is associated with heightened mortality. Furthermore, we have developed a scoring system to identify patients at greatest risk for this complication. The UIRI accurately predicts a patient’s likelihood of having an unplanned intubation event within 30 days of a surgical procedure. A previous analysis of unplanned intubation in surgical patients aged >65 years identified numerous risk factors with modest levels of association, with reoperation being the most significant one, but did not create a scoring system. A recent study by Ramachandran et al. improved on this previous effort; after identifying 17 risk factors, present individual risk factors were added together to separate patients into risk classes. However, their analysis focused on early unplanned intubation within the first 3 postoperative days, which only accounts for half of the population at risk. Although the risk of unplanned intubation is highest in the early postoperative period, 50% of all events occur after the first 3 postoperative days, with events still occurring weeks after surgery (Fig. 3). The novel features of the UIRI are its inclusiveness, its good discrimination, and its simplicity, which ultimately makes it more useful in a clinical setting.

We found that age, ASA physical status, the presence of preoperative sepsis, and total operative time were the factors that were most predictive of unplanned intubation. High ASA physical status was associated with the greatest hazard, followed by total operative time >6 hours and age >80 years. Although it is intuitive that sicker patients are at higher risk for this complication, the UIRI highlights the factors that matter most for risk stratification of the “sick patient.” Furthermore, several variables that have previously been associated with increased frequency of postoperative respiratory complications did not significantly improve the discriminative power of the UIRI. Preoperative comorbidities, laboratory abnormalities, and the type of surgical procedure have been identified as significant risk factors for respiratory failure.3,4 Although the modified Charlson comorbidity index and type of surgery were both associated with a significantly increased hazard of unplanned intubation, their inclusion into the risk index did not add significantly to its discriminative power. This is likely because the inclusion of ASA physical status, which is a clinician’s assessment of a patient’s comorbidities and overall health status, was able to capture the effect of major comorbid conditions.

The risk of unplanned intubation increases linearly with the UIRI score; thus, the UIRI can be used to prognosticate for individual patients and may be useful for both clinicians and patients to better understand the risks inherent in surgery and the postoperative period. For patients at high risk for unplanned intubation, clinicians may choose to intensify monitoring or optimize pulmonary function in the perioperative period. The UIRI can aid planning and allocation of the necessary resources for this monitoring (e.g.,

Figure 3. Number of unplanned intubations (UI) by postoperative day. The number of unplanned intubation events occurring on each postoperative day for the cohort is shown.
Unplanned Intubation Risk Index

ensuring adequate staffing for recovery room nursing and respiratory therapists, ensuring the availability of monitored intensive care unit or step-down beds, and planning for surgical admissions and surgical scheduling). However, the UIRI should not be used in isolation to determine whether or not patients should remain intubated after surgery. The decision to extubate at the end of surgery should balance the risk of unplanned intubation with the risks of prolonged mechanical ventilation, such as ventilator-associated pneumonia, sepsis, and a longer period of immobility.

Most importantly, the UIRI can also be used for risk adjustment of rates of unplanned intubation across the different contributing NSQIP hospitals. In 2009, an analysis by Ghaferi et al. of variations in hospital mortality found that unplanned intubation was the only postoperative complication whose incidence was greater in very high-mortality hospitals in comparison to very low-mortality hospitals. Use of the UIRI would help determine whether hospitals with higher rates of unplanned intubation were simply taking care of sicker patients or whether they had an appropriate UIRI-adjusted rate of unplanned intubation. This information could be used at an individual hospital level for quality assurance purposes.

There are several limitations to our study. First, although the method of data collection in the NSQIP has been well validated, cases of unplanned intubation may have been missed. Second, the components of the UIRI are not modifiable risk factors, and thus, they themselves are not appropriate targets for interventions. Also, in choosing to simplify the risk index to make it more convenient to use, we lost some discriminative power by not including every possible risk factor for unplanned intubation; our AUC, at 0.79, has only moderate diagnostic accuracy. Although we have simplified the risk index, its relative complexity (4 variables with different point values assigned) may limit its clinical usefulness unless it is automatically computed and incorporated into the electronic medical records. Finally, participation in the NSQIP is voluntary and thus the study sample is unlikely representative of all surgical patients. As a result, it might be unwise to extrapolate the findings of this study into other study populations and geographical regions.

Nevertheless, the UIRI developed in this study seems to be a valuable tool for accurately predicting unplanned intubation after surgery and may potentially be used for preoperative risk stratification. Further research is needed to refine this risk index and assess its utility in improving the management of unplanned intubation and related adverse consequences in the perioperative setting.

DISCLOSURES

Name: May Hua, MD.
Contribution: This author helped design the study, conduct the study, analyze the data, and write the manuscript.
Attestation: May Hua has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.

Name: Joanne Brady, SM.
Contribution: This author helped design the study, conduct the study, analyze the data, and write the manuscript.

Attestation: Joanne Brady has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.

Name: Guohua Li, MD, DrPH.
Contribution: This author helped design the study, conduct the study, analyze the data, and write the manuscript.
Attestation: Guohua Li has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.

This manuscript was handled by: Sorin J. Brull, MD, FCARCSI (Hon).

ACKNOWLEDGMENTS

The authors acknowledge Dr. Margaret Wood for her guidance and support.

REFERENCES