Predicting Extubation Failure After Successful Completion of a Spontaneous Breathing Trial

Babak Mokhlesi MD MSc, Aiman Tulaimat MD, Ty J Gluckman MD, Yue Wang PhD, Arthur T Evans MD MPH, and Thomas C Corbridge MD

OBJECTIVE: To derive a clinical prediction rule that uses bedside clinical variables to predict extubation failure (reintubation within 48 h) after a successful spontaneous breathing trial. METHODS: This prospective observational cohort study was performed at the Northwestern Memorial Hospital in Chicago, Illinois, which is a large tertiary-care university hospital. Among 673 consecutive patients who received mechanical ventilation during a 15-month period, 122 were ventilated for at least 2 days and did not undergo withdrawal of support or tracheostomy. These patients were followed after extubation to identify those who were reintubated within 48 h (extubation failure). We used logistic regression analysis to identify variables that predict reintubation, and we used bootstrap resampling to internally validate the predictors and adjust for overoptimism. RESULTS: Sixteen (13%) of the 122 patients required reintubation within 48 h. Three clinical variables predicted reintubation: moderate to copious endotracheal secretions ($p = 0.001$), Glasgow Coma Scale score ≤ 10 ($p = 0.004$), and hypercapnia ($P_{aCO_2} \geq 44$ mm Hg) during the spontaneous breathing trial ($p = 0.001$). Using logistic regression and bootstrap resampling to adjust for overfitting, we derived a clinical prediction rule that combined those 3 clinical variables (area under the receiver operating characteristic curve 0.87, 95% confidence interval 0.74–0.94). CONCLUSIONS: With our clinical prediction rule that incorporates an assessment of mental status, endotracheal secretions, and pre-extubation P_{aCO_2}, clinicians can predict who will fail extubation despite a successful spontaneous breathing trial. *Key words: weaning, extubation failure, endotracheal secretions, hypercapnia, mental status*. [Respir Care 2007;52(12):1710–1717. © 2007 Daedalus Enterprises]

Introduction

Before extubating a patient who is receiving mechanical ventilation, the clinician must decide whether the patient is ready to breathe without assistance. This decision is based on the outcome of a spontaneous breathing trial (SBT), with either a T-piece or low-level pressure support.[1–3] If the patient tolerates spontaneous breathing, the clinician must then decide whether the patient can tolerate extubation. This decision is important, because failed extubation occurs in 10–20% of patients, and extubation failure is associated with worse patient outcomes.[4,5] Though weaning predictors and methods of conducting SBT have been thoroughly studied,[6,7] limited data are available on factors that predict extubation failure.

The amount of endotracheal secretions,[8–11] strength of cough,[8,10,12] and pre-extubation mental status[10,13] have each been shown to predict extubation outcome after a success-

Babak Mokhlesi MD MSc is affiliated with the Section of Pulmonary and Critical Care Medicine, University of Chicago Pritzker School of Medicine, Chicago, Illinois. Aiman Tulaimat MD is affiliated with the Division of Pulmonary and Critical Care Medicine, Cook County Hospital, Chicago, Illinois. At the time of this study, Ty J Gluckman MD was affiliated with the Department of Medicine, Northwestern University, Chicago, Illinois. Yue Wang PhD and Arthur T Evans MD MPH are affiliated with the Collaborative Research Unit, Stroger Hospital of Cook County, Chicago, Illinois. Thomas C Corbridge MD is affiliated with the Division of Pulmonary and Critical Care, Northwestern University Feinberg School of Medicine, Chicago, Illinois.

The authors report no conflicts of interest related to the content of this paper.

Correspondence: Babak Mokhlesi MD MSc, Section of Pulmonary and Critical Care Medicine, University of Chicago Pritzker School of Medicine, 5841 South Maryland Avenue, MC0999, Room L11B, Chicago IL 60637. E-mail: bmokhles@medicine.bsd.uchicago.edu.

Table 1. Selection of the Study Cohort

Consecutive patients who received mechanical ventilation during the 15-month study period	673
Reasons for exclusion (a priori)	
Mechanical ventilation for < 48 h	206
Death or withdrawal of support	141
Research team not notified*	80
Tracheostomy	78
Extubation performed > 6 h after SBT	24
Self-extubation	20
Transfer to another hospital	2
Eligible study subjects	122

*These patients were similar to the study cohort in age, sex, Acute Physiology and Chronic Health Evaluation (APACHE II) score, and cause of respiratory failure. Sixty-seven percent were from sites other than the medical intensive care unit.
SBT = spontaneous breathing trial

ful SBT. Patients with moderate or abundant secretions were 3–8 times more likely to fail extubation than those with few to no secretions.[9,10] Coplin et al reported that the Glasgow Coma Scale (GCS) score did not predict extubation outcome in brain-injured patients,[8] but other investigators reported that impaired mental status was predictive of extubation failure.[10,13] Patients who fail weaning may retain CO_2, due to an imbalance between respiratory muscle strength and imposed load.[14–17] Patients extubated despite developing hypercapnia ($P_{aCO_2} \geq 45$ mm Hg) during a successful SBT may have a higher mortality due to respiratory failure than do patients who do not develop hypercapnia during SBT.[18]

We prospectively assessed the relationship between several variables (including mental status, amount of endotracheal secretions, and pre-extubation P_{aCO_2}) and reintubation within 48 h after extubation, in order to develop a simple clinical prediction rule that might be useful to physicians making extubation decisions.

Methods

Patients

During a 15-month period in 1998–1999, we screened all patients who received mechanical ventilation in the medical and surgical intensive care units (ICUs) at the Northwestern Memorial Hospital, Chicago, Illinois, which is a large tertiary-care university hospital. Hospitalized patients were eligible for our study if they were extubated after receiving mechanical ventilation for at least 2 days and had not undergone a tracheostomy or withdrawal of support. Our a priori exclusion criteria are listed in Table 1. The institutional review board of Northwestern University approved the study and waived the requirement of written informed consent from participants.

Design

This was a prospective observational cohort study, so the research team did not influence any clinical decision, including the decision to wean, the method of weaning, the timing of SBTs, or the decisions to extubate or to reintubate.

Weaning and Extubation Procedures

Weaning was not guided by a protocol. Weaning was discussed during morning rounds by bedside nurses, respiratory therapists, house staff, and board-certified intensivists. According to standard practices,[3,19] patients were considered ready for SBT when they had resolution or improvement of the underlying cause of respiratory failure, had adequate gas exchange, were hemodynamically stable without vasoactive medications, and had an adequate cough during suctioning. The primary team stopped the SBT if a patient had any of the following: respiratory rate > 35 breaths/min, oxygen saturation below 90%, heart rate > 140 beats/min, a sustained 20% increase or decrease in heart rate, systolic blood pressure above 180 mm Hg or below 90 mm Hg, agitation, diaphoresis, or anxiety. Patients were extubated by the primary team if they tolerated 120 min of spontaneous breathing on positive end-expiratory pressure of 5 cm H_2O with pressure support of 5–7 cm H_2O. Patients were followed for 48 hours after extubation. The decisions to extubate and re-intubate were made solely by the primary team, without any involvement from the research team, according to the standard practices at the time.[3,19]

Data Collection

Variables collected on each patient included age, Acute Physiology and Chronic Health Evaluation (APACHE II) score,[20] duration of mechanical ventilation, hemoglobin and blood chemistries, arterial blood gas values 1 h into the SBT, time of extubation, use of paralytics or systemic corticosteroids during the period of mechanical ventilation, and negative inspiratory pressure measured before the initiation of SBT. Additionally, endotracheal secretions were assessed by the patient's nurse prior to extubation, and mental status (GCS score) was assessed by the research team at the time of extubation.[21] Assessment of the verbal component of the GCS score was based on the patient's orientation and ability to mouth or write words. The nurses documented in the medical record the amount of endotracheal secretions, according to the following semiquantitative scale: no secretions, minimal secretions (suctioning required every 2–4 h), moderate secretions (suctioning required every 1–2 h), or copious secretions (suctioning required several times per hour). For the pur-

pose of the study, the amount of secretions in the 8-hour shift preceding extubation was collected. The research team did not share any of the collected data with the primary team. However, these data (nurses' estimation of endotracheal secretions, GCS score, and P_{aCO_2} during SBT) were readily available to the primary team. Data on cuff leak test prior to extubation was not collected by the research team, and it was left to the discretion of the primary team to measure it if clinically indicated.

Data Analysis

The primary outcome of interest was extubation failure, defined as reintubation within 48 h of extubation. For bivariate comparisons between the outcome and each of the predictors, we used the Wilcoxon test for all continuous and ordinal variables, and the chi-square of Fisher's exact test for categorical variables. A variable's strength for independently predicting extubation failure was assessed by calculating adjusted odds ratios and 95% confidence intervals using multivariable logistic regression models.

We selected potential predictor variables to study based on our collective clinical experience and information from other studies. The variables included age, APACHE II score, history of chronic obstructive pulmonary disease, duration of mechanical ventilation, type of ICU, time or shift of extubation, hemoglobin, creatinine, serum phosphorus, arterial blood gas values measured during SBT, use of paralytics or systemic corticosteroids, negative inspiratory pressure, volume of endotracheal secretions, and GCS score at the time of extubation.[4] Variables with p values ≤ 0.2 in univariate analysis were entered into a logistic regression model in order to avoid missing potentially important associations.[22] In the logistic regression model, differences were considered significant if p was $\leq 0.05\%$.

Because the aim was to provide a simple predictive model,[23] we dichotomized P_{aCO_2} (≥ 44 mm Hg or < 44 mm Hg), GCS score (≥ 10 or < 10), hemoglobin (≥ 10 g/dL or < 10 g/dL), and duration of mechanical ventilation (≥ 8 d or < 8 d), because there was no difference in the predictive ability of the model compared to using continuous versions of these variables.[23] We assessed for clinically sensible interactions between predictors and tested for colinearity. Calibration of the final model—agreement between predicted and observed probabilities—was assessed with the Hosmer-Lemeshow test. Standard formulas were used to calculate likelihood ratios. Receiver operating characteristic curves were generated to determine the relative accuracy of different models for predicting extubation outcome.

A risk in building a predictive model from a small data set is "overfitting," which is building a model that fits the idiosyncrasies of the study sample but does less well when tested on an independent sample from the target population. A model with excessive overfitting, therefore, would be invalid. Given the absence of an independent sample to externally validate the model, we internally validated the predictive model by performing bootstrap resampling. Bootstrapping involves creating a large number of different samples, each with a sample size identical to the original (n), by randomly selecting n subjects (with replacement) from the original sample. The modeling strategy is then applied to each of these new bootstrapped samples. Bootstrapping has been shown to provide nearly unbiased estimates of predictive accuracy and is much more efficient than alternative procedures such as split-sample testing or cross-validation. In addition, the "over-optimism" of the original model (caused by overfitting) can be estimated with bootstrapping. The measure of over-optimism describes to what extent the model coefficients exaggerate their true predictive power. These coefficients can then be shrunk, based on the degree of over-optimism, so that they more accurately reflect valid relationships. Thus, shrinkage of coefficients is a method to correct for overfitting. By convention, a shrinkage factor of $\leq 10\%$ in model coefficients is considered small.[22–25] Accordingly, we used bootstrap resampling procedures to assess the internal validity of multiple models and to adjust for overfitting or over-optimism. One thousand random bootstrap samples were drawn with replacement from the full sample. The regression coefficients were estimated in each sample, and the performance of the models was evaluated by calculating the area under the receiver operating characteristic curves. The difference between the performance estimated in the bootstrap samples and the performance in the original sample provided a measure of the amount of over-optimism inherent in the predictive model and allowed for "shrinking" the model coefficients to more realistic values.[26,27] For the data analysis we used statistics software (SPSS 12, SPSS, Chicago, Illinois, and STATA 9, StataCorp, College Station, Texas).

Results

Of 673 consecutive patients who required mechanical ventilation, 122 were eligible for the study after tolerating a 2-h SBT (see Table 1): 76 from the medical ICU, 23 from the surgical ICU, 16 from the neurosurgical ICU, and 7 from the coronary care unit. The baseline characteristics of the patients are provided in Table 2. The duration of mechanical ventilation was similar between the 4 ICUs. The reintubation rate was 10.5% in the medical ICU, 14% in the coronary care unit, 22% in the surgical ICU, and 12.5% in neurosurgical ICU ($p = 0.6$). The primary indications for mechanical ventilation were: pneumonia or acute respiratory distress syndrome (40 patients), upper-airway edema or need for airway protection (22 patients), heart

Table 2. Baseline Characteristics of the Study Cohort*

Variable	Entire cohort (n = 122)	Successful Extubation (n = 106)	Failed Extubation (n = 16)	p
Male (n, %)	58 (47)	50 (47)	8 (50)	0.8
Age (mean ± SD y)	60 ± 19	61 ± 19	57 ± 22	0.5
APACHE II score (mean ± SD)	21 ± 8	21 ± 8	20 ± 10	0.6
History of COPD (n, %)	14 (11.5)	11 (10.4)	3 (18.8)	0.3
Duration of mechanical ventilation (d, 25th–75th percentile)	5 (3–8)	6 (3–8)	4 (3–6)	0.09
Use of paralytics during ventilation (n, %)	10 (8)	9 (9)	1 (6)	0.8
Use of corticosteroids during ventilation (n, %)	43 (35)	37 (35)	6 (38)	0.8
Hemoglobin (g/dL)	10 ± 1.3	10.1 ± 1.3	9.4 ± 1.2	0.05
Hemoglobin ≥ 10 g/dL (%)	49	52	31	0.17
Serum creatinine (mg/dL)	1.96 ± 2.21	1.97 ± 2.18	1.86 ± 2.46	0.9
Serum phosphorus (mg/dL)	3.58 ± 1.56	3.64 ± 1.60	3.24 ± 1.42	0.4
Arterial blood gas values during SBT				
pH	7.41 ± 0.06	7.42 ± 0.06	7.39 ± 0.06	0.07
p_{aCO_2} (mm Hg, 25th–75th percentile)	42 (36–47)	41 (36–46)	47 (44–49)	0.006
P_{aO_2} (mm Hg, 25th–75th percentile)	89 (78–102)	90 (78–102)	88 (82–100)	0.5
P_{aCO_2} ≥ 44 mm Hg (%)	39	33	81	0.001
Negative inspiratory force (cm H_2O)	38.4 ± 11.5	38.6 ± 11.5	37.2 ± 11.2	0.67
Glascow Coma Scale score ≤ 10 (n, %)	12 (10)	6 (6)	6 (38)	0.001
Endotracheal secretions (n, %)				
None or minimal secretions	90 (74)	84 (79)	6 (37)	0.001
Moderate or copious secretions	32 (26)	22 (21)	10 (63)	0.001

*For continuous and ordinal variables the values are mean ± SD. If the data were not normally distributed, the values are medians (25th–75th percentiles).
APACHE = Acute Physiology and Chronic Health Evaluation
P values were calculated by comparing successful extubation vs failed extubation.

Table 3. Adjusted Odds Ratios for Variables That Predict Extubation Failure, Via Logistic Regression Without and With Bootstrapping (Shrunk Odds Ratio)

Variable	OR (95% CI)	Shrunk OR (95% CI)	p
Pre-extubation hypercapnia (P_{aCO_2} ≥ 44 mm Hg)	13 (3–59)	11 (3–44)	0.001
Glasgow coma scale score ≤ 10	13 (2–72)	11 (2–53)	0.004
Moderate or copious secretions	12 (3–51)	10 (3–39)	0.001

OR = odds ratio
CI = confidence interval

failure or cardiac arrest (18 patients), postoperative respiratory failure (17 patients), exacerbation of obstructive airways disease (12 patients), and other causes (13 patients). The weaning methods included: gradual reduction of pressure support (57%), synchronized intermittent mandatory ventilation (SIMV) (39%) (predominately in the surgical and neurosurgical ICUs), and daily SBTs with a T-piece (4%).

Sixteen patients (13%) required reintubation within 48h of extubation. The reasons for reintubation, as determined by the primary team, included: secretions (3 patients), progression of the underlying process (3 patients), upper-airway edema (2 patients), depressed mental status (2 patients), respiratory muscle fatigue (2 patients), pulmonary edema (2 patients), atelectasis (1 patient), and unclear reasons (1 patient). Median time to reintubation was 13 h (interquartile range 4–33 h).

Univariate analysis found 5 variables that were different ($p \leq 0.2$) between the patients who required reintubation and those who were successfully extubated: hemoglobin, P_{aCO_2}, amount of endotracheal secretions, GCS score, and duration of mechanical ventilation (see Table 2). Age, sex, weight, APACHE II score, type of ICU, time or shift of extubation, creatinine, serum phosphorus, pH, P_{aO_2}, use of paralytics or systemic corticosteroids, and negative inspiratory pressure were similar between patients who failed extubation and those who were successfully extubated.

In a multivariable logistic regression model, 3 variables independently predicted reintubation within 48h of extubation: moderate or copious endotracheal secretions, GCS score

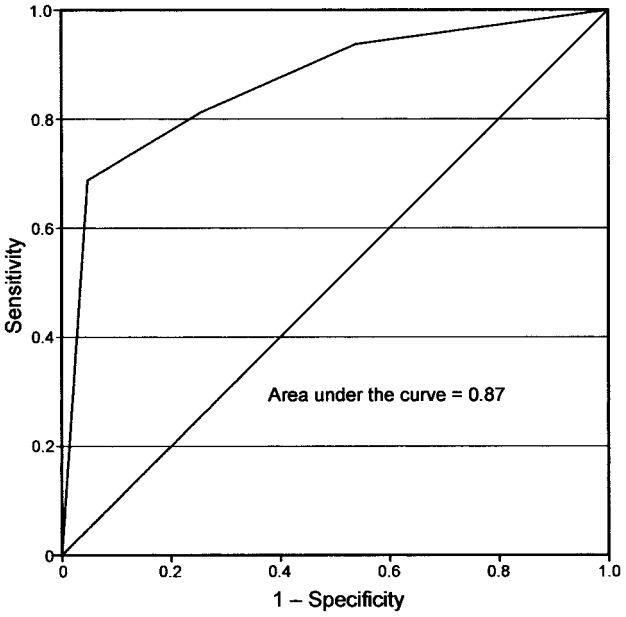

Fig. 1. Receiver operator characteristic curve for the two variable model (hypercapnia and the presence of either moderate to copious endotracheal secretions, low GCS score, or both). The expected ability to discriminate in another population would be less because of "over-optimism" of the modeling procedure. The curve is generated by plotting sensitivity against false-positive rate (1 − specificity). The curve for a test with no discriminatory value would appear as a diagonal line, whereas a useful test has a receiver operating characteristic curve that rises rapidly and reaches a plateau close to the upper left corner. The area under the curve is expressed as a proportion of the total graph (box).

<10, and pre-extubation hypercapnia ($P_{aCO_2} \geq 44$ mm Hg). There were no clinically or statistically significant interactions (Table 3).

The best model included 2 variables: pre-extubation hypercapnia and a dichotomous composite variable that combined information about mental status and amount of secretions—no or minimal secretions *and* GCS score > 10 versus moderate or copious secretions *and/or* GCS score ≤ 10. This model had the best combination of accuracy and reliability, with a low shrinkage factor (7%), which suggests that extubation failure could be predicted reliably in other patient populations (reasonably small over-optimism). The model was well-calibrated and had a high concordance between predicted and observed probabilities (Hosmer-Lemeshow test, $p = 0.70$). The area under the receiver operating characteristic curve, which provides a measure of the model's accuracy, was 0.87 after applying the shrinkage factor (Fig. 1). Alternative 2-variable predictive models had areas under the curve nearly as good but with higher shrinkage factors (Table 4).

For patients with hypercapnia and either moderate or copious secretions or a GCS score ≤ 10, the extubation failure rate was high, 69% (observed 11/16). When all 3 risk factors were absent, the extubation failure rate was low, 2% (observed 1/50). For all other combinations of risk factors, the extubation failure rate was moderate, 7% (observed 4/56). Figure 2 represents the shrunk predicted probabilities and the 95% confidence intervals for extubation failure in these groups of patients.

Discussion

Deciding when to extubate a patient requires clinical judgment that balances the potential benefits of early extubation against the potential harms and costs of failed extubation. Although a successful SBT is a standard criterion for extubation, it is imperfect, and clinicians have always incorporated other factors into their decision making. The present study demonstrates that 3 clinical variables—secretions, mental status, and pre-extubation P_{aCO_2}—are critically important for predicting successful extubation.

Khamiees et al found that patients with moderate or abundant secretions are 8 times more likely to fail extubation than those with no or minimal secretions (risk ratio 8.7, 95% confidence interval 2.1–35.7).[9] Salam et al reported that patients who accumulate secretions at a rate of > 2.5 mL/h are 3 times more likely to have unsuccessful extubation than patients who accumulate secretions at a lower rate.[10] Our findings are consistent with those studies; patients with moderate or copious secretions, based on the need to suction at least every 1–2 h, were more likely to fail extubation. In clinical practice, physicians typically rely on nurses or respiratory therapists to estimate the amount of endotracheal secretions. Asking the patient's nurse to quantify the amount of endotracheal secretions according to the frequency of suctioning is simple, practical, generalizable, and, based on our results, valuable.

In a recent large study that evaluated the risk factors for extubation failure in patients following a successful SBT, Frutos-Vivar et al did not find an association between increased endotracheal secretions or depressed mental status and reintubation rate.[28] However, the protocol and enrollment period had begun after several publications that reported increased risk of extubation failure in patients with copious endotracheal secretions and depressed mental status. As discussed by Frutos-Vivar et al, it is possible that the clinicians who were caring for those patients had already incorporated the information on copious secretions and depressed mental status into their decision making to delay extubation. Moreover, the frequency of abundant secretions in their study was less than previously observed by other investigators, and therefore the study may have not had adequate power to detect an association between secretions and reintubation.

Table 4. Summary of 2-Variable Models

Variables in the Model	Area Under ROC Curve*	95% CI	Shrunk 95% CI	Shrinkage Factor (%)
Hypercapnia, secretions and/or low Glasgow Coma Scale score	0.87	0.74–0.94	0.74–0.93	7
Hypercapnia, secretions	0.83	0.70–0.93	0.67–0.91	13
Secretions, low Glasgow Coma Scale score	0.81	0.69–0.92	0.66–0.90	18
Hypercapnia, low Glasgow Coma Scale score	0.77	0.62–0.89	0.50–0.87	5

*The areas under the receiver operator characteristic (ROC) curves reported above were calculated based on the shrinkage model. The shrinkage factor measures the degree of over-optimism in estimating the model coefficients, based on bootstrap resampling, and is used to "shrink" the coefficients to more realistic (less extreme) values.
CI = confidence interval (95% CI calculated with 1,000 bootstrap samples)

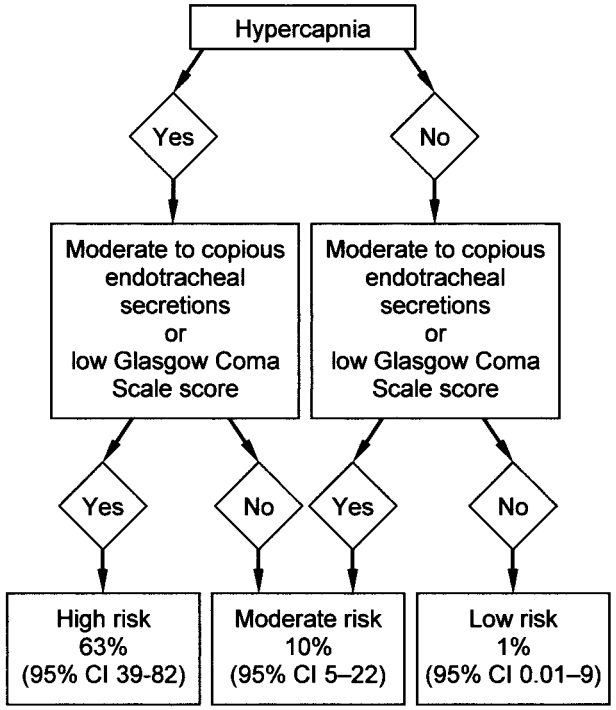

Fig. 2. Clinical prediction rule for extubation failure after successful spontaneous breathing trial. The estimated probabilities and the 95% confidence intervals have been "shrunk" toward the overall mean, to compensate for model overfitting, in order to more realistically estimate probabilities when the model is applied in other populations. Extubation failure was defined as reintubation within 48 h of extubation.

This contradiction between earlier and later studies was highlighted in a study by Krishnan et al, in which weaning according to a protocol was not superior to routine care in a well-organized and well-staffed tertiary-care university hospital.[29] This suggests that after a certain period of time, physicians are proficient at extracting principles that emerge from research studies and incorporating them into their everyday practice.[30]

Three prospective studies have evaluated the integration of neurologic measurements into the extubation decision.[8,10,13] In one report, a GCS score of ≤ 8 delayed extubation but did not preclude successful extubation in brain-injured patients.[8] In contrast, another study demonstrated that in neurosurgical patients a GCS score ≥ 8 was associated with successful extubation in 75% of cases, compared to 33% success among patients with a GCS score < 8.[13] Those 2 reports are limited because they included patients from neurosurgical ICUs. Salam et al demonstrated that a measure of neurologic status that assessed the patient's ability to perform 4 simple tasks was an independent predictor of extubation outcome.[10] But, to our knowledge, this finding has not been replicated. The present study confirms that a depressed mental status is associated with a higher risk of extubation failure in patients who successfully complete an SBT. Although the GCS score is a standard measure with good performance characteristics verified in multiple settings, it does have drawbacks for mechanically ventilated patients and does not necessarily measure the neurologic function that underlies the physiology of airway reflexes. Therefore, it is possible that another method of assessing mental status, such as the validated Richmond Agitation-Sedation Scale (RASS), might outperform the GCS. But it is unlikely that the RASS could do too much better, because the RASS and GCS score are so highly correlated ($r > 0.9$).[31] Unfortunately, the validity of RASS in mechanically ventilated patients had not been demonstrated during the enrollment period of our study.

Tobin et al found that during weaning, P_{aCO_2} was higher in patients who failed the trial (56 ± 4 mm Hg) than in patient who were successfully extubated (42 ± 2 mm Hg).[14] Thus, failure is likely when there is an imbalance between the load imposed on the respiratory muscles and the muscles' capacity to respond, leading to alveolar hypoventilation.[15] Yang and Tobin later identified $P_{aCO_2} > 50$ mm Hg as a sign of failed weaning.[1] Moreover, patients extubated despite hypercapnia during SBT ($P_{aCO_2} \geq 45$ mm Hg) have a higher rate of death due to respiratory failure, compared to patients who do not develop hypercapnia during SBT.[18] In contrast, Esteban and colleagues reported that routine standard monitoring during SBTs—which included respiratory frequency, heart rate, systolic blood pressure, and arterial oxygen saturation measured by pulse oximetry—did not allow them to predict which patients will ulti-

mately require reintubation within 48 h.[32,33] However, in those studies routine standard monitoring during the SBT did not include the measurement of P_{aCO_2} while breathing spontaneously or the assessment of endotracheal secretions.

We derived the P_{aCO_2} threshold of 44 mm Hg from its receiver operating characteristic curve (not shown), which fortuitously is also the upper limit of normal and is similar to the P_{aCO_2} level in patients who failed weaning in the study by Dunn et al.[16] In our study, hypercapnia was the most sensitive predictor of extubation failure, but the median P_{aCO_2} among those who failed extubation was < 50 mm Hg, which was the limit used by Yang and Tobin.[1] As frank hypercapnia ($P_{aCO_2} > 50$ mm Hg) is a sign of weaning failure, milder hypercapnia (P_{aCO_2} 44–50 mm Hg) appears to be a marker of increased risk for reintubation. Patients who fail extubation may be breathing at their maximum capacity during the pressure-support SBT and then may require reintubation when the support is removed at extubation. If so, then these patients might have been frankly hypercapnic if the SBT had been performed without ventilator assistance (ie, with only a T-piece). This hypothesis is supported by the increase in P_{aCO_2}, from 46 mm Hg before extubation to 49 mm Hg after extubation, among patients who required reintubation ($p = 0.02$), compared to the P_{aCO_2} decrease, from 41 mm Hg before extubation to 40 mm Hg after extubation, among patients who were successfully extubated.

Our study has several limitations. First, we did not have an external validation set. However, we have prospectively validated the importance of 3 predictors that had been identified by other investigators in other settings.[9,10] Our study was conducted before these findings were reported in the literature, and therefore it is possible that the presence of copious secretions, weak cough, or depressed mental status were not explicitly incorporated into the extubation decision.

Second, over-optimistic estimates of predictive performance is a likely problem with our predictive model because of the small size of the data set and the low event-per-candidate variable ratio. Therefore, some correction factor needs to be applied to give a more realistic estimate of odds ratios and the discriminating ability of the model. Using bootstrap resampling procedure, we were able to estimate the degree of over-optimism, which was small (7%). The shrunk estimates are more realistic for what will happen in another independent sample but will not obviate the need for validating the prediction model in an independent sample. Furthermore, if the patients in the independent sample are much different than our patient population, the model's discrimination might be even worse.

Third, we did not document the reasons that the primary team ended the SBT. Although weaning was not protocolized, the standard clinical practice in the institution was the approach outlined by Esteban et al.[3,32,33] The extubation failure rate of 13% is consistent with reported failure rates in cohorts where formal criteria were used both to assess the success of SBTs and the decision to re-intubate.[3,32,33]

Fourth, although the research team did not share any of the collected data with the primary team, the primary team was not blinded to the estimation of endotracheal secretions by the nurses, GCS score, and P_{aCO_2} during SBT. Despite the fact that these variables were readily available to the primary team, apparently it did not affect their decisions to extubate patients, given the findings of increased extubation failure in patients with copious secretions, depressed mental status, or presence of hypercapnia during the SBT. As stated above, it is important to reiterate that our study was conducted before other investigators had reported the increased risk of extubation failure in the presence of copious secretions, weak cough, or depressed mental status.

Another limitation to the generalizability of our data is that 39% of the patients were weaned on SIMV. This finding is consistent with the data reported by Esteban et al that in the years 1997 and 1998, 30–37% of weaning attempts were performed on SIMV, with or without pressure support, despite the evidence that SIMV prolongs the duration of weaning from mechanical ventilation.[34,35] Moreover, Esteban et al reported that the reintubation rates were similar between pressure-support weaning (19%), SIMV weaning (14%), and once-daily or intermittent T-piece trials (15–23%).[3] Although the variables used in our clinical prediction model are not intuitively altered by the weaning mode, this remains unproven. Finally, we did not measure interobserver agreement in the estimation of secretions or GCS score.

Conclusions

In summary, extubation success requires a patent airway with adequate ventilation. Our clinical prediction rule quantifies the importance of these 2 requirements. The 2 determinants of airway patency—mental status and airway secretions—are easily assessed at the bedside. Adequacy of ventilation is also easily measured via P_{aCO_2}. Although blood gas analyses are not routinely performed at many institutions during the SBT, the P_{aCO_2} appears to be a strong predictor of extubation failure. A longer observation period in the ICU may be prudent after extubating patients who have one or more of these risk factors. In cooperative patients who develop mild hypercapnia, noninvasive positive-pressure ventilation could be an alternative to a delay in extubation.[18]

REFERENCES

1. Yang KL, Tobin MJ. A prospective study of indexes predicting the outcome of trials of weaning from mechanical ventilation. N Engl J Med 1991;324(21):1445–1450.
2. Brochard L, Rauss A, Benito S, Conti G, Mancebo J, Rekik N, et al. Comparison of three methods of gradual withdrawal from ventilatory support during weaning from mechanical ventilation. Am J Respir Crit Care Med 1994;150(4):896–903.
3. Esteban A, Frutos F, Tobin MJ, Alia I, Solsona JF, Valverdu I, et al. A comparison of four methods of weaning patients from mechanical ventilation. N Engl J Med 1995;332(6):345–350.
4. Epstein SK. Decision to extubate. Intensive Care Med 2002;28(5): 535–546.
5. Epstein SK. Extubation failure: an outcome to be avoided. Crit Care 2004;8(5):310–312.
6. Meade M, Guyatt G, Cook D, Griffith L, Sinuff T, Kergl C, et al. Predicting success in weaning from mechanical ventilation. Chest 2001;120(6 Suppl):400S–424S.
7. MacIntyre NR, Cook DJ, Ely EW Jr, Epstein SK, Fink JB, Heffner JE, et al. Evidence-based guidelines for weaning and discontinuing ventilatory support: a collective task force facilitated by the American College of Chest Physicians; the American Association for Respiratory Care; and the American College of Critical Care Medicine. Chest 2001;120(6 Suppl):375S–395S.
8. Coplin WM, Pierson DJ, Cooley KD, Newell DW, Rubenfeld GD. Implications of extubation delay in brain-injured patients meeting standard weaning criteria. Am J Respir Crit Care Med 2000;161(5): 1530–1536.
9. Khamiees M, Raju P, DeGirolamo A, Amoateng-Adjepong Y, Manthous CA. Predictors of extubation outcome in patients who have successfully completed a spontaneous breathing trial. Chest 2001; 120(4):1262–1270.
10. Salam A, Tilluckdharry L, Amoateng-Adjepong Y, Manthous CA. Neurologic status, cough, secretions and extubation outcomes. Intensive Care Med 2004;30(7):1334–1339.
11. El Solh AA, Bhat A, Gunen H, Berbary E. Extubation failure in the elderly. Respir Med 2004;98(7):661–668.
12. Smina M, Salam A, Khamiees M, Gada P, Amoateng-Adjepong Y, Manthous CA. Cough peak flows and extubation outcomes. Chest 2003;124(1):262–268.
13. Namen AM, Ely EW, Tatter SB, Case LD, Lucia MA, Smith A, et al. Predictors of successful extubation in neurosurgical patients. Am J Respir Crit Care Med 2001;163(3 Pt 1):658–664.
14. Tobin MJ, Perez W, Guenther SM, Semmes BJ, Mador MJ, Allen SJ, et al. The pattern of breathing during successful and unsuccessful trials of weaning from mechanical ventilation. Am Rev Respir Dis 1986;134(6):1111–1118.
15. Jubran A, Parthasarathy S. Hypercapnic respiratory failure during weaning: neuromuscular capacity versus muscle loads. Respir Care Clin N Am 2000;6(3):385–405.
16. Dunn WF, Nelson SB, Hubmayr RD. The control of breathing during weaning from mechanical ventilation. Chest 1991;100(3):754–761.
17. Prechter GC, Nelson SB, Hubmayr RD. The ventilatory recruitment threshold for carbon dioxide. Am Rev Respir Dis 1990;141(3):758–764.
18. Ferrer M, Valencia M, Nicolas JM, Bernadich O, Badia JR, Torres A. Early noninvasive ventilation averts extubation failure in patients at risk: a randomized trial. Am J Respir Crit Care Med 2006;173(2): 164–170.
19. Ely EW, Baker AM, Dunagan DP, Burke HL, Smith AC, Kelly PT, et al. Effect on the duration of mechanical ventilation of identifying patients capable of breathing spontaneously. N Engl J Med 1996; 335(25):1864–1869.
20. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985; 13(10):818–829.
21. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. Lancet 1974;2(7872):81–84.
22. Steyerberg EW, Eijkemans MJ, Harrell FE Jr, Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med 2000; 19(8):1059–1079.
23. Steyerberg EW, Eijkemans MJ, Harrell FE Jr, Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. Med Decis Making 2001;21(1): 45–56.
24. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15(4):361–387.
25. Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat Sci 1986;154–177.
26. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. Stat Med 1990;9(11):1303–1325.
27. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Assoc 1983;78:316–331.
28. Frutos-Vivar F, Ferguson ND, Esteban A, Epstein SK, Arabi Y, Apezteguia C, et al. Risk factors for extubation failure in patients following a successful spontaneous breathing trial. Chest 2006;130(6): 1664–1671.
29. Krishnan JA, Moore D, Robeson C, Rand CS, Fessler HE. A prospective, controlled trial of a protocol-based strategy to discontinue mechanical ventilation. Am J Respir Crit Care Med 2004;169(6): 673–678.
30. Tobin MJ. Of principles and protocols and weaning. Am J Respir Crit Care Med 2004;169(6):661–662.
31. Ely EW, Truman B, Shintani A, Thomason JW, Wheeler AP, Gordon S, et al. Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation-Sedation Scale (RASS). JAMA 2003;289(22):2983–2991.
32. Esteban A, Alia I, Gordo F, Fernandez R, Solsona JF, Vallverdu I, et al. Extubation outcome after spontaneous breathing trials with T-tube or pressure support ventilation. Am J Respir Crit Care Med 1997;156(2 Pt 1):459–465.
33. Esteban A, Alia I, Tobin MJ, Gil A, Gordo F, Vallverdu I, et al. Effect of spontaneous breathing trial duration on outcome of attempts to discontinue mechanical ventilation. Am J Respir Crit Care Med 1999;159(2):512–518.
34. Esteban A, Anzueto A, Alia I, Gordo F, Apezteguia C, Palizas F, et al. How is mechanical ventilation employed in the intensive care unit? An international utilization review. Am J Respir Crit Care Med 2000;161(5):1450–1458.
35. Esteban A, Anzueto A, Frutos F, Alia I, Brochard L, Stewart TE, et al. Characteristics and outcomes in adult patients receiving mechanical ventilation: a 28-day international study. JAMA 2002;287(3): 345–355.