

# Accuracy and Reliability of Extubation Decisions by Intensivists

Aiman Tulaimat MD and Babak Mokhlesi MD MSc

**BACKGROUND:** Mortality increases when extubations fail. Although predictors of extubation failure have been evaluated, physicians' reasoning to extubate a patient has received minimal attention. We hypothesized that the accuracy and reliability of physicians' extubation decisions are low. **METHODS:** We sent surveys to 55 physicians in the divisions of pulmonary and critical care medicine of 3 teaching hospitals in Chicago, Illinois. The survey comprised 32 clinical vignettes of real patients who were extubated after they tolerated a spontaneous breathing trial (16 failed extubations). Unaware of the outcomes of extubation, the physicians were asked if they would extubate each patient, and to give reasons if they opted not to. We quantified the agreement between and accuracy of the physicians' decisions, determined the patient characteristics that influence the extubation decision, and described the tradeoffs leading to that decision. **RESULTS:** Completed surveys were obtained from 45 physicians (82%). The physicians postponed extubation in 37% of the cases. Agreement between any 2 physicians was fair (mean  $\pm$  SD phi  $0.34 \pm 0.15$ ) and was highest between attending physicians from the same institution ( $0.37 \pm 0.15$ ). In deciding to extubate a patient, 33% of the physicians relied on the breathing pattern on pressure support ventilation, 49% relied on the acid-base status, 13% relied on the mental status, and 8% relied on the amount of secretions. The accuracy of the physicians' extubation decisions was low (area under the receiver operating characteristic curve 0.35). The sensitivity of the physicians identifying the patients who were successfully extubated was 57%, and the specificity was 31%. A model that comprises the same variables that influenced the physicians was more accurate in predicting extubation outcome (area under the receiver operating characteristic curve 0.88). **CONCLUSIONS:** For a decision made on an almost daily basis in intensive care units, physicians' extubation decisions are inaccurate and only fairly reliable. *Key words:* extubation; mechanical ventilation; physicians; decision making; intensive care unit; survey. [Respir Care 2011;56(7):920–927. © 2011 Daedalus Enterprises]

## Introduction

Interventions that reduce the duration of mechanical ventilation improve the outcomes of patients with respiratory

failure.<sup>1-3</sup> As a result, physicians attempt to reduce the ventilator support (weaning) as soon as the patient improves, and remove the artificial airway (extubation) once the patient tolerates breathing spontaneously.<sup>4</sup> The rate of extubation failure ranges between 2% and 25%.<sup>5</sup> The mortality of these

---

Dr Tulaimat is affiliated with the Division of Pulmonary and Critical Care Medicine, John H Stroger Jr Hospital of Cook County, Chicago, Illinois. Dr Mokhlesi is affiliated with the Section of Pulmonary and Critical Care, University of Chicago, Chicago, Illinois.

The authors have disclosed no conflicts of interest.

Correspondence: Aiman Tulaimat MD, Division of Pulmonary and Critical Care Medicine, John H Stroger Jr Hospital of Cook County, 1900 W Polk Street, Room 1402, Chicago IL 60612. E-mail: atulaimat@cookcountyhhs.org.

DOI: 10.4187/respcare.00851

---

SEE THE RELATED EDITORIAL ON PAGE 1050

---

patients is significantly higher than the mortality of patients who tolerate extubation,<sup>5-8</sup> so it is desirable to accurately identify patients who are at risk for re-intubation.<sup>9-14</sup>

Some of the variability in the rate of extubation failure is probably due to differences in cohort characteristics, yet differences in physician practice may also explain some of

this variation.<sup>5</sup> In particular, physicians may rely on different sets of variables, interpret variables with certain biases, or make different tradeoffs (in the case of extubation, between prolonged intubation and failed extubation).<sup>15-19</sup>

Similarly, extubation failure occurs when a susceptible patient is extubated by a provider unaware of this susceptibility. Although predictors of extubation failure have been described, the rate of extubation failure remains high.<sup>5,11-14,20,21</sup> The reasoning errors behind extubation failure have received minimal attention.<sup>22,23</sup> Given the common occurrence of extubation and the substantial mortality when it fails, it is a decision of high value.

We believe that by studying the reasoning and decision making process we might uncover cognitive biases and knowledge gaps, and that ultimately patient outcomes might improve if we remedy these biases and gaps. Therefore, we studied the reasoning process by posing the following questions. How reliable and accurate are physicians in deciding to extubate a patient? Does institution affiliation and physician experience influence the reliability of the decision? What patient characteristics influence the decision? How do physicians make the tradeoff between extubation and prolonged intubation?

### Methods

The institutional review board of John H Stroger Jr. Hospital of Cook County approved this survey study and waived the requirement for written informed consent from participants.

### Study Design

We surveyed all the attending physicians and fellows in the divisions of pulmonary and critical care medicine in 3 teaching hospitals in Chicago, Illinois. The attending physicians were board-certified in critical care medicine. The survey was composed of a cover letter, 32 clinical vignettes, and definitions of the clinical variables in the vignettes. The survey was tested for clarity, sent to the participants, and collected personally by us once completed. The cover letter read, in part:

The 32 clinical vignettes in the booklet are real patients who were mechanically ventilated for at least two days on assist control mode and did not undergo withdrawal of support or tracheostomy. These vignettes provide a constellation of clinical and laboratory data that were available to the critical care specialists caring for the patients. All patients tolerated a 2-hour weaning trial on CPAP of 5 cm H<sub>2</sub>O, with 5–7 cm H<sub>2</sub>O of pressure support, and a critical care fellow is considering to extubate them. Imagine yourself as the critical care specialist in charge. After reviewing each vignette, you will

be asked to decide whether you would extubate the patient. If you decide to postpone extubation, please indicate which of the provided clinical or laboratory variables guided your decision. Please list in the order of importance the reasons for postponing extubation and limit them to a maximum of five.

### Vignettes

The 32 vignettes (Table 1) were extracted from a prospective study of the predictors of extubation failure,<sup>11</sup> in which 122 patients were extubated after a successful spontaneous breathing trial, and 16 patients (13%) failed extubation. At the time of data collection, weaning and extubation were not protocol-driven. All 16 patients who failed extubation within 48 hours of extubation were included in the clinical vignettes. The reasons for re-intubation in the original cohort were: 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). The other 16 vignettes were randomly selected from the 106 patients who tolerated extubation.<sup>11</sup> The physicians surveyed were unaware of the results of the study or the extubation outcomes, and were not involved in the care of those patients. The vignettes included information usually available to the physician during work rounds. Each vignette included:

- Demographics: age and sex
- Medical history, including type of intensive care unit (medical, cardiac, surgical, or neurosurgical), clinical diagnoses, Acute Physiology and Chronic Health Evaluation II score, Glasgow coma score during the spontaneous breathing trial, type of nutrition, and use of paralytics or corticosteroids during mechanical ventilation
- Respiratory information, including size of endotracheal tube, reason for intubation, duration of mechanical ventilation, estimation of endotracheal secretions (none, mild, moderate, copious) by the bedside nurse on the day of extubation, breathing pattern, arterial blood gas values, and negative inspiratory pressure. The breathing pattern (tidal volume [ $V_T$ ], frequency, and frequency/ $V_T$  ratio) was evaluated 15 min after switching the patient from assist control mode to CPAP of 5 cm H<sub>2</sub>O and pressure support of 7 cm H<sub>2</sub>O. Arterial blood gases were measured 60 min after switching the patient to CPAP and pressure support.
- Laboratory information, including white-blood-cell count, hemoglobin level, and blood chemistry for that day

ACCURACY AND RELIABILITY OF EXTUBATION DECISIONS BY INTENSIVISTS

Table 1. Characteristics of Patients in the Vignettes\*

	All Patients (n = 32)	Successful Extubation (n = 16)	Failed Extubation (n = 16)
<b>Demographic</b>			
Age (y)	59 ± 19	60 ± 16	57 ± 21
Female (%)	53	56	50
<b>Clinical</b>			
Received corticosteroids (%)	28	19	38
Received paralytic medication (%)	13	19	6
Glasgow coma score ≤ 10 (%) (P = .04)†	22	6	37
APACHE II score	18 ± 9	19 ± 8	20 ± 10
<b>Respiratory</b>			
Ventilator days (d)‡	6 ± 4	7 ± 4	5 ± 3
Moderate to copious secretions (%) (P = .002)†	33	6	62
f (breath/min)	21 ± 8	20 ± 5	22 ± 10
V <sub>T</sub> (mL)	541 ± 173	527 ± 120	556 ± 216
f/V <sub>T</sub>	43 ± 24	40 ± 16	46 ± 30
Negative inspiratory force (cm H <sub>2</sub> O)	40 ± 10	42 ± 9	37 ± 11
pH	7.40 ± 0.07	7.40 ± 0.1	7.39 ± 0.06
P <sub>aCO<sub>2</sub></sub> ≥ 44 mm Hg (%) (P = .03)†	60	38	81
P <sub>aO<sub>2</sub></sub> (mm Hg)	96 ± 23	97 ± 27	94 ± 17
<b>Laboratory</b>			
White blood cell count (1,000/μL) (P = .02)‡	12 ± 6	14 ± 7	9 ± 4
Hemoglobin (g/dL) (P = .008)‡	10 ± 1	10 ± 1	9 ± 1
Albumin (g/dL)	2.0 ± 1.0	2.2 ± 0.6	1.9 ± 0.8
Bicarbonate (mEq/L)	26 ± 6	28 ± 8	26 ± 4
Blood urea nitrogen (mg/dL)	36 ± 31	39 ± 41	34 ± 19
Creatinine (mg/dL)	2 ± 3	3 ± 3	2 ± 3
Phosphate (mg/dL)	3 ± 1	4 ± 1	3 ± 1

± values are mean ± SD.

\* The individual values for each patient were included in the vignettes. The surveyed physicians did not know the extubation outcomes.

† Via chi-square test.

‡ Via t test for difference between the 2 groups.

APACHE = Acute Physiology and Chronic Health Evaluation

f = frequency (respiratory rate)

V<sub>T</sub> = tidal volume

- The physicians were asked the following question after each vignette:

Would you extubate this patient?

Yes

No

If you decide to postpone extubation, please list in the order of importance the reasons, and limit them to a maximum of five.

**Reliability and Its Relationship to Institution Affiliation and Experience**

We assessed reliability by measuring agreement between physicians. Each physician was paired with all other physicians, and agreement in each pair was measured by calculating phi (the chance independent agreement). We did not use kappa because it can underestimate agreement if

the probability of a positive rate becomes extreme.<sup>24</sup> To calculate phi we calculated the odds ratio from a 2×2 table that displays the agreement between 2 observers. The odds ratio is the odds of a positive classification by rater B when rater A gives a positive classification, divided by the odds of a positive classification by rater B when rater A gives a negative classification. Phi is calculated as:

$$\text{Phi} = (\sqrt{\text{odds ratio}} - 1) / (\sqrt{\text{odds ratio}} + 1)$$

Phi ranges from -1 (extreme disagreement) to +1 (extreme agreement). Agreement is considered almost perfect if phi is 0.8-1, substantial if phi is 0.6-0.8, moderate if phi is 0.4-0.6, fair if phi is 0.2-0.4, and poor if phi is less than 0.2.<sup>24</sup>

We compared the phi of the pairs composed of physicians affiliated with the same institution to the phi of pairs composed of physicians affiliated with different institutions. We also compared phi between attending-attending, attending-fellow, and fellow-fellow pairs with one-way analysis of variance. We evaluated the interaction between experience and institutional affiliation with a univariate linear regression analysis in which agreement was the dependent variable and the institutional affiliation and experience of the pair were independent variables.

### Accuracy of the Physicians' Decisions

We defined the physicians' consensus as the percentage of physicians who opted to extubate in each case. Then we determined the accuracy by calculating the area under the receiver operating characteristic (ROC) curve for the consensus. The accepted standard was the known extubation outcome. We compared the accuracy of the attending physicians to the accuracy of the fellows, and to the accuracy of a logistic regression model that comprises the reasons most frequently given by the physicians to postpone extubation. We also calculated sensitivity and specificity. We defined sensitivity as the percentage of successfully extubated patients the respondents correctly decided to extubate. We defined specificity as the percentage of reintubated patients for which the respondents correctly decided to postpone extubation.

### Determinants of the Extubation Decision

We determined the variables that influenced the physicians' decisions with 2 methods: deterministic and regressive. In the deterministic method we tallied the reasons (variables) most frequently given by the physicians to postpone extubation. In the regressive method we performed a 2-step analysis of each physician's responses. First we used univariate analysis to identify the patient-related variables that influenced the extubation decisions. Then we performed a logistic regression to determine which of the variables identified in the univariate analysis predicted the physician's decision to extubate. A variable was considered significant if the *P* value for the Wald statistic was  $\leq .05$ .

### Reasoning Process and Tradeoffs

We divided the decisions to postpone extubation into 2 groups: decisions based on only one reason, and decisions based on 2 or more reasons. We compared the predictive value of those 2 groups. For each physician's responses we determined the percentage of extubations postponed for only one reason (number of extubations postponed for only one reason divided by the total number of postponed

extubations) and the percentage of patients who were extubated (number of extubations divided by the total number of patients). We correlated the percentage of extubations postponed for only one reason with sensitivity, specificity, and percentage of patients extubated. We used these 4 variables (percentage of patients extubated, percentage of extubations postponed for only one reason, sensitivity, and specificity) in a cluster analysis to divide the physicians into 2 subgroups that have different approaches to extubation.

### Results

Forty-five physicians (82%) responded: 23 attending physicians, and 22 fellows. The mean experience was 13 years for the attending physicians, and 2 years for the fellows.

#### Agreement

We generated 990 data pairs: 346 pairs from the same institution, 644 pairs from different institutions, 506 fellow-attending pairs, 253 attending-attending pairs, and 231 fellow-fellow pairs. Agreement between physicians was only fair (mean  $\pm$  SD phi  $0.34 \pm 0.15$ ). Agreement between 2 physicians from the same institution ( $0.36 \pm 0.15$ , 95% CI 0.34–0.37) was higher than agreement between 2 physicians from different institutions ( $0.33 \pm 0.16$ , 95% CI 0.32–0.34, *P* = .008). Agreement between 2 attending physicians ( $0.37 \pm 0.15$ , 95% CI 0.35–0.39) was higher than agreement between 2 fellows ( $0.33 \pm 0.14$ , 95% CI 0.31–0.35) or between a fellow and an attending physician ( $0.33 \pm 0.16$ , 95% CI 0.31–0.34, *P* < .001). The institutional affiliation of the pair (same institution vs different institution) and the experience (attending physician vs fellow) of the pair were independent predictors of agreement in the multiple regression analysis. There was no interaction between these variables.

#### Reasons to Postpone Extubation

The physicians postponed extubation in 532 (37%) of the 1,440 decisions. The 4 most common reasons for postponing extubation (Table 2) were:

- Acid-base status: pH,  $P_{aCO_2}$ , or an interpretation of the acid-base status, such as acute respiratory acidosis
- Breathing pattern: respiratory rate,  $V_T$ , ratio of respiratory rate to  $V_T$ , or minute ventilation
- Mental status: Glasgow coma score
- Secretions: amount of secretions obtained in endotracheal suctioning

Those 4 categories were the first reasons given in 468 cases (88% of postponed extubations). The frequency of

## ACCURACY AND RELIABILITY OF EXTUBATION DECISIONS BY INTENSIVISTS

Table 2. Reasons Given for Postponing Extubation\*

Second Reason	First Reason			
	Acid-Base	Breathing Pattern	Secretions	Mental Status
Acid-base	NA	23 (20)	6 (7)	17 (21)
Breathing pattern	20 (11)	NA	10 (11)	3 (4)
Secretions	7 (4)	1 (1)	NA	26 (32)
Mental status	5 (3)	2 (2)	19 (21)	NA
Total (% with a second reason)	180 (18)	116 (23)	90 (39)	82 (58)

\* The pairs of values are the numbers and (in parentheses) percentages of vignettes in which the respondents listed the second reason (in the row) in addition to the first reason (in the column). For example, when mental status was cited as the first reason to postpone extubation, the physicians also cited acid-base status, breathing pattern, and secretions in 58% of the vignettes.

the other reasons the respondents gave for postponing extubation was too low to allow meaningful analysis.

The reasons the physicians gave were internally valid. When the acid-base status was the first listed reason to postpone extubation, the pH was lower ( $7.33 \pm 0.09$  vs  $7.39 \pm 0.04$ ,  $P < .001$ ) and the  $P_{aCO_2}$  was higher ( $63 \pm 24$  mm Hg vs  $45 \pm 8$  mm Hg,  $P < .001$ ) than when acid-base status was not the first reason. When the breathing pattern was the first reason given, the respiratory rate was higher ( $35 \pm 11$  breaths/min vs  $20 \pm 6$  breaths/min,  $P < .001$ ), and the  $V_T$  was lower ( $428 \pm 107$  mL vs  $535 \pm 199$  mL,  $P < .001$ ) than when breathing pattern was not the first reason. Similarly, when the Glasgow coma score was the first reason, the Glasgow coma score was lower ( $9 \pm 1$  vs  $14 \pm 1$ ,  $P < .001$ ) than when another reason was first. When secretions was the first reason, 100% of patients had moderate to copious secretions, compared to 26% when secretions was not the first reason ( $P < .001$ ).

### Determinants of the Extubation Decision Identified via Regression

Table 3 shows the variables that most frequently influenced the physicians. Acid-base status was the most frequent independent determinant. There were no differences in the distribution of influences between the attending physicians and the fellows.

### Accuracy of the Physicians

The area under the ROC curve for the consensus of the attending physicians was  $0.35 \pm 0.10$ , and was similar to the area under the ROC curve for the consensus of the fellows ( $0.38 \pm 0.10$ ). Both areas were statistically insignificant.

Table 3. Percentage of Physicians for Whom Acid-Base Status, Breathing Pattern, Glasgow Coma Score, and Secretions Were Independent Determinants of the Extubation Decision

	Physicians* (%)
Acid-Base	49
pH	42
$P_{aCO_2}$	7
Breathing Pattern	33
f	17
$V_T$	0
f/ $V_T$	15
Glasgow coma score	13
Secretions	8

\* There is an overlap of these independent determinants, so the percentage values total more than 100%.

f = frequency (respiratory rate)

$V_T$  = tidal volume

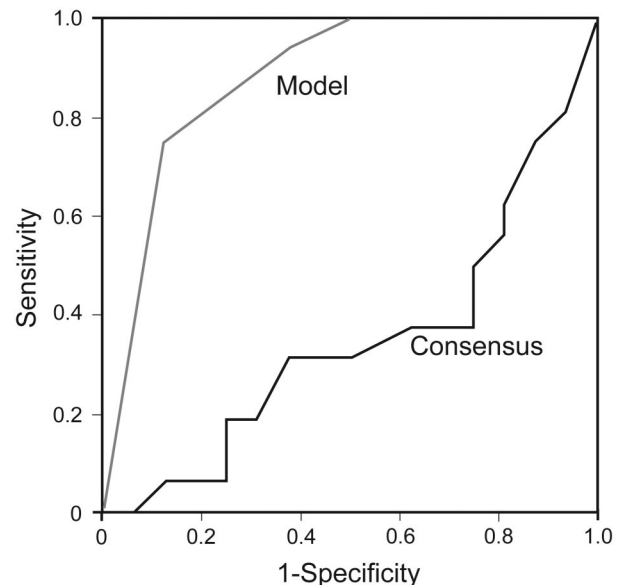


Fig. 1. Receiver operating characteristic (ROC) curves for the accuracy of extubation decisions made by attending physicians (area under the curve 0.35) versus with a logistic regression model (area under the curve 0.88) that incorporates pH,  $P_{aCO_2}$ , secretions, breathing pattern, and mental status.

nificant. In contrast, a logistic regression model that included the variables that most frequently influenced the physicians (breathing pattern, acid-base, mental status, secretions) was predictive of re-intubation and more predictive than the consensus of the attending physicians (Fig. 1). The mean sensitivity and specificity of the physicians were 57% and 31%, respectively. The sensitivity and specificity were similar between the attending physicians and the fellows.

### Number of Reasons to Postpone Extubation

The physicians postponed extubation for only one reason in 36% of the cases, and for 2 or more reasons in the remaining 64%. For example, the acid-base status was the first reason in 180 (34%) of the 532 postponed extubations, and in that subgroup of 180 postponed extubations, breathing pattern ( $n = 20$ ), secretions ( $n = 5$ ), and mental status ( $n = 7$ ) were the most common second reasons (see Table 2). The decision to postpone extubation was more frequently correct when the physicians relied on one reason than when they relied on 2 or more reasons (58% vs 39%,  $P < .001$ ). The percentage of patients who actually failed extubation was higher when acid-base status was the only reason given than if there was also a second reason (80% vs 62%,  $P = .01$ ). The same was true for secretions (40% vs 20%,  $P = .07$ ). That relationship was absent if mental status or breathing pattern was the first reason.

### Tradeoffs Made by the Physicians

The cluster analysis identified 2 physician approaches to extubation: a liberal approach and a conservative approach. The 38 physicians with a liberal approach extubated on average 68% of the cases, had a sensitivity of 62% and a specificity of 27%, and gave one reason in 40% of the postponed extubations. The 7 physicians with a conservative approach extubated 38% of the cases, had a sensitivity of 29% and a specificity of 52%, and gave one reason to postpone extubation in 27% of the postponed extubations.

### Discussion

In our survey responses, the physicians' extubation decisions were unreliable and inaccurate, despite their reliance on some variables that are theoretically or experimentally valid. This is an important finding because it concerns a decision that can affect the risk of death.<sup>25</sup>

The weak agreement between the physicians regarding extubation can probably explain some of the wide variation in the reported rates of extubation failure. The agreement was highest among attending physicians from the same institution, which suggests that experience and institutional biases make practice habits similar. Despite that, agreement was far from good, and these changes might not be clinically relevant. This analysis might also be limited by dependence between the observations. The observed low agreement between physicians may be due to biases in the choice and interpretation of influential variables, absence of a standard practice, or that our method underestimated agreement (see below).

Acid-base status had the strongest influence on the extubation decision. It was the most frequently cited first

reason to postpone extubation and the most frequently identified determinant of the extubation decision. These findings confirm the report by Salam and colleagues that abnormalities of blood gas values were responsible for 50% of the decisions to postpone extubation.<sup>22</sup> We believe that the physicians rely on the acid-base status because of its theoretical validity, despite the lack of experimental validation, and its absence from standard monitoring during weaning in clinical trials.<sup>2,8,10</sup>

The elements of the breathing pattern on the ventilator (respiratory rate and  $V_T$ ) obtained during weaning was the second most influential variable. The breathing pattern is a predictor of extubation success when it is measured with a spirometer before the spontaneous breathing trial.<sup>26</sup> The accuracy of the breathing pattern drops when it is assessed during spontaneous breathing trials on CPAP, with or without low-level pressure support, or when it is used to predict the outcome of a patient who tolerated a spontaneous breathing trial.<sup>20,27,28</sup> Our data suggest that the respondents ignored these facts and relied on the breathing pattern obtained on CPAP with low-level pressure support during a successful spontaneous breathing trial.

Several investigators have evaluated the accuracy of secretions and mental status in predicting re-intubation.<sup>11-13</sup> Less than 13% of our respondents were influenced by secretions and mental status. We speculate that they were skeptical about using secretions or mental status because the published studies that have evaluated them are few and small.

The accuracy of the physicians' extubation decisions was low. Based on the calculated sensitivity and specificity, the rate of failed extubations if this sample was real would be 55%, which is much higher than the reported maximum extubation failure rate of 25%. Therefore, we speculate that what ultimately keeps the extubation failure rate low in clinical practice is not necessarily the clinical acumen of critical care physicians, but the fact that most patients are ready to be extubated by the time the physicians decide to do so. Prior studies found that computerized weaning protocols significantly decreased the duration of mechanical ventilation and led to earlier extubation, without increasing re-intubation rate, compared to usual care based on written guidelines and a systematic approach to weaning.<sup>29,30</sup> Such computerized weaning algorithms decrease variability in extubation practice and could ultimately improve patient outcomes if they were based on valid variables.

We also found that there are 2 groups of physicians. One group (7 physicians) postponed extubation in two thirds of the patients and based that decision on 2 or more reasons in 72% of the cases. These physicians probably perceive the risk associated with extubation failure to be higher than the risk associated with postponed extubation. The other group (38 physicians, 84%) extubated the ma-

jority of the patients, and when they postponed extubation they frequently gave only one reason. These physicians probably perceive the risk associated with postponing extubation to be higher than the risk associated with extubation failure. The liberal extubation group seems to adhere to the rule to extubate unless you have one strong reason not to. However, it is important to note that the accuracy (area under the ROC curve) was similar between the liberal and conservative groups.

Some would contend that our method was unsound because the physicians' accuracy and agreement would have been higher if they had the opportunity to examine the patients. We chose the vignette method as a pragmatic approach to study physicians' extubation decisions, and we believe our method was valid for the following reasons. First, in the study by Ely and colleagues, the physicians were not present at the bedside at the time of the spontaneous breathing trial, and were informed by written communication that a patient passed the trial.<sup>2</sup> Our vignette design was more comprehensive because, in addition to providing the information that the patient had successfully completed a spontaneous breathing trial, it provided the breathing pattern, arterial blood gases, tracheal secretions, and mental status. Second, we know that the data available to the surveyed physicians were adequate to correctly classify the patients with some accuracy, because the logistic regression model based on the same variables that influenced the physicians was accurate. Also, bringing 45 physicians to examine actual patients would be impossible. Therefore, the vignette design is a reasonable and a pragmatic method to answer the questions we posed. Moreover, vignettes with hypothetical cases have been used in the critical care literature to understand physician behavior.<sup>31,32</sup> Our vignettes were from real patients, and the outcome of interest was known to us.

A second criticism of this study is that the population we chose was predestined to show poor predictive accuracy and that the rate of extubation failure in the survey is higher than in clinical practice. However, we believe that the mix did not influence our calculation of reliability and accuracy. Furthermore, we believe that including only 32 cases helped us to obtain a relatively high response rate, and that including all 122 original cases would have reduced the response rate. It is unclear if the results would have been different if all the physicians had responded to the survey.

### Conclusions

The accuracy and reliability of the physicians' extubation decisions was low. The respondents relied on variables of limited value in predicting extubation outcome, such as breathing pattern on low-level pressure support in

patients who successfully tolerated a spontaneous breathing trial, or on variables insufficiently validated as predictors of extubation outcome (acid-base status), and ignored other variables such as mental status and secretions. The role of these variables in the extubation decision needs further study, and, ideally, simple rules of thumb or computerized algorithms should be developed to reliably prevent extubation failure without prolonging the duration of mechanical ventilation.

### REFERENCES

1. 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.
2. 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.
3. Kress JP, Pohlman AS, O'Connor MF, Hall JB. Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. *N Engl J Med* 2000;342(20):1471-1477.
4. Tobin MJ. Advances in mechanical ventilation. *N Engl J Med* 2001;344(26):1986-1996.
5. Epstein SK. Decision to extubate. *Intensive Care Med* 2002;28(5):535-546.
6. Epstein SK. Etiology of extubation failure and the predictive value of the rapid shallow breathing index. *Am J Respir Crit Care Med* 1995;152(2):545-549.
7. Epstein SK, Ciubotaru RL, Wong JB. Effect of failed extubation on the outcome of mechanical ventilation. *Chest* 1997;112(1):186-192.
8. 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. Spanish Lung Failure Collaborative Group. *Am J Respir Crit Care Med* 1999;159(2):512-518.
9. 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.
10. Ely EW, Meade MO, Haponik EF, Kollef MH, Cook DJ, Guyatt GH, et al. Mechanical ventilator weaning protocols driven by nonphysician health-care professionals: evidence-based clinical practice guidelines. *Chest* 2001;120(6 Suppl):454S-463S.
11. Mokhlesi B, Tulaimat A, Gluckman TJ, Wang Y, Evans AT, Corbridge TC. Predicting extubation failure after successful completion of a spontaneous breathing trial. *Respir Care* 2007;52(12):1710-1717.
12. 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.
13. Salam A, Tilluckdharry L, Amoateng-Adjepong Y, Manthous CA. Neurologic status, cough, secretions and extubation outcomes. *Intensive Care Med* 2004;30(7):1334-1339.
14. 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.
15. Eisenberg JM. Sociologic influences on decision-making by clinicians. *Ann Intern Med* 1979;90(6):957-964.

16. Schull MJ, Ferris LE, Tu JV, Hux JE, Redelmeier DA. Problems for clinical judgment: 3. Thinking clearly in an emergency. *CMAJ* 2001; 164(8):1170-1175.
17. Tu JV, Schull MJ, Ferris LE, Hux JE, Redelmeier DA. Problems for clinical judgment: 4. Surviving in the report card era. *CMAJ* 2001; 164(12):1709-1712.
18. Redelmeier DA, Cialdini RB. Problems for clinical judgment: 5. Principles of influence in medical practice. *CMAJ* 2002;166(13): 1680-1684.
19. Redelmeier DA, Ferris LE, Tu JV, Hux JE, Schull MJ. Problems for clinical judgment: introducing cognitive psychology as one more basic science. *CMAJ* 2001;164(3):358-360.
20. Tobin MJ, Jubran A. Variable performance of weaning-predictor tests: role of Bayes' theorem and spectrum and test-referral bias. *Intensive Care Med* 2006;32(12):2002-2012.
21. 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.
22. Salam A, Smina M, Gada P, Tilluckdharry L, Upadya A, Amoteng-Adjepong Y, et al. The effect of arterial blood gas values on extubation decisions. *Respir Care* 2003;48(11):1033-1037.
23. Milic-Emili J. Is weaning an art or a science? *Am Rev Respir Dis* 1986;134(6):1107-1108.
24. Meade MO, Cook RJ, Guyatt GH, Groll R, Kachura JR, Bedard M, et al. Interobserver variation in interpreting chest radiographs for the diagnosis of acute respiratory distress syndrome. *Am J Respir Crit Care Med* 2000;161(1):85-90.
25. Epstein SK, Ciubotaru RL. Influence of gender and endotracheal tube size on preextubation breathing pattern. *Am J Respir Crit Care Med* 1996;154(6 Pt 1):1647-1652.
26. 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.
27. El-Khatib MF, Jamaledine GW, Khoury AR, Obeid MY. Effect of continuous positive airway pressure on the rapid shallow breathing index in patients following cardiac surgery. *Chest* 2002;121(2):475-479.
28. Sassoon CS, Light RW, Lodia R, Sieck GC, Mahutte CK. Pressure-time product during continuous positive airway pressure, pressure support ventilation, and T-piece during weaning from mechanical ventilation. *Am Rev Respir Dis* 1991;143(3):469-475.
29. Bouadma L, Lellouche F, Cabello B, Taille S, Mancebo J, Dojat M, et al. Computer-driven management of prolonged mechanical ventilation and weaning: a pilot study. *Intensive Care Med* 2005;31(10): 1446-1450.
30. Lellouche F, Mancebo J, Jolliet P, Roeseler J, Schortgen F, Dojat M, et al. A multicenter randomized trial of computer-driven protocolized weaning from mechanical ventilation. *Am J Respir Crit Care Med* 2006;174(8):894-900.
31. Escher M, Perneger TV, Chevrolet JC. National questionnaire survey on what influences doctors' decisions about admission to intensive care. *BMJ* 2004;329(7463):425.
32. Jain M, Canham M, Upadhyay D, Corbridge T. Variability in interventions with pulmonary artery catheter data. *Intensive Care Med* 2003;29(11):2059-2062.

This article is approved for Continuing Respiratory Care Education credit. For information and to obtain your CRCE (free to AARC members) visit [www.RCJournal.com](http://www.RCJournal.com)

