

Extubation outcome after a successful spontaneous breathing trial: A multicenter validation of a 3-factor prediction model

YANG LIU¹, YU MU², GUO-QIANG LI¹, XIN YU³, PEI-JUN LI⁴, ZHI-QI SHEN⁵,
HAO-XUN WANG⁶ and LU-QING WEI¹

¹Medical Intensive Care Unit, Pingjin Hospital, Logistics College of The Chinese People's Armed Police Forces, Tianjin 300162; ²Coronary Care Unit, Tianjin Chest Hospital, Teaching Hospital of Tianjin Medical University, Tianjin 300051; ³Surgical Intensive Care Unit, Pingjin Hospital, Logistics College of The Chinese People's Armed Police Forces, Tianjin 300162; ⁴General Intensive Care Unit, Tianjin Chest Hospital, Teaching Hospital of Tianjin Medical University, Tianjin 300051; ⁵Coronary Care Unit, General Hospital of Chinese People's Armed Police Forces, Beijing 100039; ⁶General Intensive Care Unit, Xizang Corps Hospital, Teaching Hospital of Tibet University, Lhasa, Tibet 850000, P.R. China

Received August 27, 2014; Accepted August 3, 2015

DOI: 10.3892/etm.2015.2678

Abstract. The aim of the present study was to validate, and if necessary update, a predictive model previously developed using a classification and regression tree (CART) algorithm for predicting successful extubation (ES) using a new cohort. This prospective cohort study enrolled adults admitted to 10 intensive care units, who had successfully passed a spontaneous breathing trial (SBT) and were considered ready for extubation. After extubation, the patients were followed up for 48 h. The primary outcome measure was ES, defined as the ability to maintain spontaneous unassisted breathing for

>48 h after extubation. The 3-factor CART model was applied to patients in this cohort. The predicted probability of ES for each patient in this validation cohort was calculated based on the original CART model using the Laplace correction method. The performance was assessed by discrimination and calibration. A decision curve analysis was used to assess the clinical net benefit (NB). Extubation failure (EF) occurred in 90/530 patients (17%). Among the 90 patients, 72 (13.6%) were reintubated, while 18 patients remained on rescue noninvasive ventilation within 48 h after extubation. The original CART model showed high discrimination but only moderate calibration with predicted probabilities that were systematically lower than expected. The original CART model was updated, and the updated model preserved excellent discrimination (area under the receiver operating characteristic curve, 0.91; 95% confidence interval, 0.87 to 0.93), but exhibited near-perfect calibration (calibration slope, 1; intercept, 0). Between threshold probabilities of 50 and 80%, the NB of using this updated model is significantly improved compared with the current strategy. The updated CART model may be used to estimate the predicted probability of ES after a successful SBT for individual patients. Applying this model appears to produce a substantial clinical consequence with regard to potential reduction in unexpected EFs.

Correspondence to: Professor Lu-Qing Wei, Medical Intensive Care Unit, Pingjin Hospital, Logistics College of The Chinese People's Armed Police Forces, 211 Chenglin Road, Tianjin 300162, P.R. China
E-mail: lqwei2009@163.com

Abbreviations: APACHE II, acute physiology and chronic health evaluation II; ARDS, acute respiratory distress syndrome; ATC, automatic tube compensation; AUC, area under the receiver operating characteristic curve; CART, classification and regression tree; CI, confidence interval; COPD, chronic obstructive pulmonary disease; DCA, decision curve analysis; ES, successful extubation; EF, extubation failure; FN, false-negative; FP, false-positive; ICU, intensive care unit; MV, mechanical ventilation; NB, net benefit; NIV, non-invasive ventilation; NPV, negative predictive value; OR, odds ratio; $P_{0.1}$, airway occlusion pressure; PPV, positive predictive value; P_t , threshold probability; RSBI, rapid shallow breathing index; $RSBI_1$, RSBI at 1 min of SBT; $RSBI_{30}$, RSBI at 30 min of SBT; $\Delta RSBI_{30}$, ratio of $RSBI_{30}$ to $RSBI_1$ as a percentage; SBT, spontaneous breathing trial; TN, true-negative; TP, true-positive

Key words: decision trees, endotracheal extubation, calibration, decision curve analysis, model updating

Introduction

Discontinuation of mechanical ventilation is an essential issue for critically ill patients and clinicians; however, controversy continues concerning the best approach for conducting this task. The current recommendation is to perform weaning using an evidence-based two-step approach, in which daily evaluation for readiness to wean is followed by a spontaneous breathing trial (SBT) (1,2). Currently, it is recommended that patients who successfully pass the SBT should be extubated if neurological status, excessive secretions, adequate cough and airway obstruction are not problematic (1,3). However, the rate

of extubation failure is high (13.5 to 22%) in patients extubated on the basis of the current strategy (4-6). Delayed and premature discontinuation of mechanical ventilation (MV) have been associated with increased mortality (1,7,8). Therefore, there is an urgent requirement for improved methods of identifying patients who are likely to undergo extubation successfully. Extensive investigations have been conducted to identify predictors of extubation outcome; however, none of the predictors reported to date have demonstrated high accuracy (2,9).

Research into weaning pathophysiology indicates that patients show substantial alterations in numerous physiological variables over time during the SBT (10). Thus, mathematically-based prediction models, which include multiple variables, may provide improved predictive outcomes compared with traditional predictors (11). Recently, a number of clinical models for predicting extubation outcome have been reported in the literature (12-14). Certain models have exhibited good discrimination; however, none has been externally validated. External validation is crucial to determine generalizability, and it is necessary to establish the ability of a model to predict outcome in different settings prior to its use in clinical practice being recommended (15,16).

In a previous study, data from a cohort of mechanically ventilated elderly patients were prospectively analyzed, and used to develop a predictive model using a classification and regressive tree (CART) algorithm (17), also known as a decision tree (18), to predict extubation outcome in patients following a successful SBT. This CART model selected three discriminators, which are readily available in the majority of intensive care units (ICUs), and showed a good discrimination with an area under the receiver operating characteristic curve (AUC) of 0.94.

The aim of the present study was to validate and, if necessary, update this previously developed CART model (17) in a multicenter cohort of adult patients admitted to 10 Chinese ICUs. Furthermore, the present study aimed to determine whether the clinical benefit achieved using this model surpasses the current strategy of extubating all patients who have successfully completed a SBT. Therefore, a net benefit analysis was also conducted (19,20), which is a novel method to quantify the clinical usefulness of a predictive model.

Materials and methods

Study setting and patients. This prospective, validation cohort study was conducted between April 2013 and October 2013 in ten Chinese ICUs in six tertiary hospitals (Pingjin Hospital, Tianjin, China; Tianjin Chest Hospital, Tianjin, China; Institute of Traumatic Brain Injury and Neurology, Tianjin, China; General Hospital of Chinese People's Armed Police Forces, Beijing, China; Xizang Corps Hospital, Lhasa, Tibet Autonomous Region; and Langfang Fourth Peoples' Hospital, Bazhou, China). Consecutive patients >18 years of age that were on MV for >48 h were included if they were considered to be able to undergo an SBT, on the basis of the readiness criteria that are listed in Table I (2). Exclusion criteria are detailed in Fig. 1. The Institutional Review Board at each of the participating sites approved this observational study and written informed consent was obtained from the patients' next of kin.

Table I. Readiness criteria for the spontaneous breathing trial.

1	Improvement in underlying conditions
2	Adequate oxygenation, indicated by $\text{PaO}_2 > 60$ mmHg at $\text{FiO}_2 \leq 0.4$ with an extrinsic positive end-expiratory pressure < 8 cm H_2O
3	Cardiovascular stability (absence of active myocardial ischemia, absence of vasopressor use or < 5 mg/kg/min dopamine or dobutamine; heart rate < 130 beats/min)
4	Body temperature $< 38^\circ\text{C}$
5	Hemoglobin > 8 g/dl
6	Awake or easily arousable
7	Adequate coughing during suctioning and did not require suctioning more often than every 2 h
8	Blood pH ≥ 7.3

PaO_2 , partial pressure of oxygen in arterial blood; FiO_2 , fraction of inspired oxygen.

Study protocol. Patients eligible for enrollment, in a semirecumbent position, were submitted to 60-min SBT immediately. The ventilator (Evita-4 or XL; Dräger, Lübeck, Germany) mode was set to 100% automatic tube compensation (ATC) plus 5 cm H_2O positive end-expiratory pressure and other settings remained unchanged. If no indications of SBT failure (Table II) were observed during the SBT, the trial was considered successful. If the patient had adequate mental status, and the ability to cough and expectorate, extubation was performed. Alternatively, if manifestations of SBT failure were detected during this period, SBT was terminated and MV reinstated using the original settings. Immediate reintubation was performed in the presence of a major clinical event (21,22) (Table II). Rescue therapy with non-invasive ventilation (NIV) was used to avoid reintubation in patients with respiratory failure after extubation (21,22) if immediate reintubation was not necessary (Table II). Respiratory therapists applied NIV (BiPAP Vision; Respironics, Inc., Murrysville, PA, USA) using the S/T mode, following standard procedures (21). The final decisions, regarding whether to conduct extubation, rescue NIV or reintubation, were made solely by the primary team, with no involvement from the research team, according to the clinical protocols of the respective institutions.

Study outcome. After extubation, the patients received follow-up for 48 h. The primary outcome measure was successful extubation (ES), defined as the ability to maintain spontaneous unassisted breathing for >48 h after extubation. Reinstitution of either NIV or invasive MV within 48 h of extubation was considered an extubation failure (EF).

Data collection and definitions of candidate predictors for ES. The original CART model has been described in detail in a previous study (17). The following candidate predictors were required to calculate the probability of ES in the validation set: Rapid shallow breathing index (RSBI) at 1 min of SBT (RSBI_1), change in RSBI at 30 min (ΔRSBI_{30}) of SBT and product of airway occlusion pressure ($P_{0.1}$) and RSBI at 30 min of SBT ($P_{0.1} \times \text{RSBI}_{30}$). $P_{0.1}$ was measured using the

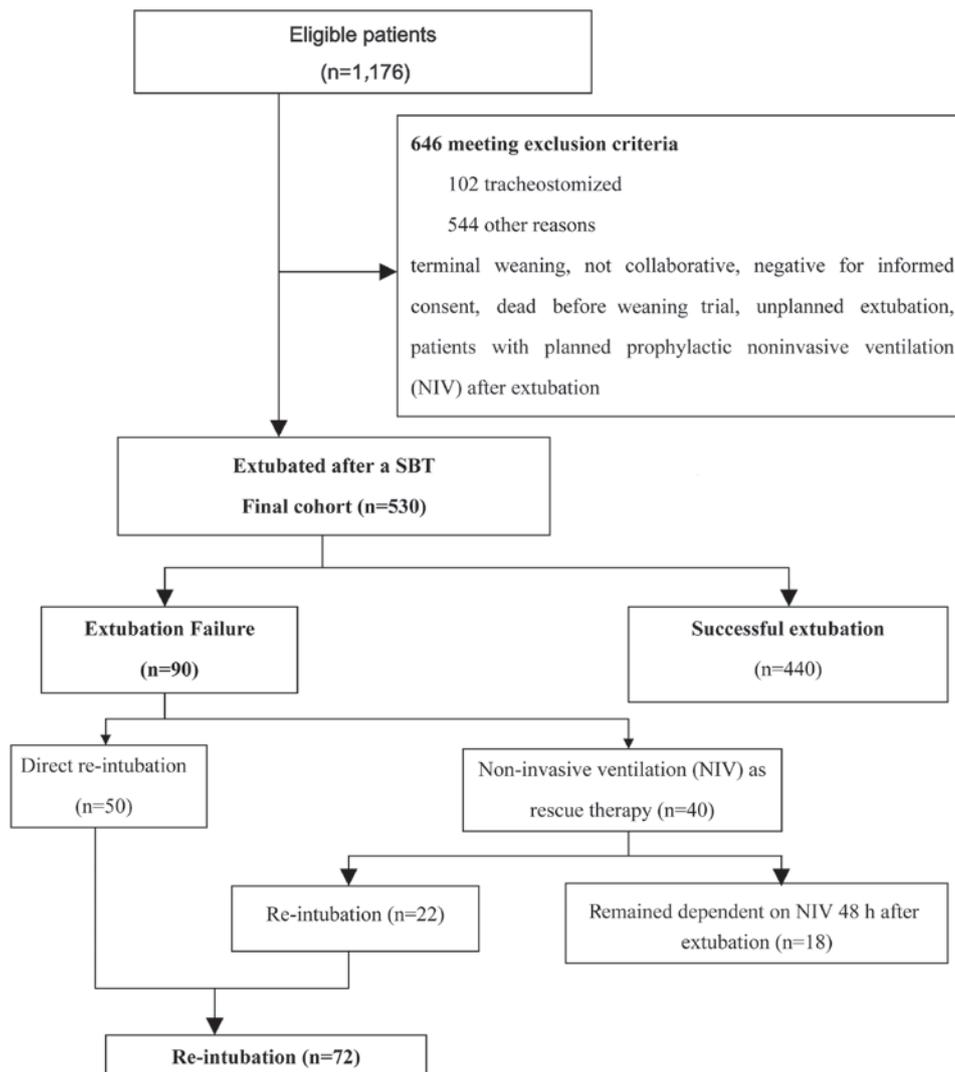


Figure 1. Flow diagram of the study protocol, patient distribution and outcome. SBT, spontaneous breathing trial.

ventilator (Evita-4; Dräger), according to previously described methods (23). RSBI was calculated by dividing respiratory rate (f) by tidal volume (V_t , in liters) and the values were displayed on the ventilator. At least three measurements were obtained, separated by an interval of ≥ 15 sec, and the mean value was used for analysis. Differences in RSBI at 30 min during the 60-min SBT (ΔRSBI_{30}) were assessed using the ratio of RSBI_{30} to RSBI_1 , expressed as a percentage.

Demographic characteristics (such as age, weight, gender and comorbid conditions), acute physiology and chronic health evaluation II (APACHE-II) score on admission, reasons for MV, and duration of MV prior to SBT were recorded. Electrocardiogram, heart rate (HR), arterial blood pressure and oxygen saturation by pulse oximetry (SpO_2) were continuously monitored. Arterial blood samples were collected prior to SBT and at 30 min after extubation, and were immediately analyzed using an ABL 520 blood gas analyzer (Radiometer Medical ApS, Copenhagen, Denmark).

Statistical analysis. Data were presented as the mean \pm standard deviation for continuous and ordinal variables, or medians (interquartile range) if the data were not normally distributed.

For bivariate comparisons, the Wilcoxon test was used for all continuous variables, and Fisher's exact test for categorical variables. As the present study aimed to test the models and single predictors in a new population, the threshold values of the three predictors alone to predict ES were the same as those used or cited in previous relevant studies (2,17,24,25): $P_{0.1} \times \text{RSBI}_{30} \leq 328 \text{ cmH}_2\text{O} \cdot \text{breaths}/\text{min}/\text{l}$; $\Delta\text{RSBI}_{30} \leq 105\%$; $\text{RSBI}_1 \leq 105 \text{ breaths}/\text{min}/\text{l}$.

Model validation. The original CART model was initially validated (17) by assessing its discrimination and calibration (agreement between predicted and actual probabilities of ES) properties in the validation sample. The probability of ES at each terminal node was calculated using the Laplace correction method (26,27) using the derivation cohort (17). Subsequently, the predicted probability of ES was obtainable for each patient in the validation cohort (P_{ori}), as each patient was ultimately allocated to one of the five terminal nodes of the original CART model. Discrimination was assessed by calculating the AUC and the 95% confidence interval (CI) derived from bootstrapping 1,000 samples. In addition, the sensitivity, specificity, positive predictive value (PPV) and negative predictive value

Table II. Criteria for SBT failure, rescue therapy with NIV and immediate reintubation.

Criteria for SBT failure (≥ 1 criterion required)

1. Respiratory rate, >35 breaths/min
2. $\text{SpO}_2 < 90\%$ in spite of increasing FiO_2 to 50% or $\text{PaCO}_2 > 50$ mmHg (or increased by ≥ 8 mmHg)
3. Significant changes in heart rate, >140 beats/min or a 20% increase or decrease from baseline
4. Systolic arterial pressure, ≥ 180 or ≤ 90 mmHg
5. Significant arrhythmia, increased premature ventricular beats of >4 beats/min or new onset of sustained supra-ventricular rhythm (>30 sec)
6. Agitation, anxiety or diaphoresis
7. Clinical signs suggestive of respiratory muscle fatigue, increased work of breathing, or both, such as use of respiratory accessory muscles, paradoxical motion of the abdomen, or retraction of the intercostal spaces.

Criteria for post-extubation respiratory failure that prompted immediate rescue therapy with NIV (≥ 1 criterion required)

1. $\text{PaCO}_2 > 45$ mmHg or $>20\%$ increase from pre-extubation with $\text{pH} < 7.35$
2. $\text{SpO}_2 < 90\%$ with inspired O_2 fraction > 0.5
3. Respiratory frequency, >35 /min
4. Decreased consciousness, agitation or diaphoresis
5. Clinical signs suggestive of respiratory muscle fatigue and/or increased work of breathing, such as the use of respiratory accessory muscles, paradoxical abdominal motion, or retraction of the intercostal spaces

Major clinical events that prompted immediate reintubation (≥ 1 event required)

1. Respiratory or cardiac arrest
2. Respiratory pauses with loss of consciousness or gasping for air
3. Massive aspiration
4. Persistent inability to remove respiratory secretions
5. Heart rate, <50 beats/min with loss of alertness
6. Severe hemodynamic instability (with a systolic blood pressure <90 mmHg despite adequate volume challenge, the use of vasopressors, or both)
7. Lack of improvement in signs of respiratory-muscle fatigue despite use of NIV
8. Lack of improvement in pH or in the partial pressure of carbon dioxide or in the $\text{SpO}_2 (<90\%)$ despite the administration of O_2 and NIV
9. Neurological deterioration
 - a. Psychomotor agitation inadequately controlled by sedation
 - b. Decreased consciousness, rendering the patient unable to tolerate NIV

SBT, spontaneous breathing trial; NIV, non-invasive ventilation; SpO_2 , peripheral capillary oxygen saturation; FiO_2 , fraction of inspired oxygen; PaCO_2 , partial pressure of carbon dioxide in arterial blood.

(NPV) of the model were computed. A true-positive (TP) test result was defined as one that predicted ES and the extubation was successful. A true-negative (TN) test result was defined as one that predicted EF and extubation was unsuccessful. Model calibration was assessed by generating plots of actual (y-axis) and predicted probabilities (x-axis) of ES across quantiles of predicted probabilities (28).

Model updating. To improve the calibration, the original CART model was updated following standardized procedures. These included updating or adjusting the intercept, followed by more extensive updates if necessary, including model revision (20,29). Initially, the simple recalibration method was used. In brief, a logistic regression model was fitted with the predicted probability of the original CART model (P_{ori}) as the only covariate in the validation set:

$$y \sim a + \beta_{\text{overall}}^* \log[\text{odds}(\hat{y})] \quad (\text{Formula 1})$$

where y is the outcome in the validation cohort ($\text{ES}=1, \text{EF}=0$), a is the updated intercept, β_{overall} is the calibration slope, and \hat{y} is the predicted probability of the original CART model (P_{ori}). Therefore, the updated model's predicted probabilities in validation set (P_{up}) were calculated as follows:

$$\log[\text{odds}(P_{\text{up}})] = a + \beta_{\text{overall}}^* \log[\text{odds}(\hat{y})] \quad (\text{Formula 2})$$

Formula 2 can be further transformed into Formula 3:

$$P_{\text{up}} = \frac{e^{(a + \beta_{\text{overall}}^* \log[\text{odds}(\hat{y})])}}{1 + e^{(a + \beta_{\text{overall}}^* \log[\text{odds}(\hat{y})])}} \quad (\text{Formula 3})$$

This updating method, known as the simple recalibration method, preserves the tree structure, but updates the predictions to obtain calibration-in-the-large (updating of intercept), and compensates for any overfitting that may have occurred at model development (β_{overall}) (20).

Table III. Primary causes of extubation failure and reintubation [n (%)].

Cause	Reintubation (n=72)	Extubation failure (n=90)
No improvement in signs of muscle fatigue ^a	25 (34.7)	37 (41.1)
Hypoxemia	27 (37.5)	31 (34.4)
No improvement in respiratory acidosis	5 (6.9)	7 (7.8)
Neurological deterioration ^b	6 (8.3)	6 (6.7)
Excess respiratory secretions	4 (5.6)	4 (4.4)
Hemodynamic instability ^c	5 (7.0)	5 (5.6)

^aSuch as the use of respiratory accessory muscles, paradoxical abdominal motion, or retraction of the intercostal spaces. ^bDefined as the presence of at least one of the following: i) Psychomotor agitation inadequately controlled by sedation; and ii) decreased consciousness, rendering the patient unable to tolerate noninvasive ventilation (NIV). ^cWith a systolic blood pressure <90 mmHg despite adequate volume challenge, the use of vasopressors, or both. Extubation failure includes re-intubated patients and patients who remained on NIV rescue therapy within 48 h following extubation.

In this method, there is no need to adjust the effects of the individual predictors (such as $P_{0.1} \times RSBI_{30}$, $\Delta RSBI_{30}$, $RSBI_1$), which is likely to be the primary advantage of the current method compared with other more extensive updating methods (20). The performance of the updated model (discrimination and calibration) was assessed using the aforementioned methods.

Clinical usefulness. Decision curve analysis (DCA) (19) was used to quantify the clinical usefulness of the updated CART model. Prediction models often provide a result in continuous form, such as the probability of an event from 0 to 100%. To evaluate such a model using decision-analytic methods, the analyst is required to dichotomize the continuous result at a given threshold probability (P_t). At each P_t , the net benefit (NB) of using the model may be calculated using the following formula (19):

$$NB = (TP - wFP) / N \quad (\text{Formula 4})$$

In this formula, TP and FP are the number of patients with true- and false-positive classification, respectively, N is the total number of patients and w is a weight equal to the odds of the threshold probability [$P_t / (1 - P_t)$], being essentially the relative harm of a false-positive (FP) and a false negative (FN) classification. In the current model, patients with predicted probability of ES above and below the P_t were classified as positive and negative, respectively. Thus, the key aspect of DCA is that a single P_t may be used to categorize patients as positive or negative and to weight FP and FN classifications (19). In practice, it is often difficult to define the optimal P_t precisely. A decision curve was constructed, which addresses this problem by assessing the NB of using a model in supporting decision-making across a full range of P_t s (19,30). To compare the NB of the updated model with extubating patients according to the current strategy at a specified P_t , the difference in NB (ΔNB) and corresponding bootstrap CI were computed. If the 95% CI did not include zero, then a statistically significant difference at the 0.05 level was declared (30).

The analyses were performed using SPSS software, version 21.0.0 (IBM SPSS, Armonk, NY, USA), MedCalc

software, version 12.7.0.0 (MedCalc, Mariakerke, Belgium) and R software, version 3.0.1 (The R Foundation for Statistical Computing, Vienna, Austria).

Results

Participating ICUs. The median profile of the 10 participating ICUs was an 18-bed ICU in a tertiary hospital with 708 beds; ICUs included a median of 57 patients in the study. No ICU contributed <5% or >15% of the total sample. During the 6-month study period, 4,539 patients were admitted to the 10 ICUs.

Patient characteristics. As shown in Fig. 1, 530 patients were finally extubated after a successful SBT. A total of 72 patients (13.6%) were re-intubated and 18 patients remained on NIV as rescue therapy within 48 h after extubation. The overall EF rate was 17%, as EF was defined as a requirement for any invasive or noninvasive ventilatory support. The reasons for reintubation and EF are presented in Table III. The case-mix in the present validation sample differed from that of the derivation cohort. Patients in the validation cohort were younger, and presented with an increased APACHE-II score (mean difference, 1.5; 95% CI, 0.45 to 2.58) compared with the derivation cohort. The median duration of MV in the validation set was 8 days (6,11), a significant reduction ($P < 0.001$) compared with 10 days in the derivation set (9,14). Relative to the derivation sample, the external validation sample represented a more diverse population in terms of disease spectrum and settings (two medical ICUs, two surgical ICUs, three general ICUs, one neurosurgical ICU and two coronary care units). Furthermore, the validation sample included a higher proportion of patients with acute respiratory distress syndrome (ARDS; 16.2 vs. 6.6%, $P = 0.02$) and a decreased proportion of patients with chronic obstructive pulmonary disease (COPD; 8.5 vs. 35.2%, $P < 0.001$) compared with the derivation sample (Table IV). Respiratory variables during SBT are presented in Table V.

Model validation. In the validation cohort, the original CART model showed high discrimination (AUC, 0.91; 95% CI, 0.87 to 0.93), but moderate calibration, with a substantial mismatch

Table IV. Demographics and clinical features of the validation cohort.

Parameter	Validation cohort (n=530)
Age (years)	67 (60,72)
Gender [male (%)]	315 (59.4)
APACHE II score	21.5±4.9
PSV (cmH ₂ O)	11.3±1.8
MBP (mmHg)	89.7±8.8
Duration of MV (days)	8 (6,11)
PaO ₂ /FiO ₂ (mmHg)	261±78
PaCO ₂ (mmHg)	41±7
HR (beats/min)	84±15
Extubation failure [n (%)]	90 (17.0)
Reason for MV [n (%)]	
COPD	45 (8.5)
Pneumonia	122 (23.0)
Septic shock	73 (13.8)
ARDS	86 (16.2)
Congestive heart failure	70 (13.2)
Postoperative acute respiratory failure	36 (6.8)
Multiple trauma without brain injury	45 (8.5)
Cardiac arrest	4 (0.8)
Neurological disease	49 (9.2)

Data expressed as the mean ± standard deviation or as the median (interquartile range) if the data were not normally distributed. APACHE II, acute physiology and chronic health evaluation; PSV, pressure support ventilation; MBP, mean blood pressure; MV, mechanical ventilation; PaO₂, partial pressure of oxygen in arterial blood; FiO₂, fraction of inspired oxygen; PaCO₂, partial pressure of carbon dioxide in arterial blood; HR, heart rate; COPD, chronic obstructive pulmonary disease; ARDS, acute respiratory distress syndrome.

between predicted and actual probabilities over the entire range of probabilities (Fig. 2A). A substantial improvement in the calibration was observed when the simple recalibration method (20) was used, obviating a clear requirement for more extensive updating methods. The calibration plot of the updated CART model exhibited an intercept of 0 and a calibration slope of 1, suggesting near-perfect calibration (Fig. 2B). The predicted probability of ES for an individual patient could be estimated using the updated CART model (Fig. 3).

The updated CART model clearly exhibited similar discriminative ability to the original CART model in the validation set, with identical AUC significantly higher compared with any single predictors alone ($P < 0.001$; Table VI). If a predicted probability > 0.49 (determined by Youden index) was used as cut-off point for predicting ES, the updated CART model had a sensitivity of 93.4% (95% CI, 90.7-95.4), a

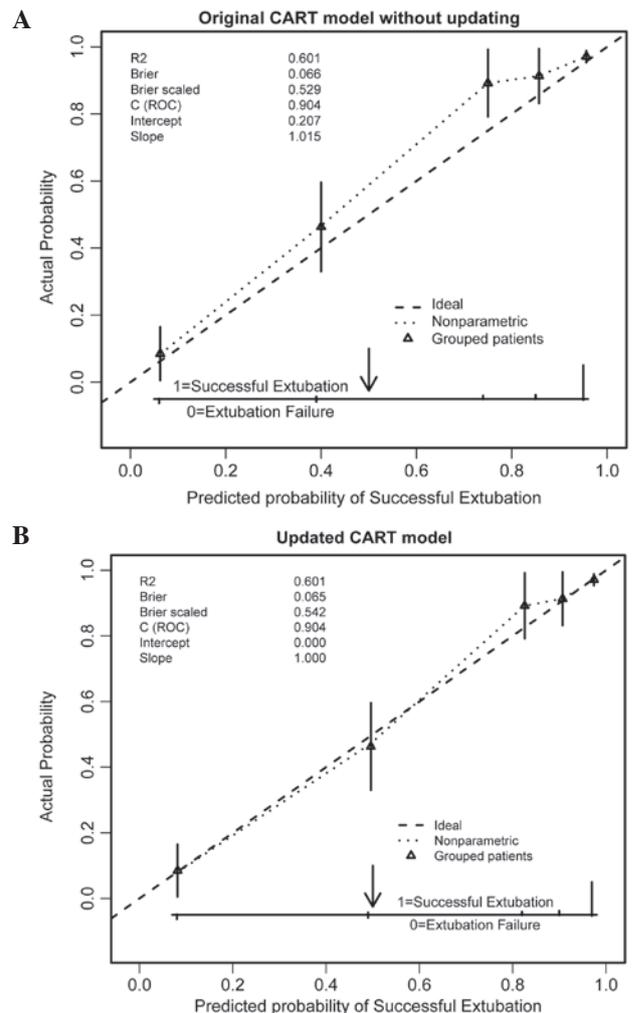


Figure 2. Calibration plots of the CART models for predicting successful extubation in the validation cohort (n=530). Predicted probabilities are on the x-axis and observed outcomes on the y-axis. The triangles indicate the actual probabilities by quintiles of predicted probabilities, with vertical lines representing 95% confidence intervals. The diagonal dashed line represents the ideal calibration (intercept=0, slope=1). (A) Original CART model without updating and (B) the CART model updated with the simple recalibration method. Note that the original CART model's predicted probabilities in the validation set are systematically too low and the resulting updated model showed excellent calibration. CART, classification and regression tree.

specificity of 80.3% (95% CI, 70.6 to 87.1%) and a diagnostic accuracy of 91.1% (95% CI, 88.7 to 93.6) (Table VI).

DCA. The results of the DCA analysis are presented in Fig. 4 and Table VI. The updated CART model displayed an improved NB compared with the 'current strategy', Δ RSBI₃₀, RSBI₃₀ and $P_{0.1} \times$ RSBI₃₀ across a range of clinically plausible Pts (50 to 80%).

Discussion

The accurate prediction of extubation outcome remains challenging. Using a multicenter prospective cohort of mechanically ventilated patients in 10 Chinese ICUs, the previously developed CART model (17) was externally validated to predict ES after a successful SBT, and to propose an updated version of this model to improve calibration. The

Table V. Baseline characteristics and respiratory variables during the SBT in the validation cohort (n=530).

Parameter	Extubation failure (n=90)	Extubation success (n=440)	P-value
Age (years)	69 (65.74)	67 (59.72)	0.04
APACHE II score	21.4±5.4	21.5±4.8	0.447
MBP (mmHg)	88.6±9.3	89.8±8.7	0.17
HR (beats/min)	86.6±15.7	82.8±15.8	0.035
Duration of MV (days)	9 (7,12.3)	8 (6,11)	0.05
1 min of SBT			
RSBI ₁ (breaths/min/l)	91±32	75±36	<0.001
P _{0.1} × RSBI ₁ (cmH ₂ O·breaths/min/l)	336±160	217±145	<0.001
30 min of SBT			
RSBI ₃₀ (breaths/min/l)	107±36	69±31	<0.001
P _{0.1} × RSBI ₃₀ (cmH ₂ O·breaths/min/l)	443±163	218±132	<0.001
ΔRSBI ₃₀ (%)	125±59	99±36	<0.001
60 min of SBT			
RSBI ₆₀ (breaths/min/l)	104±57	67±34	<0.001
P _{0.1} × RSBI ₆₀ (cmH ₂ O·breaths/min/l)	420±251	211±133	<0.001

Data expressed as the mean ± standard deviation or as the median (interquartile range) if the data were not normally distributed. SBT, spontaneous breathing trial; APACHE II, acute physiology and chronic health evaluation; MBP, mean blood pressure; HR, heart rate; P_{0.1}, airway occlusion pressure; P_{0.1} × RSBI, product of airway occlusion pressure and RSBI; RSBI_n, rapid shallow breathing index at n min; ΔRSBI₃₀, change in RSBI at 30 min of SBT, expressed as a percentage of RSBI₃₀ to RSBI₁.

Table VI. Performance of the original and updated models and other single predictors in predicting successful extubation in the validation cohort (n=530).

Variable	RSBI ₃₀	ΔRSBI ₃₀	P _{0.1} × RSBI ₃₀	CART model	
				Original	Updated
AUC	0.69 (0.65 to 0.73) Ref.	0.70 (0.65 to 0.73) z=0.09, P=0.93	0.83 (0.79 to 0.86) z=5.15, P<0.001	0.89 (0.87 to 0.93) ^{c,d} z=7.75, P<0.001	0.89 (0.87 to 0.93) ^{c,d} z=7.75, P<0.001
Sensitivity ^a	85.5 (81.9 to 88.4)	72.7 (68.4 to 76.7)	76.4 (72.2 to 80.1)	93.4 (90.7 to 95.4)	93.4 (90.7 to 95.4)
Specificity ^a	52.2 (42 to 62.2)	65.6 (55.3 to 74.6)	88.9 (80.7 to 93.9)	80.3 (70.6 to 87.1)	80.3 (70.6 to 87.1)
PPV ^a	89.7 (86.5 to 92.3)	91.2 (87.7 to 93.7)	97.1 (94.8 to 98.4)	95.8 (93.5 to 97.3)	95.8 (93.5 to 97.3)
NPV ^a	42.3 (33.6 to 51.6)	33.0 (26.5 to 40.1)	43.5 (36.5 to 50.7)	71.3 (61.8 to 79.2)	71.3 (61.8 to 79.2)
NB (Pt=50%) ^b	0 (-0.68 to 0.83)	0 (-0.02 to 0.01)	0 (-0.47 to 0.57)	8.0 (4.86 to 12.7)	8.0 (4.86 to 12.7)
NB (Pt=75%) ^b	14.5 (6.6 to 22.2)	10.7 (2.0 to 19.4)	25.6 (15.2 to 35.2)	35.3 (26.2 to 44.2)	35.3 (26.2 to 44.2)
NB (Pt=80%) ^b	23.3 (10.8 to 27.7)	21.8 (6.3 to 27.6)	40.7 (22.5 to 45.7)	48.9 (32.4 to 53.3)	48.9 (32.4 to 53.3)

^aThreshold values of each single predictor for predicting successful extubation were: RSBI₃₀, ≤105 breaths/min/l; ΔRSBI₃₀, ≤105%; and P_{0.1} × RSBI₃₀, ≤328 cmH₂O·breaths/min/l end-diastolic volume. ^bCompared with the current strategy: Patients who successfully pass the spontaneous breathing trial (SBT) should be extubated if neurological status, excessive secretions, adequate cough and airway obstruction are not issues. ^cP<0.001 vs. ΔRSBI₃₀ and ^dP<0.001 vs. P_{0.1} × RSBI₃₀. The optimal threshold values of the two models to predict successful extubation were: Predicted probability >0.4 for the original CART model and predicted probability >0.49 for the updated CART model. Values in parentheses are 95% confidence intervals. RSBI₃₀, rapid shallow breathing index at 30 min of SBT; ΔRSBI₃₀, change in RSBI at 30 min of SBT, expressed as the percentage of RSBI₃₀ to RSBI₁; P_{0.1} × RSBI₃₀, product of airway occlusion pressure and RSBI at 30 min of SBT; CART, classification and regression tree; AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value; NB, net benefit; Pt, threshold probability; Ref., reference value for AUC comparisons.

updated CART model outperformed the use of the single predictors alone and the current strategy with regard to the clinical NB.

Validations of predictive models are necessary to verify their generalizability to new sites, as performance in the original data may be optimistic; however, temporal and

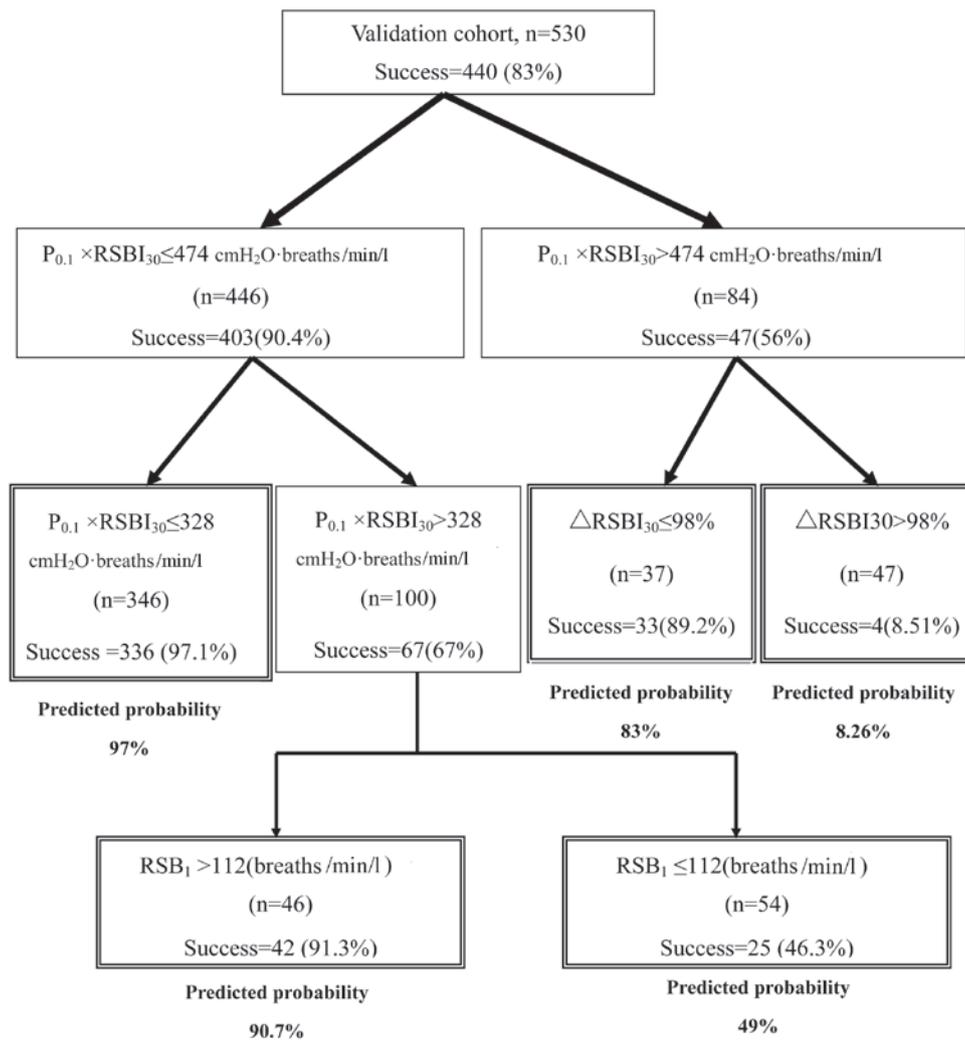


Figure 3. Updated CART model to estimate the predicted probability of successful extubation after a successful spontaneous breathing trial (SBT) of individual patients. Double-framed boxes indicate terminal nodes. The percentage of patients extubated successfully in all nodes are displayed. The predicted probability of successful extubation is shown below each of the five terminal nodes. Success, successful extubation; failure, extubation failure; RSBI, rapid shallow breathing index; $RSBI_1$, $RSBI (f/V_T)$, where f , respiratory rate and V_T , tidal volume) calculated at 1 min of the SBT; $P_{0.1}$, airway occlusion pressure; $P_{0.1} \times RSBI_{30}$, product of $P_{0.1}$ and $RSBI$ calculated at 30 min of SBT; $\Delta RSBI_{30}$, $(RSBI_{30}/RSBI_1 \times 100)$, assessed as the percentage of the $RSBI$ at 30 min ($RSBI_{30}$) to $RSBI$ at 1 min of SBT ($RSBI_1$); CART, classification and regression tree.

external validation studies are limited (16). Prediction models that are not validated are typically not sufficiently developed for clinical application (18). Models predicting ES are usually submitted only to internal validation by the bootstrapping (14) or split-sample method (31). Results are often accepted without sufficient regard to the importance of external validation (16,20), limiting the generalizability of a prediction model to future settings. To the best of our knowledge, the current study is the first to validate a prediction model for extubation outcome based on a multi-center population.

It has been suggested that a formal validation study should consist of an adequate sample of different but related patients compared with the derivation population. This standard required for the validation of a prediction model (20) has been fulfilled in the present study. In this study, the group of patients used for the validation of the CART model differed in numerous aspects, such as age, APACHE-II score and prevalence of disease, from the group of patients used for derivation. This is known as a difference in case-mix. However, this difference

in case-mix, may result in the poor calibration of the original model in the validation population (20,29). As demonstrated in Fig. 2A, the original CART model systematically underestimated the probabilities of ES. This type of miscalibration, with a slope ~ 1 and a non-zero intercept, is a typical result in external validation studies, as described by Steyerberg (20). This result indicates that certain patient characteristics (such as age, APACHE-II score and prevalence of disease), which were not included in the prediction model, were distributed differently between the derivation and validation set (20,29). In such cases, a simple recalibration method is often sufficient and preferable (20,29). As expected, the model updated with this simple recalibration method showed excellent calibration (Fig. 2B), indicating that the updated CART model was able to provide accurate predicted probabilities of ES in the new population.

Newly collected patient data are often used to develop a new prediction models instead of validating an existing model (29,32,33) if a model exhibits worse performance

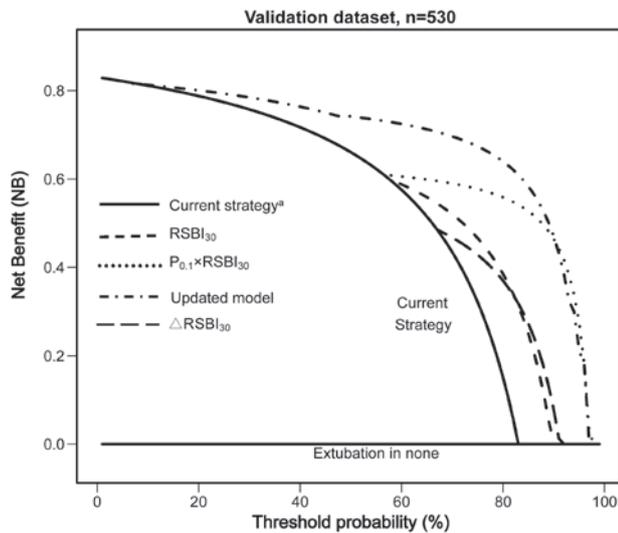


Figure 4. Decision curve analysis (DCA) for prediction of successful extubation (ES; n=530). The net benefit (NB) of the updated CART model and single predictors is plotted against the threshold probability (P_t); patients with predicted probability of ES above the threshold are classified as positive. The zero reference line (extubation in none) represents the strategy of assuming all patients will fail the extubation and thus extubating none of them, whilst the line labeled current strategy represents the extubation of all patients having successfully passed the SBT if neurological status, excessive secretions, adequate cough and airway obstruction were not issues. Δ NB is the difference in net benefit between a specific model (or a specified single predictor) and the current strategy at a single P_t . SBT, spontaneous breathing trial; CART, classification and regression tree; $RSBI_{30}$, rapid shallow breathing index at 30 min of SBT; $\Delta RSBI_{30}$, change in $RSBI$ at 30 min of SBT, expressed as a percentage of $RSBI_{30}$ to $RSBI_1$.

in another population. This may lead to a loss of previous scientific information captured in the previous (development) study, which is counterintuitive to the notion that inferences and guidelines to enhance evidence-based medicine should be based on as much information as possible (34). A superior alternative to redeveloping new models in each new patient sample is to update existing prediction models and adjust or recalibrate them to the local circumstances or setting of the validation sample under investigation (29,34,35). Validation and updating, as conducted in the present study, may lead to more stable and generalizable prediction models, since updated models combine the information captured in the original model with information from new individuals (32,36,37). Therefore, validation and updating may be preferable to the development of new models.

Notably, the updated CART model produced an identical AUC (0.91) to the original one. This similar discriminatory ability may be explained by the simple re-calibration method employed to update the original CART model (20). This updating strategy preserves the tree structure, but re-estimates the predictions of the outcome in each of the five terminal nodes of the original CART model. This updating method involves refitting the intercept to produce an average predicted probability equal to the observed overall event rate and, if necessary, the calibration slope (overall weights of the predictors) (20,29,32). This method would not alter the relative ranking of the predicted probabilities (20,29,34,38), and thus would not affect a model's discrimination. However, the updated CART model, which exhibited near-perfect calibration, remains preferable compared

with the original. Furthermore, the evaluation of calibration is more important if model predictions are used to inform patients or physicians to make clinical decisions (20,30,32).

The original and updated CART models were able to discriminate effectively, with an identical AUC (0.91) that was significantly higher compared with any single predictive factor alone (such as $RSBI$, $P_{0.1} \times RSBI_{30}$) in this validation cohort. In the present study, EF occurred in 17% (90/530) of the extubated patients who completed the SBT and were judged appropriate by their physician for the extubation. If a predicted probability >0.49 (determined by Youden index) was used as cut-off point for predicting ES (Table VI), the updated CART model correctly predicted extubation outcome in 91.1% of patients that successfully passed the SBT, with a sensitivity of 93.4, a specificity of 80.3 and AUC of 0.91, identifying 72/90 EF patients. In total, 11/18 EF patients were not identified by the model; however, EF in these cases occurred for reasons other than acute respiratory failure or respiratory distress, such as neurological deterioration or excess respiratory secretions. Indeed, the etiology of EF influences outcome, with the highest mortality for respiratory failure or respiratory distress (8). Among the EF patients in the present study, 83% exhibited acute respiratory failure or respiratory distress as the reason for EF, which is consistent with a prior study (8). The updated CART model identified 91% of patients with EF resulting from respiratory failure or respiratory distress, which may be clinically useful.

The CART model possesses a number of clear advantages. First, the updated CART model (Fig. 3) is practical as it is simple to use, easy to remember with a graphical presentation, does not require any calculations and may be applied anywhere in an ICU. By contrast, traditional models based on logistic regression are complicated to use, require extensive calculation and may not be practical for application in ICUs that have a large workload. Second, the CART model consists of three clearly defined predictors that do not require any subjective interpretation. Third, the generalizability of the CART model is supported by its external validation and updating in diverse clinical settings. The more numerous and diverse the settings in which the model is demonstrated to be valid, the more likely it is that the model will be applicable to an untested setting (15). Indeed, the updated model combines the information captured in the original model with information from new individuals. Hence, the updated CART model is based on data used in the development and validation studies, further improving its stability and generalizability (20,29). Fourth, the patients included in the present study sample represent a broad disease spectrum, including the majority of diseases leading to invasive MV, including postoperative acute respiratory failure, multiple trauma, cardiac arrest, neurological disease, congestive heart failure, ARDS, septic shock, pneumonia and COPD.

As the CART model has been validated and updated in a new patient population with a varying case-mix, the final CART model for predicting ES appears to be robust, stable, and reliable. We classify our model that was validated across diverse medical settings as level 2 on the hierarchy (15). Level 2 should be, ideally, followed by impact evaluation (level I). The current data suggest that strict application of the updated CART model in the studied patient population would reduce the incidence of EF from 17 to 3.4%, but increase the occur-

rence of unnecessarily delayed extubation from 0 to 5.4%. As unnecessarily delayed extubation and EF are associated with increased mortality, traditional metrics, such as sensitivity, specificity and AUC, are unable to determine whether the CART model is sufficiently accurate to be beneficial in a clinical context (19). Decision-analytic methods incorporate consequences and, in theory, may resolve this ambiguity. For example, if a clinician assumes that the harm associated with an unexpected EF (FP classification) is equal to the harm of unnecessary delay in extubation (FN classification), then $w=1$ and $P_t=0.5$. The NB of 0.08 at a P_t of 50% can be interpreted in terms that extubating patients according to our updated CART model, compared with using the current strategy, leads to the equivalent of a net 8.0 TP results per 100 patients without an increase in the number of FP classifications. The NB formula (19) calculates that this is the equivalent of a net 8.0 fewer false-positive results per 100 patients. In other words, use of the updated CART model may lead to the equivalent of a 8.0% reduction in the number of patients experiencing EF, with no increase in the number of patients with extubation being delayed unnecessarily. It may be difficult to define this threshold in practice; therefore, a full range of P_t s were considered. The updated CART model is improved compared the current strategy across a wide range of P_t s (50 to 80%).

It is crucial to emphasize that the CART model is not able to predict EF caused by airway protection issues. Adding factors of airway competence (such as cough strength and secretion volume) may have improved the performance of the model. However, in recent studies of extubation outcome (6,8,12,22,24,39), the authors reported that retained secretions and/or weak cough accounted for only 6.2% (range, 0 to 9.9%) of the EFs, which is consistent with the present results. The low frequency of abundant secretions in these studies suggests that a number of years after the increased risk of EF in the presence of copious secretions and weak cough was initially reported, physicians are proficient at extracting principles that emerge from research studies and incorporating them into everyday clinical practice. Thus, adding these quantitative measures of cough strength and quantity of endotracheal secretions is unlikely to produce a substantial improvement in the performance of the updated CART model.

A number of limitations of the present study require consideration. First, as with all observational studies on extubation outcome, a selection bias may have occurred, as the predefined criteria for reintubation included certain subjective indices, including 'agitation', 'massive' and 'decreased consciousness'. However, decisions regarding reintubation in the ICUs were made by board-certified intensive care specialists using the standard policy in the institutions. Furthermore, the reintubation and EF rates (17%) in the present study are consistent with those observed in previous studies (4,6,8). Although never subjected to rigorous cost/benefit analyses, reintubation rates of 5 to 20% are generally considered reasonable, and are typical for the majority of effectively managed ICUs (3,40). That the CART model reintubation rate (13.6%) falls within this range suggests that the impact of this selection bias on the results was minimal. A second limitation was that the present study was conducted in a limited number of centers within a single country. The level of generalizability of the present results to other institutions that have differences in case-mix is not clear.

However, the CART model has been validated and adjusted in a new population with a wide spectrum of disease across diverse medical settings, further improving its stability and generalizability. Therefore in theory the CART model may be applicable to other clinical settings with a high degree of confidence (15). However, external validation is a dynamic process and a number of issues remain unresolved. For example, it is not obvious when a model may be considered sufficiently validated and updated. The extent to which this process of model validation and adjustment requires continuation prior to clinical application, will depend on the context (32), and general rules are as yet unavailable. Considering the limited number of centers involved in the present validation study, the updated CART model may require further repeated validation studies in hospitals or institutions of other geographical areas, preferably from different countries, to improve its stability and generalizability. Third, a selection bias may have been present, as the use of ATC has the potential to allow marginal patients to tolerate the SBT, who may subsequently develop ventilatory failure following extubation. However, Elsasser *et al* observed that the built-in commercial ATC may provide adequate inspiratory tube compensation with minimal overassistance (41). This conclusion is further supported by the observation that the reintubation rate in the present study (13.6%) is comparable with that reported from other trials assessing outcome after SBTs (4,6,8).

The original CART model discriminated effectively between patients that were extubated successfully following an SBT and those who were not; however, in the present validation set, calibration was modest. The model was improved using a simple re-calibration method. The updated CART model represents a simple approach that may be used by clinicians to estimate the predicted probability of ES following a successful SBT for individual patients. The clinical consequence of applying the updated CART model appears to be substantial in terms of the potential reduction in unexpected EF. Future studies are required to further validate and, if required, update the CART model in hospitals or institutions of other geographical areas, within China or across countries, that have differences in case-mix, as the reliable prediction of the probability of ES after a successful SBT is of substantial practical value.

Acknowledgements

The authors thank the nursing staff, attending physicians and respiratory therapists at each of the participating sites for their cooperation in this study. The authors also wish to thank Professor Mao-Ti Wei for his statistical guidance on the decision curve analysis.

References

1. Boles JM, Bion J, Connors A, *et al*: Weaning from mechanical ventilation. *Eur Respir J* 29: 1033-1056, 2007.
2. MacIntyre NR, Cook DJ, Ely EW Jr, *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* 120 (Suppl 6): 375-395, 2001.
3. Macintyre NR: Evidence-based assessments in the ventilator discontinuation process. *Respir Care* 57: 1611-1618, 2012.

4. Esteban A, Alia I, Tobin MJ, *et al*: Spanish Lung Failure Collaborative Group: Effect of spontaneous breathing trial duration on outcome of attempts to discontinue mechanical ventilation. *Am J Respir Crit Care Med* 159: 512-518, 1999.
5. Frutos-Vivar F, Esteban A, Apezteguia C, *et al*: Outcome of reintubated patients after scheduled extubation. *J Crit Care* 26: 502-509, 2011.
6. Perren A, Previsdomini M, Llamas M, *et al*: Patients' prediction of extubation success. *Intensive Care Med* 36: 2045-2052, 2010.
7. Coplin WM, Pierson DJ, Cooley KD, Newell DW and Rubenfeld GD: Implications of extubation delay in brain-injured patients meeting standard weaning criteria. *Am J Respir Crit Care Med* 161: 1530-1536, 2000.
8. Thille AW, Harrois A, Schortgen F, Brun-Buisson C and Brochard L: Outcomes of extubation failure in medical intensive care unit patients. *Crit Care Med* 39: 2612-2618, 2011.
9. MacIntyre N: Ventilator discontinuation process: Evidence and guidelines. *Crit Care Med* 36: 329-330, 2008.
10. Jubran A and Tobin MJ: Pathophysiologic basis of acute respiratory distress in patients who fail a trial of weaning from mechanical ventilation. *Am J Respir Crit Care Med* 155: 906-915, 1997.
11. Jubran A, Grant BJ, Laghi F, Parthasarathy S and Tobin MJ: Weaning prediction: Esophageal pressure monitoring complements readiness testing. *Am J Respir Crit Care Med* 171: 1252-1259, 2005.
12. Frutos-Vivar F, Ferguson ND, Esteban A, *et al*: Risk factors for extubation failure in patients following a successful spontaneous breathing trial. *Chest* 130: 1664-1671, 2006.
13. Bien MY, Shui Lin Y, Shih CH, *et al*: Comparisons of predictive performance of breathing pattern variability measured during T-piece, automatic tube compensation and pressure support ventilation for weaning intensive care unit patients from mechanical ventilation. *Crit Care Med* 39: 2253-2262, 2011.
14. Mokhlesi B, Tulaimat A, Gluckman TJ, Wang Y, Evans AT and Corbridge TC: Predicting extubation failure after successful completion of a spontaneous breathing trial. *Respir Care* 52: 1710-1717, 2007.
15. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG and Richardson WS; Evidence-based Medicine Working Group: Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. *JAMA* 284: 79-84, 2000.
16. Altman DG, Vergouwe Y, Royston P and Moons KG: Prognosis and prognostic research: Validating a prognostic model. *BMJ* 338: b605, 2009.
17. Liu Y, Wei LQ, Li GQ, *et al*: A decision-tree model for predicting extubation outcome in elderly patients after a successful spontaneous breathing trial. *Anesth Analg* 111: 1211-1218, 2010.
18. Adams ST and Leveson SH: Clinical prediction rules. *BMJ* 344: d8312, 2012.
19. Vickers AJ and Elkin EB: Decision curve analysis: A novel method for evaluating prediction models. *Med Decis Making* 26: 565-574, 2006.
20. Steyerberg EW (ed): 19 Patterns of External Validity. In: *Clinical Prediction Models: A Practical Approach to Development, Validation and Updating*. Springer, New York, pp333-360, 2009.
21. Ferrer M, Valencia M, Nicolas JM, Bernadich O, Badia JR and Torres A: Early noninvasive ventilation averts extubation failure in patients at risk: a randomized trial. *Am J Respir Crit Care Med* 173: 164-170, 2006.
22. Esteban A, Frutos-Vivar F, Ferguson ND, *et al*: Noninvasive positive-pressure ventilation for respiratory failure after extubation. *N Engl J Med* 350: 2452-2460, 2004.
23. Kuhlen R, Hausmann S, Pappert D, Slama K, Rossaint R and Falke K: A new method for P0.1 measurement using standard respiratory equipment. *Intensive Care Med* 21: 554-560, 1995.
24. Segal LN, Oei E, Oppenheimer BW, *et al*: Evolution of pattern of breathing during a spontaneous breathing trial predicts successful extubation. *Intensive Care Med* 36: 487-495, 2010.
25. Tobin MJ and Jubran A: Weaning from mechanical ventilation. In: *Principles and Practice of Mechanical Ventilation*. 2nd edition. Tobin MJ (ed) McGraw-Hill, New York, NY, pp1185-1220, 2006.
26. Perlich C, Provost F and Simonoff JS: Tree induction vs. logistic regression: A learning-curve analysis. *J Mach Learn Res* 4: 211-255, 2003.
27. Provost F and Domingos P: Tree induction for probability-based ranking. *Machine Learning* 52: 199-215, 2003.
28. Steyerberg EW, Vickers AJ, Cook NR, *et al*: Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology* 21: 128-138, 2010.
29. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE and Vergouwe Y: Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 61: 76-86, 2008.
30. Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ and Van Calster B: Assessing the incremental value of diagnostic and prognostic markers: A review and illustration. *Eur J Clin Invest* 42: 216-228, 2012.
31. Vassilakopoulos T, Routsis C, Sotiropoulou C, *et al*: The combination of the load/force balance and the frequency/tidal volume can predict weaning outcome. *Intensive Care Med* 32: 684-691, 2006.
32. Moons KG, Altman DG, Vergouwe Y and Royston P: Prognosis and prognostic research: Application and impact of prognostic models in clinical practice. *BMJ* 338: b606, 2009.
33. Reilly BM and Evans AT: Translating clinical research into clinical practice: Impact of using prediction rules to make decisions. *Ann Intern Med* 144: 201-209, 2006.
34. Moons KG, Kengne AP, Grobbee DE, *et al*: Risk prediction models: II. External validation, model updating and impact assessment. *Heart* 98: 691-698, 2012.
35. van Houwelingen HC: Validation, calibration, revision and combination of prognostic survival models. *Stat Med* 19: 3401-3415, 2000.
36. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ and Habbema JD: Validation and updating of predictive logistic regression models: A study on sample size and shrinkage. *Stat Med* 23: 2567-2586, 2004.
37. Moons KG: Criteria for scientific evaluation of novel markers: A perspective. *Clin Chem* 56: 537-541, 2010.
38. Toll DB, Janssen KJ, Vergouwe Y and Moons KG: Validation, updating and impact of clinical prediction rules: A review. *J Clin Epidemiol* 61: 1085-1094, 2008.
39. Zapata L, Vera P, Roglan A, Gich I, Ordonez-Llanos J and Betbese AJ: B-type natriuretic peptides for prediction and diagnosis of weaning failure from cardiac origin. *Intensive Care Med* 37: 477-485, 2011.
40. MacIntyre N: Discontinuing mechanical ventilatory support: Removing positive pressure ventilation vs removing the artificial airway. *Chest* 130: 1635-1636, 2006.
41. Elsasser S, Guttman J, Stocker R, Mols G, Priebe HJ and Haberthür C: Accuracy of automatic tube compensation in new-generation mechanical ventilators. *Crit Care Med* 31: 2619-2626, 2003.