RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

Holistic Ultrasound to Predict Extubation Failure in Clinical Practice

Mark E Haaksma, Jasper M Smit, Micah LA Heldeweg, Jip S Nooitgedacht,
Leila N Atmowihardjo, Annemijn H Jonkman, Heder J de Vries, Endry HT Lim,
Thei Steenvoorden, Erik Lust, Armand RJ Girbes, Leo MA Heunks, and Pieter R Tuinman

BACKGROUND: A weaning trial can be considered a stress test of the cardiorespiratory system; it increases oxygen demand and thus warrants a higher cardiac index and elevated breathing effort. We hypothesized that the combination of easily performed ultrasound measurements of heart, lungs, and diaphragm would yield good diagnostic accuracy to predict extubation failure. METHODS: Adult subjects ventilated for > 72 h with a successful spontaneous breathing trial were included. Ultrasound measurements of heart (left ventricular function), lungs (number of B-lines), and diaphragm thickening fraction were performed during a spontaneous breathing trial. The primary outcomes were sensitivity, specificity, and area under the receiver operating characteristic curve of a holistic ultrasound approach for extubation failure. Re-intubation within 48 h was considered extubation failure. RESULTS: Eighty-three subjects were included, of whom 15 (18%) were re-intubated within 48 h. The sensitivity and specificity of a holistic approach were 100% (78.2–100%) and 7.7% (2.5–17.1%), respectively, with an area under the receiver operating characteristic curve of 0.54. The sensitivity and specificity of diaphragm thickening fraction, using a cutoff value of < 30% for extubation failure were 86.7% (59.5–98.3%) and 25.4% (15.5–37.5%), respectively, with an area under the receiver operating characteristic curve of 0.61. CONCLUSIONS: In subjects ventilated for > 72 h who had a successful spontaneous breathing trial, holistic ultrasound was a weak predictor for extubation failure. (ClinicalTrials.gov registration NCT04196361). *Key words: critical care; diaphragm; extubation; holistic; lung; ultrasound.* [Respir Care 0;0(0):1–•. © 0 Daedalus Enterprises]

Introduction

The decision to extubate mechanically ventilated patients is a ubiquitous yet impactful and complex part of critical care medicine. The underlying reason is that prolonged mechanical ventilation and failure to sustain spontaneous breathing (ie, re-intubation) are associated with adverse outcomes including hospital mortality.[1] For this reason, clinicians assess a patient's readiness to be extubated through spontaneous breathing trials (SBTs). However, even after a successful SBT, approximately 13–36% of patients are re-intubated.[2-4] To improve the predictive accuracy of breathing trials, studies have developed and evaluated additional tools such as P_{aO_2}/F_{IO_2}, the Rapid Shallow Breathing Index, and diaphragm thickening measured with ultrasound, with different levels of success.[5-7]

These predictive parameters often evaluate a single organ system and fail to encompass the complex pathophysiology of extubation, which can be considered a stress test of the cardiorespiratory system because it increases oxygen demand and thus warrants a higher cardiac index and elevated breathing effort. Failure may thus result from impaired pulmonary mechanics, cardiac dysfunction, or diaphragm weakness, or from any combination of these factors. Ultrasonography offers a unique possibility for rapid bedside assessment of the aforementioned organs, with minimal additional burden for the patient and improved accuracy compared to chest radiography.[8,9] Even though several studies have linked their ultrasonographic evaluation independently to extubation outcomes, studies using a holistic approach that combines ultrasound of the heart, lungs, and diaphragm are scarce.[10,11]

Therefore, we aimed to test the hypothesis that an integrated ultrasonographic assessment of the diaphragm, heart, and lungs has good diagnostic accuracy for predicting extubation failure in subjects who had a successful SBT.[7,12,13] Preliminary analyses of this study were presented as an abstract at the Annual Congress of the European Society of

RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

Intensive Care Medicine, held October 20–24, 2018, in Paris, France.[14]

Methods

This prospective, observational cohort study was conducted in a 34-bed academic ICU (Amsterdam University Medical Centers, location VUmc, The Netherlands). The protocol was approved by the local ethics board (Medisch Ethische Toetsings Commissie, study number 2016.465) and registered on ClinicalTrials.gov (NCT04196361). The study population consisted of adult subjects ($>$ 18 y old) who were intubated and mechanically ventilated for $>$ 72 h and who had a successful SBT. Sex, age, reason for ICU admission, duration of mechanical ventilation before extubation, Sequential Organ Failure Assessment score on the day of extubation, ventilator settings, inflammatory markers, hemoglobin, and creatinine were recorded. These data were obtained at the bedside or from an electronic patient record upon completion of the exam. The last data available before extubation were used. Informed consent was obtained from subjects or their substitute decision-maker. STROBE guidelines were followed.

The study protocol was the same as the hospital's local weaning protocol, with the addition of holistic ultrasound. Subjects were evaluated on a daily basis for readiness to undergo an SBT. To be deemed eligible, all the following criteria had to be met: reason for ventilatory support reversed or at least under control, adequate neurological status to maintain airway patency, and adequate cough reflex and strength. The SBT was performed under standardized ventilator settings, ie, both pressure support and PEEP were $<$ 10 cm H_2O, with a recommended duration of 60 min. If at the end of an SBT the treating physician deemed it successful, ultrasound measurements were obtained. An SBT was considered successful if none of the following criteria were met: heart rate $>$ 140 beats/min, systolic blood pressure $<$ 80 mm Hg or $>$ 180 mm Hg, peripheral oxygen saturation $<$ 90%, breathing frequency $>$ 35 breaths/min, serious agitation, or excessive diaphoresis.[15] The final decision regarding SBT success and subsequent extubation was made at the discretion of the treating physician, who was blinded to ultrasound results and was not part of the study team. Subjects were only included for the analysis if extubation occurred within 6 h of the ultrasound measurements.

QUICK LOOK

Current knowledge

Extubation is a crucial turning point in critical care medicine, as it requires effective adaptation of the cardiorespiratory system. Failure of any organ involved (ie, heart, lungs, or diaphragm) could lead to extubation failure. Predicting this outcome is crucial, as failure is associated with worse outcomes. Ultrasound offers the possibility to assess the lungs non-invasively. However, the evidence for a holistic approach is sparse and needs further investigation.

What this paper contributes to our knowledge

In subjects with a successful spontaneous breathing trial, holistic ultrasound of heart, lungs, and diaphragm was not a good predictor of extubation failure. This highlights the importance of careful patient selection, timing with regard to spontaneous breathing trial, and consideration of ultrasound parameters used.

Extubation failure was defined as the need for re-intubation or rescue noninvasive ventilation (NIV) within 48 h after extubation; however, no cases of rescue NIV were encountered in the study. Neither high-flow nasal cannula therapy nor postextubation NIV were common practice in our ICU during the study, and thus they were not used in any enrolled subject. Subjects were followed for 48 h after extubation or until they died or were transferred to another ICU within this timeframe.

Ultrasound images were acquired with a Philips CX50 ultrasound machine (Andover, Massachusetts; 100–240 V, 2.65 A, maximum frame rate 755 frames/s, dynamic range 170 dB). A total of 6 investigators (MEH, JMS, LNA, EHTL, PRT) participated in a 2-d ultrasound course and were thereafter supervised by a physician with extensive ultrasound experience ($>$ 5 y) until sufficient expertise was reached (a minimum of 30 exams) before performing any ultrasound for this study.[16] In the first 24 subjects, the intra-rater and inter-rater reliability for diaphragm thickness and

---

The authors are affiliated with the Department of Intensive Care Medicine, Amsterdam University Medical Centers VUmc, Amsterdam, The Netherlands. Drs Haaksma, Smit, Heldeweg, and Tuinman are affiliated with the Amsterdam Leiden Intensive Care Focused Echography (ALIFE), Amsterdam, The Netherlands. Drs Haaksma, Smit, Heldeweg, de Vries, Girbes, Heunks, Tuinman and Ms Jonkman are affiliated with Amsterdam Cardiovascular Sciences Research Institute, Amsterdam University Medical Centers, Amsterdam, The Netherlands.

Dr Haaksma presented a version of this study at the Annual Congress of the European Society of Intensive Care Medicine, held October 20–24, 2018, in Paris, France.

Supplementary material related to this paper is available at http://www.rcjournal.com.

The authors have disclosed no conflicts of interest.

Correspondence: Mark E Haaksma MD, VU University Medical Centre Amsterdam, Postbox 7507, 1007MB, Amsterdam, The Netherlands. E-mail: m.haaksma@amsterdamumc.nl.

DOI: 10.4187/respcare.08679

RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

left ventricular function were calculated. This was not performed for the lung as these have already been calculated in other studies.[17,18] Researchers performed and analyzed all images at the bedside, with the exception of the diaphragm thickening fraction, which was measured directly after completion of the exam. Until completion of all measurements, researchers were blinded to the subject's extubation outcome.

Diaphragm images were acquired with a 4–12 MHz high-frequency linear probe placed on the mid-axillary line in the zone of apposition, as previously described.[19,20] Minimal depth to adequately visualize the diaphragm was chosen to obtain the highest resolution. Only the right diaphragm was studied because in the absence of unilateral paralysis, which was the case in all subjects, it can be visualized more consistently than the left side and it can be assumed that there are no significant differences.[21,22] The left side was briefly visualized to determine whether normal functionality was present. A video of at least 3 respiratory cycles was recorded in B-mode. Then the probe position was marked on the skin, the transducer was lifted from the skin and repositioned at the marked site, and a second and third video of 3 respiratory cycles were captured for a total of 9 respiratory cycles. Offline measurements of end-inspiratory and end-expiratory diaphragm thickness (Tdi) were performed in each set, using the respiratory cycle where the pleural and peritoneal line could be delineated most clearly. This yielded 3 end-inspiratory and 3 end-expiratory values for diaphragm thickness; their averaged values were used in the calculation of diaphragm thickening fraction (TF_{di}):

$$\frac{\text{mean } T_{di\ \text{at end inspiration}} - \text{mean } T_{di\ \text{at end expiration}}}{\text{mean } T_{di\ \text{at end expiration}}}.$$

Lung ultrasound was performed using a 1–5 MHz cardiac or 2–5 MHz abdominal transducer. For both transducers, tissue harmonic imaging was disabled in line with the manufacturer's suggestion. Image depth was set at 16 cm for both transducers on the basis of lung ultrasound acquisition as suggested by Lichtenstein[23] and to ensure standardization of imaging. The transducers were placed transversely on the ribcage so that the image was delineated by ribs on either side and tilted slightly outward to be perpendicular to the lung surface. Images were acquired according to the bedside lung ultrasound in emergency (BLUE) protocol, which entails a 6-point examination (2 anterior points and 1 posterior point per hemi-thorax) and evaluation of lung artifacts including A-lines (horizontal reverberation artifacts generally indicating normal lungs), B-lines (vertical reverberation artifacts indicating loss of aeration due to fluid buildup), and the shred sign (tissue-like appearance of the lung indicating total loss of aeration). The BLUE profile (A profile: A-line predominance; B profile: B-line predominance; A/B profile: even distribution of A-lines and B-lines; and C profile: tissue-like pattern) was determined based on these signs, and the total number of B-lines was counted for each view at the bedside. Additionally, images were made according to the 8-region protocol (ie, 4 quadrants per hemithorax).[24-26]

Cardiac ultrasound was performed using a 1–5 MHz cardiac probe. Settings were freely adjustable by the researcher to acquire best images. Left ventricular function was assessed in the subcostal, apical, parasternal short axis, and parasternal long axis views. Based on contractile activity, cardiac function was classified as poor, moderate, or good, as judged by the researcher.

The primary outcome was the diagnostic accuracy of a holistic ultrasound approach involving the heart, lungs, and diaphragm for detecting extubation failure, expressed as sensitivity, specificity, and area under the receiver operating characteristic curve. For this holistic ultrasound approach, TF_{di}, the total number of B-lines, and rapid visualization of left ventricular function were used a reflection of diaphragm function, pulmonary aeration status, and cardiac function. Presence of pathological findings, such as lowered contractility ($< 30\%$), presence of ≥ 8 B-lines, or impaired or poor cardiac function, were regarded as potentially predictive of extubation failure and were evaluated as such in parallel testing described by Weinstein et al.[27] Secondary outcomes were the diagnostic accuracy of TF_{di} with a previously described cutoff at $TF_{di} < 30\%$ and the correlation of TF_{di}, total number of B-lines, BLUE profile, and left ventricular function with extubation outcome.[13]

We also assessed the association between extubation outcome and several clinical parameters: duration of mechanical ventilation until extubation, ventilator settings, P_{aO_2}/F_{IO_2}, Rapid Shallow Breathing Index score, Sequential Organ Failure Assessment score, and laboratory tests including C-reactive protein, white blood cell count, creatinine, and hemoglobin. Clinical parameters were obtained at the bedside or from a chart review. The last data available before extubation were used.

Statistical Analysis

The sample size calculation was based on an estimated prevalence of 20% for extubation failure, a specificity of 94% for a holistic approach, with a margin of error of 0.06 and a confidence interval of 1.96, which yielded a sample size of 75 subjects.[2,28,29] We enrolled 8 additional subjects to allow for potential missing data or withdrawal of consent. Statistical analyses were performed using SPSS 22 (IBM, Armonk, New York). Variables were tested for normality with the Shapiro-Wilk test, evaluation of histograms, and Q-Q plots. Descriptive statistics are presented as mean ± SD, median (interquartile range), or *n* (%) as appropriate. Differences in baseline characteristics between extubation success and failure groups were tested with an independent-samples *t* test, Mann-Whitney *U* test, or chi-square test as appropriate. There were no missing data except for

RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

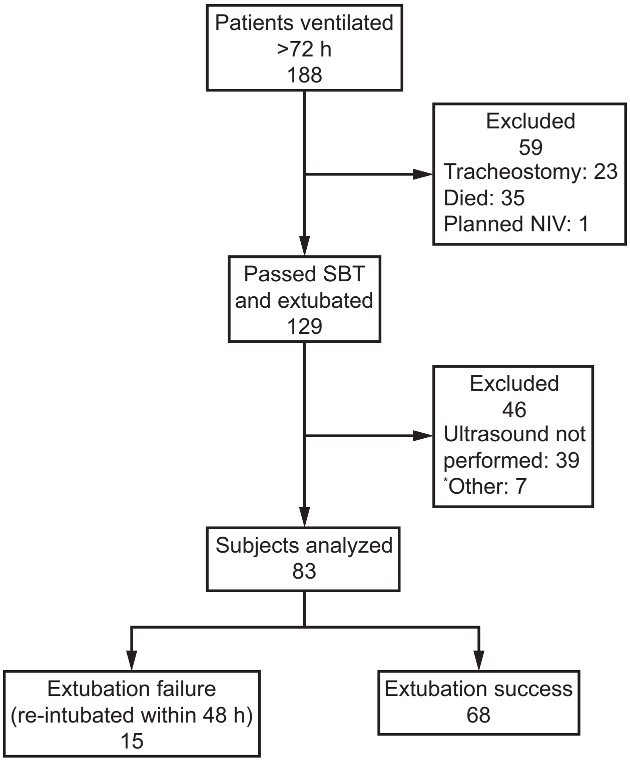

Fig. 1. Flow chart. NIV = noninvasive ventilation; SBT = spontaneous breathing trial. *Other = self extubation, strict isolation, or transfer to another hospital.

measurements of left ventricular function and the number of B-lines, due to inability to acquire images in certain ultrasound image views (see results). Missing cardiac function data were excluded from the regression analysis and for the holistic model, while for the latter the areas not visualized were counted as zero B-lines. The same was done for the subject with a measured thickening fraction of > 300%.

Combined predictive accuracy, expressed as sensitivity and specificity, was calculated as described by Weinstein et al[27] for parallel testing. Variables potentially predictive for extubation outcome were included in the logistic regression analysis. Initially, univariate logistic regression analysis was performed to assess the association of each variable with extubation outcome. For this analysis, left ventricular function was dichotomized (ie, normal vs impaired/poor) due to the low rate of poor ventricular function. Several post hoc analyses were performed. First, the ideal cutoff point (ie, the greatest area under the receiver operating characteristic curve) was determined for the number of B-lines for predicting extubation outcome, as its application as a continuous variable provided too small an effect size to be statistically relevant. In addition, the B-line average per field visualized was calculated to detect the effect of missing data due to surgical dressings. Second, the difference in TF_{di} between the successful and failed group was evaluated when dichotomized for pressure support (0 vs > 0) because we hypothesized that pressure support may have affected outcomes. Third, as a sensitivity analysis, subjects who failed extubation ($n = 3$) due to inability to clear airway secretions were removed because we hypothesized that this type of failure would have been impossible to predict by the chosen ultrasound parameters. Variables with a *P* value < .2 were included in a backward stepwise Wald regression analysis. For the intra-rater and inter-rater reliability analysis, we used measurements of diaphragm ultrasound, and a 2-way mixed intraclass coefficient (ICC) was calculated. The effect of pressure support ventilation and PEEP on the predictive ability of ultrasound-derived measurements was tested through multiple binomial regression. Statistical analyses were performed using 2-sided hypothesis tests; a *P* value of < .05 was regarded as statistically significant.

Results

The study was conducted during 3 separate intervals, which were chosen based on the availability of the scientific personnel who were trained for ultrasound measurements (September 2016 to March 2017, September 2018 to February 2019, and September 2019 to November 2019). Subject enrollment is summarized in Figure 1.

Of 83 subjects, 68 (82%) were extubated successfully and 15 (18%) were re-intubated within 48 h of extubation. Reasons for re-intubation were inability to clear airway secretions ($n = 3$), hypercapnia ($n = 4$), and hypoxemia ($n = 8$). No differences in baseline characteristics were found between the 2 groups (Table 1).

Ultrasound images of the diaphragm were successfully acquired in all subjects (100%), while images of the heart were obtained in all but 2 subjects (97.7%). Of the 664 possible lung-ultrasound images (ie, 8 per subject), 575 (86.6%) were successfully acquired. Surgical wounds and dressings or tubes were limiting factors in this regard.

The intra-rater and inter-rater reliability for diaphragm ultrasound was excellent (ICC intra-rater 1: 0.951 [95% CI 0.907–0.977], ICC intra-rater 2: 0.954 [95% CI 0.913–0.978], ICC inter-rater 1 + 2 on average of 3 measurements: 0.970 [95% CI 0.932–0.987]). The intra-rater and inter-rater reliability for heart ultrasound were identical at kappa 0.833 (95% CI 0.517–1.00). TF_{di}, the total number of B-lines, and cardiac function as measured during the SBT did not differ between the success and the failure group (Fig. 2). A statistically non-significant differences with more B-lines in the failure group was observed (*P* = .11). When TF_{di} was only analyzed in the no-support group, TF_{di} for the failure group versus the success group was 19.3% (95% CI 12–28.3) versus 23.7% (95% CI 19.2–31.2) (*P* = .11), respectively.

Table 1. Baseline Characteristics of Included Subjects

	Overall	Successful	Unsuccessful	*P*
Subjects	83 (100)	68 (82)	15 (18)	
Age, y	65 ± 15	66 ± 16	65 ± 15	.96
Male	60 (72)	51 (75)	9 (60)	.24
SOFA score				
On day of admission	10 (8–13)	10 (8–13)	10 (8–12)	.87
On day of extubation	8 (6–11)	8 (6–10)	7.5 (6–10)	.42
Duration of mechanical ventilation, h	155 (96–212)	138 (96–196)	175 (96–311)	.26
Admission diagnosis				.75
Cardiovascular	21 (25.3)	17 (25.0)	4 (26.7)	
Pulmonary	16 (19.3)	12 (17.6)	4 (26.7)	
Gastrointestinal	18 (21.7)	14 (2.6)	4 (26.7)	
Neurological	13 (15.7)	12 (17.6)	1 (6.7)	
Other	15 (18.1)	13 (19.1)	2 (13.3)	
Pressure support during SBT				.72
0	52 (62.7)	42 (61.8)	10 (66.7)	
> 0	31 (37.3)	26 (38.2)	5 (33.3)	
PEEP, cm H_2O	5.3 (4.8–7.5)	5.2 (4.8–7)	6.9 (4.9–8.1)	.17
F_{IO_2}	0.35 (0.32–0.41)	0.35 (0.31–0.40)	0.40 (0.34–0.45)	.43
Tidal volume, mL	501 (345–620)	498 (391–624)	506 (377–562)	.62
Breathing frequency, breaths/min	28 ± 7	20 ± 6	24 ± 6	.76
Rapid Shallow Breathing Index	46 ± 25	45 ± 26	51 ± 18	.36
P_{aO_2}/F_{IO_2}	204 (155–267)	203 (144–269)	221 (167–259)	.81
Lab				
C-reactive protein, mg/L	65 (38–142)	66 (39–136)	63 (11–266)	.92
White blood cell count, × 10^9/L	12 (10–16)	12 (10–15)	13 (9–20)	.37
Hemoglobin, mmol/L	5.7 (5.1–6.9)	5.6 (5.1–6.9)	6.1 (5.5–6.9)	.32
Creatinine, μmol/L	80 (60–130)	80 (59–29)	76 (61–140)	.95

Data are presented as *n* (%), mean ± SD, or median (interquartile range).
SOFA = Sequential Organ Failure Assessment

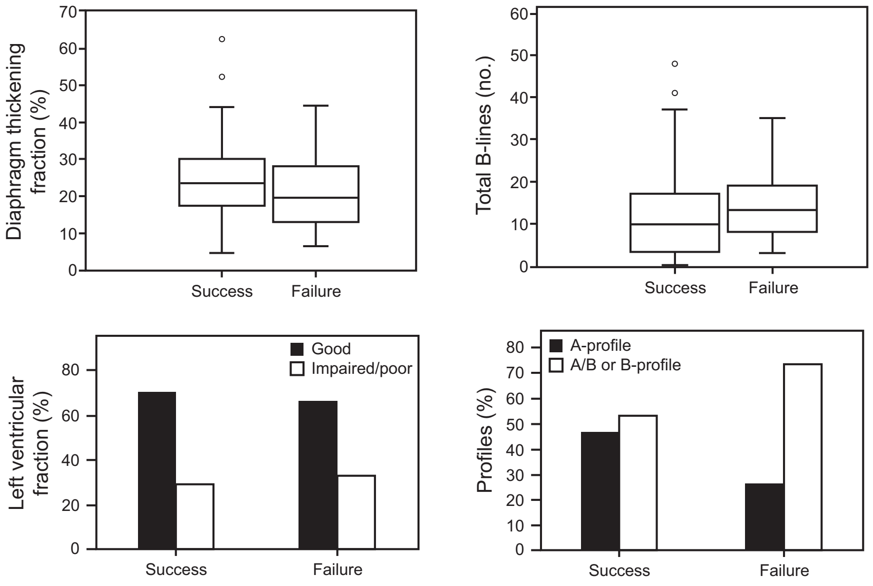

Fig. 2. Overview of ultrasound parameters.

RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

Furthermore, while not statistically significant, a large difference between BLUE profiles was seen between the 2 groups (47.1% A profile vs 26.7% A profile in the successful group and the failed group, respectively) (Table 2). The sensitivity and specificity with a TF_{di} cutoff value of $< 30\%$ for extubation failure were 86.7% (95% CI 59.5–98.3) and 25.4% (95% CI 15.5–37.5), respectively, with an area under the receiver operating characteristic curve of 0.61. The best predictive value for the total number of B-lines was found at a cutoff of ≥ 8 with an 80% (95% CI 51.9–95.7) sensitivity, 50% (95% CI 37.6–62.4) specificity, and area under the receiver operating characteristic curve of 0.63. Calculating the average number of B-lines per field made no difference in its discriminative ability.

The sensitivity and specificity for left ventricular systolic function were 33.3% (95% CI 11.8–61.6) and 74.2% (95% CI 63–84.3), respectively, with an area under the receiving operator characteristic curve of 0.54. Sensitivity and specificity for all 3 variables in a holistic approach, in which the presence of thickening fraction $< 30\%$, ≥ 8 B-lines, or impaired/poor left-ventricular function was regarded as failure, were 100% (95% CI 78.2–100) and 7.7% (95% CI 2.5–17.1), respectively, with an area under the receiver operating characteristic curve of 0.54 (Table 3). Pressure support ventilation and PEEP did not affect the predictive ability of ultrasound parameters (see the supplementary materials at http://www.rcjournal.com).

Table 2. Ultrasound Parameters of Diaphragm, Lungs, and Heart

	Overall	Successful	Unsuccessful	*P*
Subjects	82*	67	15	
Diaphragm thickness				
At end inspiration, mm	1.8 (1.5–2.4)	1.8 (1.5–2.5)	1.7 (1.5–2.3)	.68
At end expiration, mm	1.5 (1.2–2.0)	1.4 (1.2–2.0)	1.5 (1.3–1.6)	.96
Thickness fraction, %	23 (16–29)	23 (17–30)	20 (13–28)	.16
B lines	8 (4–17)	8 (3–17)	13 (8–19)	.11
BLUE profile				.30
A	36 (43.4)	32 (47.1)	4 (26.7)	
B	14 (16.9)	10 (14.7)	4 (26.7)	
A/B	33 (49.8)	26 (38.2)	7 (46.7)	
Heart function				.64
Good	58 (69.9)	48 (7.6)	10 (66.7)	
Impaired	21 (25.3)	16 (23.5)	5 (33.3)	
Poor	2 (2.4)	2 (2.9)	0 (0)	

Data are presented as *n* (%) or median (interquartile range).
* One patient had a calculated thickening fraction of > 300%, which was deemed a measurement error and was excluded from analyses involving diaphragm thickness.
BLUE = bedside lung ultrasound in emergency

When subjects who failed extubation ($n = 3$) due to inability to clear airway secretions were removed, a statistically non-significant difference in number of B-lines was found between the successful and failed group ($P = .07$), while TF_{di} and left ventricular function did not differ ($P = .20$ and $P = .41$, respectively). Univariate logistic regression analysis identified breathing frequency ($P = .12$) and white blood cell count ($P = .20$) as potential risk factors for extubation failure, but these were discarded in the multivariate backward stepwise analysis (Table 4).

Discussion

In this study, we hypothesized that a holistic ultrasound assessment including heart, lung, and diaphragmatic function would have good diagnostic accuracy for predicting extubation failure in subjects who had a successful SBT. There are 2 main findings of our study. First, a holistic ultrasound approach including measurements of the lung, heart, and diaphragm was a weak predictor of extubation failure in subjects who had a successful SBT. Second, a statistically non-significant difference with more B-lines in the failed group was observed when subjects failing to due airway secretions were excluded. The fact that holistic ultrasound was a weak predictor is somewhat surprising. Assessment of multiple organ function related to extubation failure seemed like a logical next step from evaluating only the diaphragm. Our results are also in contrast to 3 other studies that evaluated an extended ultrasound approach and reported cardiac and lung ultrasound to be relevant predictors of postextubation distress and extubation failure.[10,11,30]

Several factors could have led to these discrepant results. First, our study was performed in a general ICU population, whereas previous studies applied more restrictions to inclusion favoring either lower risk subjects (eg, no previously failed SBT, no history of severe COPD)[10] or higher risk subjects (eg, > 65 y old, underlying cardiopulmonary comorbidity).[30] This is also demonstrated by the differences

Table 3. Diagnostic Accuracy of Ultrasound Parameters

	Sensitivity	Specificity	Positive Likelihood Ratio	Negative Likelihood Ratio
Diaphragm	86.7 (59.5–98.3)	25.4 (15.5–37.5)	1.2 (0.9–1.5)	0.5 (0.1–2.0)
B-lines	80 (51.9–95.7)	50 (37.6–62.4)	1.6 (1.1–2.3)	0.4 (0.1–1.1)
Left ventricular function	33.3 (11.8–61.6)	74.2 (62–84.2)	1.3 (0.6–3.0)	0.9 (0.6–1.3)
Holistic	100 (78.2–100)	7.7 (2.5–17.1)	1.1 (1.0–1.2)	

Data are presented as percent (95% CI).

RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

Table 4. Univariable and Multivariable Regression Analysis

	Univariable	P	Multivariable	P
Ultrasound				
Thickening fraction < 30%	0.99 (0.29–3.6)	.99		
B-lines (total)	1.00 (0.99–1.10)	.36		
B-lines ≥ 8	2.00 (0.62–6.47)	.25		
BLUE-profile				
A/b – profile	1.49 (0.36–6.2)	.59		
B – profile	0.46 (0.12–1.76)	.26		
LVF (normal vs impaired/poor)	0.69 (0.21–2.32)	.55		
Scores				
SOFA	0.92 (0.77–1.11)	.39		
Rapid Shallow Breathing Index	1.01 (0.99–1.03)	.45		
Lab				
C-reactive protein, mg/L	1.00 (0.99–1.01)	.30		
White blood cell count, × 10^9/L	1.06 (0.97–1.17)	.20	1.07 (0.97–1.18)	.20
Hemoglobin, mmol/L	1.27 (0.80–2.01)	.32		
Creatinine, μmol/L	1.00 (0.99–1.01)	.76		
Other				
Breathing frequency, breaths/min	1.07 (0.98–1.17)	.12	1.07 (0.98–1.17)	.12
Duration of mechanical ventilation, h	1.00 (1.0 – 1.01)	.28		
$P_{aO_2}/F_{IO_2} < 200$	1.14 (0.37–3.50)	.82		

Data are presented as odds ratios (95% CI). Variates with a P value < .2 in the univariable regression analysis were taken into the multivariable backward stepwise regression analysis.
BLUE = bedside lung ultrasound in emergency
LVF = left ventricular function
SOFA = Sequential Organ Failure Assessment

in failure rate between the studies, ranging from 18–45%. Second, while we evaluated subjects ventilated for > 72 h, the other studies included subjects with a shorter duration of ventilation before inclusion (> 48 h). This difference underscores the previous point of a discrepancy in population in terms of risk of failure, with longer ventilation times increasing this risk. While not statistically significant, a large difference was seen in duration of mechanical ventilation between the successful and failed group. Third, while all studies (including ours) have performed measurements during an SBT, the definition and methods of an SBT vary with regard to the amount of ventilator support and PEEP. This is an essential aspect, as ventilator settings have been reported to have an important impact on respiratory physiology and seldom reflect postextubation circumstances.[31,32] This makes interpretation of ultrasound measurements more difficult and raises the question of which settings are optimal in this context. The same could be argued for conventional prediction tools such as the Rapid Shallow Breathing Index; while this study was not designed to evaluate this index, this value did not differ between groups and would therefore not have been of value to predict extubation outcome. Fourth, the timing of ultrasound with regard to the SBT also differed.

While we performed our measurements at the end of an SBT, other studies did so at the beginning[11] or at both the beginning and the end.[10,30] These differences could impact the results, as was highlighted by one of the studies that reported an increase in B-lines from start to end of an SBT.[33] For the diaphragm and heart, this could also hold true, where breathing effort and thereby TF_{di} may differ after a certain amount of time without supported ventilation. Also, different ultrasound methods and approaches were used to assess the heart, lung, and diaphragm, respectively. For example, while we calculated the total number of B-lines, another study used a loss of aeration score based on B-lines.[30] Finally, the chosen outcome to which the ultrasound measurements were correlated also differed. While we selected extubation failure within 48 h, others studies selected postextubation distress and not re-intubation per se,[10] or selected weaning and extubation failure.[30]

Taken together, we think that it is important to appreciate the differences in design between the current and previous studies on this topic. We do not necessarily believe that holistic ultrasound is a tool that lacks utility; rather we have to carefully evaluate in which population to use it, as well as when and how measurements are performed. For now, we think that caution is warranted when implementing holistic ultrasound until we find the optimal circumstances for its use. Looking ahead, future studies should also look at extending the holistic approach with the assessment of intra-abdominal fluid and of expiratory muscles, as both could impact the ability to be weaned from the ventilator.[20]

RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

Lung ultrasound was not correlated with extubation outcome, which is not in line with previous studies.[34,35] However, a difference was observed for the total number of B-lines and BLUE-profiles between the failed and successful group, albeit statistically not significant. This may be explained by the fact that 95% of our subjects ($n = 79$) received some level of PEEP during ultrasound measurements, which limits the increase in extravascular lung water during the SBT and results in a lower number of B-lines. An alternative explanation is the difference approach for quantifying B-lines. While we calculated the total number of B-lines and BLUE profiles observed during SBT as described in a recent consensus, other studies have used a scoring system based on modified BLUE profiles or the change in the number of B-lines between the beginning and end of the SBT.[25,36] Recent literature suggests that these differences in methodology as well as nomenclature are relevant and potentially modify the final outcome.[17,26,36-38] This is also highlighted by studies reporting that the change in B-line presence during SBT and subtypes of the B-profile (B1 vs B2 profile) correspond better to extubation outcome than previously described measures.[10,30] Regardless of approach, lung ultrasound might be a valuable asset in predicting extubation outcome. Thus, finding the optimal approach for it is of great importance.

An important finding of our study is that TF_{di} is not a good predictor of extubation failure in subjects who had a successful SBT. This is in line with one recent study[39] but not with others.[7,12,13] This might be explained by several factors. First, we only included subjects who had a successful SBT. In this group, diaphragm dysfunction might play a less important role than earlier in the weaning phase and in patients who failed an SBT. In patients who had a successful SBT, other factors such as airway patency, secretions, and cough strength might have a larger impact. This was noted in a recent study that reported ineffective cough as an important risk factor for extubation failure.[39] In addition, part of our study population still received minimal support during ultrasound measurements. Additional support during breathing lowers the effort needed during inspiration and will influence TF_{di}. In this regard, a post hoc analysis showed a difference in TF_{di}, albeit statistically not significant, between the failed and successful group in the no-support group. As mentioned, a recent meta-analysis reported significant differences regarding certain parameters (eg, respiratory work, effort) depending on ventilator settings, some reflecting physiological values better than others.[31] Taking this into consideration, it seems logical to standardize the ventilator settings during measurements to best reflect postextubation circumstances.

Left ventricular systolic function was not correlated with extubation outcome. This is in contrast with previous studies that reported cardiac dysfunction to be correlated with extubation failure and a difficult weaning process.[40,41] We reason that this might be attributed to the fact that we chose a rather pragmatic approach of visualizing left ventricular systolic function, whereas in the mentioned literature more elaborate parameters such as flow/velocity or flow/volume ratios were often used to calculate ejection fraction. However, to our knowledge there are no studies that demonstrate whether these truly offer a significant advantage over a pragmatic approach. In most previous studies, as in our study, only systolic function was used as surrogate for the cardiac component of extubation failure, while recent studies indicate that diastolic dysfunction might also play an important role in weaning failure.[40,41] We chose not to incorporate this into our approach for this study because it requires advanced ultrasound skills that most intensivists do not possess.

The strengths of this study are its relatively large size compared to previous studies and its heterogeneous population, as we did not exclude any subjects based on comorbidities, clinical status (eg, absence of fever, use of vasopressors), or ability to cooperate during measurements.[7,10,11,13,30] This increases its external validity and reliability. Furthermore, we only included subjects who had a successful SBT and were extubated, whereas in other studies failed SBT or palliative sedation were also regarded as failed attempts. Importantly, compared to studies also using a holistic approach, we used TF_{di} instead of motion as a surrogate for diaphragm dysfunction, which is less likely to be influenced by supported ventilation. Lastly, we used a pragmatic approach of ultrasound measurements that was relatively easy to perform, thereby improving clinical applicability.

This study has some limitations that should be acknowledged. First, ultrasound measurements of the diaphragm were only performed on the right side. This is reasonable, as various studies have reported that there are no important differences in healthy volunteers, but this might not be the case for critically ill patients receiving ventilator support.[21,22] Second, ultrasound measurements were assessed in subjects receiving minimal pressure support and PEEP. This might influence results in theory, but did not seem to be relevant in our study.[31] Third, 46 of the 129 eligible patients were not included in the study due to extubation during the evening or weekends; while theoretically possible, it seems unlikely that this had a significant impact on our population. Finally, we chose not to measure diastolic dysfunction, even though it seems to be a good predictor of weaning failure, because it requires advanced ultrasound skills and we set out to find an approach that is easily applicable by any intensivist.

Conclusions

In subjects ventilated > 72 h who had a successful SBT, holistic ultrasound was a weak predictor of successful

RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

extubation. Finding the optimal circumstances for the application of holistic ultrasound to predict extubation failure is vital.

REFERENCES

1. Béduneau G, Pham T, Schortgen F, Piquilloud L, Zogheib E, Jonas M, et al. Epidemiology of weaning outcome according to a new definition: the WIND study. Am J Respir Crit Care Med 2017;195(6):772-783.
2. Thille AW, Richard J-CM, Brochard L. The decision to extubate in the intensive care unit. Am J Respir Crit Care Med 2013;187(12):1294-1302.
3. Liu Y, Mu Y, Li G-Q, Yu X, Li P-J, Shen Z-Q, et al. Extubation outcome after a successful spontaneous breathing trial: a multicenter validation of a 3-factor prediction model. Exp Ther Med 2015;10(4):1591-1601.
4. Longhini F, Maugeri J, Andreoni C, Ronco C, Bruni A, Garofalo E, et al. Electrical impedance tomography during spontaneous breathing trials and after extubation in critically ill patients at high risk for extubation failure: a multicenter observational study. Ann Intensive Care 2019;9(1):88.
5. Yang KL, Tobin M. A prospective study of indexes predicting the outcome of trials of weaning from mechanical ventilation. N Engl J Med 1991;324(21):1445-1450.
6. Saugel B, Rakette P, Hapfelmeier A, Schultheiss C, Phillip V, Thies P, et al. Prediction of extubation failure in medical intensive care unit patients. J Crit Care 2012;27(6):571-577.
7. Farghaly S, Hasan AA. Diaphragm ultrasound as a new method to predict extubation outcome in mechanically ventilated patients. Aust Crit Care 2016;30(1):37-43.
8. Mayo P, Volpicelli G, Lerolle N, Schreiber A, Doelken P, Vieillard-Baron A. Ultrasonography evaluation during the weaning process: the heart, the diaphragm, the pleura and the lung. Intensive Care Med 2016;42(7):1107-1111.
9. Winkler MH, Touw HR, van de Ven PM, Twisk J, Tuinman PR. Diagnostic accuracy of chest radiograph, and when concomitantly studied lung ultrasound, in critically ill patients with respiratory symptoms: a systematic review and meta-analysis. Crit Care Med 2018;46(7):707-714.
10. Silva S, Ait Aissa D, Cocquet P, Hoarau L, Ruiz J, Ferre F, et al. Combined thoracic ultrasound assessment during a successful weaning trial predicts postextubation distress. Anesthesiology 2017;127(4):666-674.
11. Haji K, Haji D, Canty DJ, Royse AG, Green C, Royse CF. The impact of heart, lung and diaphragmatic ultrasound on prediction of failed extubation from mechanical ventilation in critically ill patients: a prospective observational pilot study. Crit Ultrasound J 2018;10(1):13.
12. Ferrari GDF, Elia F, Panero F, Apra F. Diaphragm ultrasound as a new index of discontinuation from mechanical ventilation. Thorax 2014;69(5):431-435.
13. DiNino E, Gartman EJ, Sethi JM, McCool FD. Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. Thorax 2014;69(5):423-427.
14. Haaksma ME, Smit JM, Hilderink B, Atmowihardjo LN, Lim E, Jonkman A, et al. Using ultrasound of heart, lungs and diaphragm to predict weaning success: a prospective observational study. (abstract). Intensive Care Medicine Experimental 2018;6. doi: 10.1186/s40635-018-0201-6.
15. Boles J-M, Bion J, Connors A, Herridge M, Marsh B, Melot C, et al. Weaning from mechanical ventilation. Eur Respir J 2007;29(5):1033-1056.
16. Touw HRW, Tuinman PR. Lung ultrasound: routine practice for the next generation of internists. Neth J Med 2015;73(3):8.
17. Haaksma ME, Smit JM, Heldeweg MLA, Pisani L, Elbers P, Tuinman PR. Lung ultrasound and B-lines: B careful! (letter). Intensive Care Med 2020;46(3):544-545.
18. Touw HR, Parlevliet KL, Beerepoot M, Schober P, Vonk A, Twisk JW, et al. Lung ultrasound compared with chest x-ray in diagnosing postoperative pulmonary complications following cardiothoracic surgery: a prospective observational study. Anaesthesia 2018;73(8):946-954.
19. Haaksma ME, Atmowihardjo L, Heunks L, Spoelstra-de Man A, Tuinman PR. Ultrasound imaging of the diaphragm: facts and future: a guide for the bedside clinician. Neth J Crit Care 2018;26(2):6.
20. Tuinman PR, Jonkman AH, Dres M, Shi Z-H, Goligher EC, Goffi A, et al. Respiratory muscle ultrasonography: methodology, basic and advanced principles and clinical applications in ICU and ED patients—a narrative review. Intensive Care Med 2020;46(4):594-605.
21. Goligher EC, Laghi F, Detsky ME, Farias P, Murray A, Brace D, et al. Measuring diaphragm thickness with ultrasound in mechanically ventilated patients: feasibility, reproducibility and validity. Intensive Care Med 2015;41(4):642-649.
22. Seok JI, Kim SY, Walker FO, Kwak SG, Kwon DH. Ultrasonographic findings of the normal diaphragm: thickness and contractility. Ann Clin Neurophysiol 2017;19(2):131.
23. Lichtenstein D. Whole Body Ultrasonography in the Critically Ill. Berlin Heidelberg: Springer; 2010:153.
24. Lichtenstein DA, Mezière GA. Relevance of lung ultrasound in the diagnosis of acute respiratory failure the BLUE protocol. Chest 2008;134(1):117-125.
25. Volpicelli G, Mussa A, Garofalo G, Cardinale L, Casoli G, Perotto F, et al. Bedside lung ultrasound in the assessment of alveolar-interstitial syndrome. Am J Emerg Med 2006;24(6):689-696.
26. Neuteboom OB, Heldeweg ML, Pisani L, Smit MR, Lagrand WK, Cherpanath TG, et al. Assessing extravascular lung water in critically ill patients using lung ultrasound – a systematic review on methodological aspects in diagnostic accuracy studies. Ultrasound Med Biol 2020;46(7):1557-1564.
27. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. AJR Am J Roentgenol 2005;184(1):14-19.
28. Buderer NMF. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. Acad Emerg Med 1996;3(9):895-900.
29. Malhotra R, Indrayan A. A simple nomogram for sample size for estimating sensitivity and specificity of medical tests. Indian J Ophthalmol 2010;58(6):519-522.
30. Bouhemad B, Mojoli F, Nowobilski N, Hussain A, Rouquette I, Guinot P-G, et al. Use of combined cardiac and lung ultrasound to predict weaning failure in elderly, high-risk cardiac patients: a pilot study. Intensive Care Med 2020;46(3):475-484.
31. Sklar MC, Burns K, Rittayamai N, Lanys A, Rauseo M, Chen L, et al. Effort to breathe with various spontaneous breathing trial techniques: a physiologic meta-analysis. Am J Respir Crit Care Med 2017;195(11):1477-1485.
32. Haaksma ME, Tuinman PR, Heunks L. Weaning the patient: between protocols and physiology. Curr Opin Crit Care 2021;27(1):29-36.
33. Ferré A, Guillot M, Lichtenstein D, Mezière G, Richard C, Teboul J-L, et al. Lung ultrasound allows the diagnosis of weaning-induced pulmonary oedema. Intensive Care Med 2019;45(5):601-608.
34. Tenza-Lozano E, Llamas-Alvarez A, Jaimez-Navarro E, Fernández-Sánchez J. Lung and diaphragm ultrasound as predictors of success in weaning from mechanical ventilation. Crit Ultrasound J 2018;10(1):12.
35. Soummer A, Perbet S, Brisson H, Arbelot C, Constantin J-M, Lu Q, et al. Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress. Crit Care Med 2012;40(7):2064-2072.

RESPIRATORY CARE Paper in Press. Published on April 13, 2021 as DOI: 10.4187/respcare.08679

36. Volpicelli G, Elbarbary M, Blaivas M, Lichtenstein DA, Mathis G, Kirkpatrick AW, et al. International evidence-based recommendations for point-of-care lung ultrasound. Intensive Care Med 2012;38(4):577-591.
37. Mongodi S, Mojoli F, Stella A, Godi I, Via G, Tavazzi G, et al. Lung ultrasound aeration assessment: comparison of two techniques. (abstract). Crit Care 2015;19(Suppl 1):P222.
38. Mongodi S, Stella A, Orlando A, Mojoli F. B-lines visualization and lung aeration assessment: mind the ultrasound machine setting. Anesthesiology 2019;130(3):444.
39. Vivier E, Muller M, Putegnat J-B, Steyer J, Barrau S, Boissier F, et al. Inability of diaphragm ultrasound to predict extubation failure. Chest 2019;155(6):1131-1139.
40. Roche-Campo F, Bedet A, Vivier E, Brochard L, Mekontso Dessap A. Cardiac function during weaning failure: the role of diastolic dysfunction. Ann Intensive Care 2018;8(1):2.
41. Papanikolaou J, Makris D, Saranteas T, Karakitsos D, Zintzaras E, Karabinis A, et al. New insights into weaning from mechanical ventilation: left ventricular diastolic dysfunction is a key player. Intensive Care Med 2011;37(12):1976-1985.