A test that accurately predicted readiness to wean from mechanical ventilation would be useful in intensive care units (ICUs), but is not currently available. Weaning patients from the ventilator involves several stages: first, recognizing patients who are ready to start weaning; secondly, progressively decreasing the support provided by the ventilator; and thirdly, the process of extubation and disconnection from the ventilator.¹

Despite an extensive published literature, the role of respiratory variables as weaning predictors remains controversial, in part because of the low methodological quality of many previous studies.²⁻⁵ Reliable predictors could improve the success of nurse- or respiratory therapist-led weaning, which can reduce ventilation times.⁶ ⁷ Systematic reviews of factors associated with weaning success or failure highlight the need to define carefully the question asked of a physiological test.⁸ ⁹ Specifically, is the test intended to predict a successful reduction in respiratory support, passing a spontaneous breathing trial (SBT), successful extubation, or a composite of these factors? Most published studies failed to conceal test data from clinicians and were open to selection bias.¹⁰ ¹¹ Research using broad non-selected groups of critically ill patients, concealing the test data where possible, is lacking.

We undertook a pragmatic controlled study to answer the question: can respiratory variables predict early weaning in unselected general intensive care patients?
outcomes after reduction in ventilator support in unselected general intensive care patients? We wished to examine variables that have been used before, with contradictory results, such as tidal volume (Vt) and minute volume (MV), respiratory rate (f), and rapid shallow breathing index (f/Vt). We also evaluated inspired–expired oxygen concentration difference (I–E)O2) and end-tidal carbon dioxide concentration (Pco2), which have not been extensively studied but are available with modern monitoring.

Methods

Setting and patients

The study took place in the general ICU of the Royal Infirmary of Edinburgh, Scotland, an 18-bedded ICU that admits adult medical and surgical patients, excluding cardiac surgery. Inclusion criteria were: (i) the patient had received >24 h of mechanical ventilation since ICU admission and (ii) met a simple pragmatic checklist of readiness to start reduction of respiratory support (Table 1). We took these criteria from our previous study, which showed that this checklist could predict eventual successful weaning.1 We excluded patients aged <16 yr, those re-admitted or re-ventilated, those with a primary neurological diagnosis, those planned for terminal care or treatment limitation, and those with a significant leak around the tracheal tube cuff during ventilation (invalidating metabolic and spirometry measurements). We also excluded patients in whom maximum inspiratory support via pressure support ventilation (PSV) was already <15 cm H2O (irrespective of PEEP level), as a semi-arbitrary cut-off value for a low level of mechanical support at baseline. Because we screened all patients from ICU admission, we expected this cut-off to mainly exclude patients with short periods of mechanical support and had confirmed this in earlier work. The study was approved by the local Ethics Committee and consent was obtained from the patient’s nearest relative.

Study protocol

All admissions were assessed for eligibility. Relatives of eligible patients were approached for consent. If consent was obtained, the patient was assessed each day until the weaning checklist was fulfilled. The patient then had a standardized weaning trial as soon as possible after fulfilling the checklist (usually within a few hours).

Standardized weaning trial

In order to evaluate the value of the ventilation and gas exchange variables as a predictive test, we aimed to subject all patients to the same weaning trial at study entry. We anticipated that some patients would be unable to tolerate reduction of all mechanical support in this standardized manner, so we agreed a priori by consensus some ‘abandonment criteria’ to minimize patient distress and ensure patient safety. Before baseline measurements, the level of sedation was adjusted if necessary, so that the patient was awake, co-operative, and not agitated. Sedation then remained unchanged during the trial unless changes were clinically indicated. Tracheal suction was avoided for at least 60 min before starting the study to minimize instability in the physiological measurements. All patients were monitored with continuous ECG, invasive arterial pressure, pulse oximetry, end-tidal carbon dioxide partial pressure (Pco2), expired Vt (Vtexp), expired MV (MVexp), respiratory rate (f), and [I–E]O2. The rapid shallow breathing index was subsequently calculated as the respiratory rate/tidal volume ratio (f/Vt). Respiratory measurements were made using a commercially available spirometry device, the M-COVX module (GE Healthcare, Helsinki, Finland). Data were downloaded to a laptop using a customized software from the monitoring systems.

If PSV was not already being used, patients were switched to this mode. A baseline blood gas analysis was performed at the level of PSV subsequently set. The weaning trial was carried out by research staff and was blinded from the responsible clinicians. The trial comprised a progressive reduction in PSV of 5 cm H2O every 10 min until the trial was either stopped by the investigator because an abandonment criterion was reached or the patient successfully achieved 5 cm H2O continuous positive airways pressure (CPAP) ventilation for 10 min continuously. The predefined ‘abandonment criteria’ were: (i) SaO2 <92% for more than 1 min (on any FiO2 value), (ii) heart rate more than 150 beats min⁻¹ for more than 1 min, (iii) evidence of myocardial ischaemia on the ECG, (iv) new cardiac arrhythmia (new atrial fibrillation, any ventricular arrhythmia other than ventricular ectopics, heart rate <45 beats min⁻¹), and (v) appearing stressed or agitated to an extent deemed unsafe by an investigator, the nurse with the patient, or both.

A blood gas sample was obtained at the end of the trial or at the time of abandonment. The ventilator settings were returned to the pre-trial levels, irrespective of the patient’s performance during the trial, without informing the clinicians responsible for the patient of the study outcome. Decisions about weaning the patient were subsequently made by the clinical team, from whom the trial data were concealed. A log was kept to confirm that all patients received a daily

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cut-off value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PaO2 to FiO2 ratio</td>
<td>&gt;25 kPa</td>
</tr>
<tr>
<td>PEEP</td>
<td>&lt;10 cm H2O</td>
</tr>
<tr>
<td>Haemoglobin</td>
<td>&gt;7 g dl⁻¹</td>
</tr>
<tr>
<td>Axillary temperature</td>
<td>Between 36 °C and 38.5 °C</td>
</tr>
<tr>
<td>Plasma K⁺ concentration</td>
<td>Between 3 and 5.0 mmol litre⁻¹</td>
</tr>
<tr>
<td>Plasma Na⁺ concentration</td>
<td>Between 128 and 150 mmol litre⁻¹</td>
</tr>
<tr>
<td>Inotrope dose</td>
<td>Reduced or unchanged over previous 24 h</td>
</tr>
<tr>
<td>Spontaneous respiratory rate</td>
<td>&gt;6 bpm on current level of support</td>
</tr>
</tbody>
</table>
weaning assessment and ongoing attempts to reduce ventilator support until successfully weaned.

The standardized weaning trial was carried out on only one occasion, as close as possible to the time that the patient first met the checklist criteria. Patients were then followed daily until the study endpoint was fulfilled. The time to achieve the study endpoint was defined as the time in hours from the start of the standardized weaning trial to the time of the start of the 24 h period when a predefined composite endpoint was achieved (see below).

**Study endpoint**

In order to evaluate the predictive value of different cut-off values for the respiratory variables of interest, we needed to define an endpoint that could be converted into a binary outcome. By consensus, we therefore agreed *a priori* that fulfilling one of the following three criteria 24 h after the weaning trial was undertaken represented a low level of mechanical support from the ventilator, such that our patients could subsequently be defined as ‘achievers’ or ‘non-achievers’:

1. a level of CPAP ventilation of ≤10 cm H2O, absolute pressure support ≤10 cm H2O for more than 24 h sequentially starting within 24 h of the weaning trial, or both or
2. extubation within 24 h of the weaning trial which subsequently lasted more than 24 h or
3. a combination of (1) and (2) that started within 24 h of the weaning trial and lasted more than 24 h in total. For example, a period of CPAP ventilation of ≤10 cm H2O lasting 6 h followed by successful extubation lasting >18 h.

This endpoint was chosen rather than extubation, because some patients remain intubated for reasons other than the need for mechanical support such as excessive secretions or for airway protection. Achieving one of the three endpoints was considered to indicate that the patient achieved the work of breathing with low levels of mechanical support or unassisted for a continuous 24 h period.

**Outcome**

We chose an outcome which divided the population into patients who achieved the study endpoint starting within 24 h from the start of the standardized weaning trial (‘achievers’) or did not (‘non-achievers’). This was chosen as an indication of ‘rapid weaners’ (<24 h from the test) and ‘delayed weaners’ (more than 24 h from the test), which we considered a clinically relevant way of dichotomizing the population for analysis. A test that predicted this outcome could be useful to guide nurse- or respiratory therapist-led ventilator adjustments.

**Study power**

We chose f/Vt, [1–E]O2, and ECO2 as our main variables of interest. From our previous study, we estimated a pre-test probability of 40% patients not achieving the study endpoint 24 h after the weaning trial.1 We aimed to have data for 30 patients who did not achieve the study endpoint within 24 h in order to have 10 events of interest for each of the main variables of interest. We estimated that 75 patients would need to be studied. Assuming a 50% recruitment rate, we estimated that this would require 30–35 weeks to complete.

**Statistical analysis**

**Comparison of achievers and non-achievers**

The distribution of continuous data was tested for normality and parametric or non-parametric tests used as appropriate. Categorical values were compared by χ² test, except when small size required Fisher’s two-tailed exact test. A P-value of <0.05 was considered statistically significant.

For all analyses, the physiological data we used as a predictive test were the mean value for the last 5 min of the weaning trial (CPAP 5 cm H2O, unless an abandonment criterion was met, when the last 5 min before abandonment was used). A true positive occurred when a patient’s test predicted the study endpoint within 24 h from the end of the weaning trial, and this occurred. Conversely, a true negative occurred when a patient’s test predicted failure to achieve the study endpoint within 24 h from the end of the weaning trial and the patient failed to achieve it. False positive and negative rates were calculated as appropriate. Standard formulae were used to calculate sensitivity, specificity, positive and negative predictive values, and likelihood ratios of the variables under examination.

We calculated the observed pre-test probability for the study cohort as the proportion of patients who actually achieved the study endpoint within 24 h. This was then used as a reference value to see if any of the test variables could significantly alter this probability if used as a predictive test. We calculated the changes to pre-test probability using the positive and negative likelihood ratios for different cut-off values for the variables examined, as recommended by Jaeschke and colleagues.12 Large changes to the observed pre-test probability when applying a potential test variable, for example, a particular cut-off value for expired tidal volume or respiratory frequency, indicate a potentially clinically useful test.

We used two approaches to explore the predictive value of different cut-off values. First, values were chosen that divided the patients into three groups which included ~30% of the population in each of the three ranges of values. Large changes in pre-test probability for a particular variable cut-off indicated a potentially clinically useful test within this range. Secondly, receiver operating characteristic (ROC) curves were generated to determine the capacity of the indices to predict the study endpoint. The probability of identifying the study endpoint correctly was calculated using the area under the ROC curve (AUC), assuming that this was not normally distributed, under non-parametric assumptions. The cut-off with best predictive power was calculated, and the test performance for this cut-off was calculated.
Results

Patients

Six hundred and one patients were admitted during the study period: 146 were never ventilated and 224 required ≥24 h of mechanical ventilation. Of the 232 patients who received more than 24 h of mechanical ventilation, 97 fulfilled the checklist of readiness to start reduction in support and were eligible for the study. A flow diagram accounting for patients admitted during the study period is shown in Figure 1.

Seventy-three of the eligible patients were successfully recruited to the study (75% enrolment rate). Of the 73 patients included in the study, 32 (17 males) achieved our study endpoint and 41 (26 males) failed to achieve the study endpoint within 24 h of the weaning test. This gave a pre-test probability for the study cohort of achieving the study endpoint within 24 h of the weaning test of 44%. We achieved the intended outcome for the power calculation, namely >30 patients failing to achieve the study endpoint within 24 h.

Comparison of achievers and non-achievers

The characteristics of achievers and non-achievers at study entry are shown in Table 2. The patients in the non-achiever group were older, had higher illness severity at ICU admission (APACHE II score), and had higher organ failures (SOFA) at study entry. There was a non-significant trend for non-achievers to have higher sedation scores at the time of study entry. The non-achievers had a clinically important and statistically significantly greater H+ before starting the standardized weaning trial; all other respiratory and acid–base values were similar between achievers and non-achievers.

In the 73 patients studied, the weaning test had to be abandoned in 19 before achieving CPAP ventilation at 5 cm H₂O. The reasons for stopping are shown in Table 3. Although more patients who met the criteria for stopping were subsequent non-achievers, this did not reach statistical significance. Of the 73 patients, 30 had undergone a tracheostomy at the time of inclusion; of these, seven (21%) were achievers. Of the patients without tracheostomy at enrolment, 24 were achievers (60%).

The physiological variables at the end of the weaning trial for achievers and non-achievers are compared in Table 4. Using univariate comparison, only four variables distinguished the groups before applying test cut-off values. The mean $\text{S}_\text{a}O_2$ was slightly less in the non-achiever group.

---

**Table 2**

<table>
<thead>
<tr>
<th>Category</th>
<th>Achievers</th>
<th>Non-Achievers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>32</td>
<td>41</td>
</tr>
<tr>
<td>APACHE II Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOFA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sedation Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H^+$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3**

<table>
<thead>
<tr>
<th>Reason for Stopping</th>
<th>Achievers</th>
<th>Non-Achievers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abandoned before achieving CPAP</td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td>Tracheostomy at enrolment</td>
<td>30</td>
<td>70</td>
</tr>
</tbody>
</table>

**Table 4**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Achievers</th>
<th>Non-Achievers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean $\text{S}_\text{a}O_2$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Fig 1. Flow diagram for all patients admitted during the study period.
relevant greater [I–E]O2 difference in the non-achiever baseline. There was a statistically significant and clinically end of the trial, most likely a reflection of the difference at analysis remained higher in the non-achiever group at the 330 excess of hypoxaemia. The H+ observed were not clinically important, and there was no although this reached statistical significance, the differences observed were not clinically important, and there was no excess of hypoxaemia. The H+ concentration on blood gas analysis remained higher in the non-achiever group at the end of the trial, most likely a reflection of the difference at baseline. There was a statistically significant and clinically relevant greater [I–E]O2 difference in the non-achiever group. In addition, there was a strong trend towards a higher P\textsubscript{ECO2} among non-achievers. All other respiratory variables were similar between the groups.

When cut-off values were used to include ∼30% of patients in each cut-off group, none of the cut-off ranges had clinically important predictive power to distinguish achievers from non-achievers (see Supplementary material for full data set). Using this approach, the only variables that had statistically significant predictive value were [I–E]O2 difference and P\textsubscript{ECO2}, but the maximum change to the pre-test probability of 44% was <20% for all cut-off ranges (see Supplementary material).

Using ROC analysis, we found no predictive value for MV\textsubscript{exp}, TV\textsubscript{exp}, f, or f/Vt. The AUCs were as follows: MV\textsubscript{exp} 0.59 [95% confidence interval (CI): 0.46–0.72; P=0.18], TV\textsubscript{exp} 0.51 (0.28–0.65; P=0.86), f 0.46 (0.33–0.59; P=0.54), and f/Vt 0.48 (0.35–0.62; P=0.79). The AUC for P\textsubscript{ECO2} was 0.63 (0.50–0.76; P=0.05) and for [I–E]O2 was 0.64 (0.51–0.77; P=0.03). The ROC curves for P\textsubscript{ECO2} and [I–E]O2 are shown in Figure 2. These indicated statistically significant, but clinically poor discrimination between achievers and non-achievers.

Using the best cut-off from sensitivity analysis, the predictive power of [I–E]O2 and P\textsubscript{ECO2}, together with the non-significant variables, was poor (see Supplementary material for full data set). The sensitivity and specificity of the [I–E]O2 and P\textsubscript{ECO2} cut-offs were poor. The positive and negative likelihood ratios were statistically significant, but only generated modest changes to the pre-test probability that were similar to those observed with the best arbitrary cut-offs (∼20%).

**Discussion**

We found that tidal and minute volume, respiratory rate, and rapid shallow breathing index values at the end of a
Sensitivity study. We excluded patients weaned within 24 h of ICU admission or who needed low levels of mechanical support when screened, because we thought they were less likely to benefit from weaning tests, and are best managed using simple weaning protocols. We further refined the timing of our weaning test by using a previously validated screening checklist, which predicted a high probability of eventual weaning success. We believe our population was typical of most mixed medical/surgical intensive care patients requiring extended intensive care. Bias was reduced by a high recruitment rate and complete patient follow-up, concealing test data from clinicians making weaning decisions, making daily weaning assessments independently from the research team, and defining a composite endpoint before the analysis. We used a progressive reduction in PSV rather than an SBT, because we believed that the physiological data would be more stable with this approach. SBTs are widely used as the primary method of assessing weaning in some health-care systems, and our data may not be valid for that approach. Our protocol was designed to enable recommended statistical techniques for assessing potential test variables to be used; a weakness of this approach was that the practice undertaken was not identical to routine care, but represented the best compromise to improve methodological quality. We also only assessed weaning outcomes during the first 24 h after the weaning test. It would be possible to assess the relation between the test results and weaning outcomes at a later time point, but the poor performance observed for all variables did not, in our view, make this worthwhile. We also did not assess the relation to other patient outcomes, such as ICU stay and mortality, which would require a larger study.

A systematic review of previous studies of weaning tests found features of poor design, which we addressed in this study. We excluded patients weaned within 24 h of ICU admission or who needed low levels of mechanical support. Even the best cut-off values for \( V_t \) expired and \( \text{E}^\text{CO}_2 \) had statistically significant predictive value, but the test performance also indicated very modest clinical value. This was evident from the small changes to the observed pre-test probability of 44% in the study population. Even the best cut-off values for \( \text{E}^\text{CO}_2 \) and \( [\text{I}–\text{E}] \text{O}_2 \) could only change clinical certainty by about 20% if used as a predictive test compared with not using these data.

A systematic review of previous studies of weaning tests found features of poor design, which we addressed in this study. We excluded patients weaned within 24 h of ICU admission or who needed low levels of mechanical support.

### Table 4
Comparison of achievers with non-achievers at the end of the weaning trial. Data are presented as mean (so) unless otherwise indicated. \( \text{H}^+ \), hydrogen ion concentration; \( \text{Pa}_\text{O}_2 \), oxygen partial pressure; \( \text{Pa}_\text{CO}_2 \), carbon dioxide partial pressure; \( \text{Sa}_\text{O}_2 \), oxygen saturation; \( \text{Fi}_\text{O}_2 \), fraction of inspired oxygen; \( [\text{I}–\text{E}] \text{O}_2 \), difference between inspired–expired oxygen; \( \text{PE}_{\text{CO}_2} \), end-tidal carbon dioxide; \( \text{f/Vt} \), frequency–tidal volume ratio.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Achievers (n=32)</th>
<th>Non-achievers (n=41)</th>
<th>Mean difference (95% CIs)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{H}^+ ) (nmol litre (^{-1} ))</td>
<td>41.2 (5.3)</td>
<td>50.5 (9.6)</td>
<td>9.3 (5.5 to 3.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>( \text{Pa}_\text{O}_2 ) (kPa)</td>
<td>12.1 (11.5)</td>
<td>11.5 (0.12)</td>
<td>0.6 (–2.3 to 1.1)</td>
<td>0.47</td>
</tr>
<tr>
<td>( \text{Pa}_\text{CO}_2 ) (kPa)</td>
<td>6.3 (3.2)</td>
<td>7.2 (2.8)</td>
<td>0.9 (–0.5 to 2.3)</td>
<td>0.12</td>
</tr>
<tr>
<td>( \text{Sa}_\text{O}_2 ) (%)</td>
<td>98 (2.6)</td>
<td>96 (2.9)</td>
<td>2.0 (–2.6 to –0.05)</td>
<td>0.04</td>
</tr>
<tr>
<td>( [\text{I}–\text{E}] \text{O}_2 ) (%)</td>
<td>5.8 (1.4)</td>
<td>6.7 (1.9)</td>
<td>0.9 (0.1–1.7)</td>
<td>0.03</td>
</tr>
<tr>
<td>( \text{PE}_{\text{CO}_2} ) (kPa)</td>
<td>5.1 (1.1)</td>
<td>5.6 (1.2)</td>
<td>0.5 (–0.01 to 1.1)</td>
<td>0.05</td>
</tr>
<tr>
<td>Minute volume expired (litre min (^{-1} ))</td>
<td>11 (3.7)</td>
<td>10 (3.8)</td>
<td>1.0 (–2.7 to 0.9)</td>
<td>0.31</td>
</tr>
<tr>
<td>Tidal volume expired (ml)</td>
<td>540 (150)</td>
<td>530 (170)</td>
<td>10 (–85.3 to 64.3)</td>
<td>0.78</td>
</tr>
<tr>
<td>Respiratory rate (bpm)</td>
<td>21.4 (8.0)</td>
<td>20.4 (8.2)</td>
<td>1.0 (–4.8 to 2.9)</td>
<td>0.62</td>
</tr>
<tr>
<td>( \text{f/Vt} ) (bpm litre (^{-1} ))</td>
<td>40.5 (21.5)</td>
<td>43.7 (35.7)</td>
<td>3.2–11.1/17.0</td>
<td>0.65</td>
</tr>
</tbody>
</table>

When compared with the reference line, both \( [\text{I}–\text{E}] \text{O}_2 \) and \( \text{PE}_{\text{CO}_2} \) demonstrated high sensitivity and specificity. The best cut-off was chosen as the value with the highest Youden index, which is widely quoted as a useful weaning test, but the systematic review showed a range of likelihood ratios. The predictive value of this test, even when using different cut-off values, varied widely between studies reflecting different test conditions, patient populations, and endpoints. Our data suggest that \( \text{f/Vt} \) at the end of a controlled reduction

![Fig 2 ROCs for \([\text{I}–\text{E}] \text{O}_2 \) and \( \text{PE}_{\text{CO}_2} \) at the end of the weaning trial as predictors of the study endpoint within 24 h. The solid line \([\text{I}–\text{E}] \text{O}_2 \) and dotted line \( \text{PE}_{\text{CO}_2} \) show the sensitivity and specificity for each cut-off. The best cut-off was chosen as the value with the best combination of predictive power and specificity. Diagonal segments are produced by tie.]
in PSV has very limited predictive value for early successful reduction in ventilator support in mixed intensive care populations, irrespective of the cut-off value chosen. Conti and colleagues\(^4\) carried out a similar study to ours, but used data during the first 2 min of an SBT. They concealed test data from responsible clinicians and used a similar analysis to ours. Their study endpoint differed from ours as it included extubation and spontaneous breathing for more than 48 h. Although their patients had a higher pre-test probability of success (73% vs 44%), they found no predictive value from \(V_{\text{texp}}\), \(M_{\text{Vexp}}\), f, \(f/Vt\), or several other variables. The equivocal findings of the systematic review, the Conti study, and our current study do not support the routine use of these measures to predict weaning outcome in unselected general intensive care patients.

We report the first evaluation of \([I - E]O_2\) and \(P_{CO_2}\) as weaning predictors. These measures performed slightly better and had statistically significant predictive power at the end of a weaning trial. Specifically, when we used multiple cut-off values for \([I - E]O_2\), low values were useful in predicting weaning success (positive likelihood ratio 2 for cut-off \(\leq 5.4\%\)), and high values predicted weaning failure (negative likelihood ratio 2 for cut-off \(> 6.4\%\)). Similarly, \(P_{CO_2}\) values \(< 5.1 \text{ kPa}\) at the end of the weaning trial had a positive likelihood ratio of 2. These tests might have some clinical value, but in our population changed the pre-test probability of 44% by only about 20%. Changes of this magnitude are unlikely to be of sufficient value to justify routine use in ICUs as a clinical test.

Our findings do provide some support for the hypothesis that a combination of increased oxygen demand from increased work of breathing and the metabolic stress of weaning trials, together with failure to achieve adequate alveolar ventilation to maintain normocapnia, makes weaning success unlikely. Several previous observational studies found associations between greater oxygen consumption and weaning failure. Older studies assessed the oxygen cost of breathing during weaning by comparing oxygen consumption during volume-controlled ventilation with consumption during unsupported breathing, and found that large differences were associated with delay or failure to wean.\(^{17-19}\) Oh and colleagues\(^{20}\) described greater increases in oxygen consumption in patients failing to wean, and also noted larger increases in catecholamines during weaning trials in these patients. This observation supports the idea that metabolic stress, from a combination of increased respiratory work, anxiety, and sympathetic activation, leads to increased oxygen demand, which results in weaning failure if cardiorespiratory reserve is insufficient. In such circumstances, a progressive increase in oxygen extraction would lead to a greater \([I - E]O_2\) difference. In our study, there was no significant difference in \([I - E]O_2\) between the achievers and non-achievers at baseline [mean difference 0.28% (95% CI \(-0.22\) to 0.78%)], but this became statistically significant at the end of the trial [mean difference 0.88% (95% CI 0.09 – 1.67, \(P = 0.03\)]. A similar trend was observed for \(P_{CO_2}\). The most marked difference between the groups at baseline was in \(\text{H}^+\) concentration, probably showing greater severity of illness in the non-achievers. This persisted at the end of the weaning trial, but the difference increased, in parallel with the increased difference in \(P_{CO_2}\). An association between an increased oxygen extraction ratio and weaning failure was also found in a study where mixed venous oxygen saturation decreased progressively.\(^{21}\) Although oxygen consumption was available from the M-COVX module that we used to measure \([I - E]O_2\) and \(P_{CO_2}\), we did not use this measure because it is calculated by integration of the gas fraction and tidal volume measurements on a breath-by-breath basis and is not validated at high respiratory rates often present during weaning.\(^{22}\) Further work is needed to assess the validity of such derived measures under clinical conditions, or whether they are clinically useful in any ICU populations.

In conclusion, in a cohort of general ICU patients with a pre-test probability of successfully reducing mechanical ventilator support within 24 h at 44%, \(V_{\text{texp}}\), \(M_{\text{Vexp}}\), f, and \(f/Vt\) values at the end of a weaning trial did not predict early weaning outcome. The \([I - E]O_2\) and \(P_{CO_2}\) values at the same time had statistically significant predictive power and changed pre-test probability by \(\approx 20\%\) when the best cut-off values were applied. The clinical value of these variables as a test was limited, and further work is needed to show if these measures can inform clinically relevant decisions or alter patient outcomes.

**Supplementary material**

Supplementary material is available at *British Journal of Anaesthesia* online.

**Acknowledgements**

The authors thank the medical and nursing staff of the Royal Infirmary General Intensive Care Unit for their support and cooperation during this study.

**Conflict of interest**

This study was supported through a collaborative grant from GE Healthcare. The design, conduct, analysis, and writing of the study were entirely the work of the authors. GE Healthcare did not place any restriction on the use of the data or approve the manuscript before submission. None of the authors has received direct financial support from GE Healthcare and none is an employee of the company.

**Funding**

This study was supported by a grant from GE Healthcare, Helsinki, Finland, and by the Edinburgh Critical Care Research Group.

**References**

6 Cook DJ, Meade MO, Perry AG. Qualitative studies on the patient’s experience of weaning from mechanical ventilation. Chest 2001; 120: 469S – 73S
8 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 2001; 120: 375S – 95S
13 Epstein SK. Weaning from mechanical ventilation. Respir Care 2002; 47: 454 – 66
15 Epstein SK. Etiology of extubation failure and the predictive value of the rapid shallow breathing index. Am J Respir Crit Care Med 1995; 152: 545 – 9
18 McDonald NJ, Lavelle P, Gallacher WN, Harpin RP. Use of the oxygen cost of breathing as an index of weaning ability from mechanical ventilation. Intensive Care Med 1988; 14: 50 – 4
19 Shikora SA, Benotti PN, Johannigman JA. The oxygen cost of breathing may predict weaning from mechanical ventilation better than the respiratory rate to tidal volume ratio. Arch Surg 1994; 129: 269 – 74