Measurement of a Baseline Minute Ventilation for the Calculation of Minute Ventilation Recovery Time: Is a Subjective Method Reliable?

Christopher W Seymour MD, Jason D Christie MD MSc, Christina Gaughan MSc, and Barry D Fuchs MD

BACKGROUND: Minute ventilation recovery time is a new predictor of extubation outcome that uses a subjective method for the determination of baseline minute ventilation (V˙E) during its measurement. The purpose of the current study is to evaluate the inter-rater reliability of this subjective method for determining baseline V˙E. METHODS: Three critical-care physicians served as independent readers. Each was trained with 5 practice V˙E trends, using the published method for determining baseline V˙E, defined as the lowest, stable nadir lasting 15–30 min prior to the final weaning trial before extubation. Readers then determined baseline V˙E prospectively from an 8-hour V˙E trend for 19 patients who were weaning from mechanical ventilation in the surgical intensive care unit of a tertiary care hospital. Each V˙E trend was an objective recording of V˙E every 15 min for 8 hours, immediately prior to the final weaning trial before extubation. RESULTS: There was excellent inter-rater reliability between trained readers for determination of a subjective V˙E baseline. Baseline V˙E was within 1 L/min for 15/19 patients (79%). Intra-class correlation across the 3 readers was 0.92 (p < 0.01). Tukey’s test revealed no significant variability between readers (p > 0.5), and Spearman correlations between all reader pairs were significant (p < 0.01). CONCLUSION: After minimal training, readers can reliably determine a subjective baseline V˙E. This study validates the original methodology for determining baseline V˙E, an essential step in the measurement of minute ventilation recovery time. Key words: minute ventilation, extubation, spontaneous breathing trial, monitoring, weaning, mechanical ventilation. [Respir Care 2005;50(4):468–472. © 2005 Daedalus Enterprises]
tially very useful as an extubation parameter, the methodology used to determine $V_{ERT}$ has not been rigorously validated. In particular, baseline $V_E$ has been defined as the subjective nadir in a patient’s total $V_E$ within the 24 hours preceding the final SBT, while breathing on resting ventilator settings. In the original study of Martinez et al., this value was determined by a single investigator who retrospectively reviewed the trend of $V_E$, thus, the reliability of this method has not been established.

Validation of this initial step of $V_{ERT}$ measurement is essential because it represents the greatest potential source of variation in the measurement of this parameter; the subsequent determination of the time required for $V_E$ to return to this baseline is objective. The current study was performed prospectively to evaluate the inter-rater reliability of determining baseline $V_E$ between 3 independent readers previously trained in the published methodology, in their subjective determination of baseline $V_E$.

**Methods**

This research protocol was approved by the Investigational Review Board of the Hospital of the University of Pennsylvania, with the requirement for written informed consent of all study patients. This was a prospective cohort study.

**Patients**

All mechanically ventilated patients were evaluated prospectively in the surgical intensive care units (trauma, cardiac, and general ICUs) of the Hospital of the University of Pennsylvania, from July to September 2001. Patients were included if they were > 18 years of age and were receiving mechanical ventilation by endotracheal tube for > 24 hours postoperatively. Patients were excluded if they had a tracheostomy, were receiving noninvasive mechanical ventilation, or were undergoing nonprotocol weaning. Daily screening for potential subjects was conducted on every weekday morning by one or more investigators who traveled with the surgical critical care service on morning rounds, for 3 months. Thirty-two patients provided informed consent for the study, of whom 19 were included (Fig. 1).

All patients were weaned from mechanical ventilation by a standardized hospital protocol performed by nurses and respiratory therapists. This protocol was implemented 13 months prior to the start of the current study, using pre-established standardized assessments to screen for weaning and extubation readiness. SBTs performed on 5 cm H$_2$O continuous positive airway pressure with pressure support of 7 cm H$_2$O were used for the initial SBTs. If unsuccessful, patients were subsequently weaned by pressure-support to a minimum value of 7 cm H$_2$O. After completing a successful SBT, the decision to extubate was made by the physician of record.

**Data Collection**

The inter-rater reliability of baseline $V_E$ between 3 board-certified critical care physician readers was evaluated by the following methodology. First, readers reviewed the published method for determining baseline $V_E$. Readers were trained using a practice set from 5 patients, each with 8-hour $V_E$ trends modified from actual data, and jointly agreed upon a value. A sample training $V_E$ trend is shown in Figure 2. After completing the training, readers then independently determined a baseline $V_E$ for all study patients, blinded to other readers.

$V_E$ was collected using the VueLink Interface, obtained from Philips Medical Systems (Andover, Massachusetts). The VueLink module interfaces the Puritan Bennett 7200 ventilators to bedside cardiac monitors (Hewlett Packard, Houston, Texas), allowing for the recording of $V_E$ per minute. Data were displayed in 1-minute averages, with $V_E$ estimated from the sum of exhaled volumes for up to 8 mandatory or spontaneous breaths over the previous 1-min period. Following patient consent, modules were connected at the bedside, and data were collected for 8 hours prior to the final SBT before extubation. The original $V_{ERT}$ method describes a retrospective survey of 24 hours of data prior to the SBT to determine baseline $V_E$. These data were...
generated using direct download by Tram-net Interface (Marquette Electronics, Milwaukee, Wisconsin), which connected Puritan Bennett 7200 ventilators to Tramscope bedside monitors (Marquette Electronics, Milwaukee, Wisconsin). Trends were constructed from data points measured every 15 min. We chose to evaluate the V\(\dot{E}\) trend during the most proximate 8-hour period immediately preceding the SBT in the current study, suspecting that “fast-track” weaning in surgical patients may preclude the recording of the complete V\(\dot{E}\) trend for review. V\(\dot{E}\) was downloaded from bedside monitors at 15-min intervals, providing a trend of 32 measurements of V\(\dot{E}\). Each measurement was the average generated that minute, and was not the average of the prior 15-min interval.

To characterize the critically ill patients who were enrolled, demographic data were obtained, including age, gender, Acute Physiology and Chronic Health Evaluation (APACHE) II score, surgery service, and the duration and mode of mechanical ventilation at the time baseline V\(\dot{E}\) measurements were collected.

Statistical Analysis

Demographic data of the study group are presented as mean ± standard deviation or median [interquartile range], dependent on normality, as assessed by the Shapiro Wilk test. To detect similarity between all 3 investigators in their baseline V\(\dot{E}\) measurements, the intraclass correlation coefficient was derived, similar to a kappa test for continuous measures. Investigators’ readings were individually compared using Spearman correlations. To detect significant differences between the baseline V\(\dot{E}\) measurements, the Tukey’s studentized range test was used. A p value of < 0.05 was considered statistically significant. All statistical analyses were performed using commercially available software (NCSS version 2000, Kaysville, Utah; SAS version 8.02, Cary, North Carolina).

Results

Nineteen patients were enrolled, with a median age of 68 years, with APACHE II score of 19 ± 4.4, and primarily female gender (79%). Surgical service was cardiac (63%), trauma (16%), and vascular or general (11% for both). The median time on mechanical ventilation was 5 days [interquartile range 2–8 d] prior to extubation. Patients were more likely to be on pressure-support (84%) than on volume-controlled ventilation at the time baseline V\(\dot{E}\) was recorded.

Determination of baseline V\(\dot{E}\) using the published methodology originally described for the measurement of V\(\dot{E}\)RT demonstrated excellent inter-rater reliability between 3 independent, blinded readers. All physician readers chose a baseline V\(\dot{E}\) within 1 L/min in 15 of 19 patients (79%), and Tukey’s test revealed no statistically significant differences between readers (F = 0.42, p = 0.66). The intra-class correlation coefficient was 0.92 (p < 0.01). Scatter plots of the 3 investigators are illustrated in Figure 3; the Spearman correlation coefficients for these investigator comparisons are shown in Figure 3, and all met statistical significance (p < 0.01). Figure 4 demonstrates data from a patient in whom a consensus within 1 L/min for the estimate of baseline V\(\dot{E}\) was not found.
Discussion

$\dot{V}_{E}RT$ has recently shown to be a potential predictor of extubation outcome. Given the adverse consequences associated with extubation failure and lack of validated predictors, further investigation of this new index is important. The objective of this study was to validate the method originally described for determining baseline $\dot{V}_{E}$, the subjective component in the measurement of $\dot{V}_{E}RT$. This study demonstrates a high inter-rater reliability between 3 investigators who determined a baseline $\dot{V}_{E}$ using this method. Inter-observer variability was not statistically significant, nor were raters’ differences deemed clinically important.

The time required to train readers to determine baseline $\dot{V}_{E}$ is less than 15 min. However, this measurement is not easily performed at the bedside without $\dot{V}_{E}$ recorded continuously from the ventilator for an extended period prior to the final SBT. In our study, the VueLink interface was used to record $\dot{V}_{E}$ directly from the ventilator, similar to the Tram-net interface used by Martinez et al. These devices are not used routinely in ICUs, but they provide objective, averaged values, which can be downloaded from bedside monitors. In most ICUs, $\dot{V}_{E}$ is not recorded continuously and can only be assessed by visual inspection of the ventilator display. Conceivably, one could observe trends in $\dot{V}_{E}$ over time and estimate $\dot{V}_{E}RT$; however, the validity of this approach has never been evaluated. As more ICUs gain direct data acquisition from multiple monitoring devices, future modification of $\dot{V}_{E}RT$ measurement may allow for more automated, and perhaps more accurate, methods for determining baseline $\dot{V}_{E}$. These methods may include a single measurement of $\dot{V}_{E}$ prior to an SBT, or automated averaging of $\dot{V}_{E}$ measured continuously.

The measurement of $\dot{V}_{E}$ and the assessment of baseline $\dot{V}_{E}$ in this study is not without potential sources of error. Patient agitation, nursing intervention, and/or level of sedation may dramatically change the spontaneous respiratory rate and tidal volume of these patients, disrupting the consistency in the trend displayed to raters. In addition, outlier measurements were not removed during the construction of $\dot{V}_{E}$ trends. The use of raw, nonmanipulated data in the current study attempts to simulate the true variability of the data available to bedside physicians who may wish to measure a baseline $\dot{V}_{E}$. Because the ventilator displays exhaled $\dot{V}_{E}$ calculated from multiple patient breaths, it is also subject to patient-related error, such as apneas, agitation, ventilator-setting change, or nursing intervention. These sources of error may explain the discon-
dance in baseline $V_E$ (> 1 L/min) found in 4 of 19 patients. As shown in Figure 4, there may be multiple time periods during the 8-hour trend where $V_E$ remains stable, albeit at different values. This variability suggests that a single assessment may be more practical when determining a baseline value to be used when measuring $V_{E\text{RT}}$.

Conclusions

The published technique for determining baseline $V_E$ subjectively is reliable and, thus, valid for the subsequent measurement of $V_{E\text{RT}}$. Further study of the utility or performance of $V_{E\text{RT}}$ in other clinical settings may be accomplished using this method. Alternatively, the development of a simpler and more objective method to measure baseline $V_E$ may be preferable.

REFERENCES