Reducing the time period of steady state does not affect the accuracy of energy expenditure measurements by indirect calorimetry

Marina M. Reeves,1,2 Peter S. W. Davies,3 Judith Bauer,2 and Diana Battistutta1
1Centre for Health Research, Queensland University of Technology, Brisbane, Queensland 4059; 2The Wesley Research Institute, Brisbane, Queensland 4066; and 3Children’s Nutrition Research Centre, Royal Children’s Hospital, University of Queensland, Brisbane, Queensland 4029, Australia
Submitted 13 November 2003; accepted in final form 9 March 2004

Reeves, Marina M., Peter S. W. Davies, Judith Bauer, and Diana Battistutta. Reducing the time period of steady state does not affect the accuracy of energy expenditure measurements by indirect calorimetry. J Appl Physiol 97: 130–134, 2004.—Achievement of steady state during indirect calorimetry measurements of resting energy expenditure (REE) is necessary to reduce error and ensure accuracy in the measurement. Steady state is often defined as 5 consecutive min (5-min SS) during which oxygen consumption and carbon dioxide production vary by ±10%. These criteria, however, are stringent and often difficult to satisfy. This study aimed to assess whether reducing the time period for steady state (4-min SS or 3-min SS) produced measurements of REE that were significantly different from 5-min SS. REE was measured with the use of open-circuit indirect calorimetry in 39 subjects, of whom only 21 (54%) met the 5-min SS criteria. In these 21 subjects, median biases in REE between 5-min SS and 4-min SS and between 5-min SS and 3-min SS were 0.1 and 0.01%, respectively. For individuals, 4-min SS measured REE within a clinically acceptable range of ±2% of 5-min SS, whereas 3-min SS measured REE within a range of −2–3% of 5-min SS. Harris-Benedict prediction equations estimated REE for individuals within ±20–30% of 5-min SS. Reducing the time period of steady state to 4 min produced measurements of REE for individuals that were within clinically acceptable, predetermined limits. The limits of agreement for 3-min SS fell outside the predefined limits of ±2%; however, both 4-min SS and 3-min SS criteria greatly increased the proportion of subjects who satisfied steady state within smaller limits than would be achieved if relying on prediction equations.

MEASUREMENTS OF RESTING ENERGY expenditure (REE) by indirect calorimetry are becoming increasingly popular, because indirect calorimeters are more widely available and inaccuracies of prediction methods are well acknowledged. To compare measurements and ensure rigor of results, measurement methods need to be standardized and rigorously described.

For tests to be considered valid and to reduce error, a stable measurement period (“steady state”) should be achieved, after which the measurement can be terminated. In the literature to date, various definitions of steady state have been reported or are often not specified (3, 5, 21, 25). To determine the optimal criteria, which best agree with 24-h measurements of energy expenditure, McClave et al. (17) tested different measurement protocols and steady-state criteria in a sample of 22 mechanically ventilated patients. Their results indicated that steady state defined as 5 consecutive 1-min intervals where oxygen consumption (VO2) and carbon dioxide production (VCO2) change in this period is not defined (14–17). Stevens et al. (19) reported using a 3-min steady-state period, although the degree to which VO2 and VCO2 change in this period is not defined (14–17). Additionally, studies that have validated steady-state criteria have only validated at the group level. No studies have presented data on the level of individual prediction, that most useful to clinical application.

This study aimed to compare the agreement between measurements of REE by using 5-min steady-state criteria with steady-state criteria over a shorter period. Energy expenditure measurements using steady-state criteria over a shorter time period may be more likely to be achieved by patients and therefore provide valid, useable results.

METHODS

Subjects. Patients with cancer and healthy subjects participating in another trial were included in the study. Patients were diagnosed with a solid tumor, excluding breast, prostate, or brain, and were recruited from a radiotherapy treatment center. Healthy subjects were recruited from a convenience sample and were group matched by gender, age, weight, and height to cancer patients. No subjects had undergone surgery within the month before the study.

Measurement protocol. Data collection procedures were identical for all subjects. Measurements of energy expenditure were conducted under outpatient conditions. That is, REE was measured not basal. Participants fasted overnight for at least 12 h and attended the clinic on wakening. Measurements were conducted between 7 AM and 9 AM. Before commencement of the measurements, subjects rested quietly for 30 min, during which time the indirect calorimeter was calibrated. Subjects sat quietly in a reclining chair for the duration of the measurement. Subjects were asked to remain awake and motionless.

Height was measured without shoes to the nearest 0.5 cm with a stadiometer. Weight and fat-free mass were measured without shoes or heavy clothing to the nearest 0.1 kg by foot-to-foot bioelectrical...
impedance analysis (model 300GS, Tanita, Tokyo, Japan). REE was estimated from the Harris-Benedict prediction equations (11) using weight, height, age, and gender. For subjects with a body mass index of \( \geq 29 \text{ kg/m}^2 \), an adjusted weight was used in the prediction of REE (1, 10).

REE. REE was measured by breath-by-breath respiratory gas exchange with the VMax 229 (SensorMedics, Yorba Linda, CA) indirect calorimetry device, using a mouthpiece and noseclip. A mass-flow sensor measured volume and airflow, and it was calibrated before each measurement by using a certified 3-liter calibration syringe. Calibration was achieved when measured stroke volume was within \( \pm 3\% \) of syringe volume. Expired gas was analyzed for oxygen concentration by using a paramagnetic oxygen analyzer and for carbon dioxide concentration by using a nondispersive infrared analyzer. Gas analyzers were calibrated before each measurement using three known standard gas concentrations (16% \( \text{O}_2 \), 4% \( \text{CO}_2 \); 26% \( \text{O}_2 \), 0% \( \text{CO}_2 \); room air 20.94% \( \text{O}_2 \), 0.05% \( \text{CO}_2 \)). Calibration was complete when gas analyzers measured oxygen and carbon dioxide concentration within \( \pm 5\% \) of expected. Measurements were ceased once a 5-min steady-state period was achieved or after 30 min, whichever occurred first.

The indirect calorimeter was connected to an IBM-compatible personal computer for management and storage of data by using the VMax Vision software for Windows (version 05.2A, SensorMedics). Respired carbon dioxide quotient (RQ), defined as \( \frac{\text{VCO}_2}{\text{VO}_2} \), was calculated by the software. \( \text{VO}_2 \) and \( \text{VCO}_2 \) were converted to REE by using the abbreviated Weir equation (28):

\[
\text{REE} = \text{VO}_2(3.941) + \frac{\text{VCO}_2(1.106)}{\text{VE}}
\]

where REE was measured in kilocalories per day and \( \text{VO}_2 \) and \( \text{VCO}_2 \) in liters per day.

**Steady-state criteria.** Three steady-state criteria were defined. The criteria were defined by a change in \( \text{VO}_2 \), \( \text{VCO}_2 \), RQ, and minute ventilation (\( \text{Ve} \)) of \( \leq 10\% \) but differed by the time period in which these variables were stable: first 5 min (5-min SS), first 4 min (4-min SS), first 3 min (3-min SS).

The indirect calorimeter software identified the breaths that were in the steady-state period as defined by the 5-min criterion. Averaged minute \( \text{VO}_2 \), \( \text{VCO}_2 \), RQ, and REE were recorded for each minute interval in the 5-min SS period determined by the commercial software. REE measured by the 5-min SS criterion was the average of the recorded REE measurements over the 5-min period. Similarly, REE measured by 4-min SS and 3-min SS were the average of REE during the first four and three 1-min intervals, respectively, in the referent 5-min period.

If 5 min of steady state were not achieved, the stored data were identified for the first 4-min SS or 3-min SS period, if achieved, and the data were similarly recorded for each averaged minute.

**Statistical analyses.** Statistical analyses were carried out with SPSS for Windows (version 11.0.1, 2001, SPSS, Chicago, IL) statistical software package. Normally distributed continuous variables are presented as means \( \pm \) SD. Biases between the steady-state criteria were not normally distributed. As such variables are presented as median (range or 2.5th to 97.5th percentile) and corresponding nonparametric tests were conducted. Characteristics of subjects who achieved 5-min SS and those who did not were compared for both clinical importance and statistical significance by independent sample \( t \)-tests and Fisher’s exact test.

Differences in measured REE between the three steady-state criteria were first assessed by Wilcoxon’s signed-rank tests. To compare our results with that of other studies, which cite correlation coefficients, Spearman’s rank correlation was used to determine the strength of the relationship of 5-min SS with 4-min SS and with 3-min SS criteria. Median bias, limits of agreement, and plot of bias against average of two measurements by using the Bland-Altman plotting approach were used to describe agreement at the individual level and assess whether the bias was consistent across the entire range of measurements (2). Spearman’s rank correlation was used to assess whether there was any trend in the bias with increasing REE measurements.

The data were initially analyzed separately for cancer patients and healthy subjects to determine whether the relationship differed on the basis of disease status. There was no significant difference between cancer patients or healthy subjects for the median bias and limits of agreement, for the comparison of either the 5-min SS and 4-min SS data or the 5-min SS and 3-min SS data (data not shown). As such, data presented are for the combined sample of cancer patients and healthy subjects.

The authors decided a priori that a difference in REE between the steady-state criteria of \( \geq 2\% \) would suggest that the limited-time steady-state criteria would not be appropriate for acceptably measuring energy expenditure in clinical settings and therefore should not be used as the basis for terminating REE measurements. Interpretations of the significance of results have been based on assessing differences for both clinical and statistical meaningfulness (95% significance level, 2 tailed).

**Ethical issues.** The study received ethical approval from the participating university and hospital institutions’ Ethics Committees. Subjects were provided with written and oral information before providing informed consent to participate in the study.

**RESULTS**

Measurements of REE were performed on 39 spontaneously breathing subjects (22 cancer patients, 17 healthy subjects). Only 21 (54.5%) subjects met 5-min SS criteria, 27 met 4-min SS criteria, and all but one subject met the 3-min SS criteria. Data used for the comparison of time periods were therefore based on the 21 subjects who reached 5-min SS. Retrospective power calculation indicated that the sample size of 21 was adequate to detect a minimum difference of 2%, on the basis of an observed standard deviation of 3.75%, with 90% power at the 95% significance level (2 tailed).

Some physical characteristics of subjects who achieved 5-min SS and those who failed to achieve are shown in Table 1. There were no clinically or statistically significant differences between the two groups with the exception of gender (Fisher’s exact test, \( P = 0.055 \)). The two groups did not differ with respect to measured RQ and \( \text{Ve} \), but they differed significantly in measured absolute \( \text{VO}_2 \), \( \text{VCO}_2 \), and REE. When adjusted for differences in fat-free mass between the two groups, \( \text{VO}_2 \), \( \text{VCO}_2 \), and REE were no longer significantly different (data not shown).

Table 1. Physical characteristics of subjects

<table>
<thead>
<tr>
<th></th>
<th>5-min SS</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>Age, yr</td>
<td>61±12</td>
<td>63±12</td>
</tr>
<tr>
<td>Men/women</td>
<td>16/5</td>
<td>8/10</td>
</tr>
<tr>
<td>Height, cm</td>
<td>171±12</td>
<td>166±9</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>80.4±16.3</td>
<td>73.1±15.5</td>
</tr>
<tr>
<td>Body mass index, kg/m(^2)</td>
<td>27.4±3.9</td>
<td>26.6±5.3</td>
</tr>
<tr>
<td>Fat-free mass, kg</td>
<td>55.1±11.1</td>
<td>49.4±9.8</td>
</tr>
<tr>
<td>( \text{VO}_2 ),* ml/min</td>
<td>240±47</td>
<td>197±61*</td>
</tr>
<tr>
<td>( \text{VCO}_2 ),* ml/min</td>
<td>170±54</td>
<td>140±43*</td>
</tr>
<tr>
<td>RQ(^\dagger)</td>
<td>0.71±0.08</td>
<td>0.72±0.07</td>
</tr>
<tr>
<td>( \text{Ve} ), l/min</td>
<td>7.7±2.5</td>
<td>7.5±1.9</td>
</tr>
<tr>
<td>REE,* kcal/day</td>
<td>1,597±304</td>
<td>1,373±311†</td>
</tr>
</tbody>
</table>

Values are means \( \pm \) SD, \( n \), no. of subjects. \( \text{VO}_2 \), oxygen consumption; \( \text{VCO}_2 \), carbon dioxide production; RQ, respiratory quotient; \( \text{Ve} \), minute ventilation; REE, resting energy expenditure; SS, steady state. *Based on maximum steady-state data available. †Compared with 5-min SS, \( P < 0.05 \).

J Appl Physiol • VOL 97 • JULY 2004 • www.jap.org
The median REE, bias, and limits of agreement for the different steady-state criteria are shown in Table 2. On the basis of Wilcoxon signed-rank tests, there was no statistically significant difference between REE measured by 5-min SS or 4-min SS (Z = −0.643, P = 0.52) or between REE measured by 5-min SS or 3-min SS (Z = −0.521, P = 0.60). The median bias between 5-min and 4-min SS was negligible, −1.3 kcal (0.1%) and −0.1 kcal (<0.01%) between 5-min and 3-min SS.

The correlations between REE by 5-min SS and 4-min SS and REE by 5-min SS and 3-min SS were ρ = 0.99 and ρ = 0.98, respectively, (P < 0.001 for both). Although this indicates almost perfect correlation, this is a measure of the strength of the covariation of these two measurements and not a measure of the level of agreement. Statistically, high correlation would be expected by definition of the two variables.

Because indirect calorimetry is most commonly used to measure the REE of individuals, the data were assessed for agreement at the individual level. If it is assumed that our sample is representative of the population, it would be expected that the limits of agreement would define an acceptable normal range of expected values of REE. Hence, for good agreement, the limits of agreement (2.5th to 97.5th percentile) should be in a narrow range within the clinically acceptable limits.

Although the median bias for both the 4-min SS and 3-min SS was very close to zero and well within the acceptable level, the limits of agreement were wider. Bland-Altman plots (2) were used to assess the spread of the bias over the complete range of REE measurements (depicted as average of 2 measurements) and to determine whether there was a significant relationship. Figures 1 and 2 show the plots for 5-min SS and 4-min SS data and for 5-min SS and 3-min SS data, respectively. The use of 4-min SS criteria produced limits of agreement of REE measurements that were 1.2% below, up to 2.0% above, REE measured by 5-min SS criteria. REE measurements that used 3-min SS criteria produced values that are 2.2% below, up to 3.4% above REE measured by 5-min SS criteria.

Figure 2 indicates that there was a slight tendency toward underestimation of REE measurement by 3-min SS criteria compared with 5-min SS criteria with increasing REE. Correlation analysis indicated a statistically significant correlation between median bias and average of measurements for 5-min SS and 3-min SS data (ρ = −0.45, P = 0.04). Whether the data were analyzed with and without (ρ = −0.36, P = 0.118) the subject at extreme REE measurement, the correlations were similarly low.

The mean difference between REE measured by 5-min SS and REE predicted by the Harris-Benedict equations was −52 kcal (3.3%) with limits of agreement of ±400 kcal (−28.3–21.8%). There was no significant correlation or pattern between the difference and the average of predicted and measured REE.

**DISCUSSION**

Achievement of a steady-state (stable measurement) period during indirect calorimetry measurements ensures a greater level of accuracy in short-term measurements of energy expenditure. McClave et al. (17) have indicated that steady-state criteria defined as 5 min of energy expenditure measurement during which time VO2 and VCO2 change by ≤10%, agrees best with 24-h measurements of energy expenditure. Investigators often recommend that if steady state is not achieved, the measurement should be aborted and the data discarded (7). Because the criteria defined by McClave et al. are quite stringent, a significant number of subjects fail to meet these criteria. In the present study, 18 (46%) subjects failed to reach steady state. This level of data loss would have an impact on the use of indirect calorimetry in both research and clinical application. This study therefore aimed to assess whether shorter time periods under the same criteria (coefficient of variation ≤10%), which are more easily achieved by subjects,
would agree with the standard criteria to clinically acceptable standards.

The acceptable degree, to which REE measured by the shorter time periods could differ from the standard criteria, was determined a priori to be 2%. This error was to allow for the normal individual variation in energy expenditure, which for basal metabolic rate is in the order of 3–5% (9, 12, 20, 26).

Results for the group indicated that there was negligible difference between the shorter time periods (4-min SS and 3-min SS) and the standard 5-min SS period. To date, studies comparing different lengths of REE measurement and steady-state criteria have only analyzed group data based on paired t-tests and correlation analysis (13, 17, 22). On the basis of this group-level analysis, our data would also lead to the conclusion that both 4-min SS and 3-min SS criteria are acceptable for measuring REE. These analyses, however, do not assess the level of agreement between the two measures for individuals, and it is the individual level that is important for clinical application of the measurements.

As expected, shortening the time period defining steady state increased the variability of REE measurements. However, analysis for agreement for individuals indicated that the 4-min SS period is an acceptable surrogate for 5-min SS criteria, with measurements of REE falling approximately within the clinically acceptable range of ±2%, whereas 3-min SS produced measurements of REE within approximately −2–3%.

One might argue that the a priori defined limits of ±2% are too rigorous. As already noted previously, this limit was based on allowing for partial variation due to intravariability of patient, which is in the order of 3–5% (9, 12, 26). Because the 3-min SS criterion produced limits of agreement that were no higher than 3%, this reduced time period may still be considered to be within an acceptable clinical range in some clinical contexts.

In this study, just over one-half of the subjects reached steady state by 5-min criteria and were included in the subsequent analysis. Although a large number were excluded from the analysis, there appeared to be no difference in the characteristics of patients included in the analysis and characteristics of the total sample. Significantly fewer women satisfied the 5-min SS criteria and were included in the subsequent analysis; however, there is no evidence to suggest a gender difference in the ability to meet steady-state criteria or variation in REE measurements. We were unable to identify any differences in gas-exchange parameters between subjects who achieved 5-min SS and those who failed to.

Significantly more subjects failed to reach steady state in this study than that reported in the study by McClave et al. (17). Subjects in the present study, however, were spontaneously breathing and measured under outpatient conditions compared with mechanically ventilated subjects (17). Frankenstein (8) has reported that energy expenditure measurements on lucid, spontaneously breathing patients will generally take longer than mechanically ventilated patients due to awareness and the need to relax. In this study, reducing the time period of steady state increased the proportion of subjects who reached steady state: 27 (69%) reached 4-min SS and 38 (97%) reached 3-min SS.

It is important to note that although achievement of steady state is necessary to ensure stability in gas-exchange parameters, this alone is not sufficient to confidently and accurately measure true REE. A number of factors, both biological and technical, may influence indirect calorimetry measurements. Errors may be introduced by air leaks, incorrect calibration of the calorimeter, involuntary hyper- or hyperventilation, fluctuating levels of fractional inspired oxygen concentration (FiO2), or acid-base disturbances (16). Measured RQ can assist in determining test validity because a well-documented physiological range for RQ exists (0.67–1.30) (4, 16). The validity of indirect calorimetry tests that produce RQ values that fall outside this range should be questioned, regardless of whether steady state was achieved. In this study, one subject (5%) who met 5-min SS had an RQ outside of this range (0.63). McClave and colleagues (16) observed similar results in their study, in which 8% of subjects had a RQ outside the physiological range. Levels of FiO2 that may affect the accuracy of indirect calorimetry measurements (>60–80%) are generally of concern only with mechanically ventilated patients. All subjects in our study were spontaneously breathing with an FiO2 of ~20%.

To ensure greater confidence in indirect calorimetry measurements, careful calibration of flow sensors and gas analyzers at regular intervals and before each measurement should be undertaken, standard conditions for testing should be adhered to (e.g., length of fasting, rest period), steady state should be achieved, and data should be assessed for physiological validity.

In this study, the indirect calorimeter used a mouthpiece for the collection of expired air. Ventilated hoods are often considered the “gold standard” in portable collection systems (7, 29). Mouthpieces have been shown to increase V̇E through increases in tidal volume (29). The effect of such increases on VO2 and RQ, however, is unclear, with some suggestion that energy expenditure is not greatly influenced (18, 24). Although use of the mouthpiece may increase variability and therefore may also account for the lower number of subjects who achieved 5-min SS, our stability criteria were based on a combination of VO2, VCO2, RQ, and V̇E ensuring greater confidence in achievement of a true stable measurement period.

If measurement of REE is not possible, or steady state is not reached during REE measurements, estimation of the patients’ REE by prediction equations is necessary. There is sufficient literature indicating that prediction methods produce considerable error compared with measured energy expenditure when predicting REE of healthy individuals and even more so in patients with disease or injury (6, 23, 27). In this study, the Harris-Benedict equations (11) predicted the REE of individual subjects within large limits of ±20–30%. Therefore, the small amount of error (±2–3%) allowed for by using 4-min SS or 3-min SS criteria is negligible compared with estimations of REE from prediction equations.

In summary, we conclude that reducing the time period of steady state to 4 min produced measurements of REE for individuals that were within clinically acceptable, predetermined limits. Reduction of the time period further to 3 min produced measurements of individuals’ REE that were just outside the a priori defined limits, which in certain contexts may be clinically acceptable. Measurement of REE using both 4-min SS and 3-min SS criteria will produce estimates of REE in a greater proportion of subjects that could not be achieved if relying on prediction equations. As the time period is reduced, variation in measurements of REE increase; therefore, an
understanding of this variation and careful interpretation of measurements and choice of time period is warranted.

ACKNOWLEDGMENTS

The authors thank the technicians at the Queensland Respiratory Laboratory, The Wesley Hospital, for conducting the measurements.

REFERENCES


