Validity of the BodyGem calorimeter and prediction equations for the assessment of resting energy expenditure in overweight and obese Saudi males

Ali M. Almajwal, MSc, PhD, Peter G. Williams, MPH, PhD, Marijka J. Batterham, MSc, PhD.

ABSTRACT

Objectives: To assess the accuracy of resting energy expenditure (REE) measurement in a sample of overweight and obese Saudi males, using the BodyGem device (BG) with whole room calorimetry (WRC) as a reference, and to evaluate the accuracy of predictive equations.

Methods: Thirty-eight subjects (mean±SD, age 26.8±3.7 years, body mass index 31.0±4.8) were recruited during the period from 5 February 2007 to 28 March 2008. Resting energy expenditure was measured using a WRC and BG device, and also calculated using 7 prediction equations. Mean differences, bias, percent of bias (%bias), accurate estimation, underestimation and overestimation were calculated.

Results: Repeated measures with the BG were not significantly different (accurate prediction: 81.6%; %bias 1.1±6.3, p>0.24) with limits of agreement ranging from +242 to -200 kcal. Resting energy expenditure measured by BG was significantly less than WRC values (accurate prediction: 47.4%; %bias: 11.0±14.6, p=0.0001) with unacceptably wide limits of agreement. Harris-Benedict, Schofield and World Health Organization equations were the most accurate, estimating REE within 10% of measured REE, but none seem appropriate to predict the REE of individuals.

Conclusion: There was a poor agreement between the REE measured by WRC compared to BG or predictive equations. The BG assessed REE accurately in 47.4% of the subjects on an individual level.

Obesity has become a global epidemic and the prevalence of both overweight and obesity is still increasing in Saudi Arabia.13 Weight-reduction programs usually aim to establish an achievable energy intake goal for weight reduction, therefore clinicians may need to assess individuals’ total energy expenditure (TEE). Resting energy expenditure (REE) represents the majority (60-75%) of TEE in sedentary people,4 thus accurate assessment of REE is an important factor in planning interventions for overweight and obese individuals. Indirect calorimetry techniques such as whole room calorimetry (WRC) and metabolic carts are considered the “gold standard” for determining energy requirements. However, the equipment is expensive and the technique requires highly skilled personnel. Recently, researchers have developed hand-held calorimetry devices useful in clinical settings due to its portability, low cost, and small size. Only a limited number of studies have validated their use and results have been inconsistent.16 Furthermore, none of the previous studies were conducted with obese subjects. Given the uncertainty about the accuracy of these devices, clinicians usually use predictive equations to assess energy requirements. Predictive equations have generally been developed in healthy people. For some equations, overweight and obese people were included, but their relative contribution to the final equation often remains unclear. It is also uncertain how well predictive equations derived from one population can be applied to a different population, and many studies show that the race has an influence on metabolic rate.17–19 Most of the predictive equations have been derived from American and European populations; no equation has been developed or validated for the Saudi population. The goals of this study were to assess the validity of the BodyGem™ (BG) and predictive equations for estimating REE in overweight and obese Saudi subjects and to provide clinicians with recommendations regarding the use of these techniques.

Methods. Overweight or obese Saudi male participants aged 20–34 years were recruited for the study, conducted at the University of Wollongong (UOW) in Australia during the period from February 2007 to March 2008. Participants were recruited by email advertisements through the Saudi Students Association. A phone interview was used to screen applicants and informed consent was obtained. Healthy participants with body mass index (BMI) >25 kg/m² and aged more than 18 years were included. Participants were not on any particular diet and had maintained a steady weight for the previous 3 months (varying ±3 kg from the initial weight). Participants with diabetes or major illness or chronic diseases that affect REE were excluded. Study approval was given by the UOW Human Research Ethics Committee.

Height was measured in an upright position without shoes and weight and percent body fat were measured using scales with a bioelectrical impedance component (Tanita TBF-622). The measurements of REE were performed on 2 mornings, once with the WRC and 2 measures with the same BG device (HealtheTech Inc., Golden, CO, USA) separated by approximately a one week period. The protocol for measurements required that participants had been fasting, not smoking, and avoiding coffee and tea for at least 10 hours. Subjects had to refrain from drinking alcohol or doing any physical activity in the 24 hours measurements preceding.

The UOW WRC facility, described in detail elsewhere,20 consists of 2 separate air-tight, ventilated and air-conditioned chambers (3 x 2.1 x 2.4 m). Each chamber has a bed, desk, chair, hand basin TV/VCR, computer, phone and toilet. The room was calibrated every day using fresh air, span and nitrogen gases and approximately every 3 weeks by means of the methanol burning test. Oxygen consumption (VO₂) and carbon dioxide production (VCO₂) were converted to REE using the abbreviated Weir formula:21

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

where REE is measured in kcal/day and VO₂ and VCO₂ in L/d.

For REE measurements, each subject was asked to stay in the chamber for approximately 2 hours and measurement was started in the second hour with the subject lying quietly on a bed. The first hour should be sufficient to reach steady-state levels in the WRC. The WRC gives the measurement for O₂ consumption and CO₂ production at 10-minute intervals. The measurements were considered only when respiratory quotient (RQ), O₂ consumption and CO₂ production were stable for at least 3 consecutive readings (coefficient variation (CV) <10%).

Measurement of REE using the BG was taken immediately after the measurement of REE in the WRC while the subject was still lying down. Subjects were asked to wear a nose clip and to breathe into a disposable plastic mouth-piece for approximately 10 minutes. BodyGem devices usually begin to collect data when the first breath is detected and continue until a steady state is reached or stop after 12 minutes.5 Carbon dioxide production is not measured by BG.
devices and an RQ of 0.85 is assumed. These devices are autocalibrated prior to each measurement. For the second REE measurements with the BG, participants were asked to rest for 25 minutes before measurements were taken. This time period was sufficient to reach steady-state levels.

The Weir equation and the assumed 0.85 RQ value was used to calculate VO₂ and the following equation was derived to be used since the BG does not measure

$$\text{VO}_2 (\text{L/d}) = \text{REE (kcal)/4.8811}$$

To assess the source of error in BG, REE was estimated using the mean measured oxygen by BG and the actual (measured by WRC) and fixed (0.85) RQs. Also REE was estimated using mean oxygen measured by WRC and BG and the actual RQ. The following equation was used: 21

$$\text{REE (kcal/d)} = (3.941 * \text{VO}_2) + (0.85 * 1.106 * \text{VO}_2)$$

Where VO₂ is measured in L/d.

Seven prediction equations were used to calculate REE for each subject. 22-28 These equations are commonly used in clinical practice particularly in Saudi Arabia. 29, 30

Resting energy expenditure was calculated using actual body weight. For the Harris-Benedict equation only, all subjects with a BMI more than 30 had an adjusted weight substituted in the equation: Adjusted weight (kg) = [(actual weight – ideal weight) * 0.25] + ideal weight. 31 With the Harris-Benedict equation, REE was calculated by 3 different versions of that formula that are commonly used clinically in Saudi Arabia: using actual body weight, adjusted body weight, and with adjustment of the REE by a factor of 1.1. 29

The data are reported as mean ± SD. The level of statistical significance was set at p < 0.05. A p-value less than 0.05 was considered statistically significant. The data were analyzed using the Statistical Package for Social Sciences (Version 15.0, 2006) (SPSS Inc., Chicago, IL). Estimation between 90% and 110% of the REE measured by WRC was considered an accurate estimation; estimations <90% or >110% of the measured REE were classified as underestimation or overestimation.

### Results

Thirty-eight Saudi male participants completed the study (Table 1). Three did not repeat the measurement of BG due to discomfort caused by the device. The mean age was 26.8 ± 3.7 years and BMI was 31.0 ± 4.8 kg/m². There were no significant differences between overweight and obese subjects in terms of the accuracy of the BodyGem or predictions equations. Therefore, results are presented for the whole study group.

Paired t tests and Bland-Altman analysis indicated that BG is a reliable device for measurement of REE, with good between-test reproducibility. The mean REE values from the 2 BG measurements were not statistically different (1907 ± 381 versus 1886 ± 394 kcal/day) (p>0.24) (Table 2). The mean difference between the 2 readings was 21 ± 110 kcal, with limits of agreement ranging from +242 to -200 kcal (Figure 1), within the

### Table 1 - Subjects characteristics (n=38)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean ± SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td>26.8 ± 3.7</td>
<td>20-34</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>1.7 ± 0.1</td>
<td>1.6-2.0</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>92.0 ± 18.1</td>
<td>66.3-174.0</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>31.0 ± 4.8</td>
<td>25.1-50.0</td>
</tr>
<tr>
<td>Body fat (%)</td>
<td>28.8 ± 5.3</td>
<td>19.0-41.0</td>
</tr>
</tbody>
</table>

### Table 2 - Accuracy of resting energy expenditure (REE), oxygen consumption (VO₂) and respiratory quotient (RQ) measured by whole room calorimetry (WRC) and BodyGem (BG).

<table>
<thead>
<tr>
<th>Accuracy of REE</th>
<th>BG1 (n=38)</th>
<th>BG2 (n=35)</th>
</tr>
</thead>
<tbody>
<tr>
<td>REE (kcal/day)</td>
<td>1907 ± 381</td>
<td>1886 ± 394</td>
</tr>
<tr>
<td>Bias</td>
<td>21 ± 110</td>
<td></td>
</tr>
<tr>
<td>%Bias</td>
<td>1.1 ± 6.3</td>
<td></td>
</tr>
<tr>
<td>REE (kcal/day)</td>
<td>1896 ± 383</td>
<td>2069 ± 336*</td>
</tr>
<tr>
<td>Bias</td>
<td>173 ± 262</td>
<td></td>
</tr>
<tr>
<td>%Bias</td>
<td>11.0 ± 14.6</td>
<td></td>
</tr>
<tr>
<td>VO₂ (ml/min)</td>
<td>209.8 ± 54.5</td>
<td>289.8 ± 52.1*</td>
</tr>
<tr>
<td>Bias</td>
<td>20 ± 42</td>
<td></td>
</tr>
<tr>
<td>%Bias</td>
<td>9.1 ± 16.4</td>
<td></td>
</tr>
<tr>
<td>RQ</td>
<td>0.85 (fixed value)</td>
<td>0.81 ± 0.15</td>
</tr>
<tr>
<td>Bias</td>
<td>- 0.4</td>
<td></td>
</tr>
<tr>
<td>%Bias</td>
<td>- 4.7</td>
<td></td>
</tr>
</tbody>
</table>

Data are expressed as mean±Standard Deviation.

Bias = measured by WRC - measured by BG,

%Bias = was measured by WRC and BG multiply by 100/measured by BG.

*Significantly different from BG, p=0.0001 (P value from paired t test).
Assessment of REE in obese Saudi males ... Almajwal et al

250 kcal margin of acceptable error. The correlation coefficients of the 2 BG measurements were statistical significant (r =0.96, p=0.0001). When individual BG values were compared, 81.6% were within the acceptable 10% difference.

Mean REE and oxygen consumption were significantly different between WRC and BG, with a mean difference of 173±262 kcal/d and 20.0±41.9 ml/min, respectively. Limits of agreement of REE ranged from 697 to -351 kcal. These values were clinically unacceptable. Bland-Altman plots did not show any specific trend (Figure 2). However, BG underestimated REE by 11.0±14.6 and VO2 by 9.1±16.4%. Mean RQ measured by WRC (0.81±0.15) was significantly lower than the fixed RQ (0.85) used to calculate REE by BG.

To assess the source of errors in BG, REE was calculated using the oxygen uptake as measured by BG (269.8 ml/min) and the actual (0.81) or fixed (0.85) RQ values and this result in an estimated REE of 1896 and 1879 kcal/d, respectively. The use of the fixed RQ in the modified Weir equation introduced a mean error of 1.0% or 18 kcal/d. Similarly, REE was calculated using the actual RQ (0.81) and the measured oxygen by WRC (289.8 ml/min) and BG (269.8 ml/min) and this result in an estimated REE of 2019 and 1879 kcal/d, respectively. Therefore, measurement of VO2 by BG introduced a mean error of 7.4% or 140 kcal/d. Logistic regression analysis did not show any significant relationships between the accuracy of BG and other factors such as age, weight, height, BMI and body fat percentage. All predictive equations were highly correlated with the REE measured by WRC. Paired t test indicated that Schofield, WHO and Harris-Benedict equations were the most accurate equations (Table 3). The mean numerical bias and the mean percentage bias estimates for Schofield (wt) was 20±22 kcal (underestimation: 1.0±10.4%), Schofield (wt & ht) 28±22 kcal (underestimation: 1.0±10.4%), WHO (wt) 13±22 kcal (underestimation: 1.0±10.3%), WHO (wt & ht) 15±21 kcal (underestimation: 1.0±10.2%), and Harris-Benedict was 60±197 kcal (underestimation: 3.0±9.9%). The prediction accuracies were 60.5%, 65.8%, 63.2%, 65.8% and 68.4%, respectively.

The Harris-Benedict equation showed the highest percentage of accurate prediction (68.4%), but with higher percentages of underestimation (28.9%) and bias (3%) compared to Schofield and WHO equations. When adjusted body weight was used with an adjustment factor of 1.1, the percentage of accurate prediction decreased (39.5 and 47.4%, respectively).

The 5 other equations were not accurate according to the paired t test. The Owen, Mifflin and Bernstein equations underestimated REE in all subjects; the Ireton equation overestimated REE for all subjects with a very low accurate prediction (13%).

Accuracy was also evaluated using the Bland-Altman analysis. Considering the numerical and bias percentage, the lowest values were shown by the WHO (wt) equation; however, the extent of error was between +450

Figure 1 - Bland-Altman plot of differences in resting energy expenditure (REE) measures between 2 BodyGem (BG) measurements with the same device in overweight and obese Saudi males (n=35). Solid line indicates mean difference between the 2 measures and dashed lines indicate mean ± 2 standard deviation (limits of agreement).

Figure 2 - Bland-Altman plot of differences in resting energy expenditure (REE) measures between one Whole Room Calorimetry (WRC) measure and the mean of 2 BodyGem (BG) measurements in overweight and obese Saudi males (n=38). Solid line indicates mean difference between the 2 measures and dashed lines indicate mean ± 2 standard deviation (limits of agreement).
kcal and -425 kcal daily. The Harris-Benedict equation had the highest accurate prediction percentage, but the error range was +454 kcal to -335 kcal/day. The errors across all equations were more than the acceptable 250 kcal maximum (Figures are not shown). From visual examination, it appears to be more variation as the mean REE increased.

**Discussion.** The results showed that the mean bias between REE measured by BG was neither statistically nor clinically different. Therefore, the device can be considered to be reliable (consistent), but mean measurements of REE with the BG differed significantly, both statistically and clinically, from the WRC results and underestimated REE by 11% (about 173 kcal/d). Five studies have investigated the use of BG for REE measurements and all demonstrated within instrument reliability.\(^5^,6,13,33,34\) Herring\(^34\) found that 2 of 3 BG devices used gave consistently lower values (12% and 9%) compared to the metabolic carts. Similarly, Blanton et al.\(^35\) found that the BG device significantly underestimated REE by about 95 kcal. Two other studies reported that BG significantly overestimated REE.\(^6,13\) Both authors adjusted their findings, considering the energy expenditure due to holding the device during the measurements. After adjustment the difference became non-significant and the authors concluded that the BG provides valid and reliable measurements of REE. Only one study has reported that BG is an accurate and reliable device without any adjustment.\(^5\)

One systematic review indicated that the hand-held calorimeters are reliable and valid for the measurement of REE.\(^35\) However, that conclusion was based only on 4 studies in which the Douglas-bag system was used as the reference method. The present study supports the conclusion in a more recent review\(^16\) that the majority of studies with BG devices report significant differences between the BG measurements and standard metabolic carts. Studies that evaluated other portable hand-held calorimeters such as MedGem,\(^5,7,12,36-38\) Fitmate,\(^39\) VO2000 calorimeter\(^40\) and MOXUS modular VO\(_2\) system\(^14\) have also provided inconsistent results.

Several possible reasons might explain these differences in studies of the measurement of REE by BG and WRC. Firstly, each technique uses different equations to calculate REE. The WRC measures both VO\(_2\) and VCO\(_2\), whereas the BG measures only VO\(_2\) and assumes that the RQ of the subject is 0.85. In this study REE was calculated using the measured VO\(_2\) by BG and RQ measured by WRC. The results showed that the difference in REE due to the assumption of the fixed RQ is relatively small (1.0% or 18 kcal/d). Secondly, each technique measures REE over periods of different duration. The BG runs the test for 10 minutes; the WRC measurement lasted approximately one hour giving more time for subjects to reach a steady state. Thirdly, each technique uses different collection systems and gas analysis, which may introduce other errors.

The REE predicted by Schofield, WHO and Harris-Benedict equations was not significantly different from REE measured by WRC. For individuals, the limits of agreement of REE from the WRC compared to 7 predictive equations were well outside clinically acceptable levels (error >250kcal). The HB equation was the most accurate of the predictive equations and accurate prediction by Schofield (wt & ht) (65.8%) and

### Table 3 • Accuracy of Resting Energy Expenditure (REE) measured by BodyGem (BG) and REE predicted from prediction equations.

<table>
<thead>
<tr>
<th>Tools</th>
<th>REE (kcal/day)</th>
<th>Bias</th>
<th>%Bias</th>
<th>Accurate estimation</th>
<th>Under-estimation (%)</th>
<th>Over-estimation (%)</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole room calorimetry (WRC)</td>
<td>2069 ± 336</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BG</td>
<td>896 ± 383*</td>
<td>-11.0 ± 14.6</td>
<td>47.4</td>
<td>47.4</td>
<td>5.3</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Harris Benedict (HB)(^22)</td>
<td>2010 ± 267</td>
<td>-3.0 ± 9.9</td>
<td>68.4</td>
<td>28.9</td>
<td>2.6</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>HB*11</td>
<td>2211 ± 294*</td>
<td>-6.9 ± 9.0</td>
<td>47.4</td>
<td>5.3</td>
<td>47.4</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>HB (adjusted body weight)</td>
<td>1827 ± 150*</td>
<td>-13.0 ± 14.7</td>
<td>39.5</td>
<td>57.9</td>
<td>2.6</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Owen(^26)</td>
<td>1818 ± 184*</td>
<td>-14.0 ± 12.1</td>
<td>60.5</td>
<td>39.5</td>
<td>0.0</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Schofield (wt)(^24)</td>
<td>2050 ± 232</td>
<td>-1.0 ± 10.4</td>
<td>60.5</td>
<td>18.4</td>
<td>21.1</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Schofield, (wt &amp; ht)(^24)</td>
<td>2042 ± 231</td>
<td>-1.0 ± 10.4</td>
<td>65.8</td>
<td>18.4</td>
<td>15.8</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Millin(^28)</td>
<td>1867 ± 207*</td>
<td>-11.0 ± 10.3</td>
<td>63.2</td>
<td>36.8</td>
<td>0.0</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>WHO (wt)(^31)</td>
<td>2057 ± 235</td>
<td>-1.0 ± 10.3</td>
<td>63.2</td>
<td>15.8</td>
<td>21.1</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>WHO (wt &amp; ht)(^35)</td>
<td>2055 ± 222</td>
<td>-1.0 ± 10.2</td>
<td>65.8</td>
<td>15.8</td>
<td>18.4</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Ireton(^32)</td>
<td>2557 ± 161*</td>
<td>+19.0 ± 9.1</td>
<td>13.2</td>
<td>0.0</td>
<td>86.8</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Bernstein(^33)</td>
<td>1585 ± 246*</td>
<td>-31.0 ± 12.8</td>
<td>0.0</td>
<td>100</td>
<td>0.00</td>
<td>0.82</td>
<td></td>
</tr>
</tbody>
</table>

Bias - REE measured by WRC – REE measured by BG or predicted by applied equations

%Bias - [(REE measured by WRC – REE measured by BG or predicted by equations) / REE measured by BG or predicted by equations] * 100

Accurate estimation: percentage of all subjects whose REE was within 90% to 110% of measured REE by WRC

Underestimation: percentage of all subjects whose REE was less than 90% of measured REE by WRC

Overestimation: percentage of all subjects whose REE was more than 110% of measured REE by WRC

*Statistically difference between REE measured by WRC and REE measured by BG or predicted by equations on a two-tailed paired t-test (p=0.0001)

r - Pearson’s Correlation Coefficient between REE measured by WRC and REE measured by BG or predicted by applied equation

WHO (wt & ht) (65.8%) equations were also very close to the HB equation (68.4%). The Mifflin and Owen equations were less accurate and both equations tend to underestimate REE in overweight and obese people, as has been reported elsewhere. When equations developed in obese populations were used (Bernstein and Ireton), measured REE was significantly different from the predictions, which is consistent with previous reports. Several statistical methods were included in this study to examine the accuracy of BG and prediction equations in the assessment of REE. To study validity, the aim should be to examine whether the outcomes are equal. Paired t-test, regression analysis and correlation coefficient can be used, but tell us little about ‘is the same’. Bias and bias percent are essential, however over- and underestimations counter balance, therefore provide no information on ‘is the same’. Accurate estimation is a good measure but may not be appropriate to be used in this project due to the small sample size. However, paired t test, bias, %bias and accurate estimation indicated that HB, Schofield, and WHO equations are more accurate than other equations.

There are several limitations to this study. First, it included only men, and further investigation is needed in women. Secondly, the sample size is small and included young subjects (<35 years) only. Thirdly, the study included only overweight and obese Saudi subjects. This is the group in which the measurement of REE is likely to be of most relevance, however, the use of normal weight subjects could produce different results. Fourthly, random order for the measurement of REE using WRC and BG were not undertaken. The primary reason is that the WRC requires one hour to reach steady-state level and therefore, for convenience the first measurement of REE using BG was undertaken after the measurement of REE using WRC.

In conclusion, this project indicates that the HB, Schofield and WHO equations predict REE more accurately than the BG device. However, their accuracy was not clinically acceptable on an individual level. Without more studies to clarify these issues, the evidence from this single small study is insufficient to recommend changes in practice. Until then dietitians and other health professionals in Saudi Arabia may need to continue using the available international predictive equations for the assessment of individual energy requirements, while being aware of their inherent inaccuracy.

Acknowledgment. We would like to thank Mr. Greg Teuss for his technical assistance.

References


---

**Case Reports**

Case reports will only be considered for unusual topics that add something new to the literature. All Case Reports should include at least one figure. Written informed consent for publication must accompany any photograph in which the subject can be identified. Figures should be submitted with a 300 dpi resolution when submitting electronically or printed on high-contrast glossy paper when submitting print copies. The abstract should be unstructured, and the introductory section should always include the objective and reason why the author is presenting this particular case. References should be up to date, preferably not exceeding 15.