Issues relating to normalization of body fat content in men and women

MI Goran¹, DB Allison² and ET Poehlman³

¹Energy Metabolism Research Unit, Department of Nutrition Sciences, University of Alabama, Birmingham, Alabama 35294; ²Obesity Research Center, St Luke’s Roosevelt Hospital Center, Columbia University College of Physicians; ³Baltimore VA Medical Center, Division of Geronontology/GRECC, Department of Medicine, University of Maryland, Baltimore, Maryland 21201, USA

OBJECTIVE: Body mass index (BMI), % body fat, and the fat:lean ratio are ratios frequently used as obesity indices. Ratios are based on an assumption that the regression between the numerator (e.g. fat mass) and the denominator (e.g. body mass) has a zero intercept. As shown in the companion paper, non-zero intercepts cause several problems when ratios are used to adjust data and analysis of covariance (ANCOVA) is frequently the preferred statistical tool. The purpose of this paper is to examine whether BMI, % body fat and the fat:lean ratio meet the necessary criteria for suitable obesity indices using gender comparisons as an example.

RESULTS: In 720 healthy men and women, BMI was higher in men (24.9 ± 3.3 vs 23.4 ± 3.2 kg/m², P < 0.001), but fat mass, % fat and the fat:lean ratio were higher in women (14.8 ± 8.4 vs 19.1 ± 8.1 kg fat; 18.6 ± 8.7 vs 29.9 ± 9.7% body fat; 0.24 ± 0.14 vs 0.46 ± 0.20 for the fat:lean ratio; P < 0.001). Body mass (BM) was correlated with height² in men (r = 0.40) and women (r = 0.36) with equivalent regression slopes (17.1 ± 1.9 vs 15.6 ± 2.3 kg per m² in women), but the intercepts were different from zero (24.1 kg in men, 20.7 kg in women). When BM was adjusted for height² using ANCOVA, men remained significantly heavier than women (74.4 ± 11.0 vs 68.8 ± 11.6 kg; P < 0.001). Fat mass (FM) was significantly correlated with BM in males (r = 0.64) and females (r = 0.78) but the regression slopes were different (0.49 ± 0.03 vs 0.71 ± 0.03 kg of fat per kg body mass in females; P < 0.05) and the intercepts were different from zero (-23.2 ± 2.2 kg in males; -24.8 ± 2.1 kg in females). FM adjusted for BM was significantly higher in women (11.7 ± 25.6 kg). FM was inversely correlated with fat free mass (FFM) in males (r = -0.17) and females (r = -0.20), with similar regression slopes (-0.16 ± 0.05 vs -0.26 ± 0.08 kg of FM per kg of FFM in women) and the intercepts were significantly different from zero (24.8 ± 3.0 kg in males; 30.7 ± 3.6 kg in females). When FM was adjusted for FFM, there was no significant difference between men (16.3 kg) and women (17.0 kg).

CONCLUSIONS: It is concluded that: (a) the presence of significant intercepts does not support the use of ratios as obesity indices and regression based models should be considered; and, (b) the direction and magnitude of the difference in obesity index between men and women changes with different normalization approaches.

Keywords: body fat; body mass index; ratios; obesity indices

Introduction

Ratios are frequently used in body composition research as obesity indices. For example, body mass index (the ratio of body mass to height squared) is an obesity index used in epidemiological studies and in the absence of more accurate techniques for measuring body fat content. Percent body fat has traditionally been used as an index of relative adiposity, where percent body fat is the ratio of fat mass to body mass multiplied by 100. A third example is the fat:lean ratio (fat mass divided by fat free mass) which is an attempt to adjust fat mass relative to fat free mass. Historical¹ and more recent²³ studies suggest that using ratios to normalize, or adjust, biological data may lead to spurious comparisons if the assumptions inherent in the ratio method are not met. Specifically, using ratios to adjust a dependent variable (e.g. fat mass, the numerator in the ratio) for a potential covariate (e.g. body mass, the denominator in the ratio) assumes that the regression between the dependent variable and the covariate is linear and has a zero intercept (see accompanying paper by Allison et al.). Biological relationships however rarely regress to a zero intercept.¹ As shown in our accompanying paper, there are numerous problems associated with using ratios when deviations in the underlying assumptions occur, resulting in difficulties with data interpretation.

Analysis of covariance, a regression based approach, has been suggested as an alternative statistical technique to adjust physiological data including resting metabolic rate,² peak oxygen consumption,³ and musculo-skeletal size.⁴ We have recently shown that the assumptions of the ratio method are not met for either resting metabolic rate or peak VO₂ (where body mass or fat free mass are often used as denominators), thus leading to spurious conclusions when comparing individuals with large differences in body composition.¹³ In the case of resting metabolic rate, when using a ratio with fat free mass as the denominator, women appear to have a higher resting metabolic rate than men, but when data are adjusted using analysis of covariance, women have a lower resting metabolic rate than men.⁵ Similarly, when total energy expenditure is divided by resting energy expenditure, children appear to have a lower activity index than adults, but when total energy expenditure is adjusted for resting energy expenditure using a regression approach, children and adults have a similar total energy expenditure.⁶ This discrepancy in data interpre-
tation is explained in these examples by the failure of the ratio approach to take into account non-zero intercepts for the regression between resting metabolic rate and fat free mass and between total energy expenditure and resting energy expenditure. In the presence of a positive intercept, statistical bias is introduced when ratios are used to adjust the data such that subjects with a higher value of the covariate (i.e., fat free mass) appear to have a lower ratio. This concept has been discussed in detail elsewhere as well as in the accompanying paper.

Issues relating to expression of body composition have not previously been considered, although the aforementioned issues apply since the traditional body composition/obesity indices (e.g., body mass index; percent body fat; the fat:lean ratio) are ratios. In the present paper, we examine whether the underlying assumptions inherent in the use of these ratios as mathematical indices to compare fatness between cohorts are satisfied. As an example, we focus on examining gender differences in body fatness using data from 720 healthy men and women.

Methods

Subjects
Data from 427 healthy Caucasian men (17–90 years) and 293 healthy Caucasian women (18–88 years) were analyzed. The data were collected from ongoing studies at the Clinical Research Center at the University of Vermont, and some data have been reported elsewhere. All studies were approved by the Committee on Human Research for the Medical Sciences at the University of Vermont.

Measurement of body composition
Body mass was measured in light clothing and without shoes on a scale to the nearest 0.01 kg and was followed by a measurement of height to the nearest 0.5 cm via a fixed wall-mounted metric ruler. Whole body density was estimated by averaging six measures of underwater weight, with simultaneous measurement of residual lung volume by helium dilution as previously described. Percent body fat was estimated from the Siri equation and fat free mass was calculated as body weight minus fat mass. The test-retest reliability of body composition measures in our laboratory has a coefficient of variation of 5% and an intraclass correlation coefficient (ICC) of 0.98 using equation ICC (3,1) from Shrout and Fleiss.

Statistics
Differences in subject characteristics between men and women were tested using one way ANOVA. Association between pairs of variables was examined using the Pearson correlation coefficient and by regression analysis. Body mass was examined as a function of height squared, and fat mass was analyzed as a function of body mass, and fat free mass using analysis of covariance. Our criteria for implementing analysis of covariance were the presence of a significant correlation between fat mass and the covariate and equality of the regression slopes for men and women. Slopes of the regression lines were compared by examining the significance of the interaction term between gender (the grouping variable) and the covariate being examined. For the regression between fat mass and body mass, the regression slopes were different in men and women. In this case the Johnson-Neyman technique was used, as described by Pedhazur. The Johnson-Neyman technique computes the range of body mass that is associated with significant gender-related differences in fat mass and calculates adjusted means. Statistics were computed using SAS for Windows (Carey, North Carolina). The level of statistical significance was set at a probability of P ≤ 0.05 for all tests. Observed variables are reported as mean ± standard deviation and regression slopes and intercepts are reported as mean ± standard error.

Results

We examined data from 427 healthy men (46 ± 20 years; 1.77 ± 0.07 m in height; 77.6 ± 10.9 kg body mass) and 293 healthy women (50 ± 18 years; 1.63 ± 0.06 m in height; 62.2 ± 9.0 kg body mass). Body mass index was significantly higher in men (24.9 ± 3.3 vs 23.4 ± 3.2 kg/m²; P < 0.001), but fat mass, percent body fat, and the fat:lean ratio were significantly higher in women (14.8 ± 8.4 vs 19.1 ± 8.1 kg fat; P < 0.001; or 18.6 ± 8.7 vs 29.9 ± 9.7% body fat; 0.24 ± 0.14 vs 0.46 ± 0.20 for the fat:lean ratio; P < 0.001).

We first examined the underlying assumptions of using body mass index as a ratio by examining the relationship between body mass and height squared. As shown in Table 1, body mass was significantly correlated with height squared in men (r = 0.40; P < 0.001) and women (r = 0.36; P < 0.001) with similar regression slopes (17.1 ± 1.9 kg per m² in men; 15.6 ± 2.3 kg per m² in women) although the intercept was significantly different from zero in both men and women.

Table 1 Regression between fat mass and various independent variables in men and women

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable = Fat mass (kg)</th>
<th>Males</th>
<th></th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>Slope (s.e.)</td>
<td>Intercept (s.e.)</td>
<td>r</td>
</tr>
<tr>
<td>Body mass, kg</td>
<td>0.64</td>
<td>0.49 (0.03)</td>
<td>-23.2 (2.2)</td>
<td>0.78</td>
</tr>
<tr>
<td>Fat free mass kg</td>
<td>-0.17</td>
<td>-0.16 (0.05)</td>
<td>24.8 (3.0)</td>
<td>-0.20</td>
</tr>
</tbody>
</table>

s.e. = standard error
Table 2 Comparison of obesity indices between men (n = 427) and women (n = 293) using various data normalization approaches

<table>
<thead>
<tr>
<th></th>
<th>Men (s.d.)</th>
<th>Women (s.d.)</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Absolute data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM, kg</td>
<td>77.6 (10.9)</td>
<td>62.2 (9.0)*</td>
<td>Men are 25% heavier than women</td>
</tr>
<tr>
<td>FM, kg</td>
<td>14.8 (8.4)</td>
<td>19.1 (8.1)*</td>
<td>Women have 29% more FM than men</td>
</tr>
<tr>
<td>2. Ratios</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>24.9 (3.3)</td>
<td>23.4 (3.2)*</td>
<td>BMI is 6% higher in men</td>
</tr>
<tr>
<td>% Body fat</td>
<td>18.6 (8.7)</td>
<td>29.9 (9.7)*</td>
<td>% Body fat is 61% higher in women</td>
</tr>
<tr>
<td>FM:FFM</td>
<td>0.24 (0.14)</td>
<td>0.46 (0.20)*</td>
<td>Ratio of fat to lean tissue is 92% higher in women</td>
</tr>
<tr>
<td>3. Regressions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM adjusted for BM (kg)</td>
<td>11.7</td>
<td>25.6*</td>
<td>After taking differences in BM into account women have 120% more FM than men</td>
</tr>
<tr>
<td>BM adjusted for height squared (kg)</td>
<td>74.4</td>
<td>66.8*</td>
<td>After taking differences in height into account men are 11% heavier than women</td>
</tr>
<tr>
<td>FM adjusted for FFM (kg)</td>
<td>16.3</td>
<td>17.0</td>
<td>After taking differences in FFM into account women have 4% more FM than men</td>
</tr>
</tbody>
</table>

s.d. is standard deviation; body composition was assessed from underwater weighing; BM = body mass; FM = fat mass; FFM = fat-free mass; *significant effect of gender (P < 0.001)

(24.1 ± 6.0 kg) and women (20.7 ± 6.2 kg). When body mass was adjusted for height squared using analysis of covariance, adjusted body mass was significantly higher in men compared to women (74.4 vs 66.8 kg; P < 0.001). Thus, when the body mass index approach is implemented using an analysis of covariance approach (i.e. adjusting weight for height squared), men remain heavier than women.

We next examined the nature of the relationship between fat mass and body mass to examine the underlying assumptions of using percent body fat as a ratio to adjust body fat content. Fat mass was highly correlated with body mass in males (r = 0.64; P < 0.001) and females (r = 0.78; P < 0.001). As shown in Table 1, the slopes for the regression between fat mass and body mass were significantly different in men and women and intercepts were similar although both significantly different from zero. Comparison of fat mass between men and women using conventional analysis of covariance with body mass as a covariate could not be implemented because of the inequality of the regression slopes (Table 1). Therefore we used the Johnson-Neyman technique to examine fat mass as a function of body mass in men and women. Below 25 kg in body mass (which is well below the lowest value in our data set) there was no significant difference in fat mass between men and women. However, above 25 kg in body mass, men had significantly less fat mass than women. Fat mass adjusted for body mass was 11.7 kg in men and 25.6 kg in women (P < 0.0001).

Finally, we examined the nature of the relationship between fat mass and fat free mass to examine the underlying assumptions of using the fat/lean ratio as an obesity index. Fat mass was correlated (inversely) with fat free mass in both males (r = -0.17; P < 0.001) and females (r = -0.20; P < 0.001). As shown in Table 1, there was no significant difference in the regression slope for fat mass and fat free mass between men and women and the intercepts were both significantly different from zero. When fat mass was adjusted for fat free mass using analysis of covariance, there was no significant difference in the adjusted means for fat mass between males (16.3 kg) and females (17.0 kg).

Table 2 summarizes the difference in various obesity indices between men and women using various normalization approaches. As seen in the Table, the direction and magnitude of the difference between men and women varies depending on the normalization approach selected.

Discussion

The purpose of this paper was to examine data normalization techniques to express and compare body fat content in men and women. Specifically, we examined the underlying assumptions of three obesity indices which are ratios (i.e. body mass index, percent body fat, and the fat/lean ratio). In all three cases, non-zero intercepts were observed when the numerator of the ratio was examined as a function of the denominator by regression analysis. The presence of these intercepts suggests that alternative statistical approaches should be considered to normalize the numerator for the effect of the denominator in the ratio. Specifically, analysis of covariance (a regression based approach), has previously been suggested as a means to statistically adjust a dependent variable to more accurately reflect the nature of the relationship with a covariate. Analysis of covariance examines the regression between a dependent variable (e.g. fat mass) and a covariate (e.g. body mass or fat free mass). A relative ‘high’ or a ‘low’ fat mass is defined by the residual between an individual value for fat mass and that value predicted by the regression equation relating fat mass to the covariate. However, use of the analysis of covariance method to compare two or more groups requires the slopes of the regression lines to be similar, otherwise the Johnson-Neyman technique or variations thereof should be used.

In the present paper we use the example of gender comparisons of body fat to illustrate the appropriate normalization procedure. In the first case we selected body mass as a normalizing variable, and examined the regression between fat mass and body mass and found significant intercepts for men and women (see Table 1). Therefore the simple ratio approach was not considered appropriate. We therefore employed a model based on analysis of covariance (ANCOVA). A key assumption of ANCOVA is homogeneity of the regression slopes in the groups being examined.
(i.e. the regression slope between fat mass and body mass should be equal for men and women). However the slopes were significantly different between men and women (Table 1). Therefore the Johnson-Neyman technique was used (an extension of ANCOVA suitable for use with heterogenous regression slopes which computes 'regions of significant differences'). The Johnson Neyman analysis showed that fat mass controlled for body mass was significantly higher in women above 25 kg of body mass. Although the Johnson-Neyman technique does not assume homogenous regression slopes, it does require the usual regression assumptions (i.e. normally distributed, homoscedastic residuals). However, in this case the residuals were significantly skewed among females and significantly heteroscedastic across sex and weight. We therefore attempted to refit the model after transforming fat mass to normality via a Box-Cox transformation. The transformation employed (fat mass\(^{0.48}\)) did not improve the extent to which the residuals fitted the regression assumptions. The next alternative in this case was to employ the alternative model described in Equation 2 of the companion paper by Allison et al. In this model the dependent variable is the ratio of fat mass to body mass, and the co-variate is the inverse of body mass. This model may be a useful alternative since the regression slopes were not significantly different in men and women \((P = 0.29)\), and the residuals were homoscedastic and not significantly skewed.

The presence of a significant (negative) intercept for the relationship between fat mass and body mass is problematic and introduces bias into data comparisons when body fat is expressed as a ratio with body mass as the denominator. This point is illustrated in Figure 1 where four hypothetical male subjects are depicted with increasing body mass and fat mass. All four subjects lie on the regression line as defined by fat mass (kg) = 0.5 * body mass (kg) -23 kg (the regression line derived for men in the present study), and by definition, fat mass relative to body mass is equal for all subjects. However, there are large differences in relative percent body fat, particularly at the lower values of fat and body mass. In addition, at the other extreme of the regression line, there are only small increases in percent body fat for similar increases in body mass and fat mass. Thus, percentage of body fat does not increase in a linear fashion with uniform step increases in fat mass. Finally, it should be noted that according to the observed mathematical relationship between body mass and fat mass, the predicted theoretical maximum for percent body fat is 50% at infinite body mass. In summary, the shortcomings of the ratio approach are summarized by the fact that the subject with the highest body mass and fat mass (39% body fat) appears 'fatter' than the smaller subject with the lowest body mass and fat mass (4% body fat) even though both subjects lie on the same regression line. This apparent discrepancy is explained by the fact that adjusting fat mass by body mass using a ratio fails to take into account the non-zero intercept that is present in the equation relating fat mass to body mass. The underlying explanation of this phenomenon is explained in the accompanying theoretical paper by Allison et al.

We considered using fat free mass as an alternative co-variate to normalize fat mass. In examining the relationship between fat mass and fat free mass we were surprised by the nature of the inverse correlation. In intervention studies in humans (e.g. exercise, fasting, overfeeding), there is usually concordance in the direction of change that occurs when fat and fat free mass change. The inverse relationship observed between fat and fat free mass may be a reflection of the cross-sectional nature of our cohort depicting leaner (increased fat free mass) and thinner (decreasing fat mass) subjects. Nevertheless, our data were used to illustrate that in a given cohort of women who appear fatter than a cohort of men based on absolute fat mass and relative percent body fat, body fat is similar between the two cohorts when fat mass is adjusted for differences in fat free mass. In other words, the proportional increase in fat mass in women is not any different than that predicted by the decreased fat free mass.

It should be pointed out that fat free mass is not a perfect co-variate for fat mass because of the inter-dependence of the two variables (both are strongly related to body mass and fat free mass is derived from body mass minus fat mass). However, the same argument applies when selecting body mass as a co-variate for fat mass, since fat mass is a component of body mass. For this reason we considered using height as a potential covariate to adjust fat mass, since fat mass and height are measured independent of one another. However, the lack of correlation between fat mass and height implies that adjusting for height would have negligible impact on the data. Thus, there is a clear absence of an 'ideal' covariate to normalize body fat mass.

One striking result of our analysis is the discrepancy of the magnitude and direction of the gender difference in body fat depending on the normalization approach used (Table 2). For example, on an absolute basis, women have 29% more fat than men, although relative percent body fat is 61% higher in women, and when adjusted for body mass, fat mass is 120% higher in women. Even if it is assumed that the regression approach is correct, discrepancies are observed depending on the choice of the co-variate. Thus, body mass adjusted for height squared is 11% higher in
on how obesity is expressed. Determination of the most appropriate way to express obesity ultimately depends on how obesity is defined in functional terms. As a simplistic example, if obesity is defined as the threshold of body fatness beyond which risk of disease increases, then, from a practical standpoint, the obesity index which has the strongest relationship with negative health outcome may be the most valuable. For example, in Table 3, fat mass adjusted for fat free mass is the index of body fat that has the highest correlation with known risk factors such as dietary fat intake, blood pressure and triglyceride concentration.

In summary, we have used body composition data from 720 healthy men and women to examine several issues relating to expression of body composition and obesity indices. The example of gender-related differences in body composition was used to illustrate the point that direction and magnitude of differences in obesity indices between men and women depends on how the data are expressed (see Table 2). Thus, in our cohort of 720 healthy men and women, fat mass, relative percent body fat, the fat:lean ratio, and fat mass adjusted for body mass were significantly higher in men. On the other hand, body mass index and body mass adjusted for height squared were significantly higher in men, and fat mass adjusted for fat free mass was similar between men and women. Our analysis does not lend mathematical support in favor of using ratios which are traditionally used as obesity indices (i.e. body mass index, percent body fat, and the fat:lean ratio) because of the presence of non-zero intercepts in the regressions relating the denominator to the numerator for the ratio in question. Although we have used gender differences to elucidate limitations in data normalization approaches for body fat, our analysis may have widespread implications for the expression of body fat content and other obesity-related variables which are often expressed as ratios (e.g. waist to hip ratio; total energy expenditure to resting energy expenditure activity factor; intra-abdominal fat to subcutaneous fat ratio), as similar issues apply.

Acknowledgements
This work was supported in part by National Institutes of Health Grants R29 HD 28720 (MIG), R01 AG 07857 (ETP), RCDA AG 00564 (ETP), P30 DK 26687 (DBA), The United States Department of Agriculture (MIG), and a Career Development Award from the American Diabetes Association (DBA).

---

Table 3 Univariate correlations between obesity indices and cardiovascular disease risk factors in the entire sample of 720 healthy men and women

<table>
<thead>
<tr>
<th></th>
<th>Fat intake</th>
<th>Supine blood pressure</th>
<th>Diastolic blood pressure</th>
<th>Cholesterol</th>
<th>Triglyceride</th>
<th>HDL</th>
<th>LDL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body mass index</td>
<td>ns</td>
<td>0.22</td>
<td>0.27</td>
<td>0.26</td>
<td>0.34</td>
<td>ns</td>
<td>0.27</td>
</tr>
<tr>
<td>Fat mass</td>
<td>0.097</td>
<td>0.27</td>
<td>0.14</td>
<td>0.47</td>
<td>0.41</td>
<td>0.13</td>
<td>0.48</td>
</tr>
<tr>
<td>% Fat</td>
<td>0.096</td>
<td>0.28</td>
<td>0.15</td>
<td>0.47</td>
<td>0.41</td>
<td>0.14</td>
<td>0.48</td>
</tr>
<tr>
<td>Fat mass adjusted for body mass</td>
<td>0.10</td>
<td>0.27</td>
<td>0.14</td>
<td>0.47</td>
<td>0.41</td>
<td>0.13</td>
<td>0.48</td>
</tr>
<tr>
<td>Fat mass adjusted for fat free mass</td>
<td>0.15</td>
<td>0.33</td>
<td>0.24</td>
<td>0.38</td>
<td>0.45</td>
<td>-0.14</td>
<td>0.42</td>
</tr>
</tbody>
</table>

All correlations are significant (P < 0.05) unless stated otherwise.
References


