Percent body fat via DEXA: comparison with a four-compartment model

GRANT E. VAN DER PLOEG,1 ROBERT T. WITHERS,1 AND JOE LAFORGIA2
1Exercise Physiology Laboratory, School of Education, Flinders University, and 2School of Pharmaceutical, Molecular, and Biomedical Sciences, University of South Australia, Adelaide 5001, Australia
Submitted 16 May 2002; accepted in final form 18 September 2002

Van der Ploeg, Grant E., Robert T. Withers, and Joe Laforgia. Percent body fat via DEXA: comparison with a four-compartment model. J Appl Physiol 94: 499–506, 2003; 10.1152/japplphysiol.00436.2002.—This study compared body composition by dual-energy X-ray absorptiometry (DEXA; Lunar DPX-L) with that via a four-compartment (4C; water, bone mineral mass, fat, and residual) model. Relative body fat was determined for 152 healthy adults [30.0 ± 11.1 (SD) yr; 75.10 ± 14.88 kg; 176.3 ± 8.7 cm] aged from 18 to 59 yr. The 4C approach [20.7% body fat (%BF)] resulted in a significantly (P < 0.001) higher mean %BF compared with DEXA (18.9% BF), with intra-individual variations ranging from −2.6 to 7.3% BF. Linear regression and a Bland and Altman plot demonstrated the tendency for DEXA to progressively underestimate the %BF of leaner individuals compared with the criterion 4C model (4C %BF = 0.862 × DEXA %BF + 4.417; r² = 0.952, standard error of estimate = 1.6% BF). This bias was not attributable to variations in fat-free mass hydration but may have been due to beam-hardening errors that resulted from differences in anterior-posterior tissue thickness.

hydrodensitometry; isotopic dilution; multicompartment body composition models; dual-energy X-ray absorptiometry

THE MEASUREMENT OF HUMAN BODY COMPOSITION to quantify nutritional and health status has become increasingly important as evidence accumulates identifying relative body fat as a significant predictor of mortality (2, 32). In addition, the fat-free mass (FFM) is frequently used to standardize or index physiological variables, such as resting metabolic rate and power. The methods regularly employed to assess body composition are limited by the generalized assumption(s) that must be applied across the entire population. Hence, because true percent body fat (%BF) is presently unmeasurable, there is a need for a highly accurate reference that can be universally applied across the entire population.

Until recently, the hydrodensitometric model was regarded as the “gold standard” for body composition assessment. This model partitions the body into two compartments of constant densities [fat mass = 0.9007 g/cm³ and FFM = 1.100 g/cm³] and assumes that the relative amounts of the FFM components [water, protein, bone mineral (BM), and non-BM] are fixed (5). Hydrodensitometry is clearly inappropriate for individuals who deviate from these fixed and/or assumed values (e.g., children, elderly, blacks, obese), and its application is, therefore, somewhat limited (21, 31, 40).

Notwithstanding these limitations, two-compartment body composition models (including hydrodensitometry, hydrometry, and total body potassium) in combination form the basis for multicompartment models that provide improved accuracy over the former because they control for the biological variability in one or more of the FFM components. Furthermore, many investigators believe that multicompartment models provide the criterion or gold standard measurement of %BF (16, 17, 37) because they are not known to be age, gender, race, and health status dependent. Nevertheless, despite their increasing popularity, these models are costly, laborious, and inconvenient for many clinical purposes. For these reasons, dual-energy X-ray absorptiometry (DEXA) has rapidly gained acceptance as a reference method for body composition analysis. Originally designed to determine bone density, DEXA technology has subsequently been adopted for the assessment of whole body composition, which has enabled rapid, noninvasive %BF estimates with minimal radiation exposure to be obtained. DEXA also has the advantage of being a three-compartment model that quantifies fat, soft lean tissue, and BM, and also yields regional as well as total body values. However, DEXA is not without limitations, and, although a precise measurement of body composition is provided, there are still considerable concerns about its validity, especially at extremes of tissue depth and hydration level (18, 20, 26, 27).

Wang et al. (37) have questioned the specific role of DEXA in clinical evaluations and in research studies and stated “the errors of the DXA [DEXA] method are still of concern if it were to be used as the criterion.” Although researchers have previously compared DEXA with multicompartment models (6, 11, 28), much of this work has been conducted on a restricted population. For example, comparisons have been made by

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
using young (6, 28) or old groups (6), and a further investigation (11) examined a wide age range but excluded those engaged in vigorous physical training. In fact, no investigation has sampled a large uninterrupted age group that has included very lean athletes. The aim of this investigation was, therefore, to test the hypothesis that DEXA underestimates the %BF of lean individuals compared with the criterion multicompartment [four compartment (4C)] model.

**METHODS**

**Subjects.** A heterogeneous sample of 152 healthy adult Australian men (n = 118) and women (n = 34), 18–59 yr of age, volunteered for this project, which was approved by the Flinders Medical Centre’s Committee on Clinical Investigation. Informed consent was obtained in accordance with the established protocol for human subjects. All experiments were conducted in the morning when the subjects were postabsorptive, euhydrated, and had not exercised for 36 h. They were requested to void before testing in an attempt to eliminate any flatus in the gastrointestinal tract.

**Anthropometry.** Height was determined to within 1 mm by using a wall stadiometer, and body mass was measured to the nearest 0.1 kg by a calibrated electronic scale (model 89). Gluteal or hip girth (n = 89 only) was recorded at the level of the greatest posterior protuberance of the buttocks (25) by using a steel tape (model W600PFT, Lufkin, executive thinline).

**Body composition.** The 4C body composition model involves measurements of body density (BD), total body water (TBW), and BM mass (BMM) by hydrodensitometry, isotopic dilution, and DEXA, respectively. These procedures have been fully described previously (36, 39). Briefly, BD was determined by hydrodensitometry with the associated gas in the respiratory system measured by either helium dilution (residual volume) or oxygen dilution (approximate functional residual capacity). TBW was estimated from saliva samples by deuterium dilution (40 mg $^2$H$_2$O/kg dose) by using an isotope ratio mass spectrometer with a 4% correction factor applied for isotopic exchange with nonaqueous hydrogen (30). BMM was obtained by multiplying the BM content or bone mass (Quetelet’s body mass index or body mass index (BMI), and gluteal girths (n = 89) were used as indicators) were evaluated by linearly regressing each of these variables on the %BF differences between the two methods. The 0.05 level was used for all tests of statistical significance.

**RESULTS**

The descriptive statistics for the sample are presented in Table 1. There was no significant difference (P = 0.292) between scale mass (75.10 ± 14.88 kg) and DEXA-determined mass (75.15 ± 14.90 kg) with an interclass correlation coefficient of 0.999 [standard error of estimate (SEE) = 0.56 kg] and individual differences ranging from −1.99 to 1.52 kg (0.05 ± 0.56 kg). The small gender difference for %BF is because 22

**Table 1. Descriptive statistics for sample**

<table>
<thead>
<tr>
<th></th>
<th>Men (n = 118)</th>
<th>Women (n = 34)</th>
<th>Combined (n = 152)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>31.1 ± 11.7</td>
<td>26.1 ± 7.8</td>
<td>30.0 ± 11.1</td>
</tr>
<tr>
<td>Height, cm</td>
<td>179.2 ± 6.7</td>
<td>166.5 ± 6.9</td>
<td>176.3 ± 8.6</td>
</tr>
<tr>
<td>Mass, kg</td>
<td>80.85 ± 12.79</td>
<td>57.83 ± 6.09</td>
<td>75.10 ± 14.88</td>
</tr>
<tr>
<td>DEXA mass, kg</td>
<td>80.16 ± 12.75</td>
<td>57.76 ± 6.13</td>
<td>75.15 ± 14.90</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>24.9 ± 3.4</td>
<td>20.9 ± 2.1</td>
<td>24.0 ± 3.6</td>
</tr>
<tr>
<td>%BF, 4C</td>
<td>20.5 ± 7.7</td>
<td>21.5 ± 6.6</td>
<td>20.7 ± 7.4</td>
</tr>
<tr>
<td>%BF, DEXA</td>
<td>18.6 ± 8.7</td>
<td>19.8 ± 7.6</td>
<td>18.9 ± 8.4</td>
</tr>
<tr>
<td>Body density, g/cm³</td>
<td>1.0575 ± 0.0184</td>
<td>1.0550 ± 0.0155</td>
<td>1.0570 ± 0.0178</td>
</tr>
<tr>
<td>FFM density, g/cm³</td>
<td>1.1067 ± 0.0045</td>
<td>1.1067 ± 0.0060</td>
<td>1.1067 ± 0.0048</td>
</tr>
<tr>
<td>FFM hydration *</td>
<td>72.3 ± 1.0</td>
<td>72.6 ± 1.2</td>
<td>72.4 ± 1.0</td>
</tr>
<tr>
<td>%BM/MM/FFM</td>
<td>5.66 ± 0.42</td>
<td>5.90 ± 0.62</td>
<td>5.71 ± 0.48</td>
</tr>
<tr>
<td>TBW, kg</td>
<td>45.58 ± 5.41</td>
<td>33.01 ± 4.53</td>
<td>42.77 ± 7.44</td>
</tr>
<tr>
<td>BMM, kg</td>
<td>3.554 ± 0.403</td>
<td>2.660 ± 0.330</td>
<td>3.354 ± 0.538</td>
</tr>
<tr>
<td>Gluteal girth, cm²</td>
<td>101.43 ± 6.89</td>
<td>92.62 ± 4.12</td>
<td>100.44 ± 7.19</td>
</tr>
</tbody>
</table>

Values are means ± SD, with range given in parentheses. DEXA, dual-energy X-ray absorptiometry; %BF, percent body fat; BMI, body mass index; 4C, four-compartment model; FFM, fat-free mass; BMM, bone mineral mass; TBW, total body water. * Determined from 4C model; *n = 89 (79 men and 10 women).

The descriptive statistics for the sample are presented in Table 1. There was no significant difference (P = 0.292) between scale mass (75.10 ± 14.88 kg) and DEXA-determined mass (75.15 ± 14.90 kg) with an interclass correlation coefficient of 0.999 [standard error of estimate (SEE) = 0.56 kg] and individual differences ranging from −1.99 to 1.52 kg (0.05 ± 0.56 kg). The small gender difference for %BF is because 22
of the 34 women were participating in vigorous physical training programs.

An interclass correlation coefficient of 0.952 and SEE of 1.6% BF were obtained for the relationship between the 4C model and DEXA %BF (Fig. 1). A dependent t-test revealed a significant difference \( (P < 0.001) \) between the means of the two body composition methods, with individual variations ranging from \(-2.6 \) to \(7.3\% \) BF \( (1.8 \pm 2.0\% \) BF). A Bland and Altman plot (Fig. 2; Ref. 4) demonstrates a definite trend over the range of %BF measurements \( (r^2 = 0.247) \) with a slope and intercept significantly different (both \( P < 0.001 \)) from zero. This result combined with Fig. 1 indicates that DEXA tends to progressively underestimate the body fat of leaner individuals compared with the 4C body composition model.

Figures 3 and 4 show the relationships of age, FFM hydration and tissue thickness [i.e., body mass, BMI, and gluteal girth \( (n = 89) \)] to the differences between the two body composition methods. All of these variables, except FFM hydration \( (r^2 = 0.004; P = 0.420) \), display significant trends (i.e., slope \( \neq 0 \)). Figure 3 reveals an age-wise increase in the %BF difference between the 4C and DEXA methods, whereas Fig. 4 demonstrates a significant inverse trend \( (P = 0.002) \) for these differences as the three surrogates of tissue thickness increase.

**DISCUSSION**

The accuracy, precision, and bias of %BF estimates derived by DEXA relative to those obtained from the 4C body composition model (fat, TBW, BMM, and residual) were examined. This study is unique because the large healthy sample \( (n = 152) \) spans a wide range
for age (18–59 yr) and %BFs (4.6–37.8% BF by DEXA) with an emphasis on very lean individuals [32 subjects (21%) with <10% BF by DEXA]. Although the correlation between the two measurements of %BF is extremely high ($r^2 = 0.952$), the results indicate that there is a significant difference between the two methods, with considerable intraindividual variations (range 2.6–7.3% BF) and a mean difference of 1.8% BF ($P < 0.001$). Linear regression of %BF by the 4C model on DEXA %BF departed significantly from the line of identity, with both the slope and intercept being statistically different ($P < 0.001$) from 1.0 and 0.0, respectively (Fig. 1). Furthermore, the bias between the two methods decreased with increasing body fat with a SEE at the mean of 1.6% BF. For example, when DEXA body fat was 10, 20, and 30%, predicted 4C body fat was 13.0, 21.6, and 30.3%, respectively. These body fat scores in combination with Fig. 1 demonstrate that DEXA tends to progressively underestimate the body fat of leaner individuals compared with the 4C model. This point is reinforced by the Bland and Altman plot (Fig. 2) for the %BF results that displays a correlation of $-0.498$ ($P < 0.001$) and a negative slope that is significantly different ($P < 0.001$) from zero.

The large-scale study by Gallagher et al. (11) similarly reported that DEXA, regardless of gender, underestimated the %BF of lean individuals compared with a 4C model. They obtained SEEs of 3.1 and 2.8% BF for women ($n = 282$) and men ($n = 475$), respectively. To our knowledge, Gallagher et al. (11) are the only investigators to publish equations for predicting the 4C model %BF from DEXA body fat for a large sample of healthy adults over a wide age range that, unlike the present study, excluded those engaged in vigorous physical training programs. Their multicompartment model was similar to the one employed in this study except that TBW was determined by tritium dilution, and Lunar DPX software version 3.6 was used. Withers et al. (39) and van der Ploeg et al. (35) also observed that DEXA yielded lower %BF values (2.9–4.1% BF)
than the 4C model for lean male and female athletes. These findings have been supported by animal studies (24, 34) that have shown for lean pigs that %BF by direct chemical analyses is greater than the corresponding DEXA score.

Prior et al. (28) also reported the relationship between the two methods (i.e., DEXA and 4C model). But their main focus was to determine whether DEXA accuracy in young adults is affected by gender, race, athletic status, and musculoskeletal development so their sample was restricted to young women (n = 81) and men (n = 91) with mean ages of 20.7 and 21.2 yr, respectively. Contrary to our findings on subjects spanning a much larger age range, they observed no significant (P = 0.10) difference between the means of the two body composition methods despite individual variations from −9.9 to 7.5% BF. They attributed these differences to FFM hydration (r = −0.51) and body

Fig. 4. Relationship of %BF difference to body mass (A), body mass index (BMI; B), and gluteal girth (C).
thickness ($r = -0.34$), with BMI used as an indicator of the latter. However, their FFM hydrations ranged from $\sim 65$ to $80\%$, which is much greater than the $70$–$76\%$ for the present study. In addition, a Hologic DEXA scanner (Waltham, MA, model QDR-1000W, software version 5.71) was utilized and blood samples were analyzed for deuterium concentration by using an infrared spectrophotometer, which is not as precise a technique as using an isotope ratio mass spectrometer (29). This may account for their large range of hydration values.

There was a highly significant correlation ($r^2 = 0.999$) and no significant difference ($P = 0.292$) between scale mass and DEXA mass, with a slope and intercept for the relationship of $0.998$ (slope $= 1; P = 0.639$) and $0.113$ (intercept $= 0; P = 0.630$), respectively. Excellent correlations like these have often been erroneously interpreted as support for the validity of the DEXA technique (12). However, our results indicate that DEXA can correctly reconstruct body mass but is unable to accurately resolve this mass into components for body composition analyses (i.e., fat mass and soft lean tissue), especially for lean individuals.

Comparisons between multicompartiment body composition models and DEXA are not new, but previous studies have concentrated mainly on individuals who depart from a healthy adult body composition. However, despite the use of various 4C models, the majority of studies on healthy adults have demonstrated that DEXA tends to underestimate FFM (0.4–4.2\% BF) compared with the multicompartiment approach (7, 9, 10, 39). Furthermore, Bergsma-Kadijk et al. (3) and Clasey et al. (6) examined the influence of age on the two body composition methods. Both noted that DEXA underestimates BF and that the individual variability associated with this method may limit its use in research settings. In fact, Bergsma-Kadijk et al. showed that DEXA has considerable mean and individual biases compared with a 4C model in both younger (19–27 yr) and older (65–78 yr) women with underestimations of 3.1 and 5.3\% BF, respectively.

The majority of the studies cited in the previous paragraph attributed the observed differences to biological variations in FFM hydration that DEXA was unable to accommodate. However, our results (Figs. 3 and 4) show that the BF difference between the 4C model and DEXA is independent of the FFM hydration but is significantly affected by age, body mass, BMI, and gluteal girth. Furthermore, this is supported by theoretical and practical evidence that indicates that the measurement of body composition by DEXA appears to be relatively unaffected by variations in hydration status (26, 27). In fact, Pietrobelli et al. (27) calculated errors of $<1.0\%$ BF due to hydration changes of 1–5\%. Hence, although pure water is theoretically scanned as 95\% soft lean tissue and 5\% fat (22), deviations in FFM hydration impact little on the resultant BF via DEXA. The influence of age on BF difference may be explained by bias in the sample toward younger individuals that perhaps could have been eliminated if the number of older subjects recruited was increased. The effects of body mass, BMI, and gluteal girth in Fig. 4 may be because leaner people tend to be lighter and may, therefore, be displaying differences in BF rather than body shape. Alternatively, these indicators of body proportions may show that DEXA measurements are sensitive to anterior-posterior tissue thickness resulting in a systematic difference between lean and obese individuals. The magnitude of this error or bias may be attributed to X-ray beam hardening that causes the measured attenuation values to increase with decreasing tissue thickness (13, 20).

Polyenergetic X-rays are more susceptible to beam hardening, i.e., the preferential attenuation of low-energy X-rays by tissue. Goodsit (13) demonstrated that the largest errors occurred at tissue thicknesses of $<5$ cm and $>20$ cm with corresponding under- and overestimations of true fat (attenuation values are higher and lower, respectively). He suggested that the effects of beam hardening could be overcome if the calibration standard and tissue were matched for thickness. However, Lunar DEXA systems estimate total tissue thickness from the detected high-energy X-ray transmission (22) and accordingly apply a correction factor to the soft tissue attenuation values. Nevertheless, if this correction factor is inaccurate and DEXA results are affected by tissue thickness, then this has real implications for longitudinal studies because one would be unsure whether an observed difference is due to beam hardening errors or to the intervention under investigation. But, as indicated by Jebb et al. (18), most subjects measured with the use of DEXA will have a tissue thickness whose depth has relatively little effect on the measured body composition. Furthermore, inaccuracies in body composition results at small tissue thicknesses may have occurred because of deadtime losses derived from pulse pile-up problems (14).

The preceding discussion is based on the assumption that the differences between the two methods of body composition analysis are due to errors in the estimate of BF by DEXA. Although multicompartiment models are theoretically more valid than two-compartment models because they control for the biological variability in some of the FFM components, there is some concern that this extra validity may be offset by the propagation of error associated with the measurement of BD, TBW, and BMM (39). Our propagated error approximates 0.6\% BF, which is considerably less than the error due to biological variability in the FFM density of 3.8\% BF for hydrodensitometry (33), and therefore is not a significant problem. Nevertheless, there is still concern regarding some of the assumptions that are made when body composition is estimated from this model (38). In addition, the 4C model is not completely independent of BF via DEXA because the DEXA-derived BM is used to calculate the former. However, notwithstanding these limitations, the 4C model is presently one of the best available in vivo body composition models.
In conclusion, the present findings on a large heterogeneous sample, with a preponderance of lean subjects, demonstrate that DEXA %BF values display marked individual differences from those obtained by the criterion 4C model. Furthermore, these discrepancies were not associated with FFM hydration but were linked to variations in surrogates of anterior-posterior tissue thickness that resulted in larger DEXA %BF underestimations for leaner individuals. The bias was also different in younger individuals, indicating age-dependent differences in body composition estimations. Hence, although the advent of DEXA has been instrumental in the advancement of research examining bone density, its application as a reference method for body composition assessment may be questionable. The effect of tissue thickness on the resolution of non-bone in both thin and thick tissue regions must be resolved if this technology is to fulfill its potential. However, because these errors are of technical origin, they can presumably be minimized through software revisions and/or hardware modifications (19). Nevertheless, the use of DEXA to evaluate body composition is appealing because it uniquely provides whole body and regional measurements and has advantages in that it is easy to administer, is atraumatic, does not rely on the assumptions of the two-compartment models, and yields data on BM content. In summary, despite the eagerness of some to embrace DEXA as the new gold standard in body composition analysis, some fine tuning is required before it can reach this point.

REFERENCES


