Evaluating body fat in girls and female adolescents: advantages and disadvantages of dual-energy X-ray absorptiometry


ABSTRACT

Background: Within the past 10 y, dual-energy X-ray absorptiometry (DXA) has become one of the most widely used methods of measuring human body composition. However, DXA has not been fully evaluated against an independent criterion method of measuring body fatness in young females.

Objective: Our objective was to determine the bias and agreement between DXA and a 4-compartment model in predicting the percentage of fat mass (%FM) in a multiethnic group of young females.

Design: The %FM values measured by DXA of 73 white, 43 African American, 14 Hispanic, and 11 Asian females with a mean (±SD) age of 13.0 ± 1.9 y were compared with the 4-compartment values, which were based on measurements of body density, body water, and bone mineral content.

Results: The %FM values measured by the 2 methods were correlated at r = 0.90 with an SEE of 3.3%; Bland-Altman analysis indicated an average bias of 3.9%. After nullification of the average bias, an individual estimate of %FM by DXA could be underestimated or overestimated by 6.7% when compared with the 4-compartment value.

Conclusions: DXA is an appropriate method for estimating body composition in a group of young females because its bias and limits of agreement are independent of age, ethnicity, and body fatness. However, the limits of agreement of 6.7% could cause an individual FM value to be underestimated or overestimated by 28% relative to the 4-compartment value. Therefore, DXA may not be the optimal method of measuring the body fatness of young females. Am J Clin Nutr 2002;76:384–9.

KEY WORDS Body fat, percentage of fat mass, fat-free mass, girls, female adolescents, densitometry, isotope dilution, dual-energy X-ray absorptiometry, 4-compartment model

INTRODUCTION

Since the development of dual-energy X-ray absorptiometry (DXA) in the early 1990s (1), DXA has emerged as one of the most widely accepted methods of measuring body composition in human subjects. The popularity of DXA can be attributed partly to its speed, ease of performance, and low radiation exposure (2). However, DXA has not been evaluated in comparison with a criterion method of measuring body fat in young females, despite a pattern of disturbing increases in excessive weight, particularly among African Americans and Hispanics (3).

Recently, on the basis of data collected on children and adolescents in the 1999 National Health and Nutrition Examination Survey (NHANES), officials from the National Center for Health Statistics (NCHS) voiced concern that the increase in obesity among America’s youth that had begun in the 1980s appears to be continuing (4). Because body weight in adulthood is strongly associated with body weight in adolescence (5), the NCHS interpreted the most recent NHANES data as suggesting the likelihood of another generation of adults who may be at risk for overweight and obesity-related health conditions. Because body weight and body mass index are poor indicators for assessing the true degree of adiposity (6), it is essential to determine the percentage of body fat with the most accurate methodology.

The aim of this study was to evaluate the level of agreement between DXA and the 4-compartment criterion model in estimating body fat in a multiethnic group of young females.

SUBJECTS AND METHODS

Human subjects

A total of 141 young females (73 whites, 43 African Americans, 14 Hispanics, and 11 Asians) participated in the study. The young females, who were between 9 and 17 y of age, were recruited from schools in the greater Houston metropolitan area. All subjects were healthy and nondiabetic at the time of the study. The Institutional Review Board for Human Research at Baylor College of Medicine approved the protocol. All subjects and their parents gave written, informed consent. The study was part of a larger study to define the body composition and energy metabolism of pregnant teenagers. None of the young females in the present study were pregnant.


by guest on 30 July 2018

Downloaded from https://academic.oup.com/ajcn/article-abstract/76/2/384/4689517
Sexual maturity evaluation

Sexual maturity was determined from a physical examination by a physician according to the Tanner stages of classification (7).

Anthropometric measurements

On admission to the Children’s Nutrition Research Center, each subject’s body weight was measured to the nearest 0.1 kg with an electronic scale (Scale-Tronix, Wheaton, IL), and height was measured to the nearest 0.1 cm with a stadiometer (Holtain Ltd, Crosswell, Crymych, United Kingdom). One investigator (JES) made all the anthropometric measurements.

Dual-energy X-ray absorptiometry

Fat mass (FM), fat-free mass (FFM), and bone mineral content (BMC) were measured by DXA (Hologic QDR-2000W, software version 5.56; Hologic, Inc, Waltham, MA) with the pencil beam. The Hologic instrument uses a single scan mode for all subjects. DXA scans were made all the anthropometric measurements. Sexes were prepared for oxygen isotope ratio measurements by collecting samples for all adults before they drank 1.25 g 18O (Isotec Inc, Miamisburg, OH)/kg body wt. A second plasma sample was collected 3 h later. The plasma samples were prepared for oxygen isotope ratio measurements by gas-isotope-ratio mass spectrometry (18). TBW was calculated as follows

\[ \text{TBW (kg)} = (d \times A \times E_0)(a \times E_d \times 10^3) \]

where \(d\) is the dose of H218O in grams, \(A\) is the amount of laboratory water in grams used in the dose dilution, \(a\) is the amount of H218O in grams added to the laboratory water in the dose dilution, \(E_0\) is the increase in 18O abundance in parts per thousand in the laboratory water after the addition of the isotopic water, and \(E_d\) is the increase in 18O abundance in parts per thousand in the 3- h postdose plasma sample. H218O was used instead of H216O to determine TBW because 1) the 18O dilution method yields a more accurate estimate of TBW than does H216O (19), 2) the analytic accuracy and precision of stable oxygen isotope ratio measurements are much higher than those of stable hydrogen isotope ratio measurements (18), and 3) the analytic procedure for the preparation of the plasma sample for stable oxygen isotope ratio measurements is much simpler than that for stable hydrogen isotope ratio measurements (20).

Four-compartment criterion method

The reference values for percentage of fat mass (%FM) were obtained by using a 4-compartment model as follows (14)

\[ \%\text{FM} = \left( \frac{2.747}{D} - 0.727 \times \left( \frac{\text{TBW}}{W} \right) + 1.146 \times \left( \frac{\text{BMC}}{W} \right) - 2.0503 \right) \times 100 \]

where \(D\) is body density in grams per milliliter measured by underwater weighing (15) with the use of the “force cube” transducer method (16) and correction for residual lung volume by underwater weighing (15) with the use of the “force cube” transducer method (16) and correction for residual lung volume by underwater weighing (15) with the use of the “force cube” transducer method (16) and correction for residual lung volume.

TBW is total body water in kilograms and will not be presented here.

FM and BMC have been extensively described in the literature (1, 8–13) and will not be presented here.

Statistical analysis

Linear regression analysis and the Bland-Altman procedure (21) were used to compare %FM values obtained by DXA (%FM\text{DXA}) with those obtained by the 4-compartment method (%FM\text{4C}). In the regression analysis, %FM\text{DXA} was plotted against %FM\text{4C}. The deviation of the slope from 1.0 and the deviation of the intercept from zero of the regression line were evaluated by using the critical \(r\) values with the corresponding degrees of freedom. The error of the DXA method was defined by the SEE of the regression line. With the Bland-Altman procedure, differences between the 2 methods were plotted against the averages of the 2 methods. Regression analysis was used to examine the relation between the differences and averages. If the slope was not significant (\(P > 0.05\)), the bias (mean difference between methods) and the 95% limits of agreement [bias \(\pm (2 \times \text{SD of the differences})\)] were computed. All statistical analyses were performed with SPSS for WINDOWS (version 8; SPSS Inc, Chicago).

RESULTS

Mean values for age, sexual maturation, physical characteristics, and body composition of the 141 young females are shown in Table 1. Because of the small number of young Hispanic and Asian females who participated in the study, Tanner-stage data are presented only for those with breast and pubic hair development at Tanner stage ≥3. Most of the young females were at Tanner stage ≥3. One-way analysis of variance and Tukey’s test showed that the young African American females were significantly heavier and taller than the young white females (\(P < 0.05\)). The young African American females also had significantly higher body mass indexes, BMC, TBW, FFM, and BMC/FFM than did the young white females (\(P < 0.05\)). Within each ethnic group, the FFM, BMC/FFM, TBW/FFM, and FM estimated by DXA were significantly different from those estimated by the 4-compartment model (\(P < 0.05\)). On average, DXA overestimated %FM by 3.9 ± 3.4% (\(P < 0.01\)) relative to the 4-compartment model.

The results of the linear regression analysis comparing %FM\text{DXA} and %FM\text{4C} are shown in Figure 1. %FM\text{DXA} was significantly correlated (\(r = 0.90, P < 0.01\)) with %FM\text{4C}. The slope (0.95) of the regression line was not significantly different from unity (\(P > 0.20\)). The intercept (5.08%), however, was significantly different from zero (\(P < 0.01\)), suggesting a systematic bias between the 2 methods. The regression analysis yielded an SEE of 3.3% for %FM\text{DXA}.

With the use of the Bland-Altman procedure, differences between %FM\text{DXA} values and %FM\text{4C} values were plotted against the average %FM values obtained by the 2 methods (Figure 2). The differences between the 2 methods were not a function of %FM (\(r = 0.13, P = 0.13\)). The results indicated that DXA, on average, overestimated %FM by 3.9%. On an individual basis, DXA could overestimate %FM by 10.6% (upper limit of agreement) or underestimate %FM by 2.8% (lower limit of agreement). Similar results were obtained when the analyses were segregated by race (bias ± SD: whites, 4.2 ± 3.7%; African Americans, 3.7 ± 2.8%; Hispanics, 3.0 ± 3.8%; Asians, 3.9 ± 2.8%). These biases were not significantly different (\(P > 0.28\)). Univariate analysis of variance indicated that the differences between the 2 methods were not affected by race (\(P = 0.70\), age (\(P = 0.71\)), or sexual maturation (\(P > 0.13\)).
A previous study showed that DXA is a precise method for body-composition measurement (25), and in fact, several authors have considered DXA to be a reference method for this purpose (26–28). These enthusiastic reports, however, have been countered by other studies showing that body-composition measurements made by DXA, particularly those of body fat, can be significantly affected by bone maturation (29, 30), age (8), sex (9), tissue thickness (31, 32), skeletal content of FFM (10), choice of instrumentation (11, 33–35), and choice of software (34, 35). On the basis of the data reported in the latter studies, several authors have disagreed with different superscript letters are significantly different, P < 0.05 (one-way ANOVA and Tukey’s test).

\[ \%FM_{\text{DXA}} = 0.95 \times \%FM_{\text{MC}} + 5.08 \]

\[ r = 0.90, P < 0.01, \text{SEE} = 3.3\% \]

FIGURE 1. Results of linear regression analysis comparing the percentage of fat mass values obtained by dual-energy X-ray absorptiometry (\%FM_{\text{DXA}}) of 141 young females with those obtained by the 4-compartment method (\%FM_{\text{MC}}). The symbols represent the \%FM values of each study subject (○, whites; ●, African Americans; △, Hispanics; +, Asians). The solid line represents the line of identity (slope = 1 and intercept = 0), and the dotted line represents the regression line given by the equation. The asterisk indicates that the intercept is significantly different from zero (P < 0.01).

### TABLE 1

<table>
<thead>
<tr>
<th>Age (y)</th>
<th>Whites (n = 73)</th>
<th>African Americans (n = 43)</th>
<th>Hispanics (n = 14)</th>
<th>Asians (n = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanner stage ≥3 (%)</td>
<td>12.7 ± 1.9a</td>
<td>13.5 ± 1.7</td>
<td>12.8 ± 2.0</td>
<td>14.0 ± 2.3</td>
</tr>
<tr>
<td>Breast</td>
<td>67.1</td>
<td>100</td>
<td>78.6</td>
<td>90.9</td>
</tr>
<tr>
<td>Pubic hair</td>
<td>64.8</td>
<td>97.7</td>
<td>78.6</td>
<td>90.9</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>48.0 ± 13.0b</td>
<td>57.2 ± 13.9</td>
<td>51.1 ± 16.3b</td>
<td>44.9 ± 9.2b</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.54 ± 0.11b</td>
<td>1.60 ± 0.07a</td>
<td>1.56 ± 0.12b</td>
<td>1.53 ± 0.09b</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>20.1 ± 4.0b</td>
<td>22.4 ± 4.9</td>
<td>20.6 ± 3.4b</td>
<td>19.0 ± 2.6b</td>
</tr>
<tr>
<td>Body density (g/mL)</td>
<td>1.0375 ± 0.0162</td>
<td>1.0362 ± 0.0158</td>
<td>1.0340 ± 0.0143</td>
<td>1.0437 ± 0.011I</td>
</tr>
<tr>
<td>BMC (kg)</td>
<td>1.53 ± 0.44b</td>
<td>1.95 ± 0.43a</td>
<td>1.65 ± 0.54ab</td>
<td>1.58 ± 0.35ab</td>
</tr>
<tr>
<td>TBW (kg)</td>
<td>26.0 ± 5.8b</td>
<td>30.6 ± 5.2</td>
<td>26.4 ± 7.9b</td>
<td>24.8 ± 4.8b</td>
</tr>
<tr>
<td>FFM_{\text{DXA}} (kg)</td>
<td>34.1 ± 7.2b</td>
<td>40.2 ± 6.2</td>
<td>35.3 ± 8.6b</td>
<td>33.0 ± 5.4b</td>
</tr>
<tr>
<td>BMC/FFM_{\text{DXA}} (%)</td>
<td>4.4 ± 0.5b</td>
<td>4.8 ± 0.6</td>
<td>4.6 ± 0.6b</td>
<td>4.8 ± 0.5b</td>
</tr>
<tr>
<td>TBW/FFM_{\text{DXA}} (%)</td>
<td>76.5 ± 5.9</td>
<td>75.9 ± 3.5</td>
<td>74.4 ± 7.6</td>
<td>75.0 ± 4.3</td>
</tr>
<tr>
<td>FFM_{\text{MC}} (kg)</td>
<td>36.1 ± 7.8b</td>
<td>42.4 ± 7.0</td>
<td>37.0 ± 10.0b</td>
<td>34.9 ± 6.2b</td>
</tr>
<tr>
<td>BMC/FFM_{\text{MC}} (%)</td>
<td>4.2 ± 0.5b</td>
<td>4.6 ± 0.5</td>
<td>4.4 ± 0.5b</td>
<td>4.5 ± 0.5b</td>
</tr>
<tr>
<td>TBW/FFM_{\text{MC}} (%)</td>
<td>72.0 ± 2.7</td>
<td>72.1 ± 2.1</td>
<td>71.0 ± 3.9</td>
<td>71.1 ± 1.6</td>
</tr>
<tr>
<td>FM_{\text{DXA}} (%)</td>
<td>27.6 ± 7.7</td>
<td>28.2 ± 8.0</td>
<td>29.3 ± 7.6</td>
<td>25.8 ± 5.8</td>
</tr>
<tr>
<td>FM_{\text{MC}} (%)</td>
<td>23.4 ± 7.2</td>
<td>24.5 ± 7.6</td>
<td>26.3 ± 7.1</td>
<td>21.9 ± 4.9</td>
</tr>
</tbody>
</table>

1. BMC, total-body bone mineral content; TBW, total body water; FFM_{\text{DXA}} and FFM_{\text{MC}}, fat-free mass by dual-energy X-ray absorptiometry (DXA) and by the 4-compartment model, respectively; and FM_{\text{DXA}} and FM_{\text{MC}}, fat mass by DXA and by the 4-compartment model, respectively. Values in the same row with different superscript letters are significantly different, P < 0.05 (one-way ANOVA and Tukey’s test).

2. ± SD.

3. The percentage of girls with breast and pubic hair development at Tanner stage ≥3.

**DISCUSSION**

Using the criteria of Lohman (22), estimates of %FM of the 141 young females had an SEE of 3.3%. However, a higher SEE is anticipated when greater variability is observed in the criterion measurement because the criteria of Lohman were based on a 76.5-kg man and a 60.0-kg woman with fat contents of 15% and 25%, respectively (22). Because the %FM_{\text{MC}} values of our 141 female subjects ranged between 8% and 42%, it would be reasonable to expect the SEE to be much higher in our group. Therefore, on the basis of the criteria proposed by Lohman (22), our results (Figure 1) indicate that DXA may be a very good method for estimating body fatness of a young, multiethnic female population.

As shown in Figure 2, on average, DXA tended to overestimate %FM by 3.9%. Because the bias remained unchanged across the entire range of %FM measures, the overestimation could be nullified by subtracting 3.9% from all DXA estimates of %FM if agreement with the 4-compartment model is the desired outcome. With this adjustment, the 95% limits of agreement became 6.7%, which means that on an individual basis, %FM could be underestimated or overestimated by 6.7% when DXA is used to estimate FM in a young, multiethnic female population. The 95% limits of agreement of DXA are an improvement over those of the skinfold-thickness equation (10%; 23) or total-body electrical conductivity method (8%; 24). More importantly, the bias and limits of agreement of DXA are not affected by body fatness or age, unlike the total-body electrical conductivity methodology (24).

A previous study showed that DXA is a precise method for body-composition measurement (25), and in fact, several authors have considered DXA to be a reference method for this purpose (26–28). These enthusiastic reports, however, have been countered by other studies showing that body-composition measurements...
already cautioned against the use of DXA as a “gold standard” for body-composition measurements (12, 36, 37).

The inaccuracy of the body-composition measurements made by DXA could be attributed to the assumption of a hydration constant of 73% for FFM (12). Although the analysis algorithm used by Hologic does not include a fixed hydration constant per se, it is indirectly assumed when the mass attenuation coefficient for the lean mass is calculated. The hydration constant of 73% for FFM might be appropriate for healthy adults but would not be appropriate for children or adolescents. Hydration of FFM has been shown to be substantially higher among children and to change with maturation (14, 38). However, subsequent reports (1, 39) have shown that under normal and even most clinical conditions, errors associated with fat estimation by DXA due to variation in soft tissue hydration are small and should not affect its accuracy.

It is possible that the propagation of the individual measurement error and technical error from each of the body density, TBW, and BMC measurements would affect our conclusions. However, in a chapter elegantly written by Siri (15), %FM could be estimated within ±1.5% by using a 3-compartment model if TBW could be measured with an error of ±1%. Because BMC accounts for a small fraction of FFM, the addition of BMC in the 4-compartment model would not alter the accuracy of the %FM estimation outlined by Siri. This conclusion is confirmed by a recent study (40) indicating that the additive errors in the multicompartiment model did not offset the improved accuracy of fat estimations over that obtained by using densitometry or body water measurement alone. Because the measurement and technical errors associated with our TBW measurements based on H218O dilution are better than ±1%, the propagated error of the %FM estimations of our subjects would be well within ±1.5%. Therefore, the measurement and technical errors associated with the body density, TBW, and BMC measurements in our study would not alter our conclusions.

It is well known that the body composition of growing children is different from that of mature adults (41). The algorithms used by the Hologic DXA instrument are based on adult proportions. Thus, for the younger or smaller children and adolescents in this study, these algorithms may be less accurate. Furthermore, the calculation of %FM used by the Hologic DXA instrument is relative to the attenuation ratios or R ratios used to define the 100%FM and 0%FM values. These choices may be less accurate when applied to children and adolescents. Another possible source of error is that the bone detection algorithms used for adults are less accurate for children, thus resulting in the misclassification of bone into the soft tissue compartment, which then alters the true lean-to-fat ratio.

It has been documented that the density of FFM (23), the skeletal content of FFM (13, 42), and body fat distribution (43–45), all of which affect the accuracy of body-composition measurements by DXA, are different between young African American and white females. Young African American females have a higher BMC (13) but lower amounts of total, visceral, and subcutaneous adipose tissue than do young white females (43–45). Asian women also have more subcutaneous adipose tissue and a different fat distribution than do white women (46). These racial differences in body composition may further complicate the accuracy of the body-composition measurements made by DXA.

Our results show that the bias and limits of agreement of DXA are independent of age, ethnicity, or %FM. The limits of agreement of 6.7% for an individual young female with an average %FFM of 76.1% (the average %FFM of all young females in Table 1) will translate into an overestimation or underestimation of FFM of 8%. However, the same limits of agreement for an individual young female with an average %FM of 23.9% (the average %FM of all young females in Table 1) will be exaggerated into an overestimation or underestimation of body fatness of 28%. Thus, although DXA, with proper calibration, is an appropriate method for estimating body fatness in a group of young females, its accuracy for an individual in this group is suboptimal. The applicability of our conclusion to DXA instruments made by other manufacturers or to DXA instruments that use different scan modes and software is not known. For example, the Lunar DXA instrument has pediatric software and uses different scan modes in different weight brackets. However, the accuracy and precision of the Lunar DXA instrument for the estimation of %FM in a multiethnic group of young females are not known and warrant further investigation.

We are indebted to the volunteers; to the staff of the Metabolic Research Unit of the Children’s Nutrition Research Center for meeting the needs of the subjects during the study; to J Hoyle in the Pediatric Department of Kelsey-Seybold West Clinic; to M desVignes-Kendrick, Director of the City of Houston Health and Human Services Department; to X Earlie, Director of Sciences of the Aldine Independent School District; to S Wooten, Principal of the Teague Middle School; to B Shargey, Dean of Instruction, and CC Collins, Principal, of the High School for Health Professions; to K Wallace for subject recruitment; to M Payau and FA Vohra for the underwater weighing measurements; to R Shyapalo and J Pratt for the dual-energy X-ray absorptiometry measurements; to L Clarke, S Zhang, and K Usuki for the isotope ratio mass spectrometric measurements; and to Children’s Nutrition Research Center Editor L Loddeke.
REFERENCES


33. Lohman TG, Boileau RA, Slaughter MH. Body composition in chil-
44. Morrison JA, Barton BA, Obarzanek E, Crawford PB, Guo SS, Schreiber GB. Racial differences in the sums of skinfolds and percentage of body fat estimated from impedance in black and white girls, 9 to 19 years of age: The National Heart, Lung, and Blood Institute Growth and Health Study. Obes Res 2001;9:297–305.