The Reliability and Convergent Validity of Field Tests of Body Composition in Young Adolescents

Michelle Ihmels, Gregory J. Welk, James J. McClain, and Jodee Schaben

Background: Advances in BIA offer practical alternative approaches to assessing body composition in young adolescents and have not been studied for comparability. Methods: This study compared reliability and convergent validity of three field tests (2-site skinfold, Omron and Tanita BIA devices) on young adolescents. Reliability was determined using intraclass correlation coefficients, convergent validity was examined by computing correlations among the three estimates, differences in estimated body fat values were evaluated using repeated-measures ANOVA, and classification agreement was computed for achieving FITNESSGRAM® Healthy Fitness Zone. Results: ICC values of all three measures exceeded .97. Correlations ranged from .74 to .81 for males and .79 to .91 for females. Classification agreement values ranged from 82.8% to 92.6%. Conclusions: Results suggest general agreement among the selected methods of body composition assessments in both boys and girls with the exception that percent body fat in boys by Tanita BIA is significantly lower than skinfold estimation.

Key Words: body composition, FITNESSGRAM, bioelectrical impedance, skinfolds

Obesity in young adolescents has become one of the top public health concerns in this country. Since the early 1970s, the prevalence of children and adolescents classified as overweight has more than doubled. It has been estimated that 16% of 6- to 19-year-olds in the United States are currently overweight according to results from the National Health and Nutrition Examination Survey.

Excessive body fat in young adolescents has been shown to be associated with higher cardiovascular risk factors including hypertension and hyperlipidemia as well as Type II diabetes. Even in the absence of these chronic diseases, obese adolescents have been shown to be at an increased risk for developing health problems as adults. A number of studies have shown that obesity patterns tend to track...
through the lifespan. Tracking analyses on the data from the Amsterdam Growth and Health Study suggested that high body fat percentages and cholesterol ratios yielded high odds ratios for risk of disease development in later life (17.7 and 22.9 for high body fat and cholesterol, respectively). Similar results were found for body fatness and cholesterol in the Bogalusa Heart Study and the Muscatine Study.

Because of its importance for the present and future health of young adolescents, body composition is included in health related fitness tests such as *FITNESSGRAM*. The recommended assessment for body composition in this battery is a two-site skinfold test, but many teachers avoid this assessment and instead use body mass index (or nothing at all). While BMI is a widely used clinical measure, estimates of body fatness would provide more meaningful information about a young adolescent’s body composition.

There are several factors that may limit the use of skinfold testing in schools. Many physical education teachers may lack the training and experience needed to accurately conduct skinfold testing. Studies have shown that a tester’s accuracy depends on proper training and that extensive practice is needed to achieve valid and reliable results. Teachers may also be reluctant to utilize skinfold testing because of its more invasive nature. Skinfold testing requires the teacher to touch the student at several sites on the body where the measurement will be taken. Some districts prevent testing of body composition because they are concerned that an emphasis on body fat testing could have negative consequences and contribute to eating disorders. Also, the additional time involved in assessing body composition is another major barrier to its use in physical education.

Bioelectric impedance devices may offer a promising alternative to skinfold testing in schools. The basic principle of bioelectric impedance analysis (BIA) is that fat and muscle differ in their ability to conduct electricity. In BIA, a low-level electrical current is passed through the participant’s body and the resistance (or impedance) is measured. This impedance value is then converted into a body fat percentage through the use of an equipment-specific (usually proprietary and unpublished) prediction equation. Several studies have investigated the overall reliability and validity of BIA in comparison to other accepted body composition assessment techniques and found BIA to be an acceptable tool. The recent development of portable “foot to foot” and “hand to hand” BIA devices (costing as little as $100) may provide an effective alternative to the use of skinfold calipers for large-scale testing. These devices only require a participant to stand on a platform while barefoot, or to hold the device in both hands, and are less invasive than skinfold testing. The assessment also requires minimal training for the tester. Because of these advantages, portable BIA devices may be especially useful in school physical education programs.

The purpose of this study was to examine the agreement between two different measures of body fatness in young adolescents. Specifically, this study was designed to assess the utility of selected BIA devices as an alternative to the more accepted skinfold testing procedures recommended in the *FITNESSGRAM* battery. Participants were measured with two portable BIA analyzers and skinfolds during two separate visits. It was hypothesized that BIA would have better reliability than skinfold measures, which depend more on the skill of the tester. It was also hypothesized that there would be no significant differences between body fat estimates from BIA and skinfold.
Methods

Participants

The sample included 132 eighth grade students (64 boys and 68 girls) from the only middle school in the local school district. Parents signed an informed consent form and the students completed an assent form prior to the start of the study. The university’s institutional review board and the local school district approved the protocols used in the study.

Procedure

The data were collected during the student’s physical education class. Four testing stations were used to collect data, with one researcher positioned at each station. The adolescents proceeded through each station in the following order: (1) height, (2) Tanita BIA, (3) Omron BIA, and (4) skinfolds. All researchers were trained experts in their anthropometry station, with skinfolds being measured by one of the highly experienced authors (GW).

Initially, height was measured using a stadiometer with the participant standing erect, without shoes, and with weight distributed evenly between both feet, heels together, arms relaxed at the sides, and the head in the Frankfort horizontal plane. For BIA assessment the participant’s height, sex, and age were entered into a Tanita body fat analyzer, Model BF-626 (Tanita Corp. of America, Inc., Arlington Heights, IL) prior to testing. Each participant stepped on the scale barefoot wearing a school-issued physical education uniform, and body fat percentage estimate was calculated from the foot-to-foot impedance. Both the weight (in kilograms to nearest 0.1) and body fat estimate (in percent to nearest 0.1) were recorded from the Tanita scale. Previous research in our lab has shown no difference in values with repeated measures, therefore only a single measure was taken.

The second BIA assessment was made using the Omron body fat analyzer, Model HBF-306 (Omron BF306, Omron Healthcare, Inc., Vernon Hills, IL). Height, weight, sex, and age were entered into the unit and participants were instructed to hold the analyzer according to the manufacture’s directions (arms extended from body and perpendicular to the floor, thumbs on top). The body fat percentage estimate from the Omron device was recorded along with the computed BMI value.

The two-site skinfold protocol from the FITNESSGRAM test battery (The Cooper Institute, Dallas, TX) was used to estimate body fatness. Lange calipers (Beta Technology, Inc., Cambridge, MD), calibrated prior to use with a standard measurement block, were used for measurements on both the right triceps and calf. The triceps measurement was made on the back of the arm midway between the elbow and acromion process of the scapula with a vertical measurement. The calf measurement was made with the right foot placed on a chair that flexed the knee at a 90-deg angle. A vertical skinfold measurement was then made just above the point of maximum calf girth on the medial side of the leg. Each measurement was taken three times and the middle value was recorded. The assessment protocol was repeated 4 days later (at the same time of day) with a subsample of 37 boys and 34 girls to examine the reliability of the measurements.
Analysis

Statistical analyses were performed using SPSS (Statistical Package for the Social Sciences for WINDOWS, 10.0; SPSS Inc., Chicago). Means and standard deviations from each assessment were reported separately for each gender group. Reliability was determined using intraclass correlation coefficients computed between the matched data from the two visits. Correlations were computed between the body composition assessments from skinfolds and BIA to provide an indication of overall agreement between the methods. Repeated-measures ANOVA was used to evaluate differences in estimated body fat (separate analyses were performed for boys and girls). The alpha level was set at .01. Bland-Altman plots were used to examine agreement between assessments across the full range of body fat estimates. Classification agreement for achieving the “Healthy Fitness Zone” on the FITNESSGRAM report (a criterion standard) was evaluated using chi-square analysis for each pairwise comparison. Based on generally accepted criteria, correlation coefficients above .80 were used to designate acceptable reliability and validity; however, the other analyses were also considered to determine the overall measurement agreement between skinfolds and BIA values.

Results

Descriptive characteristics of the participants are shown in Table 1. One boy did not complete all of the testing, so his results were excluded from the analyses. Analyses of these data indicated that boys were significantly taller and had lower percent body fat than girls ($p < .05$). There were no significant differences in weight or BMI between the sexes.

A main purpose of the study was to investigate the reliability of the BIA devices and skinfold. It was hypothesized that BIA would have better reliability. The reliability of all three measures was high, with ICC values exceeding .97. The ICC values for boys was $R = .99$ for skinfolds, $R = .98$ for Tanita, and $R = .98$ for Omron. The ICC values for girls was $R = .97$ for skinfolds, $R = .99$ for Tanita, and $R = .99$ for Omron.

Table 1  Descriptive Characteristics, Mean ($\pm SD$)

<table>
<thead>
<tr>
<th></th>
<th>Boys ($n = 63$)</th>
<th>Girls ($n = 68$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>13.4 (0.5)</td>
<td>13.4 (0.5)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>164.5 (8.8)</td>
<td>160.8 (6.9) *</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>55.8 (11.5)</td>
<td>53.7 (9.8)</td>
</tr>
<tr>
<td>BMI (kg/m$^2$)</td>
<td>20.5 (2.8)</td>
<td>20.9 (3.1)</td>
</tr>
<tr>
<td>Skinfold (% fat)</td>
<td>18.5 (7.3)</td>
<td>25.9 (6.8) *</td>
</tr>
<tr>
<td>Tanita (% fat)</td>
<td>14.4 (5.5)</td>
<td>25.0 (6.1) *</td>
</tr>
<tr>
<td>Omron (% fat)</td>
<td>17.8 (6.1)</td>
<td>25.1 (5.3) *</td>
</tr>
</tbody>
</table>

* $p < .05$
Convergent validity was evaluated using correlations, ANOVAs, Bland-Altman plots, and classification agreement. Correlations between the body composition estimates ranged from $r = .74$ to $r = .81$ for boys, and $r = .79$ to $r = .91$ for girls (see Table 2). There was reasonable predictor error for both BIA techniques using skinfolds as the reference measure as calculated with the standard error of measurement (boys: 4.3% for Tanita and 4.9% for Omron; girls: 3.6% for Tanita and 4.1% for Omron). For girls, there were no significant differences in body fat estimates by method, $F(2, 66) = 1.96, p = .148$. The three measures were within 0.9% body fat and the effect sizes (ES) were small (Skinfold vs. Tanita: ES = .06; Skinfold vs. Omron: ES = .05; Tanita vs. Omron: ES = .004).

For boys, the ANOVA analyses revealed a significant gender × method interaction, $F(2, 61) = 36.05, p < .001$. Estimates with the Tanita monitor were considerably lower (~4-5% body fat) than estimates from the skinfold and Omron procedures (Skinfold vs. Tanita: ES = .62; Omron vs. Tanita: ES = .57). The Omron was not significantly different from skinfold (Skinfold vs. Omron: ES = .11).

Bland-Altman plots were used to examine agreement between assessments across body fat estimate ranges and these are shown in Figure 1. Overall, the distribution of the points shows that there was good agreement between measures across a full range of body composition values. A test of the slope indicated significant correlations for the Tanita in boys ($r = .44, p < .0001$) and the Omron in girls ($r = .38, p < .0001$), with nonsignificance for the Omron in boys ($r = .18$) and the Tanita in girls ($r = .22$).

Classification agreement values between skinfold and Tanita assessments were 82.8% for boys and 88.2% for girls. Classification agreement between Skinfold and Omron assessments were 92.6% for boys and 87.2% for girls.

**Table 2 Correlations Among Body Composition Measures**

<table>
<thead>
<tr>
<th></th>
<th>Tanita</th>
<th>Omron</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skinfold</td>
<td>0.81</td>
<td>0.74</td>
</tr>
<tr>
<td>Tanita</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skinfold</td>
<td>0.85</td>
<td>0.79</td>
</tr>
<tr>
<td>Tanita</td>
<td>0.91</td>
<td></td>
</tr>
</tbody>
</table>

**Discussion**

This study examined the consistency and the agreement between two different measures of body fatness in young adolescents. Skinfold testing is the recommended approach for body composition assessment in the *FITNESSGRAM* battery, but the development of portable BIA devices offers a viable alternative for assessing body composition in physical education settings.

Measurement consistency is a prerequisite for any accuracy assessment. We found that all three assessment methods had good reliability, with all test-retest ICC
Figure 1A — Bland-Altman plots comparing body fat estimates from skinfolds and BIA for boys.
Figure 1B — Bland-Altman plots comparing body fat estimates from skinfolds and BIA for girls.
values exceeding .97. As a result our original hypothesis was not upheld. The high values for the BIA were somewhat expected but the values for the skinfold estimations were higher than expected. The good reliability of this measure is likely due to the fact that the measures were obtained by a trained expert in skinfold testing. It is likely that the reliability of skinfold testing would be less than the reliability of BIA if we had utilized novice testers, but no data are available in the current study to support this hypothesis.

The results of the present study show good overall agreement between estimates from BIA and skinfold; however, some differences were noted between the different BIA devices. For girls, there were no significant differences between the estimates from the two BIA devices and the skinfold body fat estimate. The pairwise correlations among the measures were high and there was general agreement on the Bland-Altman plots across the full range of body composition values. A test of the slope indicated a significant correlation for the Omron in girls, indicating the overestimation of low body fat percentages and underestimation of high body fat percentages. Collectively, this supports the utility of the BIA devices as an alternative estimate of body fatness in girls.

The results for the boys were less clear. The handheld device (Omron) yielded estimates that were not significantly different from those computed from skinfolds. However, the Tanita foot-to-foot BIA device yielded estimates that were significantly lower than the estimates from the Omron and the skinfold. A test of the slope indicated a significant correlation for the Tanita in boys, indicating the overestimation of low body fat percentages and underestimation of high body fat percentages.

The different results for the two BIA devices were not expected, but other investigators have reported inaccuracies in the Tanita estimates for males. In a study with 10- to 14-year-old boys, Parker et al.\textsuperscript{22} found considerable bias in the Tanita estimates of total fat mass (2.3 kg) and wide limits of agreement (2.3 ± 7.8 kg) in comparison with an established reference method (3 component model). Based on these results, they recommended that the Tanita device be used with caution. Lazzer et al.\textsuperscript{16} also found that the Tanita analyzer, when compared to DEXA, underestimated fat mass and percent fat mass in adolescents with android obesity. Consistent with our findings, this error was greater in males. Less work has been done with the Omron, but one previous study\textsuperscript{17} found good agreement (bias = 0.2 ± 3.0%, $p = 0.76$) between estimates from DEXA and the Omron 306, the same model used in this study.

The accuracy of BIA devices may be influenced by a variety of factors, but it is not clear why the Tanita device yielded the low body fat estimates for the boys. BIA is sensitive to both hydration levels and extremity length,\textsuperscript{14,15} and it is possible that males differ in both of these characteristics from females. Dehydration would tend to cause elevated body fat levels since electrical conduction would be reduced, but it is not clear if females or males are more likely to maintain good hydration status. We made no effort in the present study to control for hydration status, since we were interested in how the devices would perform under real-world conditions. It is possible that results would be less variable if this was controlled, but this still wouldn’t explain the differential results for the Tanita and the Omron devices.

The most likely explanation for the different results for the two BIA devices is that there are differences in the internal prediction algorithms used to estimate
body fatness (because these algorithms are not published, it is not possible to
determine how they may vary). Maturation is associated with dramatic changes in
body composition, and the predictions of body fatness may depend on the reference
population used in developing these algorithms. Older adolescent males would tend
to gain muscle mass and this may contribute to larger total water stores and lower
body fat estimates for a given impedance value. If the equations for the Tanita were
developed on an older adolescent sample (with a higher assumed body water level),
it is possible that the equations could have contributed to a systematic underestima-
tion in our younger middle-school sample (presumably with lower muscle mass
and body water).

A potential limitation of this study was the relatively low amount of ethnic
diversity in our sample. The ethnic diversity of this school district is 19.5% minori-
ties and 80.5% Caucasian. It is not clear what samples were used in the internal
prediction algorithms for the device, but it is possible that different results would
be obtained if our sample were more diverse. A further limitation of the study was
the lack of a true criterion measure such as DEXA. We observed good agreement
between the Omron and skinfolds, but it is not possible from this study to determine
the actual validity of the estimates. As stated in the Results, there was reasonable
predictor error for both BIA techniques using skinfolds as the reference measure
as calculated with the standard error of measurement (boys: 4.3% for Tanita and
4.9% for Omron; girls: 3.6% for Tanita and 4.1% for Omron). Previous work has
supported the validity of the FITNESSGRAM two-site skinfold equation.23 There-
fore, despite the lack of a true criterion, we feel it is appropriate to extrapolate
and conclude that the Omron device may provide a useful alternative to skinfold
testing in schools.

Another limitation of this study is the significant correlations that were found
in the Bland-Altman plots for Tanita in boys and Omron in girls. This indicates
that the differences between skinfold and BIA estimates vary systematically across
the range of body composition levels. Further research utilizing a true criterion
measure and a more diverse population and age range of young adolescents would
help to further evaluate the validity of portable BIA devices, but the current study
provides preliminary support for the use of the Omron.

A key issue with regard to the use of a BIA tool in FITNESSGRAM is with
classification agreement since teachers frequently report the percentage of children
attaining the Healthy Fitness Zone. Our results indicate that youth would be classi-
ified similarly with skinfolds or the Omron device 92.6% of the time in males and
87.2% of the time in females. This would appear to be reasonable agreement for use
in school based settings, considering that no two measures will agree perfectly.

A potential advantage of BIA devices in a school based setting is that they pro-
vide a more objective assessment that does not require skill for testing. When teach-
ers are not properly trained or skilled in skinfold assessments, their measurements
can be inaccurate.9–11 Also, teachers and school districts may not be comfortable
with the invasive nature of skinfold assessments. The easy assessments provided by
the BIA devices circumvent this issue entirely. Collectively, these advantages may
make BIA devices such as the Omron a better choice for field based testing.

The increased attention on problems of overweight conditions in youth1,24,25 has
placed a greater importance on accurate determinations of body fatness in youth.
School physical education programs are ideally positioned to assist with tracking
and documentation of body fat. The newest version of the FITNESSGRAM software (8.0) allows for district-wide tracking of fitness tests including body composition testing. The software can be a valuable tool for tracking young adolescents as they progress through a school system, and can contribute to increasing awareness about healthy levels of body composition in youth. Dissemination of this information through the personalized reports produced from the FITNESSGRAM software provides a way for teachers to effectively communicate with parents about the child’s current body composition status.

References


