Reliability and validity of body composition measures in female athletes

WILLA C. FORNETTI,1 JAMES M. PIVARNIK,2 JEANNE M. FOLEY, AND JUSTUS J. FIECHTNER
Departments of 1Kinesiology and 2Osteopathic Surgical Specialties, Michigan State University, East Lansing, Michigan 48824

Fornetti, Willa C., James M. Pivarnik, Jeanne M. Foley, and Justus J. Fiechtner. Reliability and validity of body composition measures in female athletes. J. Appl. Physiol. 87(3): 1114–1122, 1999.—The purpose of this investigation was to determine the reliability and validity of bioelectrical impedance (BIA) and near-infrared interactance (NIR) for estimating body composition in female athletes. Dual-energy X-ray absorptiometry was used as the criterion measure for fat-free mass (FFM). Studies were performed in 132 athletes [age = 20.4 ± 1.5 (SD) yr]. Intrasclass reliabilities (repeat and single trial) were 0.987–0.997 for BIA (resistance and reactance) and 0.957–0.980 for NIR (optical densities). Validity of BIA and NIR was assessed by double cross-validation. Because correlations were high (r = 0.969–0.983) and prediction errors low, a single equation was developed by using all 132 subjects for both BIA and NIR. Also, an equation was developed for all subjects by using height and weight only. Results from dual-energy X-ray absorptiometry analysis showed FFM = 49.5 ± 6.0 kg, which corresponded to %body fat (%BF) of 20.4 ± 3.1%. BIA predicted FFM at 49.4 ± 5.9 kg (r = 0.981, SEE = 1.1), and NIR prediction was 49.5 ± 5.8 kg (r = 0.975, SEE = 1.2). Height and weight alone predicted FFM at 49.4 ± 5.7 kg (r = 0.961, SEE = 1.6). When converted to %BF, prediction errors were ~1.8% for BIA and NIR and 2.9% for height and weight. Results showed BIA and NIR to be extremely reliable and valid techniques for estimating body composition in college-age female athletes.

Generally, high ratios of fat-free mass (FFM) to fat mass (FM) are favorable for an athlete, but too little body fat may result in the deterioration of both health and performance (14). Optimal body composition may vary among individuals in different sports (11). Body composition analysis typically includes an estimate of an individual's %body fat (%BF), which is calculated by dividing FM by total body mass (i.e., weight). Some examples of reported values for college-age women include 10–15% for female runners (35), 15–17% for female gymnasts (18), and 18–30% for female nonathletes (3).

Knowledge of body composition is also important in helping medical personnel in their constant surveillance of the athlete's physical and mental health. Radical changes in body composition can be indicative of serious health concerns. For example, the female athlete triad, consisting of disordered eating, amenorrhea, and osteoporosis, is prevalent in many athletes (28, 33, 43). The triad has serious health implications and may result in bone loss or even death (28, 43). Elite or competitive athletes participating in sports in which low body weights or lean physiques are considered advantageous are at a greater risk for developing the triad (28, 33).

Hydrometrisometry is the most widely used method for estimating %BF (45). This method has been shown to have an error of ~2–3% under the best conditions because of interindividual variability in the chemical composition of FFM (36). In hydrometrisometry, body volume is typically estimated by weighing an individual underwater (UWW). Body density (DB) is then estimated by dividing body weight (out of water) by body volume. Density is subsequently converted into %BF, from which FM and FFM can be calculated.

The original equation from which DB is converted into %BF was based on chemical analysis performed on five Caucasian human cadavers (42). On the basis of cadaver analysis, assumptions for the density of FM were 0.90 g/ml, whereas FFM was 1.10 g/ml. Inherent in these assumptions is that FFM is at a constant hydration level of 73% (29).

As technology has improved, the original assumptions for UWW have been challenged (11, 24). Biological factors such as hydration status, ratio of protein to mineral, and altered bone mineral density may contribute to errors in densitometric determination of body composition in athletes. This variability may be fairly prevalent in female athletes, depending on their training and menstrual histories (40). Therefore, UWW may not assess %BF accurately in female athletes. Furthermore, UWW may not be an appropriate criterion measure when body composition prediction models for female athletes are developed by using various “field” techniques (22). New, multicomponent models are needed to account for interindividual variability in FFM to obtain more accurate estimations of body fatness (11, 22, 46).

A more valid and precise method (compared with UWW) for measuring body composition may be dual-energy X-ray absorptiometry (DEXA). DEXA divides the body into three components: bone, fat-free and bone-free tissue, and fat. Originally, this method was designed to determine bone mineral content (BMC) and bone mineral density. More recent studies have shown that DEXA is highly reproducible for both BMC and body composition (19, 27, 44). Studies have also shown that UWW and DEXA agree well at high, moderate, and low levels of body fat (10).
The fact that DEXA incorporates BMC when %BF is being estimated may be a major advantage compared with UWW. Bone comprises <5% of FFM, but it has the highest density. Therefore, changes in BMC can have a great effect on the average density of FFM (45). In young adult women, variability in BMC has been shown to contribute significantly to variability in D_{\text{avg}} (4).

Although DEXA is extremely valid and precise for %BF measures, this instrument is expensive and is found primarily in clinical settings. Other electronic methods, such as bioelectric impedance analysis (BIA) and near-infrared interactance (NIR), have been developed to be used in the field. Both devices are lightweight, portable, and require little technical training for proper use. These two techniques are less costly than DEXA and may provide reasonably accurate results.

BIA uses an electric current to estimate total body water (TBW) by measuring the resistance and reactance of the tissues. NIR measures the optical density to indirectly assess tissue composition. These measures, along with height, weight, age, and so on, have been used to develop equations to predict FFM and %BF.

Despite the reasonable principles underlying BIA and NIR, neither technique has been thoroughly tested for reliability and validity in female athletes. Two equations for BIA are available for female athletes (15, 25). However, neither equation was based on a homogenous athletic population, and one was published only in abstract form (15). No equations specifically designed for female athletes are available for NIR. Moreover, no studies using female athletes have evaluated the validity of either BIA or NIR by using DEXA as a criterion measure.

The purpose of this investigation was to determine the reliability and validity of BIA and NIR for estimating body composition of female athletes. The DEXA technique was used as the criterion measure. In addition, comparisons were made between these new equations and height and weight alone for estimating body composition.

**METHODS**

**Subjects**

We invited athletes from all Michigan State University (MSU) women’s varsity sports to participate. These sports included basketball, crew, cross-country, field hockey, golf, gymnastics, soccer, softball, swimming and diving, tennis, track and field, and volleyball. Athletes were recruited through the MSU athletic department from each team’s coaches and athletic trainers.

The study was approved by the MSU committee on research involving human subjects, and written informed consent was obtained from each participant. Although the medical staff expressed an interest in knowing some of the study results, care was taken to ensure that the athletes understood that their results would remain confidential unless they informed us otherwise.

**Instrumentation**

**DEXA.** The DEXA instrument used as the criterion measure in the present investigation was a Hologic QDR-1000W (software 6.10). The instrument works on the physics principle that, as X-rays pass through the body, the exiting attenuated signal is exponentially related to the pathlength, tissue density, and energy of the X-ray (31). The X-ray source, mounted beneath the patient, generates a narrow, tightly collimated beam of X-rays that pass through the patient at rapidly switched energies of 70 and 140 kVp. The transmitted intensity at each energy level is measured by a radiation detector mounted on a C frame (movable arm) directly above the X-ray source. During a scan, the C arm oscillates rapidly in the transverse direction while slowly moving longitudinally. Participants lie supine on the scanning bed and are scanned from head to toe in ~20 min.

As the X-ray beam is introduced into the body, the external detector analyzes one small cross-sectional area, or pixel (1 × 1-mm area), at a time. For the analysis, each pixel in the image is determined to be one of two components: bone or soft tissue. Spatial threshold is the outline around bone and determines the mass of the bone and tissue in adjacent areas.

DEXA divides the body into three components: total body bone mineral, or ash; fat-free, mineral-free soft tissue; and fat. This is done by separating pixels into those with soft tissue only, then those with both soft tissue and bone. By using the appropriate beam ratio and/or attenuation values obtained from phantom tissue samples, the FM and FFM can be established from the measured values (31).

**BIA.** BIA (RJL 101A analyzer; Clinton Township, MI) induces a small current through the body, providing resistance and reactance (Xc) values (both measured in \( \Omega \)) from body tissues. This current measures TBW, which is indicative of FFM (12). Individuals of a given height and weight will offer different resistance to the current, on the basis of body geometry and the bioconductor volume of their bodies (26).

Procedures were performed on the basis of methods described by Heyward and Stolarczyk (11) in an ambient temperature of ~22–25°C. For the BIA method, the athlete lay supine on a mattress, with the arms abducted 45° from the legs (which are less abducted than the arms). All measurements were performed on the right side of the body, and skin electrode sites were cleaned thoroughly with an alcohol pad.

The RJL analyzer delivered an alternating current of 800 µA at a fixed 50-kHz frequency. Two electrodes induced the current source. Placements for these current-source electrodes included the dorsal surface of the wrist, so that the upper border of the electrode bisected the head of the ulna (sensor electrode, proximal), and the dorsal surface of the ankle, so that the upper border of the electrode bisected the medial and lateral malleoli. The current was detected by the sensor electrodes, which were located proximal to the source electrodes just described. The placement of the sensor electrodes included the base of the second metacarpal and phalangeal joints of the hand and foot (11). Before testing, the analyzer was calibrated according to the manufacturer’s instructions by testing the actual resistance obtained from the analyzer current being run through a 500-Ω calibration resistor. In all cases, the resistance was within calibration specifications (490–510 Ω).

**NIR.** NIR (Futrex, Gaithersburg, MD) involves a wand, emitting infrared light, that is placed over the participant’s biceps while she is seated. The assumption is that the degree of infrared light absorbed and reflected is related to both the composition of the tissues through which the light is being passed and the specific wavelength being emitted by the light.
The Futrex-5000 emits two wavelengths, 940 and 950 nm, from the wand and subsequently measures the amount of light reflected by the underlying tissues as optical densities 1 (OD1) and 2 (OD2) (6).

Measurements were made in accordance with the manufacturer’s instructions (7). Each subject was seated with her right arm extended (palm up) and resting comfortably on a table. The NIR wand was placed halfway between the antecubital fossa (of the elbow) and the axilla (the armpit). The light-emitting wand was held perpendicular to the measurement site. Caution was used to make sure that the shield flaps were pressed to the skin so that no light would penetrate and affect the optical density measurements. The same amount of pressure was exerted on the wand for each subject, because all measurements were performed by the same investigator.

To obtain the OD1 and OD2 measurements, the “Clear” button was pushed, and the numbers 881 were entered into the NIR instrument (11). Also, before each NIR measurement, the wand was calibrated with a Teflon standard 1 cm thick.

Data Collection Procedures

Testing occurred in the Rheumatology Center, located in the St. Lawrence Health Science Pavilion (East Lansing, MI). On arriving, each woman was measured for body composition by each of the three methods: DEXA, BIA, and NIR. Trained investigators operated each station, and the athletes rotated through the stations. First, each subject had her standing height and weight measured by using a precalibrated beam balance scale and stadiometer. Next, each subject performed BIA and NIR tests. The order of tests was BIA, NIR, repeat BIA, and repeat NIR. Finally, all subjects underwent a single DEXA analysis. The DEXA machine was calibrated daily to a lumbar spinal phantom for bone density, and a tissue bar was also included in every scan for soft tissue analysis of FFM.

Athletes wore the same long T-shirt with a sports bra and shorts for all methods of body composition except the DEXA analysis. For the DEXA measurement, only the long T-shirt and sports bra were worn.

Each athlete was given a set of written guidelines to adhere to before her designated testing date. These guidelines were distributed at either a team meeting or through the mail. The guidelines included (11) the following: 1) no large meals 4 h before the test; 2) no heavy exercise 12 h before the test; 3) urination immediately before the test; 4) no alcohol consumption 4 h before the test; 5) no diuretic medications 7 days before the test; and 6) consumption of liquids limited to 1% of body weight, or ~two 8-oz. glasses of water, 2 h before the test.

Statistical Analyses

Descriptive statistics (mean, SD, and range) were calculated for each varsity team studied. Variables included age (yr), height (cm), weight (kg), body mass index (BMI; kg/m²), FFM (from DEXA), and %BF (from DEXA). Those individuals who were athletes, yet not members of a varsity team, were included in a category termed “other.” They were recruited via personal communication with the investigator. Also, if <10 members of a given varsity team were tested, then the subjects’ results were added to the category noted above. Although team values were calculated for descriptive and general comparison purposes, all reliability and validity analyses were conducted on the total sample as a whole.

Reliability. BIA and NIR analysis were performed twice in each subject. Dependent variables included resistance and Xc for the BIA method (measured in Ω) and the two optical density values (OD1 and OD2) obtained with NIR (measured in nm). Intraclass correlation coefficients were determined for each BIA and NIR variable, using repeated-measures ANOVA, where Rxx = (MSs – MSr)/MSs, for a reliability estimate given both trials and an estimate given a single trial = (MSs – MSr)/(MSs + MSr), where MSs was the mean square for subjects, MSr was the mean square for error (1), and Rxx was the reliability coefficient for the test scores. Standard errors of measurement (SEM) were also calculated according to the formula $SE = S_x \sqrt{(1 - R_{xx})}$, where $S_x$ was the SD of the test scores. SEM values were calculated in both absolute and relative (%) terms.

Validity. The FFM value measured by DEXA was used as the criterion measure. Stepwise multiple regression was used to derive an equation for both BIA and NIR estimates of FFM. Predictor variables entered into the BIA equation were height, weight, resistance, and $X_c$. For the NIR equation, predictor variables included height, weight, OD1, and OD2. In all cases the resistance, $X_c$, OD1, and OD2 values used in the regression analysis represented the average of the repeat-test values.

Prediction equations for BIA and NIR were developed by using a double cross-validation technique (17). That is, the subjects were split into two samples (on the basis of odd (Odd) or even (Even) identification nos.), and an equation was developed for each, with the opposite group being used to cross-validate each equation. If the equations proved to be similar (evaluated by comparison of multiple r values, and visible inspection of graphs), groups were combined and a single equation was developed by using the entire sample.

In addition to correlation and regression techniques, error analysis was also performed. Standard errors of the estimate (SEE) were calculated (SEE = $S_x \sqrt{(1 - R^2)}$) and used as an error of prediction of DEXA-derived FFM by using the BIA and NIR estimates. Total error (TE) was also calculated and used as an indication of SD of error relative to the line of identity [TE = $\sqrt{\text{SEE}^2 + \text{SEE}_x^2}$] (23). The method of Bland and Altman (2) was used to graphically display the pairwise comparison of error results. Briefly, this was done by plotting the difference between the criterion (DEXA) and predictor (BIA or NIR) FFM values for each subject against the mean value of the two methods [(DEXA + BIA)/2, or (DEXA + NIR)/2].

An additional body composition prediction equation was developed that included only height and weight in the analysis. This equation was developed for all subjects without using a cross-validation technique.

RESULTS

A total 135 female athletes between the ages 18 and 27 yr were recruited as subjects. The data of three subjects were dropped from the study: two subjects were too tall for the DEXA scanner, and the remaining one was suspected of having an eating disorder. Therefore, the final subject total for this study was 132 athletes. All but 14 subjects were varsity athletes presently participating at MSU. The “nonvarsity” athletes included 4 ice hockey players from the MSU ice hockey club team and 10 former competitive athletes who were still physically active. Minimum activity for these additional athletes included 5 days or more of vigorous activity per week, 30 min per bout of exercise.

The total number of athletes studied represented ~65% of all members of women’s varsity teams during the length of the study. Very few potential subjects (<10) who were told about the study actually refused to...
participate. The major reason for nonparticipation was a time conflict with the available testing schedule. Three teams (basketball, golf, and tennis) were not included because of time and scheduling conflicts. Eight subjects were African-Americans, one was an Asian-American, and the rest were Caucasians.

Table 1 shows descriptive data for the subjects organized by teams. It should be pointed out that diving (n = 3), swimming (n = 6), track (sprinters and field events; n = 8), and volleyball (n = 4) team data were merged with those for the other group because of small numbers. Also, cross-country runners and track and field athletes who specialized in the 800-m run (or longer) were combined in a group called “Distance running.”

A reliability analysis is presented in Table 2. Reliability coefficients were extremely high for all measures when analyzed for both multiple and single trials. Specifically, the BIA reliability range was $R_{xx} = 0.987 - 0.997$, whereas the range for NIR values was $R_{xx} = 0.957 - 0.980$. In addition to high reliability for both multiple and single trials, SEM values were low for all variables mentioned. On a percentage basis, SEM was greater for NIR compared with BIA modalities. This was due to the higher SD of the OD1 and OD2 measures.

To assess the validity of BIA and NIR techniques, the 132 subjects were divided into two groups of 66 (Even and Odd), on the basis of whether a subject’s identification number was odd or even. The distribution of the various sports team athletes was similar in both samples. Table 3 shows BIA and NIR equations derived from the Even group. Also included in Table 3 are the cross-validation results in the cross-validation sample (Odd group). That is, subjects in the Odd group were not included in the calibration sample.

Table 2. Reliability estimates for BIA and NIR procedures

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean ± SD</th>
<th>$R_{xx}$</th>
<th>SEM</th>
<th>SEM, %</th>
<th>Single</th>
<th>SEM, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resistance ($R$)</td>
<td>548 ± 49 0.997 3.8 0.7 0.994 5.4 1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xc ($\Omega$)</td>
<td>65 ± 7 0.993 0.8 1.3 0.987 1.1 1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OD1, nm</td>
<td>0.0818 ± 0.076 0.978 0.016 19.5 0.957 0.022 27.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OD2, nm</td>
<td>0.0822 ± 0.0879 0.980 0.017 21.3 0.961 0.024 29.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

n = 132 Subjects. BIA, bioelectrical impedance analysis using resistance and reactance ($X_c$) measures; NIR, near-infrared interactance using 2 optical density measures (OD1 and OD2, respectively); $R_{xx}$, the intraclass correlation coefficient for reliability of multiple trials; Single, the reliability estimate of a single trial; SEM, standard error of measurement. Note: the mean ± SD values represent the average of 2 trials computed for each subject.
used to predict FFM by using the equation developed for the Even group. Table 4 shows the equations developed for the Odd group and cross-validated with their Even counterparts.

It is apparent from Tables 3 and 4 that the validity coefficients (r values), SEE, and TE were similar between Even and Odd samples and that the cross-validations showed similar results. Figure 1, A and B, presents a graphical representation of the closeness of the equations developed by using the Even vs. Odd group. The regression lines were virtually identical, with deviation from the line of identities being similar for both samples. Thus a single equation using all 132 subjects was developed for BIA and NIR prediction of FFM. In addition, height and weight alone were used to develop an equation. All three equations, validity coefficients, and error analyses can be seen in Table 5. The order of entry of predictor variables was weight, height, and X₃ for the BIA model and weight, OD₂, and height for the NIR model (Table 6).

**DISCUSSION**

Accurate body composition assessment is beneficial for female athletes, as it can be indicative of health and performance status. Our data provide measured (DEXA) and predicted (BIA and NIR) FFM and %BF values for the individual sports of crew, distance runners, field

Table 3. Prediction equations for FFM, using even-numbered subjects

<table>
<thead>
<tr>
<th>Method</th>
<th>Equation/Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEXA</td>
<td>Measured FFM = 49.5 ± 6.3 kg</td>
</tr>
<tr>
<td>BIA</td>
<td>FFM = (0.272 × ht) + (0.461 × wt) - (0.036 × resistance) + (0.101 × X₃) - 11.567</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.6 ± 6.2 kg, r = 0.964, SEE = 1.2</td>
</tr>
<tr>
<td></td>
<td>Cross-validation using odd-numbered subjects</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.3 ± 5.6 kg, r = 0.983, SEE = 1.1, TE = 1.2</td>
</tr>
<tr>
<td>NIR</td>
<td>FFM = (0.065 × ht) + (0.72 × wt) - (14.334 × OD₂) - 5.673</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.5 ± 6.2 kg, r = 0.984, SEE = 1.1</td>
</tr>
<tr>
<td></td>
<td>Cross-validation using odd-numbered subjects</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.6 ± 5.5 kg, r = 0.965, SEE = 1.5, TE = 1.7</td>
</tr>
</tbody>
</table>

n = 66 Even-numbered subjects. r, validity coefficient; SEE, standard error of estimate; TE, total error.

Fig. 1. A: fat-free mass (FFM) measured via dual-energy X-ray absorptiometry (DEXA; kg) vs. FFM measured via bioelectrical impedance (BIA; kg). B: FFM measured via DEXA (kg) vs. FFM measured via near-infrared interactance (NIR; kg). ○, Even-numbered subjects (n = 66); ●, odd-numbered subjects (n = 66). Note that there are 2 regression lines superimposed on each graph, representing prediction models developed from both Even and Odd subject groups.

Table 4. Prediction equations for FFM, using odd-numbered subjects

<table>
<thead>
<tr>
<th>Method</th>
<th>Equation/Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEXA</td>
<td>Measured FFM = 49.5 ± 5.8 kg</td>
</tr>
<tr>
<td>BIA</td>
<td>FFM = (0.284 × ht) + (0.38 × wt) - (0.037 × resistance) + (0.096 × X₃) - 7.85</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.4 ± 5.7 kg, r = 0.983, SEE = 1.1</td>
</tr>
<tr>
<td></td>
<td>Cross-validation using even-numbered subjects</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.6 ± 6.3 kg, r = 0.982, SEE = 1.1, TE = 1.3</td>
</tr>
<tr>
<td>NIR</td>
<td>FFM = (0.015 × ht) + (0.565 × wt) - (9.549 × OD₂) - 10.346</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.4 ± 5.6 kg, r = 0.969, SEE = 1.4</td>
</tr>
<tr>
<td></td>
<td>Cross-validation using even-numbered subjects</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.3 ± 6.3 kg, r = 0.981, SEE = 1.1, TE = 1.4</td>
</tr>
</tbody>
</table>

n = 66 Odd-numbered subjects.

Table 5. Prediction equations for FFM, using all subjects

<table>
<thead>
<tr>
<th>Method</th>
<th>Equation/Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEXA</td>
<td>Measured FFM = 49.5 ± 6.0 kg</td>
</tr>
<tr>
<td>BIA</td>
<td>FFM = (0.282 × ht) + (0.415 × wt) - (0.037 × resistance) + (0.096 × X₃) - 9.734</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.4 ± 5.9 kg, r = 0.981, SEE = 1.1</td>
</tr>
<tr>
<td>NIR</td>
<td>FFM = (0.111 × ht) + (0.641 × wt) - (12.397 × OD₂) - 8.423</td>
</tr>
<tr>
<td></td>
<td>Predicted FFM = 49.5 ± 5.8 kg, r = 0.975, SEE = 1.2</td>
</tr>
<tr>
<td></td>
<td>Height and FFM = (0.143 × ht) + (0.565 × wt) - 10.03</td>
</tr>
<tr>
<td></td>
<td>Weight Predicted FFM = 49.4 ± 5.7 kg, r = 0.961, SEE = 1.6</td>
</tr>
</tbody>
</table>

n = 132 Subjects.
hockey, gymnastics, soccer, softball, and other female athletes. The study purpose was to determine whether the BIA and NIR field techniques could provide both reliable and valid estimates of body composition in this population.

In this study, both BIA and NIR techniques were used to predict FFM compared with a value obtained from a criterion measurement instrument (DEXA) that can separate fat-free bone and soft tissue from fat tissue. BIA measures the conductivity of TBW and electrolytes (found exclusively in the FFM) with an applied radiofrequency electrical current. The NIR is designed to differentiate between the water and fat contents of the various body compartments by the amount of light that is absorbed and reflected from an infrared source. Both TBW and electrolytes have been shown to be highly correlated with FFM (20).

Although FFM may be the appropriate variable to be measured and predicted in this investigation, most body composition studies in the literature have centered around an individual’s %BF. This allows a more logical point of comparison when differences among groups that differ widely in absolute body mass are being discussed.

As can be seen in Table 1, there was little variability in either FFM or %BF measures when making comparisons either between and within our different sport teams. As a group, distance runners had the lowest values (FFM = 46.0 ± 3.8 kg; %BF = 18.3 ± 2.7%) and crew had the highest (FFM = 53.1 ± 5.4 kg; %BF = 21.9 ± 2.3%). On the whole, FFM measured by DEXA averaged 49.5 ± 6.0 kg and %BF averaged 20.4 ± 3.1% in the 132 athletes we tested. Indeed, even the 14 other nonvarsity athletes had similar values, supporting their inclusion in our sample.

Many body composition studies have been performed with female athletes. For instance, elite runners (training distance of 64.5 ± 15.8 miles/wk) have been measured at 14.3% with UWW (9). Pichard et al. (30) used DEXA to determine body composition in elite runners and reported that the average BF was 14.8%. The athletes in these two studies had somewhat lower %BF values than our distance runners, who averaged 18.3 ± 2.7%. This difference may be due to lower mileage run by college distance runners vs. highly elite runners, who are more seasoned and typically have more rigorous training schedules.

Withers et al. (46) investigated the %BF values of female soccer and softball players by using UWW. Their results showed an average of 22.0% for soccer players and 19.1% for softball athletes. These values are consistent with our data, which showed %BF values of 21.8% for soccer players and 20.9% for softball.

The body composition of gymnasts has also been studied. Sinning (39) found a mean value of 15.3% in champion gymnasts by using UWW. In contrast, Kirchner et al. (18) found an average value of 17.0% for collegiate gymnasts. Our results agree more with the study by Kirchner et al., for our gymnasts averaged 19.1% BF. A possible reason for the difference in %BF may include that Sinning’s gymnasts were all collegiate champions who were taller and lighter than subjects in our sample.

Sinning and Wilson (41) measured body composition (UWW and skinfolds) in a group of 79 women who were similar to our overall study sample. Their subjects included women who participated in a variety of inter-collegiate sports at Kent State University. The average %BF for these athletes was 20.1%, whereas the average %BF for athletes in our study was 20.4%.

Although most other studies have utilized UWW as the criterion, DEXA has been shown to correlate well with UWW for eumenorrheic female subjects (10). Also, given the BMC assessment included in all DEXA body composition estimates, it is likely to be a more widely used criterion measure than UWW in the future (11). Further contributing to the validity of DEXA, studies have shown high correlations between scale weight and DEXA (8, 32) and also DEXA vs. a four-component model (32). It is of interest to note that, in our study, DEXA values for weight were highly correlated with (r = 0.999) and only slightly above (<1%) scale weight values. This overestimation was small and should not have contributed significant error to our body composition estimates.

DEXA is believed to be a more sensitive measurement technique for body composition than UWW because it is theoretically independent of compartmental assumptions (25). However, DEXA is an expensive device and is not readily available to trainers, coaches, or even most medical staff. Therefore, other methods are frequently utilized in field settings. These techniques, if shown to be reliable and valid, are more desirable because of their reduced cost and portability, and they require little operational training.

The field techniques employed in this study included BIA and NIR. The reliability of BIA and NIR variables was assessed by calculating multiple- and single-trial reliability coefficients. Single-trial reliability for BIA was high at 0.994 and 0.987 for resistance and Xc, respectively. Also, the single-trial reliability for NIR was also high at values of 0.957 and 0.961 for OD1 and OD2, respectively. Previous investigations of the reliability of resistance in BIA studies have shown similar

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Model/Variables</th>
<th>r2</th>
<th>SEE</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIA</td>
<td>Weight</td>
<td>0.91</td>
<td>1.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>+Height</td>
<td>0.92</td>
<td>1.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>+Resistance</td>
<td>0.95</td>
<td>1.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>+Reactance</td>
<td>0.96</td>
<td>1.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NIR</td>
<td>Weight</td>
<td>0.91</td>
<td>1.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>+OD2</td>
<td>0.94</td>
<td>1.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>+Height</td>
<td>0.95</td>
<td>1.2</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

n = 132 Subjects; r2, squared value of the validity coefficient.
The high single-trial reliability coefficients found in this study indicate that only a single trial is necessary when the field techniques of BIA and NIR are utilized. This is important for the busy sports training staff faced with caring for hundreds of athletes on a daily basis. However, it is important to note that the high reliability coefficients obtained in this study were likely a function of the short interval between trials and the close attention to manufacturers’ instructions regarding measurement techniques paid by the investigator. Thus careful attention to detail regarding electrode placement, wand placement, and so on is crucial for obtaining accurate results. Despite the high single-trial reliabilities found under our specific experimental conditions, we suggest that investigators continue to perform repeat measurements whenever possible.

The accuracy of results with these field techniques depends on specific equations developed for each technique. Our equation for BIA predicted FFM and used (in order of entry) weight, height, resistance, and reactance (Table 6). Although height²/resistance is a term frequently included in BIA prediction models, it was not found to add to the variance above and beyond height and resistance entered separately. Our NIR equation also predicted FFM and included (in order of entry) the variables of weight, OD₂, and height. FFM was also highly predicted by OD₁ but slightly less so than by OD₂, with which it was highly correlated. Consequently, OD₁ was not entered into the equation (Table 6). Equations developed for specific populations (such as a similar age group and activity level) improve the accuracy of predicting body composition in a comparative population. Additionally, cross-validation of these equations is important to test for accuracy.

To illustrate the point discussed above, Sinning and Wilson (41) studied the validity of body composition analysis by using several skinfold equations in a group of female athletes (n = 79), age 17.8–22.5 yr. Only two of the nine equations tested were found to be in agreement with values obtained via UWW. The authors pointed out that the failure of all of these equations to be acceptable emphasized the need for further cross-validation studies.

In our study, we tested 132 female athletes. All had equivalently high levels of physical activity, which included daily (or twice daily) sessions of weight training and aerobic conditioning in both the on- and off-season.

Tables 3 and 4 show that all correlations were extremely high and similar for both BIA (r = 0.964–0.983) and NIR (r = 0.956–0.954) whether validation or cross-validation samples were used. Similarly, SEE and TE values were very low and consistent in all cases (1.1–1.4 kg). Additionally, a graph comparing the two equations (Fig. 1, A and B) indicated virtually identical lines for both BIA and NIR. This graph added further support for combining both samples and developing a single equation for all 132 subjects. Not surprisingly, the single-sample regression lines for BIA and NIR (Table 5) were nearly identical to all others developed from the split sample. In the future, further cross-validation of our equations should be performed by using samples collected from other laboratories.

Pichard et al. (30) used 12 BIA equations, including 2 provided by the manufacturer (RJL Systems; Detroit, MI) and 9 other equations commonly found in the literature. The best equation for elite runners was found to be the RJL Systems-2 equation, which had FFM results of r = 0.90, SEE = 2.0 kg, and TE = 2.1 kg. We tested this equation with our data (n = 132), which resulted in an average FFM of 48.7 kg, r = 0.94, SEE = 2.0 kg, and TE = 2.7 kg. Although the correlation by using the RJL equation for our subjects was high, FFM was underpredicted, and error terms were nearly twice as high as those in our BIA equation. We also tested another generalized equation developed by Kushner and Schoeller (21), who used older (age = 32.0 ± 6.6 yr) sedentary women as subjects. The results showed that...
average FFM was underpredicted at 45.4 kg, \( r = 0.94 \), SEE = 2.0 kg, and TE = 4.7 kg.

As previously mentioned, only two published BIA equations exist that have specifically included female athletes in the validation samples (15, 25). Again, we tested both equations by using our subject data. Results for the Houtkooper et al. (15) equation included average FFM = 49.3 kg, \( r = 0.935 \), SEE = 2.0, and TE = 2.2. Also, results for Lukasaki and Bolonchuk (25) were average FFM = 46.8 kg, \( r = 0.94 \), SEE = 2.0, and TE = 3.4. As was the case with the RJL Systems-2 and Kushner and Schoeller (21) equations, our prediction models had slightly higher correlation values and lower SEE and TE. One possible explanation for these differences may be that our study utilized female athletes exclusively, whereas the other equations were developed by using both active and inactive female subjects.

As stated earlier, no published research study has formulated an equation for female athletes by using NIR to assess body composition. Authors of other NIR studies suggested that researchers not use the Futrex formulas provided by the manufacturer (11). Thus we are the first to develop a valid equation for female athletes by using NIR. From our data, the NIR technique was shown to be as valid as BIA for predicting FFM in this population.

Some investigators have questioned the ability of field techniques such as BIA and NIR to add significantly to body composition estimates. Specifically, how much do resistance, \( X_a \), and OD values add to height and weight when one is developing appropriate prediction equations? Table 6 shows the variables that were entered into the prediction equations. Although weight alone accounted for most of the variance, all other variables entered added significantly to FFM prediction via BIA. However, OD\(_3\) did not explain any additional variance and was not included in the NIR model. This was due to the high multicollinearity between OD\(_1\) and OD\(_2\) measures. To further address this issue, we developed an equation based only on the heights and weights of our athletes (Table 5). Our results showed a predicted FFM value of 49.4 kg, \( r = 0.961 \), and SEE = 1.6. Given this finding, the results for BIA and NIR may appear less impressive. The predictive ability of this height and weight regression equation may be more of a function of the homogeneity of our population rather than an inadequacy of the field techniques. However, this homogeneity likely attributed to the accuracy of our BIA and NIR equations as well. Thus the limited range of FFM values in our sample contributed to the accuracy of our prediction models. However, we feel that our sample is very representative of the population of similarly aged female athletes in general. Figure 2, A and B, shows the pairwise comparison (Bland-Altman analysis) between FFM measured by DEXA and the BIA and NIR techniques. In general, the results indicate that there is no systematic increase or decrease in prediction error based on absolute FFM of a given subject. However, the two to three subjects with the most extreme discrepancy between criterion and predictor measures were at the high end of the FFM range. This occurred in both BIA and NIR prediction models.

The following example utilizing SEE analysis may put the issue of the utility of BIA or NIR vs. height and weight into perspective. For example, DEXA values averaged 62.2 kg for weight, 49.5 kg for FFM, and 20.4% for %BF. The BIA equation predicted an FFM of 49.4 kg, with SEE = 1.1. The NIR equation yielded an FFM value of 49.5 kg, with SEE = 1.2. Finally, the height and weight equation predicts FFM to be 49.4 kg, with SEE = 1.6. All three predictions appear very similar. However, by converting FFM predictions into %BF values, further comparisons can be made. For example, FFM from BIA ranges from 48.3 to 50.5 kg, which corresponds to a %BF of 22.3–18.8%. Compared with DEXA values, the SEE for BIA %BF was 1.9%. Similarly, NIR-predicted FFM is 48.3–50.7 kg, with %BF of 22.3–18.5%. Thus NIR results in a 1.8% BF error. Last, height and weight predict FFM to be 47.7–51.1 kg, which corresponds to a %BF of 23.3–17.8%. Hence, height and weight yield a 2.9% error for %BF (∼33% more error than the BIA and NIR equations).

One might argue that an increase in predictive accuracy for BF of only ∼1% is not worth the trouble and expense of using techniques such as BIA and NIR. However, the reality is that these techniques are being used all over the US and throughout the world. Practically speaking, it may be difficult to convince coaches and medical staff that this loss of accuracy is balanced by the money and time saved. This is a decision that will certainly involve more study to completely resolve. In the meantime, our data show that the BIA and NIR techniques are extremely reliable and valid for estimating body composition values in female collegiate athletes, provided that the equations given are used. The prediction models developed in this study were shown to be slightly better than using height and weight alone.

We are indebted to all the subjects for cooperation in this study. The technical assistance of Cherrie Barth, Jen Jallo, and Laurie Pruitt is sincerely appreciated.

Address for reprint requests and other correspondence: J. M. Pivarnik, Dept. of Kinesiology, 3141 Sports Circle Bldg., Michigan State Univ., East Lansing, MI 48824-1049 (E-mail: JMPIV@MSU.EDU).

Received 23 September 1998; accepted in final form 26 May 1999.

REFERENCES


