Reliability and Validity of Contemporary Bioelectrical Impedance Analysis Devices for Body Composition Assessment

Original Research

Justin J. Merrigan,¹,² Nina L. Stute,¹ Jason J. Eckerle,¹ Nicholas S. Mackowski,¹ James R. Walters,¹ Maegan L. O’Connor,¹ Kristyn N. Barrett,¹ Robert A. Briggs,¹ Adam J. Strang,¹ Joshua A. Hagen¹,²,³

¹STRONG Lab, Air Force Research Labs; Wright Patterson Air Force Base, Dayton Ohio
²The Human Performance Collaborative; The Ohio State University, Columbus Ohio
³Department of Integrated Systems Engineering, The Ohio State University

Abstract

Introduction: The aim was to determine reliability and validity of bioelectrical impedance analysis (BIA) compared to dual energy x-ray absorptiometry (DEXA).

Methods: Participants (n=93) were fasted and euhydrated (confirmed with urine specific gravity, USG) and underwent anthropometrics, DEXA scan (GE Prodigy), and three repeated trials on each BIA device (i.e., InBody 770, SECA mBCA 514, and FitTrack). A subset of participants (n=36) re-tested 12-weeks later to investigate longitudinal changes.

Results: All BIA devices had acceptable within-session reliability (coefficient of variation < 2%). Agreement with DEXA was unacceptable, poor, and moderate for FitTrack (Lin’s Concordance Correlation Coefficient, CCC=0.61), InBody (CCC=0.88), and SECA (CCC=0.91). FitTrack did not demonstrate systematic bias but had wide limits of agreement and larger underestimations occurring at higher BF%. InBody showed systematic underestimations with proportionate bias demonstrating more error at lower BF%. SECA demonstrated no bias but tended to underestimate BF% according to Bland-Altman Plots. Errors in BIA BF% estimates were not related with USG but were associated with DEXA measured BF%. The subtle 12-week changes in BF% did not agree between BIA and DEXA.

Conclusions: Caution should be taken when using BIA devices to assess BF% as devices demonstrated unacceptable agreement compared to DEXA.

Key Words: DEXA; body fat percent; Military

Introduction

The latest World Health Organization Global Health Observatory declared obesity a pandemic, with data showing more than 1.9 billion adults classified as overweight and 650 million as obese.¹ Trends in body composition of United States (US) citizens are worrisome, with 78% of US adults projected to be overweight or obese by 2030.² The largest concern is the association of high amounts of fat mass and low amounts of lean body mass with mortality, but there are many additional concerns for the US military.³ Not only will the increasing prevalence of obesity challenge military...
recruiting efforts, but the health, physical preparedness, retention, and health care costs of current US service members are also greatly hindered. Additionally, body composition estimates should be accurate and provide perspective beyond body mass index and waist measurements to provide contextual information for weight loss progress, informing nutrition education, identifying malnourishment, and preventing undesirable loss of lean body mass.

The most accurate body composition technique encompasses a 4-compartment model (4C; fat, protein, water, and mineral), but is difficult to conduct in practical settings due to time commitment, expenses, and data processing limitations. Instead, dual-energy x-ray absorptiometry (DEXA), a 3C model (i.e., fat, lean body mass, and mineral), is often treated as the criterion technique considering its strong agreement with 4C models. The DEXA technique generates X-rays at two different energies, making use of the differential attenuation of the X-ray beam to calculate the bone mineral content and soft tissue composition in the scanned regions. However, DEXA systems are prohibitively costly (upwards of $250,000 USD), require certified technicians, are large/bulky, and are relatively time consuming (up to 10 min per scan). Thus, despite many considering DEXA the gold standard for assessing total and segmental fat tissue mass, lean body mass, and bone mineral content, the alternative measures of body composition (i.e., bioelectrical impedance analysis; BIA) remain popular because they are far less expensive, portable, and quick/easy to implement.

Bio-electrical Impedance (BIA) is a 2C model that utilizes measures of conduction and resistance of electrical currents through the body to distinguish between body tissue types (i.e., fat mass and fat free mass) based on the assumption that each has varying water content that results in different electrical properties. For example, fat free mass is a good conductor (low impedance) of electrical current since it is primarily composed of water and electrolytes, while fat mass poses a higher resistance to electrical current (poor conductor, high impedance) due to minimal water content and creates increased resistance to electrical current. Thus, greater electrical resistance values detected by BIA will equate to greater fat mass (adiposity) levels. The electrical resistance values, coupled with demographic data (height, body mass, age, sex), then provides an estimation of body composition. However, it is critical to understand the reliability and validity of these tools before using them for body composition estimates that may drive intervention and administrative decisions.

Some investigations have shown strong relationships between body composition measures estimated by BIA and DEXA, but wide ranges of individual error indicates further evaluation is needed. For direct comparisons, studies found BIA to underestimate body fat percentage and fat mass, especially at higher levels of adiposity. However, conflicting evidence exists, suggesting body fat percentage and fat mass to be overestimated in leaner resistance trained individuals (men, <17%; women, <25%) with more overestimations as body fat mass increased. If the errors were systematic, one might suspect that changes over time are similar among methods, but BIA has underestimated the magnitudes of change from training interventions for fat mass and lean body mass compared to DEXA. Prior research illustrates the wide variation in BIA findings, perhaps due to lack of reliability in the measurement itself and differing patient populations across studies (i.e., confounding demographics such as age, body composition). As such, the aim of this study was to compare multiple BIA devices to DEXA.

Reliability of BIA devices was also determined, as well as comparisons of relative changes in body composition from a training intervention by BIA devices and DEXA.

Scientific Methods

Participants

Department of Defense (DoD) civilians, contractors, and service members (females, N=39; height= 65.1±2.8 cm; body mass= 73.7±14.4 kg; age= 33±7 years; males, N=51; height= 70.4±2.9 cm; body mass= 88.7±12.6; age= 35±9 years) were recruited between January and May of 2022. Participants were generally healthy and were experiencing little or no current pain or discomfort. All participants were informed of the study purpose and protocols. All participants verbally consented to testing; approved by the Institutional Review Board of the Air Force Research Laboratory (protocol; FWR202200227N). The study was performed in accordance with the ethical standards outlined in the Declaration of Helsinki.

Protocol

To determine the reliability and validity of BIA devices, several body composition analyses were conducted between January and May of 2022. Upon arriving to the laboratory, participants underwent anthropometric assessments (height, body mass), provided a urine sample (~30 to 60 mL) for hydration testing via urine specific gravity (USG: ATAGO...
Pen-Pro; Tokyo, Japan), and completed a prerequisite DEXA survey. Participants then completed one total-body DEXA scan (GE Prodigy, GE Medical Systems, Madison WI, USA). Then participants completed BIA analyses on 3 separate devices (InBody 770, SECA mBCA 515, FitTrack DARA) in random order and according to the instructions of each manufacturer. All participants completed three trials of each BIA device to determine within session reliability. A subset of participants (N=36) returned to the laboratory to complete the same testing procedures, following a 12-week training intervention, to investigate comparisons of longitudinal changes. For all sessions, participants abstained from exercise, were fasted ≥ 3hrs prior to testing, and were euhydrated upon arriving to the laboratory. Testing took place in an environmentally controlled laboratory with a controlled ambient temperature of ~23°C.

Dual-energy X-ray Absorptiometry (DXA)  
Participants underwent a full body DEXA scan using a narrow-angle fan beam technique (GE Prodigy, GE Medical Systems, Madison WI, USA). Participants wore standardized clothing with no metal parts (drawstring pants, T-shirt) and removed all jewelry. During the scan, participants laid supine, laterally centered on the table with the palms facing down and feet 45° inward. A Certified Bone Densitometry Technologist supervised all scans. Daily calibration of the equipment was performed using a manufacturer-provided quality assurance block calibration.

Bioelectrical Impedance Analysis (BIA)  
Body composition testing consisted of three repeated trials for each BIA device: InBody (InBody 770, Cerritos, CA, USA), SECA (SECA mBCA 515, SECA GmbH & Co., Hamburg, Germany), and FitTrack (FitTrack DARA, Vienna, Austria). Participants wore the same clothing as during DEXA, removed jewelry, and stood erect with bare feet on all BIA platforms. Body height, sex, age, etc. were recorded and entered the same for each participant across all devices. While preparing the devices, and to reduce possible errors from alterations in body fluid distribution, participants stood still and upright (~5-10 minutes) prior to BIA testing. For each device, body mass was measured to the nearest 0.1 kg while the impedance was calculated. Participants were instructed to remain as still as possible while undergoing testing which lasted ~60 seconds for each device.

FitTrack is a foot-to-foot BIA using 4 pressure contact electrodes on a platform scale. Thus, electrical signals were recorded from heel-to-heel and used in manufacturer algorithms to estimate body fat percentage. InBody included hand (4 electrodes) and feet (4 electrodes) components. Thus, participants stood barefoot with the soles of their feet positioned on the scale’s four electrodes while holding each handle to maintain contact with the four corresponding electrodes on the thumbs and palms of each hand. During this time, participants maintained fully extended elbows with their shoulder joint slightly abducted. SECA followed similar procedures as InBody, as they both are multi-frequency 8-point stand-on hands-to-feet BIA systems. Each uses drop in voltages as resistance between sensor electrodes at the hands and feet to determine total body water and estimate body composition values using manufacturer predictive equations. For the purposes of this study, body fat percentage was the main metric of interest.

Statistical Analysis  
Ordinary least products regressions (OLP; Model II linear regression) were used to determine whether systematic and/or proportional bias existed between BIA devices and DEXA, which would indicate that these devices cannot reliably be used interchangeably. Systematic bias (consistent under / over estimations) was present when the 95% confidence interval of the intercept did not include “0”. Proportional bias (changes in device agreement as body fat percentage increases) was present when the 95% confidence interval of the slope did not include “1.0”. Additionally, agreement was tested using Lin’s Concordance Coefficients (CCC) with 0.80 as unacceptable, 0.81-0.89 as poor, 0.90-0.94 as moderate, 0.95-0.99 as substantial, and > 0.99 as almost perfect. Agreement was visually assessed via Bland-Altman plots. Additional agreement calculations consisted of mean average error (MAE = DEXA BF% - BIA BF%), mean absolute percent error (MAPE = absolute (DEXA BF% - BIA BF%) / DEXA BF%), and root mean square error (RMSE). All statistical procedures were conducted using R, version 4.1.2 (R Core Team, Vienna, Austria; https://www.R-project.org) with alpha level set to < 0.05.

Results  
First, reliability of the 3 BIA devices was assessed by calculating the coefficient of variation for the trials run in triplicate. All BIA devices reported acceptable reliability measures (Table 1).
Table 1. Descriptive statistics and reliability for body fat percentage.

<table>
<thead>
<tr>
<th>Device</th>
<th>CoV</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual Energy X-ray Absorptiometry (DEXA)</td>
<td>30.28 ± 8.89%</td>
<td></td>
</tr>
<tr>
<td>InBody 770</td>
<td>1.73</td>
<td>26.40 ± 9.64%*</td>
</tr>
<tr>
<td>SECA mBCA 515</td>
<td>0.99</td>
<td>28.24 ± 8.75%*</td>
</tr>
<tr>
<td>FitTrack DARA</td>
<td>0.24</td>
<td>26.68 ± 7.51%*</td>
</tr>
</tbody>
</table>

Coefficient of Variation (CoV) > 10% is considered unreliable

*, indicates significant difference compared to DEXA

According to CCC, agreement magnitudes with DEXA BF% were as follows: moderate for SECA, poor for InBody, and unacceptable for FitTrack (Table 2). The MAPE for all BIA devices were >10% indicating large error rates on average (Table 2). Our results indicated 58% of measures within 10% error for SECA, 35% of measures within 10% for InBody, and 30% of measures within 10% for FitTrack when compared with DEXA BF% (Figure 1). FitTrack did not demonstrate systematic bias according to OLP (Table 2), but Bland-Altman Plots revealed underestimations with further underestimations occurring at higher BF% (Figure 2A). However, InBody showed systematic underestimations with proportionate bias demonstrating more error at lower BF% (Table 2, Figure 2B). SECA did not demonstrate systematic or proportionate bias, according to OLP (Table 2), but tended to underestimate BF% according to Bland-Altman Plots (Figure 2C).

Table 2. Agreement among DEXA, InBody, SECA, and FitTrack body fat percentage.

<table>
<thead>
<tr>
<th>Device</th>
<th>CCC (L-Cl, U-Cl)</th>
<th>MAPE</th>
<th>MAE</th>
<th>RMSE</th>
<th>(Est. L-Cl, Est. U-Cl)</th>
<th>(Slope L-Cl, Slope U-Cl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FitTrack</td>
<td>0.61 (0.50, 0.69)</td>
<td>21.8</td>
<td>6.03</td>
<td>7.64</td>
<td>(-4.97, 3.07)</td>
<td>(1.04, 1.33)</td>
</tr>
<tr>
<td>InBody</td>
<td>0.88 (0.84, 0.91)</td>
<td>14.8</td>
<td>4.02</td>
<td>4.75</td>
<td>(4.84, 7.22)</td>
<td>(0.88, 0.96)</td>
</tr>
<tr>
<td>SECA</td>
<td>0.92 (0.90, 0.95)</td>
<td>10.1</td>
<td>2.86</td>
<td>3.46</td>
<td>(-0.21, 3.11)</td>
<td>(0.96, 1.07)</td>
</tr>
</tbody>
</table>

Lin’s Concordance Correlation Coefficient (CCC) < 0.8 is unacceptable, 0.81-0.89 is poor, 0.90-0.94 is moderate, 0.95-0.99 is substantial, and > 0.99 is almost perfect agreement.

Mean absolute percent error (MAPE) > 10% is considered unacceptable.

Ordinary Least Products Regression (OLP) Estimates (Est.) not including 0 indicates systematic bias.

OLP Slope not including 1 indicates proportionate bias (i.e., error increases as values increase).
Figure 1. Histograms of % error ranges within 0-10% percent error (PE, green), 10-20% PE (yellow), and greater than 20% PE (red) for each bioelectrical impedance analysis (BIA) device (InBody left, SECA center, FitTrack right) compared to dual X-ray absorptiometry (DXA).
Urine Specific Gravity (USG) measures were made prior to body composition assessments to 1) investigate any potential association between hydration levels and error and 2) to show overall hydration status of the participants in support of the methods. There was no relationship between USG and BIA BF% estimation errors compared to DEXA (Table 3). However, DEXA measured BF% was a significant predictor of error from BIA devices compared to DEXA (Table 3). For every increase in DEXA BF% there was a 0.42 increase in error for FitTrack. For every increase in DEXA BF% there was a -0.05 decrease in error for InBody. For every increase in DEXA BF% there was a 0.06 increase in error for SECA.

<table>
<thead>
<tr>
<th>Table 3. Sources of BF% error.</th>
<th>InBody</th>
<th>SECA</th>
<th>FitTrack</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>21.12 (17.02)</td>
<td>17.19 (18.07)</td>
<td>23.90 (34.95)</td>
</tr>
<tr>
<td>Urine Specific Gravity</td>
<td>-15.36 (16.64)</td>
<td>-16.96 (17.67)</td>
<td>-32.29 (34.16)</td>
</tr>
<tr>
<td>DEXA Body Comp (%)</td>
<td>-0.05 (0.03) *</td>
<td>0.06 (0.03) *</td>
<td>0.42 (0.05) ***</td>
</tr>
<tr>
<td>R^2</td>
<td>0.03</td>
<td>0.05</td>
<td>0.32</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.02</td>
<td>0.04</td>
<td>0.31</td>
</tr>
<tr>
<td>Number of observations</td>
<td>140</td>
<td>132</td>
<td>138</td>
</tr>
</tbody>
</table>

*, indicates statistical significance at p < 0.05, while ***, indicates statistical significance at p < 0.001
Changes in BF% (DEXA $\Delta = 1.45 \pm 1.8$) from the two timepoints (Week 0 and Week 12) did not agree between BIA devices and DEXA (Table 4). Despite average errors being below 3% (BF%) for InBody and SECA, there was high proportionate bias and unacceptable agreement at the individual level (Table 4, Figure 3).

**Table 4.** Agreement among DEXA, InBody, SECA, and FitTrack body fat % changes.

<table>
<thead>
<tr>
<th>Device</th>
<th>CCC (L-CI, U-CI)</th>
<th>MAE</th>
<th>RMSE</th>
<th>(Est. L-CI, Est. U-CI)</th>
<th>(Slope L-CI, Slope U-CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FitTrack</td>
<td>0.15 (-0.06, 0.34)</td>
<td>2.94</td>
<td>4.32</td>
<td>(-9.45, -3.96)</td>
<td>(0.33, 0.67)</td>
</tr>
<tr>
<td>InBody</td>
<td>0.33 (0.03, 0.58)</td>
<td>1.36</td>
<td>1.64</td>
<td>(-4.36, 1.23)</td>
<td>(0.46, 0.90)</td>
</tr>
<tr>
<td>SECA</td>
<td>0.33 (0.00, 0.60)</td>
<td>1.78</td>
<td>2.28</td>
<td>(-4.2, 2.64)</td>
<td>(0.47, 0.97)</td>
</tr>
</tbody>
</table>

Lin’s Concordance Correlation Coefficient (CCC) < 0.8 is unacceptable, 0.81-0.89 is poor, 0.90-0.94 is moderate, 0.95-0.99 is substantial, and > 0.99 is almost perfect agreement.

Ordinary Least Products Regression (OLP) Estimates (Est.) not including 0 indicates systematic bias.

OLP Slope not including 1 indicates proportionate bias (i.e., error increases as values increase).

**Figure 3.** Bland-Altman plots of BIA (FitTrack, A; InBody, B; SECA, C) compared to DEXA changes in body fat percentage over 12 weeks.

**Discussion**

The purpose of this study was to compare popular BIA devices (FitTrack DARA, InBody 770, SECA mBCA 514) to the established criterion device (DEXA) for BF% measures within a single session and BF% changes over time. Each BIA device demonstrated acceptable reliability across 3 assessments conducted within the same session. However,
BIA devices did not reveal acceptable agreement with DEXA measured BF%. Foot-to-foot BIA (FitTrack) had the worst agreement with DEXA, while SECA demonstrated the most promise in terms of average agreement but still had wide limits of agreement according to Bland-Altman plots. The within session error from BIA compared to DEXA was not influenced by hydration (USG) but was associated with DEXA measured BF% for all devices. Lastly, despite average changes over time being generally comparable between two BIA devices (InBody, SECA), there were wide limits of error rates and unacceptable individual level agreement according to CCC and Bland-Altman plots.

In line with previous findings, BIA measures of body composition did not agree with DEXA. 12,13,16,26 Some prior research has identified significantly higher estimations of BF% when using a 4-electrode foot-to-foot BIA device in recreationally trained individuals classified as having low/moderate levels of body fatness. 16 However, in a more diverse sample of healthy individuals (19-60 years), a 4-electrode hand-to-foot BIA underestimated BF% in the high BF% group (BF% >30) but overestimated in the low BF% group (BF% <20%). 21 FitTrack, the 4-electrode foot-to-foot BIA in the current study also demonstrated further underestimations in BF% as body fat levels increased, which has also been proposed in various BIA validity studies involving overweight and obese participants (range: 36.1-54.7 BF% ). 23,33 Therefore, BIA may overestimate BF% in leaner, athletic individuals, and underestimate BF% in heavier individuals with greater BF%. Such systematic underestimations may be problematic as they could conceivably delay interventions which might have mitigated weight gain among an at-risk population. However, when using 8-electrode hands-to-feet methods, BIA has shown strong relationships with criterion measures. 21,33 For example, some found the SECA mBCA 515 to estimate fat mass more accurately compared with conventional BIA methods. 34 The current study also noted stronger agreement with DEXA for the 8-electrode InBody and SECA compared to FitTrack. However, the InBody had significant systematic underestimations compared to DEXA, with greater underestimations at lower BF%. SECA did not have systematic or proportionate bias, although Bland-Altman plots revealed a greater tendency for SECA to also underestimate BF%. When comparing changes in BF% longitudinally over 12 weeks, InBody and SECA demonstrated mean absolute error <2% compared to DEXA and no systematic bias; however, both devices yielded proportionate bias with wide error limits and unacceptable agreement at the individual level. Thus, these devices should likely not be used interchangeably with criterion BF% measures considering their wide ranges of error.

Additional rationale for a lack of validity from BIA measures could be attributed to variations in hydration status or fluid shifts during orthostatic posture. 14,19,26 For example, fluctuations in hydration can result in undulations of BF% estimates from 2-3%,35 with states of dehydration causing underestimations of fat free mass by 2.6 kg 36 and acute ingestion of 500 mL of water prior to BIA measurement causing underestimations of BF%.37 However, hydration was likely not a factor in accuracy within the current protocol according to regression analyses and considering instructions to participant required they maintain hydration which they achieved according to the range of USG values (1.018±0.014).38 Participants were also asked to refrain from water consumption and use the restroom immediately before testing to reduce any impact of hydration or weight change.39 Thus, it is likely that the error rates in the current study may be exacerbated in real-world application scenarios, such as military or athletic settings, where hydration statuses or fluid consumption may be more uncertain and difficult to confirm compared to laboratory settings.

It is important to reiterate that BF% levels influence the result of each BIA device differently. FitTrack had much larger underestimations of BF% at higher levels of body fat, while InBody had much larger underestimations of BF% at lower levels of body fat. Although SECA did not demonstrate proportionate bias according to OLP or Bland-Altman plots, regression analyses suggested that BF% estimation errors increased by a small amount (0.06%) as DEXA BF% increased by one percent. Considering heterogeneity, in terms of body composition, these proportionate biases are problematic for various military, first responder, and athletic settings. These divergent findings are worrisome as overestimating BF% in those with lower BF% or underestimating BF% in those that are considered obese may subject individuals in either case to inappropriate physical training or nutritional advice; and could potentially unrightfully exclude individuals from various opportunities.

The goal is to provide an affordable, yet accurate, assessment of body composition that can be implemented with ease across large groups of individuals. This is why Body Mass index (BMI) is still one of the most viable options for screening individuals in large cohorts for obesity related ailments despite being an inaccurate measure of body composition due to only considering body height and mass.40 Thus, many populations, such as military, have added additional assessments to verify whether individuals are at risk by incorporating additional assessments such as abdominal circumference – neck circumference, skin calipers, and waist-to-height ratio. Although BIA seems like a viable option by being cost-effective and easy to administer to large cohorts of individuals in a short period of time, the validity of such systems may not yet be ready for widespread implementation across the military or athletic settings.
with confidence in subsequent decisions making. Until these concerns are addressed, other options for widespread body composition testing should continue to be considered such as air displacement plethysmography, which has shown some promise in comparison to DEXA,\textsuperscript{41} although accuracy is still influenced by an individual’s body composition.\textsuperscript{16} Ultimately, if detailed body composition assessments are being used to govern any pragmatic (i.e., physical training, nutrition) or administrative (i.e., prerequisites for occupational specialties) decisions they should be collected with a trusted device that delivers reliable and valid results for the entire cohort. Otherwise, testing physical preparedness through means of physical fitness testing may be the most viable option since increasing levels of BF\% decreases physical capacity making it increasingly difficult to meet physical fitness standards.\textsuperscript{4,42}

**Conclusions**

Our findings display that despite showing reliable within-session results, BIA devices (FitTrack DARA, InBody 770, SECA mBCA 514) demonstrated unacceptable errors in BF\% at the individual level. Since BF\% influenced the magnitude of error for BIA devices, they may not be suitable for diverse cohorts such as many military settings. If exact body composition data are required, BIA should not be used in lieu of criterion measures such as DEXA systems. When use of such devices is necessary, caution should be taken with interpreting and evaluating BF\% results from BIA. Until these methods are improved, additional research is warranted to elucidate the utility of alternative body composition estimations (waist-to-height ratio, air displacement plethysmography) vs. the gold standard DEXA and subsequent related ailments or levels of preparedness.

**Acknowledgements**

The views expressed in this material are those of the authors, and do not reflect the official policy or position of the U.S. Government, the Department of Defense, or the Department of the Air Force. This study was financially supported by Air Force Research Laboratory (AFRL) and U.S. Special Operations Forces Acquisition, Technology, and Logistics Center.

**References**


42. Popkin BM. Is the obesity epidemic a national security issue around the globe? Curr Opin Endocrinol Diabetes Obes. 2011;18(5):328.