Interdevice Reliability of A-Mode Ultrasound to Measure Body Composition

Megan Bigler¹, Dale R. Wagner² and R. Stephan Lowry³

¹Utah State University, Athletics Department, Logan, UT, USA; ²Utah State University, Kinesiology & Health Science Department, Logan, UT, USA; ³Grantsville High School, Grantsville, UT, USA

Abstract

A-mode ultrasound is a noninvasive, rapid method for measuring subcutaneous fat thickness and estimating body fat percentage (%BF). Validity and reliability of the BodyMetrix BX2000 A-mode ultrasound has been reported; however, this study aimed to compare results from two machines to determine interdevice reliability. Ultrasound measures were repeated with two BX2000 machines at 10 body sites (chest, biceps, triceps, scapula, lower back, hip, waist, thigh, calf, axilla) on 42 males of varying age and leanness (age: 28.6±11.9 y, BMI: 25.4±4.6 kg/m²). The intraclass correlation coefficients ranged from 0.939 to 0.998 with standard errors of measurement from 0.31 to 0.58 mm of fat thickness. The largest mean difference between devices was 0.37 mm at the scapula. The difference between machines in %BF was not significant (0.34%BF; p=0.09). The interdevice reliability is similar to the previously reported test-retest reliability with no clinical significance between machines.

Keywords: subcutaneous adipose tissue, body fat, intermachine variability

Introduction

Ultrasound as a diagnostic tool to measure subcutaneous fat has recently increased in popularity. This method is fast, cost-effective, portable, and easy to use. A-mode (amplitude modulation) ultrasound provides a graphical representation of tissue thicknesses in response to reflected sound waves. Each tissue interface elicits a spike in the graph with a large amplitude recorded at the subcutaneous fat-muscle junction. One reads the depth or thickness of the subcutaneous fat layer along the x-axis (Figure 1). For
a more complete explanation of this technology, see the review by Wagner (2013).

A commercial A-mode ultrasound device commonly used for body composition assessment is the BodyMetrix BX2000 (IntelaMetrix, Inc., Livermore, CA, USA). Fat thickness measurements obtained from this device were validated against cadaver dissections (Wagner, Thompson, Anderson, & Schwartz, 2019) and high-resolution brightness modulation (B-mode) ultrasound (Wagner, Teramoto, Judd, Gordon, McPherson, & Robison, 2020). In addition to validity, test-retest, day-to-day, and interrater reliability of the BX2000 have also been evaluated (Hendrickson, Davison, Schiller, & Willey, 2019; Loenneke et al., 2014; Miclos-Balica et al., 2021; Smith-Ryan, Fultz, Melvin, Wingfield, & Woessner, 2014; Wagner, Cain, & Clark, 2016; Wagner & Teramoto, 2020). Test-retest reliability for the estimate of body fat percentage (%BF) obtained from the BX2000 was reported to be excellent with a ICCs of 0.979 (Miclos-Balica et al., 2021) to 0.996 (Wagner et al., 2016). Similarly, the day-to-day reliability was nearly as good with ICCs ranging from 0.935 to 0.980 (Loenneke et al., 2014; Smith-Ryan et al., 2014). Additionally, the interrater reliability was excellent (ICC=0.972 to 0.987) for both experienced technicians (Miclos-Balica et al., 2021; Wagner et al., 2016) and novice examiners (ICC=0.969 to 0.990) (Wagner & Teramoto, 2020).

Despite excellent ratings for test-retest, day-to-day, and interrater reliability, the interdevice reliability is unknown. Interdevice or intermachine reliability, although rarely evaluated, serves to further demonstrate that a measurement device is acceptable for intersite and multicenter body composition testing. The purpose of this study was to evaluate the interdevice reliability of A-mode ultrasound by comparing site-specific subcutaneous fat thicknesses and %BF estimates from two BodyMetrix BX2000 machines.

Methods

Participants

This study included 42 male participants, ranging from 18 to 57 years of age and varying in body type (16 elite, 18 athletic, and 8 non-athletic). The BodyView Professional software that accompanies the BodyMetrix BX2000 defines these body types as: elite being those with low body fat and good muscle definition, including “six-pack” abdominal muscles; non-athletic being clearly overweight or obese; and the athletic category includes everyone else. Inclusion of the body type classification aids the software in selecting which A-mode peak corresponds to the fat-muscle interface.

The university’s Institutional Review Board approved the study (protocol #9696). Upon arrival on testing day, participants signed a written informed consent before participating.

Procedures

Participants emptied their bladders before testing began. We measured height and weight to the nearest 0.1 cm and 0.1 kg, respectively. Height was measured with a wall-mounted stadiometer (Seca 216, Seca Corp., Ontario, CA), and weight was measured with a digital scale (Seca 849, Seca Corp., Ontario, CA). Participants wore only compression shorts for all measurements. In addition to height and weight, we entered age and body type, as defined above, into the BodyView Professional software.

Each participant had ten measurement sites (chest, scapula, axilla, triceps, waist, hip, thigh, lower back, biceps, and medial calf) marked with a surgical marker to facilitate accurate placement of the ultrasound transducer head for both devices. Site-point ultrasound measurements were taken with the participant standing, according to the manufacturer’s instructions as detailed by Wagner (2013). Participants were measured at all 10 sites with device 1, and then measurements were repeated with device 2. The BodyView Professional software requires multiple measurements at each site to obtain a site-specific fat thickness; thus, 3-5 measurements per site produced an averaged site-specific measurement for each device. The same technician took all measurements with both devices.

Statistical Analyses

Data were analyzed using SPSS version 25.0 (IBM, Inc., Chicago, IL, USA). We calculated means and standard deviations for all variables. The BodyView software uses a proprietary formula to convert the ultrasound fat thicknesses into a %BF estimation using the measurement sites of common skinfold prediction equations. The BodyMetrix conversion of the 7-site Jackson and Pollock (1978) equation was used to estimate %BF. A paired t-test was done to determine if there was a significant difference between the two devices for the estimate of %BF. We assessed interdevice reliability for each of the 10 measurement sites with a single measures intraclass correlation coefficient (ICC) with absolute agreement. Additionally, standard error of measurement (SEM) was calculated [SEM = SD x √(1-ICC)]. The standard deviation used for SEM calculation was the combined standard deviations of the two ultrasound devices. Subsequently, minimal difference (MD) was calculated as MD = SEM x 1.96 x √2. The ICC, SEM, and MD are the recommended statistics for quantifying reliability (Weir, 2005).

Results

The sample varied in age (28.6±11.9 y), height (182.4±7.6 cm), weight (84.5±16.9 kg), and body mass index (25.4±4.6 kg/m²). During preliminary screening, one participant was determined to be a statistical outlier (>3 SD) for the chest, hip, and axilla measurements. Consequently, this participant was removed from the site-specific analyses for these three sites and the %BF estimation. See Table 1 for interdevice reliability results. The ICCs were excellent for all 10 sites, ranging from 0.939 to 0.998. Additionally, the SEM were small, ranging from 0.31 mm to 0.58 mm, with MD ranging from 0.86 mm to 1.61 mm. Further, mean differences were <0.4 mm at each site and were smaller than the SEM at each site. The difference in %BF between ultrasound 1 and ultrasound 2 was only 0.34±1.24%BF and not statistically significant; t(40) = 1.76, p=0.09.

Discussion

This research represents the first interdevice reliability study of the BodyMetrix BX2000 A-mode ultrasound. The high ICCs and low SEMs at each measurement site indicate excellent agreement between the two machines. Furthermore, the ICCs for interdevice reliability were sim-
The interdevice reliability was excellent, suggesting that re-

years older. Despite the age difference between the devices,

from the same manufacturer, yet one machine was several

device was not influenced by high %BF. One limitation

of our sample were in the “non-athlete” BodyMetrix body type classification. More research may be needed on obese individuals to determine if the outlier identified in this study is an aberration or if obese individuals are subject to machine or technician error.

Both devices used were BodyMetrix BX2000 A-mode ultrasound machines. These devices were the same model from the same manufacturer, yet one machine was several years older. Despite the age difference between the devices, the interdevice reliability was excellent, suggesting that results from older devices are still comparable to newer machines.

Acknowledgements
There are no acknowledgements.

Conflict of Interest
The authors declare that there is no conflict of interest.

Received: 05 May 2021 | Accepted: 13 July 2021 | Published: 01 February 2022

References

Interdevice reliability is an important component in determining the overall usefulness of a measurement tool. High interdevice reliability gives greater confidence for comparing results between labs or testing centers. Other body composition assessment tools have undergone interdevice reliability testing. For example, interdevice reliability of the Bod Pod was carried out by evaluating the reliability of two machines in the same laboratory (Ball, 2005) as well as a multi-site comparison (Collins, Saunders, McCarthy, Williams, & Fuller, 2004). No clinically significant differences in the estimation of %BF were found in these interdevice reliability studies, and these studies helped to further solidify the Bod Pod as a viable assessment method. Similar to the Bod Pod studies, this research yielded strong correlation and agreement between the A-mode ultrasound machines.

It is important to note that there are several A-mode ultrasound machines from various manufacturers commercially available for body composition measurement. This study was limited to the BodyMetrix BX2000 device, which operated at a fixed frequency of 2.5 MHz. Other devices may operate at different frequencies or with different software algorithms. Thus, while our research suggests that the results from various clinics or laboratory settings using the BodyMetrix BX2000 can be compared, we do not recommend comparing the results from different A-mode ultrasound manufacturers. In conclusion, there was no significant difference in the estimate of %BF between the two BodyMetrix BX2000 A-mode ultrasound machines, and the interdevice reliability was high.

*Journal of Orthopaedic Trauma, 33*(9), 472-477. doi:10.1097/BOT.0000000000001498

Table 1. Interdevice Reliability Results Between Two BodyMetrix BX2000 A-Mode Ultrasound Machines

<table>
<thead>
<tr>
<th>Measurement site</th>
<th>ICC (95% CI)</th>
<th>SEM (mm)</th>
<th>MD (mm)</th>
<th>Mean difference (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest</td>
<td>0.982 (0.927-0.993)</td>
<td>0.40</td>
<td>1.11</td>
<td>0.33</td>
</tr>
<tr>
<td>Triceps</td>
<td>0.986 (0.970-0.993)</td>
<td>0.35</td>
<td>0.97</td>
<td>0.20</td>
</tr>
<tr>
<td>Biceps</td>
<td>0.939 (0.889-0.966)</td>
<td>0.56</td>
<td>1.55</td>
<td>0.14</td>
</tr>
<tr>
<td>Scapula</td>
<td>0.980 (0.936-0.992)</td>
<td>0.47</td>
<td>1.30</td>
<td>0.37</td>
</tr>
<tr>
<td>Back</td>
<td>0.998 (0.996-0.999)</td>
<td>0.31</td>
<td>0.86</td>
<td>0.09</td>
</tr>
<tr>
<td>Hip</td>
<td>0.991 (0.983-0.995)</td>
<td>0.33</td>
<td>0.91</td>
<td>0.05</td>
</tr>
<tr>
<td>Waist</td>
<td>0.998 (0.996-0.999)</td>
<td>0.52</td>
<td>1.44</td>
<td>0.26</td>
</tr>
<tr>
<td>Thigh</td>
<td>0.985 (0.971-0.992)</td>
<td>0.35</td>
<td>0.97</td>
<td>0.17</td>
</tr>
<tr>
<td>Calf</td>
<td>0.984 (0.970-0.991)</td>
<td>0.34</td>
<td>0.94</td>
<td>0.06</td>
</tr>
<tr>
<td>Axilla</td>
<td>0.971 (0.941-0.985)</td>
<td>0.58</td>
<td>1.61</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Legend: ICC - intraclass correlation coefficient; SEM - standard error of measurement; MD - minimal difference

