2020, Volume 3 (Issue 2): 4 OPEN ACCESS

# Intra- and Inter-Rater Reliability of Assessing Body Composition Using B-Mode Ultrasound in Conjunction with Artificial Intelligence Software

*Original Research*

Alexa J. Chandler[1], Samuel T. Dona[2], Harry P. Cintineo[1], Bridget A. McFadden[1], David J. Sanders [3], Robert Monaco[4], Shawn M. Arent[1, 3]

[1]*Department of Exercise Science, University of South Carolina, Columbia, SC/USA*
[2]*Tri-County Orthopedics, Cedar Knolls, NJ/USA*
[3]*Rutgers, The State University of New Jersey, New Brunswick, NJ/USA*
[4]*Atlantic Sports Health, Morristown, NJ/USA*

**Abstract**

**Introduction**: Ultrasound (US) is proposed as a portable body composition (BC) assessment tool, however technician skill-level may impact measurement reliability. The purpose of this study was to assess intra- and inter-rater reliability of two technicians (T1; T2) using brightness mode (B-mode) US to determine body fat percentage (%BF).

**Methods**: 17 adults (male: n=9, $M_{age}$=25.9±3.9 yr; $M_{BMI}$=23.0±2.5 kg/m²; female: n=8, $M_{age}$=29.6±10.8 yr; $M_{BMI}$=25.7±3.3 kg/m²) were scanned at seven anatomical locations using B-mode US. %BF was determined by modified skinfold equations. Reliability was assessed by intraclass correlation coefficients (ICC) and one-way ANOVA was used to determine differences within and between technicians.

**Results**: There was strong intra-rater reliability (T1: ICC= 0.998; CI 0.996-0.999; T2: (ICC= 0.997; CI 0.992-0.999) and strong inter-rater reliability (ICC= 0.983; CI 0.946-0.994) for %BF measures with no significant differences within or between technicians (P>0.10). Agreement between technicians was stronger when assessing females (ICC=0.992; CI: 0.963-0.998) compared to males (ICC=0.867; CI 0.430-0.970), which is also reflected in differences between technicians' individual site measurements in males only (P<0.05).

**Conclusions**: There is strong intra- and inter-rater reliability when using B-mode US to determine %BF in recreationally active adults. However, BC-specific US training may be beneficial for all technicians, even those with vast US-imaging experience.

***Key Words***: ultrasonography, anthropometrics, body fat.

*Published: April 11, 2020*

Corresponding author: SHAWN M. ARENT, SARENT@MAILBOX.SC.EDU

## Introduction

Body composition (BC) assessments are used in research, medicine, and athletics to evaluate changes in adipose and muscle mass over time by determining body fat percentage (%BF) and fat-free mass (FFM)[1,2]. These assessments are especially valuable in weight-class sports, where athletes often take extreme measures to reach their weight goals (i.e. severe dehydration), which may place both health and performance at risk if changes in both %BF and FFM are not monitored appropriately[3]. However, many practitioners and coaches do not have access to expensive, yet accurate, laboratory equipment needed for these assessments. Previously validated tools, such as dual X-ray absorptiometry (DXA) and air displacement plethysmography, are expensive, not portable, and may require high technical expertise to operate, often leaving practitioners and coaches to rely on less reliable field-based methods[4]. Unfortunately, skinfold calipers (SFC) often have poor inter- and intra-rater reliability and bioelectrical impedance analysis (BIA) is highly impacted by hydration and electrolyte status, both of which may be altered during a weight-cut[5-7]. Therefore, a reliable BC assessment tool that can be used in field settings is necessary to increase exam accessibility and feasibility. The ability to implement BC assessments in both clinical and athletic settings will allow practitioners to further evaluate health and fitness by monitoring both adipose and lean tissue changes over time.

Ultrasound (US) is proposed as a BC assessment tool with both the accuracy of laboratory devices and the portability of field tools[8]. US has been used to assess BC in clinical settings for many years, but recent technological advances have resulted in more portable and less expensive devices, making them accessible in field settings[1,8]. Many US assessment protocols are similar to those used with SFC, as both methods utilize subcutaneous adipose tissue (SAT) measurements to predict total body fat percentage (%BF). However, findings by Wagner, Cain, Clark [9] suggest higher test-retest reliability using US compared to SFC when evaluating the agreement between two technicians. The most common US devices used in BC research are amplitude- (A) mode and brightness- (B) mode[2]. A-mode US uses a single pulse-echo beam to determine SAT thickness, while B-mode US combines multiple A-mode signals to create two-dimensional (2D) tissue images[1,2]. The 2D images allow for identification of adipose-muscle and muscle-bone tissue boundaries in order to determine SAT and muscle thickness[1,2]. A-mode US devices are often used solely for BC assessments but do not appear to produce accurate results in either extremely lean individuals or overweight and obese populations[10-14]. This may limit their utility in field settings, particularly when assessing combat-sport athletes, who often possess low levels of adipose compared to the general population[3]. The increased technical aspects of B-mode US and resultant 2D images, which A-mode devices cannot produce, may improve accuracy when assessing BC in extremely lean individuals and athletes, for whom identifying muscle-adipose borders may be more difficult[15,16].

B-mode US appears to be a valid and reliable method to determine both SAT and muscle thickness in healthy adults, but, as seen with SFC exams, reliability may be dependent on the technician's previous experience using US and interpreting images[6]. Technicians must consider methodological aspects such as using proper US probe orientation, minimizing SAT compression, and applying appropriate amounts transmission gel while conducting exams[17]. Further, technicians must also be able to accurately identify adipose-muscle boundaries on each image in order to effectively measure SAT thickness, which may require training and practical experience. Less experienced technicians appear to have lower test-retest reliability compared to their more experienced counterparts and struggle to clearly assess sites with more fascia, such as suprailiac and abdomen, compared to limb sites (i.e., triceps and quadriceps)[18,19]. However, few studies have evaluated reliability of %BF, as opposed to only assessing SAT thickness measurements at different anatomical sites, using B-mode US[18-20].

Most researchers using B-mode US determine %BF by regression equations based on multicomponent model results, but using modified SFC equations may be a more practical method for clinicians and coaches[2,10-12,21,22]. Modified SFC equations have rarely been applied to B-mode US exams, although a recent study reported high between-day reliability when using this technique in a sample of overweight and obese adults[16]. Additionally, commercially available artificial intelligence (AI) software use these modified SFC equations to automatically determine %BF, which may be useful for coaches in field settings who need instantaneous results. However, there is no consensus as to the requisite US imaging experience a technician should possess in order to reliably determine %BF using B-mode US, and if multiple technicians can produce similar results to each other. Therefore, the purpose of this study was to assess the intra- and inter-reliability of technicians with varying prior US experience to determine BC using B-mode US in conjunction with AI software and modified SFC procedures. It was hypothesized that the technician with more US imaging experience would have higher intra-rater reliability compared to the less-experienced technician, but there would be strong inter-rater reliability between technicians' %BF results.

## Methods

*Participants*

Seventeen healthy adults ($M_{age}$=28±8 yr; $M_{BMI}$=24.4±3.2 kg/m²) consisting of nine males ($M_{age}$=25.9±3.9 yr; $M_{BMI}$=23.0±2.5 kg/m²) and eight females ($M_{age}$=29.6±10.8 yr; $M_{BMI}$=25.7±3.3 kg/m²) were recruited to participate in this study. Subjects reported to the laboratory where written informed consent was obtained after an investigator explained all study procedures.

Two physicians were recruited to conduct all US assessments. Both physicians use US in their daily sports medicine practice for various musculoskeletal diagnostic purposes. The first physician (T1) reported having approximately three years of prior US experience, while the second physician (T2) reported over 15 years US experience and was board certified in diagnostic and interventional musculoskeletal US. Both technicians were inexperienced at using US to assess BC and, therefore, completed approximately 50 trial exams prior to this study. Technicians also completed one virtual training session led by the operations and performance team at MuscleSound® (Denver, CO, USA), the manufacturer of the software used in this study, to ensure proper SAT imaging techniques were used. This study was approved by the Institutional Review Board at Atlantic Health Systems in accordance with the Declaration of Helsinki.

*Protocol*

Subjects reported to the laboratory in a normally hydrated having refrained from exercise for ≥12 hours. Upon arrival, height and weight were measured using a calibrated stadiometer and scale, respectively. Each technician conducted three consecutive exams ($T1_1$, $T1_2$ $T1_3$; $T2_1$, $T2_2$, $T2_3$) on each subject using the B-mode US transducer (Lumify L12-4, Philips Healthcare, B. V., Netherlands). Technicians imaged SAT at seven anatomical locations (pectoralis [PEC], triceps [TRI], subscapula [SCA], midaxilla [MID], suprailiac [ILI], abdomen [ABD], and anterior thigh [THI]) corresponding to the Jackson & Pollock SFC guidelines used by the American College of Sports Medicine[23,24].

All images were uploaded to a cloud-based AI software system (MuscleSound®) to measure SAT thickness and calculate %BF and FFM. Technicians each reviewed their own images in the AI software to ensure that the fat-muscle interface were correctly determined, while blinded to each other's results. After confirming and adjusting all measurements as needed, technicians used the preset modified Jackson-Pollock 7-site SFC equation (Equation 1A; Equation 1B) in MuscleSound® to calculate body density ($D_b$), %BF, and FFM using the Siri equation (Equation 2) for each exam[23,25,26].

Equation 1. Modified Jackson-Pollock equations.

Eq. 1A. Males.

$$D_b = 1.112 - \left[0.00043499\times\left(\sum SAT_{thickness}\times 2\right)\right] + \left[0.00000055\times\left(\sum SAT_{thickness}\times 2\right)^2\right] - (0.0002882\times age)$$

Eq. 1B. Females.

$$D_b = 1.112 - \left[0.00046971\times\left(\sum SAT_{thickness}\times 2\right)\right] + \left[0.00000056\times\left(\sum SAT_{thickness}\times 2\right)^2\right] - (0.00012828\times age)$$

Equation 2. Siri equation.

$$\%BF = [4.95 \div (D_b - 4.50)]\times 100$$

*Statistical Analysis*

All statistical analyses were conducted using SPSS v25 (IBM, Armonk, NY, USA). Intra- and inter-rater reliability was assessed by mixed-model intra-class correlation coefficients (ICC) with 95% CI. Standard error of measurement (SEM=√SD(1-ICC)) and minimum differences (MD=SEM*1.96*√2) were determined based on methodology in previous reliability research[9,12]. One-way ANOVAs were used to determine differences between trials within each technician. A one-way ANOVA was also used to determine differences between technicians regardless on trial. An alpha level of $P<0.05$ was used to determine statistical significance for all measures.

**Results**

*Intra-Rater Reliability*

Both T1 and T2 demonstrated strong intra-rater reliability and no differences between trials for %BF ($ICC_{T1}$=0.998, 95% CI=0.996-0.999, $P<0.001$; $ICC_{T2}$=0.997, 95% CI= 0.992-0.999, $P<0.001$) and FFM ($ICC_{T1}$=1.000, 95% CI=0.999-1.000, $P<0.001$; $ICC_{T2}$=0.999, 95% CI=0.998-1.000, $P<0.001$) (see Table

1). There was also high intra-rater reliability when analyzing males and females separately with no significant differences between %BF and FFM measures for either technician (see Table 1). Additionally, both technicians showed strong intra-rater reliability for each site individual site measurement and no significant differences between any trial (see Tables 2 and 3).

**Table 1. Intra-Rater Reliability Measurements**

| | | | Trial 1 | Trial 2 | Trial 3 | ICC | 95% CI | SEM | MD |
|---|---|---|---|---|---|---|---|---|---|
| **T1** | %BF | M | 13.2±4.7 | 13.3±4.3 | 13.4±4.2 | 0.994** | 0.983-0.999 | 0.33 | 0.91 |
| | | F | 27.2±7.1 | 27.7±6.6 | 27.7±7.4 | 0.996** | 0.987-0.999 | 0.52 | 1.45 |
| | FFM | M | 31.1±3.3 | 31.0±3.1 | 31.0±3.2 | 0.999** | 0.996-1.000 | 0.21 | 0.60 |
| | | F | 20.8±1.8 | 20.8±1.9 | 20.7±1.5 | 0.994** | 0.980-0.999 | 0.29 | 0.79 |
| **T2** | %BF | M | 15.0±4.3 | 14.8±4.0 | 14.8±5.0 | 0.988** | 0.958-0.997 | 0.47 | 1.31 |
| | | F | 28.1±8.2 | 28.4±7.6 | 27.9±8.1 | 0.995** | 0.982-0.999 | 0.54 | 1.49 |
| | FFM | M | 30.4±3.2 | 30.5±3.0 | 30.1±3.2 | 0.997** | 0.990-0.999 | 0.36 | 1.01 |
| | | F | 20.5±1.8 | 20.5±1.9 | 21.0±3.9 | 0.991** | 0.962-0.998 | 0.37 | 1.02 |

Measurements (mean±SD) for %BF and FFM for males (M) and females (F) by each technician. FFM is expressed in kg and site measurements are expressed in mm. **=significant correlation (P<0.01).

**Table 2. Intra-rater Reliability: T1 Site Measurements**

| | | ICC | 95% CI | SEM | MD |
|---|---|---|---|---|---|
| **TRI** | M | 0.973** | 0.919-0.993 | 0.22 | 0.62 |
| | F | 0.962** | 0.880-0.992 | 0.53 | 1.46 |
| **SCA** | M | 0.981** | 0.942-0.995 | 0.27 | 0.75 |
| | F | 0.980** | 0.935-0.996 | 0.34 | 0.93 |
| **PEC** | M | 0.976** | 0.956-0.992 | 0.31 | 0.87 |
| | F | 0.977** | 0.926-0.995 | 0.56 | 1.57 |
| **ILI** | M | 0.970** | 0.910-0.993 | 0.45 | 1.24 |
| | F | 0.951** | 0.842-0.989 | 0.56 | 1.56 |
| **ABD** | M | 0.989** | 0.966-0.997 | 0.59 | 1.64 |
| | F | 0.986** | 0.954-0.997 | 0.85 | 2.37 |
| **THI** | M | 0.974** | 0.916-0.994 | 0.30 | 0.84 |
| | F | 0.985** | 0.951-0.997 | 0.71 | 1.98 |
| **MID** | M | 0.984** | 0.952-0.996 | 0.25 | 0.68 |
| | F | 0.979** | 0.932-0.995 | 0.49 | 1.36 |

*=significant correlation (P<0.05); **=significant correlation (P<0.01). M=male; F=female.

**Table 3. Intra-rater Reliability: T2 Site Measurements**

| | | ICC | 95% CI | SEM | MD |
|---|---|---|---|---|---|
| **TRI** | M | 0.969** | 0.908-0.992 | 0.39 | 1.09 |
| | F | 0.992** | 0.968-0.998 | 0.43 | 1.20 |
| **SCA** | M | 0.854** | 0.569-0.964 | 1.03 | 2.85 |
| | F | 0.910** | 0.707-0.980 | 1.03 | 2.86 |
| **PEC** | M | 0.951** | 0.853-0.988 | 0.57 | 1.58 |
| | F | 0.977** | 0.923-0.995 | 0.58 | 1.60 |
| **ILI** | M | 0.875** | 0.617-0.969 | 0.80 | 2.21 |
| | F | 0.968** | 0.891-0.993 | 0.74 | 2.04 |
| **ABD** | M | 0.935** | 0.799-0.984 | 1.27 | 3.51 |
| | F | 0.980** | 0.933-0.996 | 1.30 | 3.61 |
| **THI** | M | 0.975** | 0.918-0.995 | 0.32 | 0.90 |
| | F | 0.983** | 0.946-0.996 | 0.63 | 1.73 |
| **MID** | M | 0.953** | 0.857-0.988 | 0.46 | 1.28 |
| | F | 0.891** | 0.653-0.976 | 1.07 | 2.98 |

*=significant correlation (P<0.05); **=significant correlation (P<0.01).

*Inter-Rater Reliability*

When analyzing the sample as a whole, there was high inter-rater reliability for %BF (ICC=0.983, 95% CI=0.946-0.994, P<0.001) and FFM (ICC=0.995, 95% CI=0.983-0.998, P<0.001) and no main effects of technician (P>0.10). When analyzing males and females separately, there was high inter-rater reliability with no main effect (P>0.10) for both %BF and FFM (see Tables 5 and 6). Additionally, there was high inter-rater reliability for each individual site measurement in both sexes, although the 95% CI ranges were narrower in females compared to males. In addition to higher inter-rater reliability in females, T1 produced smaller measurements than T2 at the TRI (P=0.006), SCA (P=0.048), and PEC (P=0.031) in males only (see Tables 5 and 6).

**Table 5. Inter-rater Reliability for Male Subjects**

| | T1 Mean±SD | T2 Mean±SD | ICC | 95% CI | SEM | MD |
|---|---|---|---|---|---|---|
| **%BF** | 13.3 ±4.4 | 14.9 ±4.9 | 0.867** | 0.430-0.970 | 1.56 | 4.32 |
| **FFM** | 31.0 ±3.2 | 30.4 ±3.1 | 0.970** | 0.844-0.993 | 1.18 | 3.26 |
| **TRI** | 4.3 ±1.4* | 6.0 ±2.3 | 0.682** | -0.266-0.929 | 1.13 | 3.13 |
| **SCA** | 5.9 ±2.0* | 7.3 ±2.4 | 0.733** | -0.047-0.938 | 1.19 | 3.29 |
| **PEC** | 4.9 ±2.1* | 6.9 ±2.5 | 0.551 | -0.346-0.887 | 1.68 | 4.64 |
| **ILI** | 6.1 ±2.6 | 6.8 ±2.1 | 0.605 | -0.724-0.911 | 1.45 | 4.01 |
| **ABD** | 14.9 ±5.8 | 13.6 ±5.0 | 0.897** | 0.581-0.976 | 1.70 | 4.71 |
| **THI** | 5.9 ±1.9 | 6.4 ±2.0 | 0.942** | 0.742-0.987 | 0.46 | 1.27 |
| **MID** | 5.6 ±2.0 | 6.2 ±2.1 | 0.940** | 0.688-0.987 | 0.49 | 1.36 |

Data are expressed as mean ±SD. FFM is expressed in kg and site measurements are expressed in mm.
*=significant difference from T2 results (P<0.05); ICC: *=significant correlation (P<0.05);
**=significant correlation (P<0.01).

**Table 6. Inter-rater Reliability for Female Subjects**

| | T1 Mean±SD | T2 Mean±SD | ICC | 95% CI | SEM | MD |
|---|---|---|---|---|---|---|
| **%BF** | 27.5 ±7.0 | 28.0 ±7.8 | 0.992** | 0.963-0.998 | 0.64 | 1.77 |
| **FFM** | 20.7 ±1.7 | 20.6 ±3.1 | 0.988** | 0.947-0.998 | 0.42 | 1.15 |
| **TRI** | 10.9 ±3.4 | 12.2 ±4.9 | 0.939** | 0.646-0.988 | 1.01 | 2.81 |
| **SCA** | 7.6 ±2.5 | 8.4 ±3.3 | 0.932** | 0.668-0.986 | 0.73 | 2.02 |
| **PEC** | 7.7 ±3.5 | 8.2 ±3.9 | 0.904** | 0.543-0.981 | 1.12 | 3.09 |
| **ILI** | 9.4 ±3.2 | 9.3 ±4.2 | 0.925** | 0.610-0.985 | 0.99 | 2.73 |
| **ABD** | 18.9 ±8.4 | 17.8 ±7.5 | 0.980** | 0.903-0.996 | 1.09 | 3.02 |
| **THI** | 13.6 ±4.6 | 13.8 ±4.9 | 0.989** | 0.949-0.998 | 0.48 | 1.34 |
| **MID** | 6.8 ±4.0 | 7.3 ±3.0 | 0.962** | 0.827-0.992 | 0.66 | 1.84 |

Data are expressed as mean ±SD. FFM is expressed in kg and site measurements are expressed in mm. *=significant difference from T2 results (P<0.05); ICC: *=significant correlation (P<0.05); **=significant correlation (P<0.01).

**Discussion**

While US is proposed as a relatively inexpensive and portable tool to assess %BF, questions remain regarding the influence of a technician's previous imaging experience on measurement reliability[17]. The strong intra-rater reliability for %BF and FFM by both technicians in the current study support previous research findings indicating technicians with minimal experience can use US to assess %BF and produce reliable results[9]. Additionally, while there was high intra-rater reliability at each anatomical measurement site by both technicians, SEM and MD values were higher for females compared to males in both technicians. T2 was able to produce more consistent results than T1 at the ABD in females, although both technicians showed the most variation at the ILI between exams in both sexes. These results partially support previous findings indicating inexperienced US technicians produce more SAT measurement error at anatomical sites with more fascia, such as the ABD and ILI, compared to those such as THI and TRI[18]. However, contrary to previous research[18], inter-rater reliability was weak at TRI, PEC, and ILI, and there were significant differences at TRI, PEC, and SCA in males. These findings support previous conclusions that SCA images often contain artifact that cause difficulty identifying the adipose-muscle boundary and lead to incorrect image analysis[18]. However, it should be noted that differences at individual site measurements did not affect overall %BF and FFM reliability in the current study.

In comparison to SFC, which often yield low inter-rater reliability and require copious amounts of practice, the current study suggests multiple technicians with minimal experience with US for body composition purposes can use US to evaluate overall BC and produce similar results to each other[6,9,18]. Despite high overall inter-rater reliability in the present study, there were higher correlations between T1 and T2 in females compared to males for %BF, but inter-rater reliability for FFM was high in both sexes. The lower reliability for %BF compared to FFM may be due to the combination of small errors in both FFM and FM measurements, thereby leading to larger error and lower reliability for total %BF. Interestingly, there was lower inter-rater reliability for anatomical site measurements in males compared to females, despite previous conclusions suggesting more SAT, rather than less, increases thickness measurement errors[18]. However, very low SAT levels may also contribute to SAT measurement error, as identifying the muscle-adipose boundaries may be challenging in these cases as well. This may explain the lower ICCs in males compared to females in the present study, as one male was relatively leaner than the others, possibly increasing overall error when assessing males.

Technical aspects of US examination may also influence inter-rater reliability, as less-experienced technicians may require additional training in both scanning technique and image analysis. Although AI software programs aim to eliminate skill-related errors in image interpretation, the software used in the present study allowed the technicians to manually identify adipose-muscle boundaries and correct any obvious mistakes in software readings, increasing likelihood of accurate BC results. In this study, T2 produced higher SAT measurements than T1 at each anatomical site. These differences may indicate tissue compression during scanning by T1, as excess transducer pressure has been shown to reduce SAT thickness readings by 25-37% depending on the measurement location[17].

The results of the current study suggest practitioners with varying degrees of US-imaging can reliably determine %BF via B-mode US in conjunction with AI software. This may be particularly important for coaches, athletic trainers, and other practitioners who would benefit from a relatively inexpensive, portable, and reliable tool to monitor health and performance aspects related to changes in adipose and muscle mass. Additionally, the AI software may be a practical solution for those who have minimal imaging experience, as less image-interpretation skills are needed to accurately perform these assessments. However, these findings cannot be extended to novice US practitioners, as both technicians in the current study had previous US-imaging experience, albeit not for body composition purposes. Future research regarding reliability and accuracy of novice technicians is needed to determine the minimum threshold of US experience one should possess in order to proficiently determine SAT thickness and %BF using B-mode US. Additionally, studies with larger sample sizes and a greater number of technicians are needed to further investigate both intra- and inter-rater reliability when using B-mode US to assess BC.

**Media-Friendly Summary**
The current findings suggest assessing BC using B-mode US and modified Jackson-Pollock SFC equations is a reliable technique that can easily be used by practitioners and coaches in field settings. AI software can be used in conjunction with these measurements to provide instantaneous results while decreasing the amount of skill required to determine %BF from US-measured SAT. However, gaining additional US technique and image-interpretation experience may further improve inter-rater reliability, as the less-experienced technician appeared to compress SAT at certain anatomical sites. However, it should be noted that although certain individual site measurements were less reliable, overall %BF and FFM estimates were not affected. While there was strong overall inter-reliability, the current best practice would be to have all subsequent BC assessments conducted by the same technician, as recommended with similar techniques such as SFC exams.

**Acknowledgements**
The authors thank all volunteers who participated in this study as well as the staff and interns at Atlantic Sports Health who assisted with data collection. The authors would also like to thank the staff at MuscleSound® for their support with this project.

**References**
1. Prado CM, Heymsfield SB. Lean tissue imaging: a new era for nutritional assessment and intervention. *JPEN J Parenter Enteral Nutr.* 2014;38(8):940-953.
2. Wagner DR. Ultrasound as a tool to assess body fat. *J Obes.* 2013;2013:280713.
3. Reale R, Burke LM, Cox GR, Slater G. Body composition of elite Olympic combat sport athletes. *Eur J Sport Sci.* 2020;20(2):147-156.
4. Meyer NL, Sundgot-Borgen J, Lohman TG, et al. Body composition for health and performance: a survey of body composition assessment practice carried out by the Ad Hoc Research Working Group on Body Composition, Health and Performance under the auspices of the IOC Medical Commission. *Br J Sports Med.* 2013;47(16):1044-1053.
5. O'Brien C, Young AJ, Sawka MN. Bioelectrical impedance to estimate changes in hydration status. *Int J Sports Med.* 2002;23:361-366.
6. Kispert CP, Merrifield HH. Interrater reliability of skinfold fat measurements. *Phys Ther.* 1987;67(6):917-920.
7. Koulmann N, Jimenez C, Regal D, Bolliet P, Launay JC, Savourey G. Use of bioelectrical impedance analysis to estimate fluid compartments after acute variations of the body hydration level. *Med Sci Sports Exerc.* 2000;32(4):857-864.
8. Ackland TR, Lohman TG, Sundgot-Borgen J, et al. Current status of body composition assessment in sport. *Sports Med.* 2012;42(3):227-249.
9. Wagner DR, Cain DL, Clark NW. Validity and reliability of A-mode ultrasound for body composition assessment of NCAA Division I athletes. *PLoS One.* 2016;11(4):e0153146.
10. Loenneke JP, Barnes JT, Wagganer JD, Pujol TJ. Validity of a portable computer-based ultrasound system for estimating adipose tissue in female gymnasts. *Clin Physiol Funct Imaging.* 2014;34(5):410-412.

11. Schoenfeld BJ, Aragon AA, Moon J, Krieger JW, Tiryaki-Sonmez G. Comparison of amplitude-mode ultrasound versus air displacement plethysmography for assessing body composition changes following participation in a structured weight-loss programme in women. *Clin Physiol Funct Imaging.* 2017;37(6):663-668.
12. Smith-Ryan AE, Fultz SN, Melvin MN, Wingfield HL, Woessner MN. Reproducibility and validity of A-mode ultrasound for body composition measurement and classification in overweight and obese men and women. *PLoS One.* 2014;9(3):e91750.
13. Pineau JC, Fillard JR, Bocquet M. Ultrasound techniques applied to body fat measurement in male and female athletes. *J Athl.* 2009;44(2):142-147.
14. Pineau JC, Lalys L, Bocquet M, et al. Ultrasound measurement of total body fat in obese adolescents. *Ann Nutr Metab.* 2010;56(1):36-44.
15. Storchle P, Muller W, Sengeis M, et al. Standardized ultrasound measurement of subcutaneous fat patterning: high reliability and accuracy in groups ranging from lean to obese. *Ultrasound Med Biol.* 2017;43(2):427-438.
16. Smith-Ryan AE, Blue MNM, Trexler ET, Hirsch KR. Utility of ultrasound for body fat assessment: validity and reliability compared to a multicompartment criterion. *Clin Physiol Funct Imaging.* 2018;38(2):220-226.
17. Toomey C, McCreesh K, Leahy S, Jakeman P. Technical considerations for accurate measurement of subcutaneous adipose tissue thickness using B-mode ultrasound. *Ultrasound.* 2011;19(2):91-96.
18. Müller W, Horn M, Furhapter-Rieger A, et al. Body composition in sport: interobserver reliability of a novel ultrasound measure of subcutaneous fat tissue. *Br J Sports Med.* 2013;47(16):1036-1043.
19. Müller W, Furhapter-Rieger A, Ahammer H, et al. Relative body weight and standardised brightness-mode ultrasound measurement of subcutaneous fat in athletes: an international multicentre reliability study, under the auspices of the IOC Medical Commission. *Sports Med.* 2019:1-18.
20. O'Neill DC, Cronin O, O'Neill SB, et al. Application of a sub-set of skinfold sites for ultrasound measurement of subcutaneous adiposity and percentage body fat estimation in athletes. *Int J Sports Med.* 2016;37(5):359-363.
21. Johnson KE, Miller B, Juvancic-Heltzel JA, et al. Agreement between ultrasound and dual-energy X-ray absorptiometry in assessing percentage body fat in college-aged adults. *Clin Physiol Funct Imaging.* 2014;34(6):493-496.
22. Bielemann RM, Gonzalez MC, Barbosa-Silva TG, et al. Estimation of body fat in adults using a portable A-mode ultrasound. *Nutr.* 2016;32(4):441-446.
23. Jackson AS, Pollock ML. Practical assessment of body composition. *Phys Sportsmed.* 1985;13(5):76-90.
24. Medicine ACoS. *ACSM's guidelines for exericise testing and prescription.* 9 ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2018.
25. Siri WE. Body composition from fluid spaces and density: analysis of methods. In: Brozek J, Henschel A, eds. *Techniques for Measuring Body Composition.* Washington, DC: National Academy of Sciences-National Research Council; 1961:223-244.
26. Sarnow P, McNamara SM, Moretto TM, Jr., Inventors; MuscleSound, LLC, Denver, CO (US), assignee. System and method for non-invasive determination of human body fat. US patent US 9,642,593 B2. Sep. 19, 2014, 2017.