Research Article

Open Access Journal of
Biomedical Science
ISSN: 2690-487X

Measurement Reliability and Precision of Anthropometric Measures for the Estimation of Whole-Body Density in Professional Football Players

Mills CD[1]*, Cooper SM[2], De Ste Croix M[1] and James D[1]

[1]*School of Sport and Exercise, University of Gloucestershire, UK*

[2]*Cardiff School of Sport and Health Sciences, Cardiff Metropolitan University, UK*

ABSTRACT

Introduction: It is not uncommon to encounter measurement reliability issues, yet when they are required to develop practical calibration models to estimate whole body density the reliability of anthropometric measures is imperative.

Materials and Methods: Data was gathered from *n*=206 male professional footballers ($\bar{x}\pm s$; age = 24.1±5.4 years, body mass = 78.8±8.4 kg and stretched stature = 180.1±7.0 cm). Reliability of *n*=28 anthropometric variables (skinfolds, girths, breadths, depths, widths, residual lung volume, air displacement plethysmography and hydrostatic weighing) was investigated by applying Technical Error of Measurement (TEM%) and the Limits of Agreement (LoA) method.

Results: TEM% indicated a range of 6.5-1.0% and 3.8-0.6% and the study's a priori criterion was set at ±3.8% for LoA and were statistically significant ($P < 0.001$). The contentious measurements of the iliac crest, supraspinale and anterior thigh skinfold, chest and hip girth and biiliocristal breadth indicated z scores values for R_2 between 0.006-0.022% and exhibited significant bias ($P=< 0.05$) and possible heteroscedasticity issues with highest *r* values of 0.611 and R_2 (%) coefficients of 37.3%.

Discussion: Test-retest values demonstrated agreement and reliability to be of practical use with this population.

KEYWORDS: Anthropometry; Measurement error; Reliability; Test-retest; Inter-tester

INTRODUCTION

The importance of reliability and measurement error within sports science is well reported, yet it is still not uncommon to encounter reliability issues ranging from equipment calibration to technical execution and repeatability Perini [1]. When estimating the impact of reliability on the outcome of a given measurement, the sport scientist must appreciate how practical and suitable these measures are Atkinson [2]. In other words, what the measurement error represents in practice. Indeed, better reliability implies better precision of measurements, although within the field of anthropometry it is not uncommon to encounter extensive amounts of random variation Hopkins [3]; Perini et al. [1]. Therefore, establishing accuracy and reliability of anthropometric measures will enable to sport scientist to be confident in making sound judgements on each variable and whether they are as error free as possible Perini [1]. For instance, if these measures had any detrimental effect, they would not provide confidence when determining which variables should be included in the development of calibration models to estimate whole body density (g ml-1).

One of the most popular methods of statistical analyses involves the estimation of intra-observer reliability and is routinely referred to in methodological studies as the test-retest method Hume [4]. This method of obtaining reliability data is often used by anthropometric rater's when conducting studies on their own ability to measure anatomical variables consistently and accurately Norton [5]. When measurements are performed in this manner,

Quick Response Code:

Address for correspondence: Mills CD, School of Sport and Exercise, University of Gloucestershire, UK

Received: September 09, 2020 **Published:** October 15, 2020

How to cite this article: Mills CD, Cooper SM, De Ste Croix M, James D. Measurement Reliability and Precision of Anthropometric Measures for the Estimation of Whole-Body Density in Professional Football Players. 2020 - 3(1) OAJBS.ID.000227. **DOI:** 10.38125/OAJBS.000227

neither the test nor the re-test will provide an unequivocally correct measure and are subject to some form of error British Standards Institution [6]. Incidences of error can generally be caused by measurement and biological variables such as gathering of data, human frailty, accuracy of measurement equipment and tools, biological variation of the participants and the ability, confidence and experience of the rater Hopkins [3]; Perini [1].

One of the most common ways of expressing measurement reliability and which identifies various sources of error is Bland and Altman's 95% limits of agreement Bland [7]. It is identified that in developing the 95% limits of agreement method, they would be able to identify and quantify the amount of agreement, that is the systematic bias, and the amount of random variation inherent in the measured data Hopkins [3]. In other words, it is possible for the rate to observe the extent to which there is error in their measures Atkinson [8]; Thomas [9]. More often than not, sport scientists' work with data measured on either an interval or a ratio scale. In such data it is common to observe a relationship between an increase in the magnitude of values and an increase in variability between values. This situation is known as heteroscedasticity. Bland [7] maintain that heteroscedasticity can be visually detected and quantified by establishing whether a positive linear relationship exists (rXY) between the absolute differences (errors/residuals) between test-retest values (Y) and the mean of the test-retest values (X) for each participant plotted on an XY scatter plot. The issue of heteroscedasticity is important here because when limits of agreement are calculated, there is an assumption that the original test or measurement data are in fact homoscedastic (the condition of equal residual variances) and that the limits of agreement will therefore remain constant throughout the range of measurements for which they were calculated.

Another appropriate statistical analysis that identifies various sources of error and measurement reliability in kinanthropometry is by means of the Technical Error of Measurement (TEM%) Perini [1]; Stewart [10]. TEM is an index of absolute reliability and it can be used to represent test and retest values respectively on a particular measurement. TEM% is often expressed relatively as a percentage and can be used to generate 68% or 95% bands of tolerance, thus providing an objective method to evaluate the competency of a rater and are reinforced by the International Society for the Advancement of Kinanthropometry (ISAK) for the accreditation of kinanthropometrist Klipstein-Grobusch et al. [11]; ISAK [12].

The issue here is one of accepting that you cannot eliminate error entirely. So, if error exists, the question remains as to whether this error is so great that it will be detrimental to what has to be said about the primary investigator measurements and/or the present study population's values. Consequently, sport scientists must strive for reliability through intensive training and periodic quality control of measurement techniques to help reach higher accuracy and hence more optimal reliability Perini [1].

The manner in which these measurement errors are best analysed and reported has been a matter of some debate amongst researchers Perini [1]. Nevill [13] have made useful contributions to this debate, with the consensus of opinion suggesting that when assessing measurement reliability, Bland [14] 95% limits of agreement method is the most appropriate statistic to report. By identifying these various sources of error, through the central focus on the Bland and Altman 95% limits of agreement method (relative reliability) and TEM% (absolute reliability) on test-retest values of anthropometric measures, can potentially establish whether a range of error of this magnitude would have any detrimental effect on the practical use of values gathered with this population of participants when developing anthropometric calibration models to estimate whole body density (g ml^{-1}) and is therefore the main aim of this study.

MATERIALS AND METHODS

Participants and Recruitment

Two hundred and six Fédération Internationale de Football Association (FIFA) registered contracted professional football players ($\bar{x}\pm s$; age = 24.1±5.4 years, body mass = 78.8±8.4 kg, stretched stature = 180.1±7.0 cm and whole body density = 1.075±0.010 g ml^{-1}) were recruited from eight professional football clubs that represented Barclays Premiership, npower Championship, npower League One, npower League Two and Blue Square Premier Leagues during the 2007-2008, 2008-2009 and 2009-2010 playing seasons. Sampling included players who were all over 18 years of age, free from disease or illness and who agreed to act as participants for the study by giving their written informed consent. Signs and symptoms of disease and diagnosed disease were determined through health screening procedure involving completion of a health screening questionnaire. Ethical approval was granted from the University of Gloucestershire's Research Ethics Committee.

PROCEDURE

Data collection procedures were followed by using robust operational procedures by ISAK. A trial-to-trial protocol with the same interval of time (≈15-20 mins) allocated between each assessment to allow changes in the compressibility of the skinfold was established to provide an index of internal consistency Stewart et al. [10]. The following measurements were selected to offer a wide range of measures to investigate the test-retest reliability of the primary investigator's ability: stretched stature (cm); sitting height (cm); skinfold thicknesses (mm, *n*= 8); girths (cm, *n*=10); breadths, depths and widths (mm, *n*=6); underwater weighing (g ml^{-1}), residual lung volume (l) (estimated from forced vital capacity) and air displacement plethysmography (body mass (kg) and body volume (l)), stretched stature and sitting height (cm). All anthropometric measurements were taken systematically and on the right-hand side of the body and, the mean value was determined for subsequent data analysis.

Hydrostatic weighing procedures followed those described in Mills et al. [15] and from a reliability point of view, underwater weight readings from attempts 1, 2 or 10 were not used Katch [16]; Mills et al. [17]. Forced vital capacity (FVC) testing procedures followed those described by Mills et al. [15] and participants were given a minimum of 3 attempts. The greatest value was then corrected for body temperature and pressure saturated (BTPS) determined by using a correction table devised by Sinning [18]. Residual lung volume was estimated by taking a constant fraction of each participant's FVC and was expressed to the nearest 0.05 l and computed using the equation proposed by Sinning [18]. The air displacement plethysmography measurement protocol was followed with step-by-step instructions displayed by the BodPod computer system.

During the assessment, participants were instructed to continue breathing normally whilst a minimum of two 50s tests were conducted to ensure consistency Biaggi et al. [19]. Once the assessments were completed, derivation of body volume, correction for residual lung volume together with measurement of

body mass, permitted the derivation of an estimate for whole body density Biaggi et al. [19].

STATISTICAL ANALYSIS

Summary results (mean, standard deviation and range) were presented for all participants and measurement values were calculated via Microsoft Office Excel version. The reliability of all directly measured variables was investigated by applying the 95% limits of agreement (LoA) method (relative reliability) and illustrated in Bland and Altman plots. Additional interpretation at this point was used to provide visual evidence of heteroscedasticity by correlating absolute differences against mean values for test-retest measures and illustrated on scatter plots, which included the slope of the best-fit line, R_2, r and P values and the distribution line to allow a visual overview of the linear relationship between the absolute differences and means values.

TEM% was used as an index of absolute reliability by establishing the degree of precision between the primary investigator against TEM% values established by a level 4 ISAK experienced kinanthropometrist (the criterion measurer) (absolute reliability). The degree of precision between two values generated by primary investigator against the criterion measurer as an index of inter-tester reliability was calculated using a programmed Microsoft Office 1998 Excel spread sheet constructed by ISAK Ross [20]; Norton, 2002). Anthropometric measurements included a comparison of n=17 common variables (n=8 skinfold thicknesses (triceps, subscapular, biceps, iliac crest, supraspinale, abdominal, anterior thigh and medial calf); n=5 girths (arm (flexed), arm (relaxed), waist, hips and calf); n=2 widths (humerus and femur), body mass and stretched stature) against the level 1 ISAK criterion. Levels 2-4 reliability standards require more advanced technical expertise than those of level 1, where minimum requirements involve a rater being able to prove that they can: i) repeat anthropometric measures with a TEM% better than ≤7.5% for skinfolds and ≤1.5% for girths, breadths and widths for level 1 accreditation and ≤5.0% for skinfolds and ≤1.0% for girths, breadths and widths and ii) establish the degree of precision between values generated by two kinanthropometrist's Perini [1].

The issue for the primary investigator was to judge, whether the identification and quantification of agreement outcomes were narrow enough for the anthropometric measures to provide practically reliable values. Therefore, the primary investigator established a priori consideration for both the Bland and Altman 95% LoA method (relative reliability) and TEM% (absolute reliability) that presented acceptable tolerable limits within the context of this study. Under review from ISAK (2001) and previous literature, the Bland and Altman 95% LoA method, a priori criteria was set at ±3.8%, $P < 0.05$ (g ml^{-1}) and TEM% < 5.0% for skinfolds and < 1.0% for girths, breadths and widths, as acceptable limits.

RESULTS

Participants' primary anthropometric measures, where within an age range between 18-38 years, body mass, stretched stature and sitting height ranging from 59.3-104.3 (kg), 162.7-201.2 (cm) and 79.5-109.4 (cm) respectively. Summary results for general characteristics of all n=206 football players can be seen in Table 1.

Results from Table 1 indicated that the iliac crest, supraspinale, abdominal and anterior thigh skinfolds (mm) had, as anticipated, the largest values and ranges. A situation previously reported by Brodie [21]; Harrison et al. [22]; Heyward [23], suggesting larger deposits of localised storage fat. The girth measurements at the chest, hip and waist had the greatest range of mean values with 99.0 ±4.8 cm, 94.0±4.5 cm and 94.0±4.5 cm respectively. Given the nature of the sport and the physiological demand placed upon the legs with a variety of football playing positions, the anterior thigh and medial calf girths showed a large range of 21.4-63.3 and 29.7-57.4 cm respectively. Breadths, depths and width values (cm) were within ranges previously reported by Casajus [24]; Loucks [25] with the anterior-posterior chest depth with a 20.7±1.8 cm and range from 16.0-33.8 cm for professional football players.

Table 1: Summary of anthropometric ($\bar{x}$ ± s) measures for (n=206) football players.

Variables	$\bar{x}$ ± s	Range
Skinfolds (mm)		
Triceps	8.3 ± 3.1	3.7 - 18.1
Subscapular	10.2 ± 2.5	6.1 - 17.7
Biceps	4.4 ± 2.0	2.1 - 11.5
Iliac crest	15.5 ± 6.2	3.8 - 39.2
Supraspinale	9.7 ± 3.9	4.1 - 26.5
Abdominal	14.6 ± 6.0	5.1 - 34.4
Anterior thigh	12.1 ± 4.4	4.5 - 29.5
Medial calf	7.0 ± 2.5	3.0 - 15.7
Girths (cm)		
Neck	38.4 ± 1.6	34.4 - 44.0
Arm (relaxed)	31.9 ± 2.2	19.8 - 37.7

Arm (flexed)	34.2 ± 2.3	29.4 - 40.2
Forearm	28.3 ± 1.7	24.1 - 39.4
Wrist	17.5 ± 0.8	15.4 - 19.9
Chest	99.0 ± 4.8	82.5 - 109.7
Waist	81.9 ± 6.3	24.0 - 98.6
Hip	94.0 ± 4.5	75.0 - 106.9
Thigh	55.7 ± 3.8	21.4 - 63.3
Calf	38.2 ± 2.5	29.7 - 57.4
Ankle	23.1 ± 1.3	18.9 - 26.0
Breadths (cm)		
Biacromial	43.4 ± 2.0	33.8 - 49.9
Biiliocristal	29.6 ± 1.7	25.0 - 33.8
Depths (cm)		
Transverse chest	30.9 ± 1.8	26.2 - 38.1
Anterior-posterior chest	20.7 ± 1.8	16.0 - 31.3
Widths (cm)		
Humerus	7.3 ± 0.7	6.2 - 10.3
Femur	9.6 ± 0.6	6.6 - 10.9

Forced vital capacities and estimated residual lung volumes ranged between 2.0-6.8 l and 0.6-2.1 l respectively. Although when reporting on professional athletes, due to recognised individual differences such as stature, race, age and the varying aerobic demands that are required for football, the estimation of residual lung volume would expect to be higher than the assumed 0.9-1.6 l in a normal healthy adult male Pesola et al. [26]; Demura et al. [27].

Body volume and body density values results range between 58.4-93.2 l and 1.050-1.100 g ml^{-1} respectively. Perhaps an indication of the range of body density values could be alluded to by the research of Schutte et al. [28]; Donnelly et al. [29]; Fields et al. [30]; Utter et al. [31]; Collins et al. [32]. These authors reported that non-Caucasian populations have fat-free masses that are denser with assumed body density values of (on average) 1.113 g ml-1 when compared with Caucasian populations body density of (on average) 1.100 g ml^{-1}.

Hydrostatic weight attempts (kg) that were conducted on all *n*=206 participants in order to assess for body mass underwater with an average of 3.84 ± 0.04 and a range of 1.12–6.22 kg). Results indicated that all participants attempted all but the final underwater weighing and that as consecutive weighing attempts continued past the sixth weighing attempt, the values began to reduce. Given the outcome from Mills et al. [17] investigation, the first two and the last three attempts were not used to determine mean underwater weight as they are known to underreport Demura et al. [33]; Slater et al. [34]. The favourable weighing attempts were the fourth, fifth and sixth as they demonstrated very small between-attempt differences and in general the highest values of underwater weighing, thereby providing the closest estimation of a 'true' underwater weight.

TEM% for stretched stature against the level 1 ISAK criterion and stretched stature and sitting height against the level 2/3 ISAK criterion were 0.09% and 0.14% respectively. TEM% for the triceps, subscapular, biceps and iliac crest measures of 1.5%, 1.2%, 1.0% and 1.8% and differences of 6.0% 6.3% 6.5% and 5.8% against the ISAK criterion for level 1 and 3.5%, 3.8%, 4.0% and 3.3% against level 2/3 ISAK criterion respectively. Results indicated that the highest values of the triceps and iliac crest skinfolds were still lower than the ISAK recommended TEM% by as much as 3.3%. The supraspinale, abdominal, anterior thigh and medial calf skinfolds were all 1.5% respectively and 3.5% difference against the ISAK criterion. The TEM% for arm (relaxed) and arm (flexed) girths against the level 1 and were 0.3% and 0.4% with an ISAK target of 1.5%, indicating a difference of 1.2% and 1.1% respectively. The TEM% indicated differences of 0.8%, 0.7%, 0.6% and 0.8%. The waist (0.2%) and hip (0.5%) girths against the level 1 ISAK criteria at 1.5% providing differences of 1.3% and 1.0% respectively, thus achieving TEM% targets for ISAK Level 1 criterion for (*n*=17 measurements). When the wrist, chest, waist and hip, thigh, calf and ankle girths, biacromial and biiliocristal breadths, and transverse chest and anterior-posterior chest depth were compared against level 2/3 ISAK criterion values results indicated that all anthropometric TEM% targets for level 2/3 (*n*=28 measurements) were well within ISAK acceptable limits.

Limits of Agreement analyses found that stretched stature and sitting height measurements indicated a bias of +0.0 cm and +0.05 cm and 95% limits of agreement of -0.1 cm to +0.1 cm and -0.0 cm to +0.2 cm respectively. Figure 1 illustrates the *n*=8 skinfold thicknesses (mm) (triceps, subscapular, biceps, iliac crest

supraspinale, abdominal, anterior thigh and medial calf) indicated a bias of a range from +0.08 to + 0.01 mm, and 95% limits of agreement of –0.2 to +0.2 mm.

The anthropometric girths (*n*=11) (cm) showed a bias of +0.02 to +0.01 cm, and 95% limits of agreement of -0.2 to +0.2 cm (Figure 2). As was to be expected the direction and size of the data scatter around the zero line suggests evidence of systematic bias in the values particularly of the chest and hip girths.

These girths can be problematic particularly due to the chest movement during inhalation and exhalation processes and with participants wearing light clothing in the hip region, thus causing greater variance in test and re-test scores.

All remaining measures of the biacromial and biiliocristal breaths (cm), transverse chest depth and anterior-posterior chest depth (cm), humerus and femur widths (cm) demonstrated a bias of +0.06 cm to +0.01 cm and 95% limits of agreement of +0.3 to -0.3 cm (Figure 3). Data scatter around the zero line for the transverse chest depth specifically, indicated systematic bias and random variation.

Overall, the differences between test and re-test values for stretched stature, sitting height, skinfolds, girths, breadths, depths and widths were normally distributed were well within acceptable limits when using the study a priori criteria.

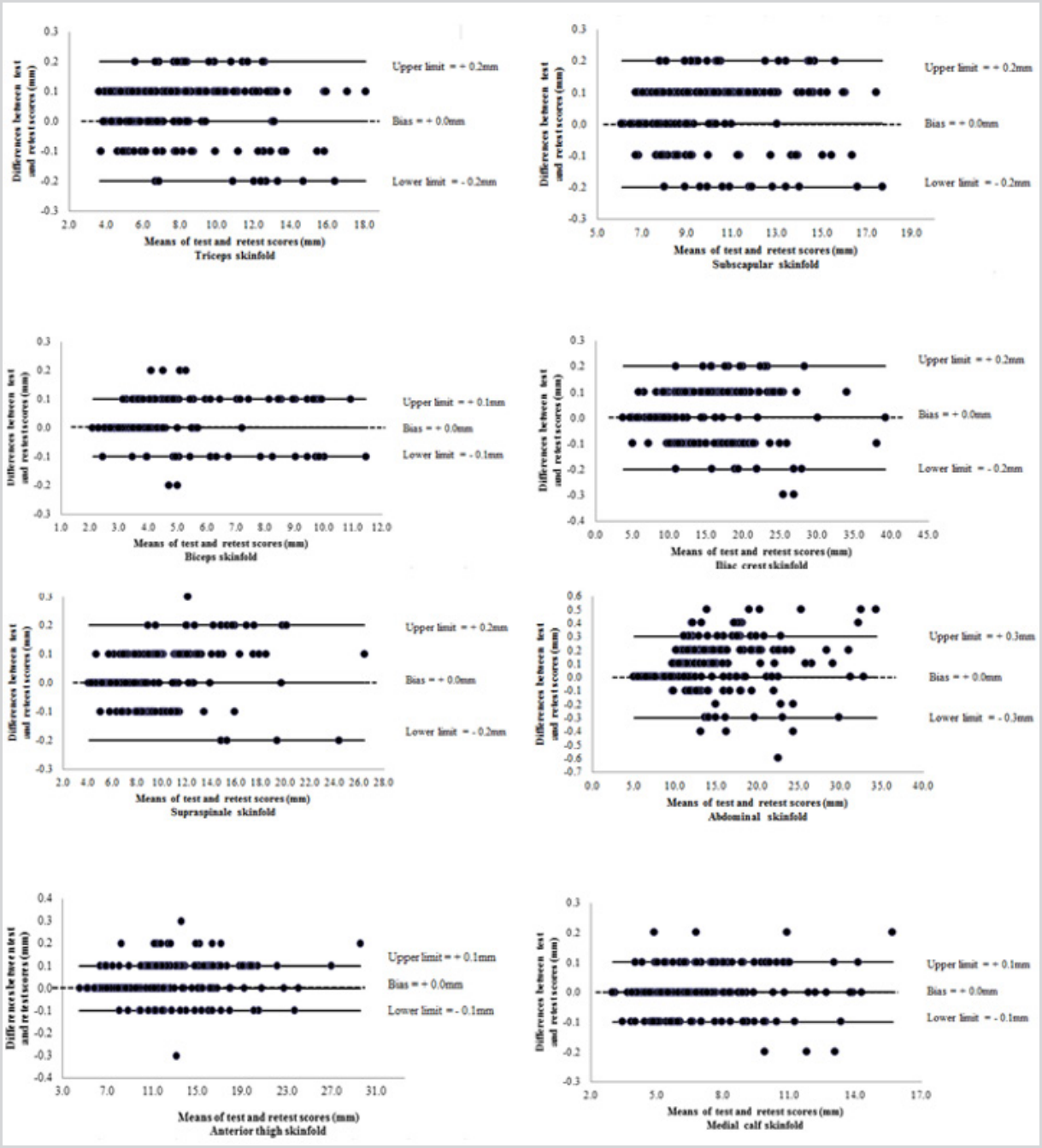

Figure 1: Bland and Altman plots summarising the 95% limits of agreement for the reliability of Triceps, Subscapular, Biceps, Iliac Crest, Supraspinale, Abdominal, Anterior Thigh and Medial Calf skinfolds (mm).

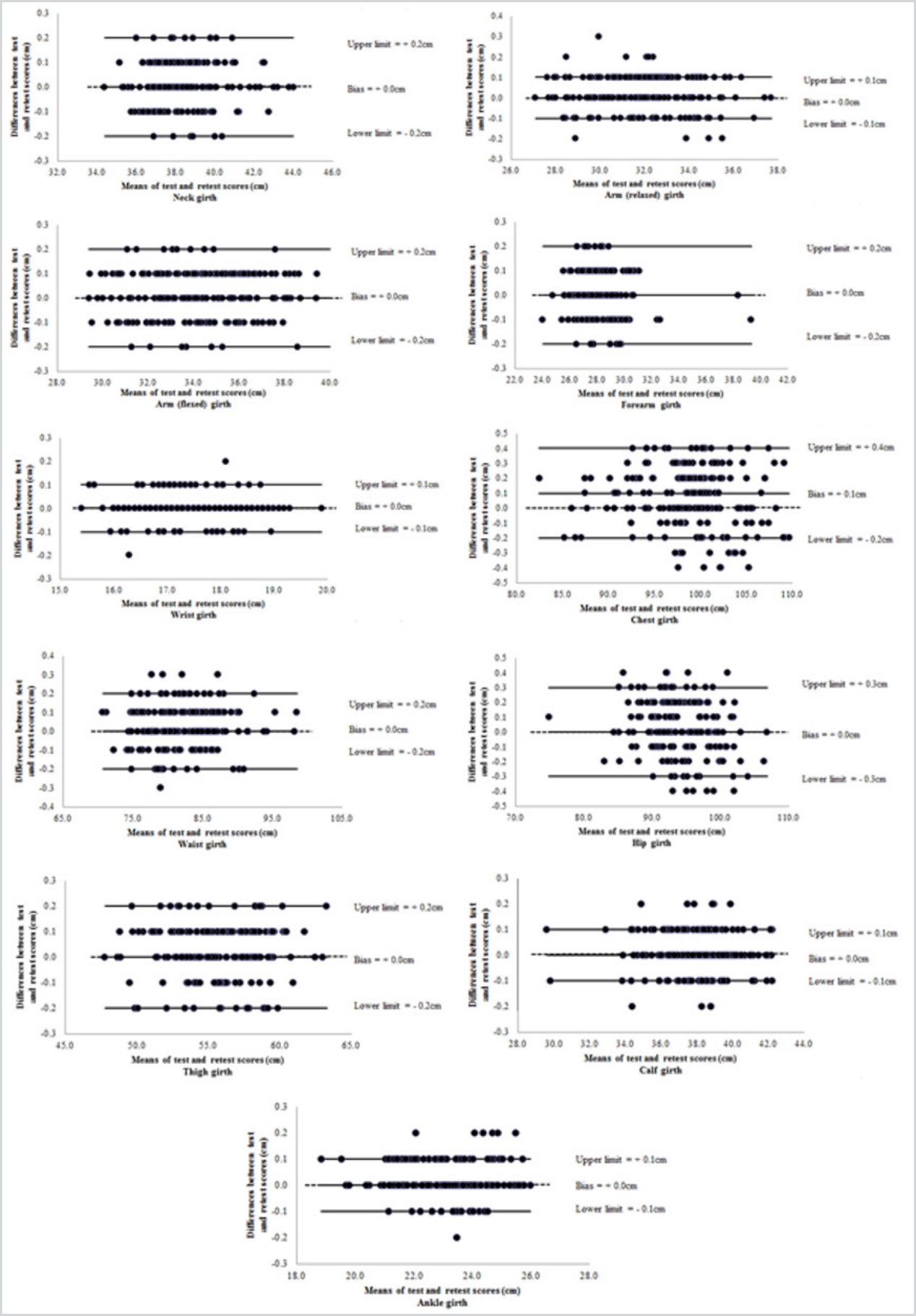

Figure 2: Bland and Altman plots summarising the 95% limits of agreement for the reliability of Neck, Arm (relaxed), Arm (flexed), Forearm, Thigh, Calf and Ankle girths (cm).

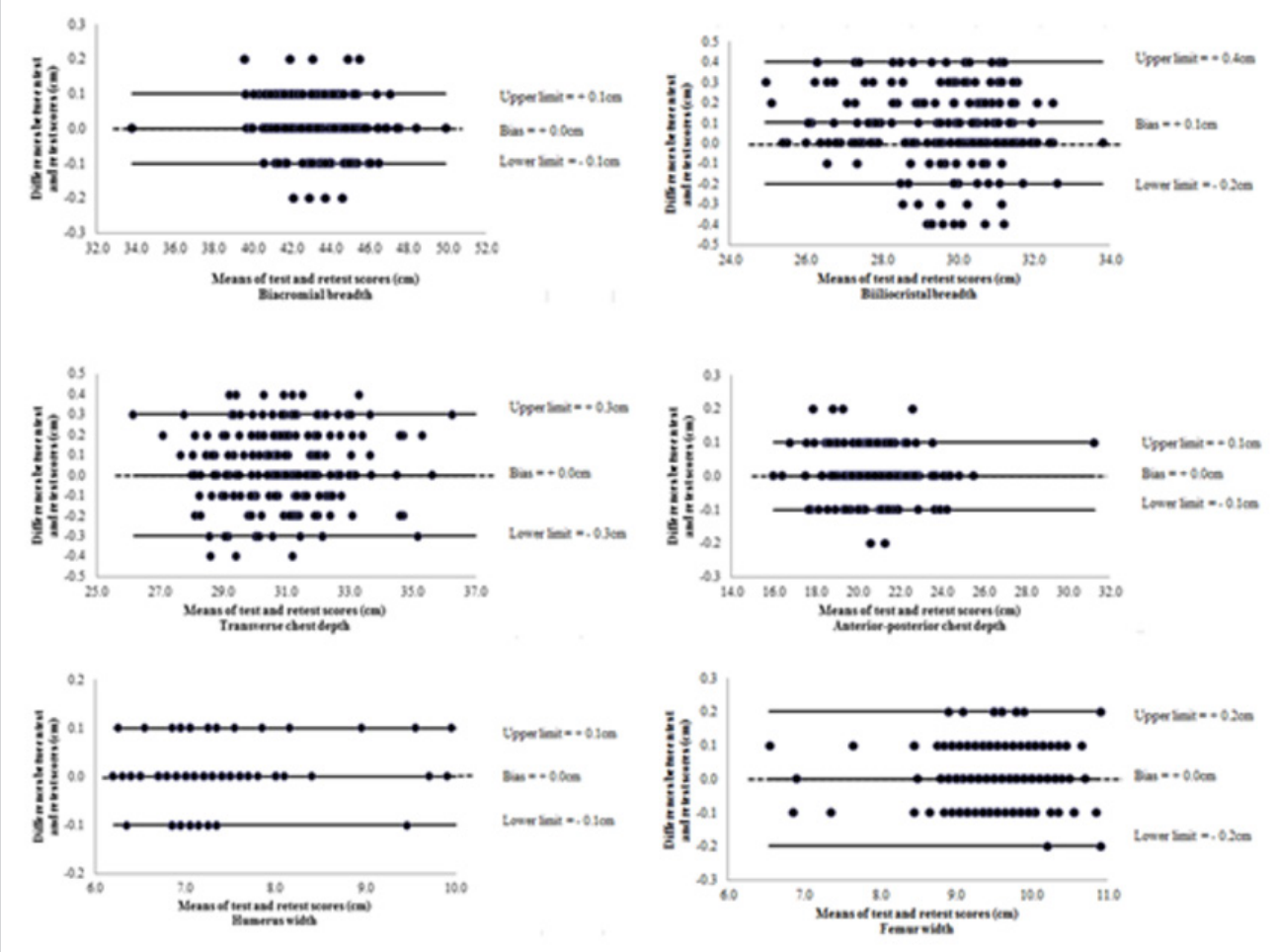

Figure 3: Bland and Altman plots summarising the 95% limits of agreement for the reliability of Biacromial and Biiliocristal Breadth, Transverse Chest and Anterior-Posterior Chest depth, Humerus and Femur width (cm).

Investigations were needed to establish heteroscedastic errors from all anthropometric measures. The extent of heteroscedasticity with r values of -0.185 and -0.073 and R_2 (%) coefficients of 3.4% and 0.5% for stretched stature and sitting height respectively. Both measures indicated statistical significance of P= <0.001 and narrow deviations from the line of best fit, suggesting very little evidence of heteroscedasticity. The triceps, subscapular, biceps and iliac crest skinfolds indicated r values of 0.331, 0.366, 0.306 and 0.500 and R_2 (%) coefficients of 10.9%, 13.9%, 30.6% and 21.2% and indicated statistical significance of *P*=< 0.001 (Table 2).

Measures for the supraspinale, abdominal, anterior thigh and medial calf skinfolds illustrated r values of 0.611, 0.483, 0.400 and 0.304 and R_2 (%) coefficients of 37.3%, 23.3%, 14.7% and 9.2% respectively (*P*=<0.001). As to be expected the triceps, subscapular, biceps, iliac crest and in particular the abdominal skinfold illustrated evidence of heteroscedasticity as there is more variance in the data values and greater error between test and re-test values with some deviations from the line of best fit (Table 2).

At this stage a decision was needed whether to log transform or not. The simple interpretation of these plots indicated that there was no issue as the values where within the 95% confidence limits between the two repeated measures and illustrated statistical significance of *P*=< 0.001> Therefore, on balance, log transformation was not warranted due to statistical analyses demonstrated excellent reliability and to be of practical use (Table 2).

DISCUSSION

It is well documented of the importance of reliability and measurement error. Indeed, better reliability suggests better precision, although within the area of body composition, it is not uncommon to encounter such random variation. These variations include anything from equipment calibration to technical execution and repeatability. In order for this study to be as applied as possible, confidence in reliability judgements on *n*=27 anthropometric measures by applying TEM% and the limits of agreement method was crucial so that sound practical judgements about professional football players' body composition could be made. The primary investigator achieved highly reliable TEM% standard values against those of the ISAK level 1 criterion indicated as high as 6.5% and as low as 1.0% differences and against TEM% level 2/3 criterion with differences as high as 3.8% and as low as 0.6%. Overall, the primary investigator achieved reliable TEM% values, indicating measurement precision and competency for *n*=27 variables anthropometric measurements well within acceptable ISAK TEM% targets Stewart et al. [35]. Statistical analysis determined via Bland and Altman's 95% limits of agreement method was used and a priori criterion set at ±3.8% as acceptable limits to determine the bias and random variation of *n*=27 anthropometric measures (Marfell-Jones, 2013 (personal communication)). Anthropometric measures illustrated obvious differences between the test and re-test values, but all were found to be normally distributed, with some evidence of systematic bias and random variation [36,37].

Table 2: Overview of R_2, *r* and *P* values for all *n*=27 anthropometric measures.

Variables	*r*	R_2 (%)	*P*
Stretched stature (cm)	-0.185	3.4	0.001
Sitting height (cm)	0.073	0.5	0.001
Skinfolds (mm)			
Triceps	0.331	10.9	0.001
Subscapular	0.366	13.9	0.001
Biceps	0.306	30.6	0.001
Iliac crest	0.5	21.2	0.001
Supraspinale	0.611	37.3	0.001
Abdominal	0.483	23.3	0.001
Anterior thigh	0.4	14.7	0.001
Medial calf	0.304	9.2	0.001
Girths (cm)			
Neck	-0.365	0.1	0.001
Arm (relaxed)	0.1	0.3	0.001
Arm (flexed)	-0.05	0.2	0.001
Forearm	0.06	0.3	0.001
Wrist	-0.126	0.1	0.001
Chest	0.09	0.8	0.001
Waist	-0.053	0.2	0.001
Hip	-0.007	0.000005	0.001
Thigh	0.017	0.3	0.001
Calf	-0.138	1.9	0.001
Ankle	0.085	0.7	0.001
Breadths (cm)			
Biacromial	-0.045	0.2	0.001
Biiliocristal	-0.047	0.2	0.001
Depths (cm)			
Transverse chest	-0.022	0.5	0.001
Anterior-posterior chest	-0.003	0.005	0.001
Widths (cm)			
Humerus	0.166	2.7	0.001
Femur	0.003	0.000006	0.001

Bland and Altman 95% limits of agreement approaches were used to determine heteroscedasticity with the n=27 anthropometric measures. The contentious measurements of the iliac crest skinfold, supraspinale skinfold, anterior thigh skinfold, chest girth, hip girth and biiliocristal breadth suggested possible heteroscedasticity issues with highest r values of 0.611 and R_2 (%) coefficients of 37.3%.

However, when investigated further, these plots did not exhibit heteroscedasticity because there is equal residual variance about the range of the values and were statistically significant (*P*=< 0.001). As a consequence, there was no need to find a cause of heteroscedasticity and resolve it by log transformation [38]. Overall, the primary investigator established accuracy and reliability of all *n*=27 anthropometric measures, and provided a sound foundation on which to develop practical and novel calibration models to estimate whole body density (g ml^{-1}) of professional football players Overall, the primary investigator established accuracy and reliability of all *n*=27 anthropometric measures, and provided a sound foundation on which to develop practical calibration models to estimate whole body density (g ml^{-1}) of professional football players and subsequently sound body composition judgements [39].

REFERENCES

1. Perini TA, Lameira de Oliveira G, Ornellas JdS, Palha de Oliveira F (2005) Technical error of measurement in anthropometry. Revista Brasileira de Medicina do Esporte 11(1): 86-90.
2. Atkinson G, Nevill AM (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Medicine 26(4): 217-238.
3. Hopkins WG (2000) Measures of reliability in sports medicine and science. Sports Med 30(1):1-15.
4. Hume P, Marfell JM (2008) The importance of accurate site location for skinfold measurement. J Sport Sci 6(12): 1333-1340.
5. Norton K (2002) Anthropometric estimation of body fat. In K Norton, T Olds (eds.) Anthropometrica. UNSW Press, Sydney, Australia.
6. British Standards Institution (1987) Precision of test methods part 1: Guide for the determination of repeatability for a standard test method by inter-laboratory tests. BS5497. British Standards Institution, London.
7. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 8(1): 307-310.
8. Atkinson G (2003) What is this thing called measurement error? In T Reilly M, Marfell Jones (Eds.). Kinanthropometry VIII: Proceedings of the 8th International Conference of the International Society for the Advancement of Ki anthropometry. Taylor and Francis, London.
9. Thomas JR, Nelson JK (2005) Research methods in physical activity. 5th Edition. Human Kinetics.
10. Stewart AD, Marfell Jones M, Olds T, Hans De Ridder J (2011) International Standards for Anthropometric Assessment. International Standards for Anthropometric Assessment. Lower Hutt, New Zealand.
11. Klipstein Grobusch, Geor T, Boeing H (1997) Interviewer variability in anthropometric measurements and estimates of body composition. Int J Epide 26(1): S194-180.
12. ISAK (International Society for the Advancement of Kinanthropometry) (2001) International standards for anthropometric assessment. National Library of Australia, Sydney, Australia.
13. Nevill AM, Atkinson G (1997) Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science. British J Sports Med 31 (4): 314-318.
14. Bland JM, Altman DG (2003) Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol. 22(1): 85-93.
15. Mills CD, De Ste Croix M, James DVB (2017a) Agreement between methods to determine procedure for maximal exhalation during Hydrostatic Weighing: A Methodological Investigation. Sport Exerc Med Open J 3(1): 20-23.
16. Katch FI, Katch VL (1980) Measurement and prediction errors in body composition assessment and the search for the perfect prediction equation. Res Quart Exerc Sport 51(1): 249-260.
17. Mills CD, De Ste Croix M, James DVB (2017b) Agreement between methods to estimate residual gas volume: A Methodological Investigation. Sport Exerc Med Open J 3(2): 63-67.
18. Sinning WE (1975) Experiments and demonstrations in exercise physiology. WB Saunders Company, Philadelphia, USA.
19. Biaggi RR, Vollman MW, Nies MA, Brener CE, Flakoll PJ, et al. (1999) Comparison of air displacement plethysmography with hydrostatic weighing and bioelectrical impedance for the assessment of body composition in healthy adults. American Journal of Clinical Nutrition 69(5): 898-903.
20. Ross WD, Marfell Jones MJ (1991) Physiological testing of the high-performance athlete. In: Kinanthropometry. 2nd Edn. pp. 223-308.
21. Brodie DA (1988a) Techniques of measurements of body composition. Part I Sports Medicine 5(1): 11-40.
22. Harrison GG, Buskirk ER, Carter JEL, Johnson FE, Lohman TG, et al. (1991) Skinfold thicknesses and measurement technique. In: Anthropometric Standardization Reference Manual. Champaign, Human Kinetics.
23. Heyward VH (2000) Nutrition and athletic performance: A position of the American dietetic association, dietitians of Canada and the American College of Sports Medicine. J Amer Diet Assoc 100(12): 1543-1556.
24. Casajūs JA, Bosco SJ (2001) Seasonal variation in fitness variables in professional soccer players. Journal of Sports Medicine and Physical Fitness 41(4): 463-469.
25. Loucks AB (2004) Energy balance and body composition in sports and exercise. J Sport Sci 22(1): 1-14.
26. Pesola GR, Sunmonu Y, Huggins G, Ford JG (2004) Measured diffusion capacity versus prediction equation estimates in blacks without lung disease. Respiration 71(5): 484-492.
27. Demura S, Yamaji S, Goshi F, Kobayashi H, Sato S, Nagasawa Y (2002) The validity and reliability of relative body fat estimates and the construction of new prediction equations for young Japanese adult males. J Sport Sci 20(2): 153-164.
28. Schutte JE, Townsend EJ, Hugg J, Shoup RF, Malina RM, et al. (1984) Density of lean body mass is greater in blacks than in whites. J App Physi. 56(6): 1647-1649.
29. Donnelly PM, Yang TS, Peat JK, Woolcock AJ (1991) What factors explain racial differences in lung volumes? The European Respiratory Journal. J Europ Soc Clini Respi Physi 4(7): 829-838.
30. Fields DA, Hunter GR, Goran MI (2000) Validation of the bod pod with hydrostatic weighing: influence of body clothing. Int J Obes Rel Meta Dis 24(2): 200-205.
31. Utter AC, Goss FL, Swan PD, Harris GS, Robertson RJ, et al. (2003) Evaluation of air displacement for assessing body composition of collegiate wrestlers. Med Sci Sport Exerc 35(3): 500-505.
32. Collins AL, Saunders S, Mc Carthy HD, Williams JE, Fuller NJ, et al. (2004) Within and between laboratory precision in the measurement of body volume using air displacement plethysmography and its effect on body composition assessment. Inter J Obes Rel Meta Dis 28 (1): 80-90.
33. Demura S, Yamaji S, Kitabayashi T (2006) Residual volume on land and when immersed in water: Effect on percent body fat. J Sport Sci 24(8): 825-833.
34. Slater GJ, Duthie GM, Pyne DB, Hopkins WG (2006) Validation of a skin fold-based index for tracking proportional changes in lean mass. British J Sport Med 40(3): 208-213.
35. Stewart AD, Sutton L (2012) Body Composition in Sport, Exercise and Health. Routledge.
36. Baumgartner TA (1989) Norm Referenced Measurement: Reliability. In MJ Safrit, TM Wood (Eds.). Measurement concepts in physical education and exercise science. Champaign, Human Kinetics, Illinois, USA.
37. Bryman A, Cramer D (2009) Quantitative Data Analysis with SPSS 14, 15 & 16. A Guide for Social Scientists, Routledge, New York, USA.
38. Dempster P, Aitkens S (1995) A new air displacement method for the determination of human body composition. Med Sci Sport Exer 27(12): 1692-1697.
39. Marfell Jones MJ, Provyn S, Clarys JP (2007) Direct depth measurement of subcutaneous adipose tissue. In: Kinanthropometry X. Proceedings of the 10th International Society for the Advancement of Kinanthropometry Conference, Held in conjunction with the 13th Commonwealth International Sport Conference. University of South Australia: Taylor & Francis Ltd, Australia.