CRITICAL ANALYSIS OF THE RELIABILITY AND PREDICTION ERROR OF A NOVEL SMARTPHONE APPLICATION FOR MEASURING BODY COMPOSITION

by

CASEY METOYER

MICHAEL R. ESCO, COMMITTEE CHAIR
MICHAEL V. FEDEWA, COMMITTEE CO-CHAIR
LEE J. WINCHESTER
MARK T. RICHARDSON
JASON M. PARTON

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Kinesiology
in the Graduate School of Education
The University of Alabama

TUSCALOOSA, ALABAMA

2023

ABSTRACT

A smartphone application was recently developed that measures body composition from a 2-dimensional image of a standing person (%$\text{Fat}_{\text{IMAGE}}$). Three experiments were performed to 1) determine the reliability of %$\text{Fat}_{\text{IMAGE}}$ in two different body positions and days, 2) determine the reliability of %$\text{Fat}_{\text{IMAGE}}$ across different cameras, room lighting levels, and colored backgrounds, and 3) assess the potential sources of error related to %Fat estimation with bioimpedance, skinfold, and IMAGE. In the first study, %$\text{Fat}_{\text{IMAGE}}$ was measured in the anterior (ANT_{Day1}) and posterior ($\text{POST}_{\text{Day1}}$) positions, then again within 48 hours of the first visit ($\text{POST}_{\text{Day2}}$). The ANT_{Day1} was significantly lower compared to the reference, whereas the $\text{POST}_{\text{Day2}}$ was not. In the second study, %$\text{Fat}_{\text{IMAGE}}$ was estimated in different megapixel cameras (12MP, 8MP, 5MP, and 0.7MP), different room lighting conditions (i.e., low=LL, medium=ML, ambient=AL, and bright light=BL), and different colored backgrounds (White=WB, Black=BB, Green=GnB, Orange=OB, and Gray=GyB). Conditions were compared to a reference (8MP, AL, and WB). The 0.7MP was significantly higher than the reference, whereas the 12MP and 5MP were not. The LL was significantly higher than the reference, whereas the BL and ML were not. The BB, GnB, and GyB were significantly higher than the WB, whereas the OB was not. In a third study, %Fat was measured via bioimpedance, skinfold, %$\text{Fat}_{\text{IMAGE}}$, and a criterion 4-compartment model. The constant error (CE) was calculated (%Fat condition – criterion) for each method ($\text{SKF}_{\text{ERROR}}$, $\text{BIA}_{\text{ERROR.}}$, and $\text{IMAGE}_{\text{ERROR}}$). The CE was correlated with markers of physical activity (PA) and muscular fitness (MF): the International physical activity questionnaire (IPAQT, IPAQV, IPAQM); Perceived functional ability questionnaire ($\text{1mile}_{\text{SUB}}$

and $3\text{mile}_{\text{SUB}}$); Physical Activity Rating questionnaire (PA-R): handgrip test (HG); and push-up test (PU). $\text{SKF}_{\text{ERROR}}$ displayed a *small* correlation with PU, while IPAQV, $1\text{mile}_{\text{SUB}}$, $3\text{mile}_{\text{SUB}}$, PA-R, and PU were associated with $\text{IMAGE}_{\text{ERROR}}$. Regression analyses indicated that only PU contributed to the model for $\text{SKF}_{\text{ERROR}}$ and $\text{IMAGE}_{\text{ERROR}}$. In conclusion, the IMAGE produced acceptable reliability across different camera, lighting, and background conditions but attention should be paid to clothing color in contrast with background color. In addition, further evaluation of MF measures is recommended, as higher PU were associated with greater $\text{IMAGE}_{\text{ERROR}}$.

DEDICATION

To my family, from my grandparents who illustrated the hard work and faith necessary to do great things. To my parents, my mother whose positivity and free spirit reminds you that every day is worth living and my father who has showed me that determination and perseverance are all one needs in life to succeed. To my brothers and their families, being defining examples of tradition, support, and belief. To my wife, the constant love, support, and honesty transcends every hardship. Lastly, to my daughter and future children, may this labor of love lay the path to your future achievements. Thank you all.

LIST OF ABBREVIATIONS AND SYMBOLS

%Fat	Percent body fat
FFM	Fat-free mass
FM	Fat mass
2C	Two-compartment
3C	Three-compartment
4C	Four-compartment
DXA	Dual energy X-ray absorptiometry
BIA	Bioimpedance analysis
BIS	Bioimpedance spectroscopy
d	Cohen's d effect size
ICC	Intraclass correlation coefficient
LOA	Limits of agreement
Mo	Total body bone mineral
R	Resistance
RLV	Residual lung volume

TBW	Total body water
Xc	Reactance
Z	Impedance
ANT_{Day1}	Anterior condition
$\text{POST}_{\text{Day2}}$	Different day condition
12MP	12-megapixel camera condition
8MP	8-megapixel camera condition
5MP	5-megapixel camera condition
0.7MP	0.7-megapixel camera condition
LL	Low light condition
ML	Medium light condition
AL	Ambient light condition
BL	Bright light condition
WB	White background condition
BB	Black background condition
GnB	Green background condition
OB	Orange background condition
GyB	Gray background condition

ACKNOWLEDGEMENTS

There are countless numbers of people who have contributed to the completion of this document, both directly and indirectly. I would like to thank my co-chair and faculty advisor Dr. Michael Esco, who encouraged this meathead to not just pursue a doctorate degree but to realize the doors it can open for me. His guidance, enthusiasm, and willingness to pursue any project has been a strong example that has helped me through this program. I would also like to thank Dr. Michael Fedewa who has helped me develop as a researcher, critical thinker, and technical writer. In addition, his constant positive support has helped me through this program, especially on sleepless nights after my daughter was born. To Dr. Winchester, my understanding of muscle physiology and biochemistry would not be the same without you. I will forever remember the propagation of an action potential through the neuromuscular junction and into the muscle because of your lectures. Your humor is unmatched. Dr. Richardson thank you for bringing your wisdom and perspective to research methods and my dissertation, it is greatly valued and respected. Dr. Parton thank you for the reminder that you can be a gym buddy as well as a mentor and very well-respected colleague in the field. Your encouragement and reminder that the journey has just begun is now being realized and is well received. I am also extremely appreciative of the lab experiences in which I was able to take part, as they exposed me to a variety of techniques and procedures that I would not have otherwise learned. Last, I am grateful for the friends and colleagues I have made during my time studying at Alabama. Early morning data collection, brainstorming, office discussions, abstract deadlines, and project collaboration

are just some of the memories that I will take with me as I move forward in my career and in life.

Thank you to everyone for making my time here an enriching and amazing experience

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:

INTRODUCTION

Obesity is a chronic condition characterized by the excessive accumulation of body fat, resulting in various metabolic, physiologic, and functional impairments. Though it is considered a disease, its presence increases risk of developing other chronic diseases such as cardiovascular, cancer, stroke, and type 2 diabetes (1-4). In recent years, the prevalence of obesity has increased dramatically, as the World Health Organization reported a 3-fold increase in the prevalence of obesity in adults over the last four decades (5). These numbers are alarming, especially considering that the COVID-19 pandemic has brought a variety of challenges related to the management of obesity. For instance, the shift toward virtual health and telemedical practices have become commonplace for countering obesity in current times.

Standard techniques for evaluating the effects of obesity, such as body mass index (BMI), waist circumference, and waist-to-hip ratio, are limited in their ability to measure fat- (FM) and fat-free masses (FFM), as well as relative adiposity (%Fat). Therefore, it is recommended that practitioners assess body composition rather than weight or girth alone (6, 7). Body composition is defined as the relative proportion of FM and FFM within the body and it is generally modeled based on those 2-compartments (8, 9). For the most accurate measurement of body composition multi-compartment models are required. However, this requires the use of sophisticated techniques such as Dual-energy x-ray absorptiometry (DXA) which is considered highly accurate, but it is relatively expensive, immobile, and exposes patients to a small amount of radiation (10). In addition, equipment required to measure body composition using DXA is typically found in radiology clinics or research laboratories, requiring trained technicians to perform the assessment. While field tools are available in mainstream health and fitness settings (i.e., skinfolds and bioimpedance), fitness professionals are often required to administer the body composition assessment. Therefore, there is a need to develop a method that will allow individuals to accurately self-administer a body composition assessment, without the need of a trained professional.

Recently, a smartphone application was developed that can estimate body composition using an automated analysis program applied to a full-body 2-dimensional image of a person standing (IMAGE) (United States Provisional Patent 62/842,826). To develop the application, total body images from DXA were taken from 188 participants scanned in the University of Alabama, Exercise Physiology Laboratory. From the images, digital measurements were taken at various anatomical locations (e.g., widest and narrowest portions of the neck, shoulders, trunk, hips, thigh, etc.) as shown in Figure 1.1. These measurements were analyzed with statistical regression equations to create algorithms that estimate total %Fat, as well as android and gynoid adiposity.

Previous validation research involving the validity of the IMAGE as an application for measuring body volume, and subsequently body composition within a 3-compartment (3C) model in 67 adult men and women (11). Results indicated that body volume measures from the IMAGE showed a *nearly perfect* correlation ($r = 0.99$, $p < 0.001$) when compared to a hydrostatic weighing criterion. When calculating %Fat via 3C model utilizing the BV from IMAGE, there a *nearly perfect* correlation ($r = 0.96$, $p < 0.001$) between the IMAGE based 3C model (21.08 ± 7.04 %Fat) and the 3C criterion (21.01 ± 7.30 %Fat) with no significant differences ($p = 0.775$) observed between methods. A subsequent study examined the validity of an IMAGE-based field 3C model which combined body volume measures from the IMAGE and total body water estimates from bioimpedance, to DXA, as well as a criterion 5-compartment (5C) model in 57 adult men and women (12). The results showed that the IMAGE-based field 3C model provided a stronger correlation ($r = 0.94$, $p < 0.001$) than DXA ($r = 0.87$, $p < 0.001$), along with smaller observed error (2.22 vs. 4.64, $p < 0.001$) when both were compared to the criterion 5C model (12). Sullivan et al. (13), also demonstrated that the IMAGE provided comparable body composition estimates (19.46 ± 6.08) as DXA (18.04 ± 6.39) when both were included in multi-compartment models ($r = 0.94$ and 0.88, respectively) in 48 adult men and women. The IMAGE system has also been used as a measure of athletic performance, as demonstrated in a recent study which assessed the relationship between performance and body composition 19 collegiate female rowers (14). Indeed, 2-kilometer rowing performance was inversely associated with FM ($r = -0.56$, $p < 0.05$) and positively associated with FFM ($r = 0.67$, $p < 0.05$) derived from the IMAGE, such that lower FM and higher FFM were associated with better performance (14).

This approach is possibly the simplest body composition prediction method available, as the automated analysis program only requires a single full-body 2-dimensional image. Thus, it has the potential to be widely used in numerous settings. For instance, practitioners could potentially use the IMAGE method to easily perform body composition assessments in clinical settings using a smartphone or tablet. In addition, the IMAGE method could potentially be used for virtual/telemedical health as remote body composition technique.

However, previous research has focused primarily on the validation of the technique in a tightly controlled laboratory setting, with minimal deviation in the IMAGE conditions. Further, the co-factors that may contribute to the observed error or bias of the IMAGE method when estimating body composition have yet to be explored. Therefore, the purpose of this dissertation was to critically examine the reliability of a 2D image smartphone application for the measurement of %Fat and identify co-factors that are associated with the observed error.

Study 1: To assess the reliability of the IMAGE for the assessment of %Fat across two different body positions (ANT vs POST) and days (DAY1 vs DAY2). We hypothesized, due to the simplicity of the application and previous research that established the validity of the method, there would be no significant differences between body positions or days.

Study 2: To assess the reliability of the IMAGE for the assessment of %Fat across the following independent conditions: 1) different megapixel cameras; 2) different room lighting; and 3) different colored backgrounds. We hypothesized that no significant mean differences would be observed between camera resolution, lighting, or background color. In addition, each condition would yield excellent reliability statistics when compared to the criterion condition.

Study 3: To determine the extent to which individual error associated with BIA, SKF, and IMAGE relates to selected markers of physical activity and muscular fitness. The associated error of the three field prediction methods was determined by comparison to a laboratory 4-compartment model. It was hypothesized that the error for each body composition field predictor would be significantly correlated with each marker of physical activity and fitness.

REFERENCES

1. Garanty-Bogacka B, Syrenicz M, Goral J, Krupa B, Syrenicz J, Walczak M, et al. Changes in inflammatory biomarkers after successful lifestyle intervention in obese children. Endokrynologia Polska. 2011;62(6):499-505.

2. Goodpaster BH, Krishnaswami S, Harris TB, Katsiaras A, Kritchevsky SB, Simonsick EM, et al. Obesity, regional body fat distribution, and the metabolic syndrome in older men and women. Archives of internal medicine. 2005;165(7):777-83.

3. Scheuer SH, Færch K, Philipsen A, Jørgensen ME, Johansen NB, Carstensen B, et al. Abdominal fat distribution and cardiovascular risk in men and women with different levels of glucose tolerance. The Journal of Clinical Endocrinology & Metabolism. 2015;100(9):3340-7.

4. Jastreboff AM, Kotz CM, Kahan S, Kelly AS, Heymsfield SB. Obesity as a disease: the obesity society 2018 position statement. Obesity. 2019;27(1):7-9.

5. Nittari G, Scuri S, Petrelli F, Pirillo I, Di Luca N, Grappasonni I. Fighting obesity in children from European world health organization member states. Epidemiological data, medicalsocial aspects, and prevention programs. La Clinica Terapeutica. 2019;170(3):e223-e30.

6. Kok P, Seidell J, Meinders A. The value and limitations of the body mass index (BMI) in the assessment of the health risks of overweight and obesity. Nederlands tijdschrift voor geneeskunde. 2004;148(48):2379-82.

7. Nevill AM, Stewart AD, Olds T, Holder R. Relationship between adiposity and body size reveals limitations of BMI. American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists. 2006;129(1):151-6.

8. Behnke AR, Feen B, Welham W. The specific gravity of healthy men: body weight÷ volume as an index of obesity. Journal of the American Medical Association. 1942;118(7):495-8.

9. Burkinshaw L. Measurement of human body composition in vivo. Progress in medical radiation physics: Springer; 1985. p. 113-37.

10. Heyward V. ASEP methods recommendation: body composition assessment. Journal of Exercise Physiology Online. 2001;4(4).

11. Fedewa MV, Sullivan K, Hornikel B, Holmes CJ, Metoyer CJ, Esco MR. Accuracy of a Mobile 2D Imaging System for Body Volume and Subsequent Composition Estimates in a Three-Compartment Model. Med Sci Sports Exerc. 2021;53(5):1003-9.

12. Sullivan K, Metoyer CJ, Hornikel B, Holmes CJ, Nickerson BS, Esco MR, et al. Agreement Between A 2-Dimensional Digital Image-Based 3-Compartment Body Composition Model and Dual Energy X-Ray Absorptiometry for The Estimation of Relative Adiposity. Journal of Clinical Densitometry. 2022;25(2):244-51.

13. Sullivan K, Hornikel B, Holmes CJ, Esco MR, Fedewa MV. Validity of a 3-compartment body composition model using body volume derived from a novel 2-dimensional image analysis program. European Journal of Clinical Nutrition. 2021:1-8.

14. Esco MR, Holmes CJ, Sullivan K, Hornikel B, Fedewa MV. Utilizing a Novel 2D Image Processing System for Relating Body Composition Metrics to Performance in Collegiate Female Rowers. International Journal of Environmental Research and Public Health. 2021;18(5):2413.

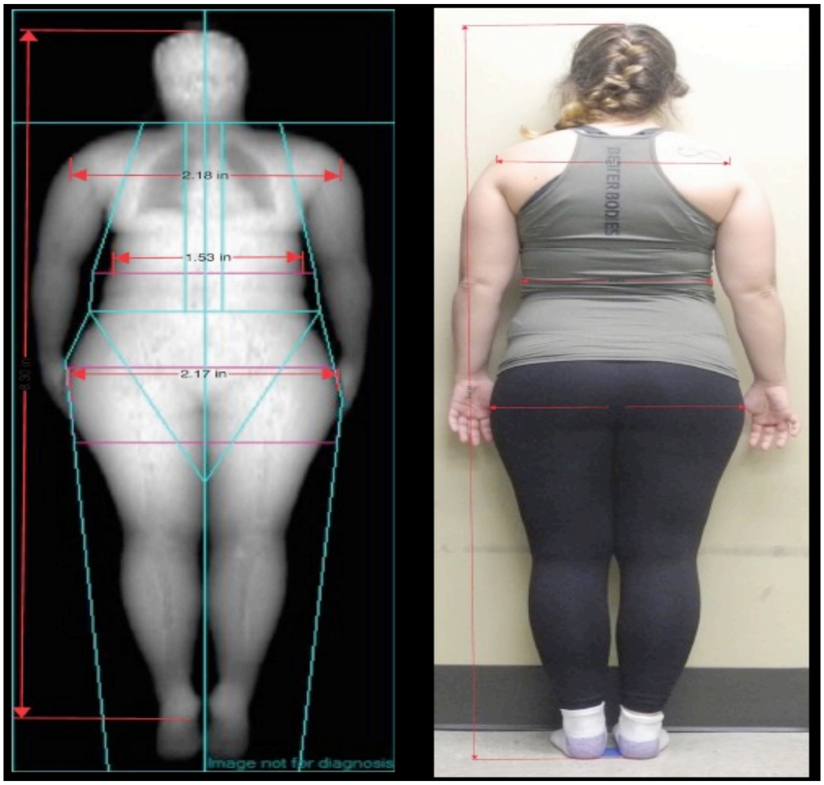


Figure 1.1. A participant within the original validation sample that was used to formulate the proprietary algorithms. The red lines in the image provide examples of the width measurements used within the algorithm.

CHAPTER 2:

RELIABILITY OF A SMARTPHONE-BASED BODY COMPOSITION APPLICATION: COMPARISON OF TWO DIFFERENT BODY POSITIONS AND DAYS

ABSTRACT

The purpose of this study was to assess the reliability of a smartphone-based application (IMAGE) for the assessment of %Fat across two different body positions ($\mathrm{ANT}_{\mathrm{Day1}}$) versus posterior ($\mathrm{POST}_{\mathrm{Day1}}$) and days ($\mathrm{POST}_{\mathrm{Day1}}$ vs $\mathrm{POST}_{\mathrm{Day2}}$) separated by 48 hours. Paired samples T-tests were used to analyze mean differences between the experimental conditions and the reference. The reference photo was completed on Day 1 with participants in the posterior position (i.e., facing away from the camera). Estimates from the $\mathrm{ANT}_{\mathrm{Day1}}$ (26.44 ± 6.19 %Fat) were considered significantly lower when compared to the reference (26.84 ± 6.07 %Fat, $p = 0.04$). However, the effect size (ES) of the difference was considered *trivial,* with *excellent* intraclass correlation coefficient (ICC = 0.992, $p < 0.001$). In addition, %Fat estimates obtained on Day 2 (26.65 ± 6.17 %Fat) were not different when compared to the reference (26.84 ± 6.07 %Fat, $p = 0.31$), with *trivial* differences between days. The ICC value (ICC = 0.993, $p < 0.001$) was also qualified as *excellent.* The overall findings of the study indicated that the IMAGE appears to produce reliable %Fat results when participants are placed in the anterior position and tested within 48-hours due to the trivial mean differences, excellent ICC values of over 0.99, and tight limits of agreement.

KEYWORDS: reliability; body composition; percent fat; 2D image; smartphone application

INTRODUCTION

Body composition is defined as the relative proportion of the tissues within the body (1, 2) and is commonly expressed as body fat percentage (%Fat), fat mass (FM) and fat-free mass (FFM). Furthermore, body composition is an important parameter of physical fitness that is strongly related to the development of hypokinetic diseases (3). For instance, the rapid rise in metabolic syndrome is a direct result of world-wide increases in adiposity (4, 5). This increase has been linked to a rise in cardiovascular disease and other degenerative conditions (i.e., Alzheimer's and Parkinson's) (5, 6). Conversely, maintenance of FFM in older adults appear to have positive longevity and health related outcomes (7). Thus, a valid, reliable, and accessible method for routine measurement of body composition in clinical and practical settings is critical.

While validity is important in the evaluation of any measurement technique, reliability is of equal importance. Reliability is an essential aspect of any techniques performance as it provides information on the consistency of the results over time and across different conditions. Bioelectrical impedance analysis (BIA) and skinfolds (SKF) are two of the most commonly used methods for measuring %Fat in the field. Though these methods have shown to be valid measures of body composition for practical purposes, concerns with reliability exist. For instance, reliability of skinfold measurement is highly dependent on technician skill level and site location (Heyward). In test-retest scenarios for %Fat measurement, skinfolds have displayed an error up to 18%. Therefore, examining the test-retest reliability of any body composition measurement tool is quintessential to determining its effectiveness.

Recently a smartphone application (IMAGE) that estimates body composition from a single two-dimensional image utilizing automated built-in algorithms has been developed. Evidence supports the validity of the IMAGE application when compared to laboratory-based criterions (8-10). For instance, the IMAGE displayed *nearly perfect* correlation with hydrostatic weighing when estimating body volume ($r=0.99$, $p < 0.001$) in a sample of 67 young adults (8). When IMAGE-based body volume was incorporated as a %Fat prediction variable within a 3-compartment (3C) model, the estimated value displayed *excellent* agreement with the laboratory 3C criterion ($r=0.96$, $p < 0.001$)(8). In addition, Sullivan et al. (10) reported an IMAGE based 3C model performed slightly better than a DXA 3C model when compared to a 5-compartment model criterion for estimating %Fat in 57 males and females. Also, FM and FFM measured using the IMAGE technique was also associated with to rowing performance in a sample of 19 collegiate athletes (11).

Previous validation research examining the agreement between the IMAGE technique and other measures of body composition were limited to single trials. In addition, the picture for the IMAGE was consistently taken with participants in the posterior position (i.e., facing away from the camera) (8-11). However, the IMAGE technique was designed for personal use, which may require testing across multiple days and with users facing the camera. To our knowledge, there have been no studies evaluating the reliability of the IMAGE across body positions and different days, which are fundamental conditions to the precision and performance of this technique. Therefore, the purpose of this study was to assess the reliability of the IMAGE for the assessment of %Fat across two different body positions (ANT vs POST) and days (DAY1 vs DAY2). Based upon the simplicity of the smartphone application, we hypothesized that no significant differences would be observed between body positions or days.

METHODS

Participants

Participants (n = 32, 68.7% female, 84.3% White/Caucasian) were recruited for this study. Descriptive characteristics of the study sample are presented in Table 1. Written informed consent was obtained from each participant before data collection. The Institutional Review Board at the University of Alabama approved the study protocol and data collection procedures.

Procedures

Body mass was measured, to the nearest 0.1 kg, with a calibrated digital scale (Tanita BWB-800, Tanita Corporation, Tokyo, Japan). Barefoot standing height was measured to the nearest 0.1 cm with a manual stadiometer (SECA 213, Seca Ltd., Hamburg, Germany). The recorded values for height and weight were manually entered into the smartphone application for each participant.

2D Image Processing System / Reference Condition

Participants wore form-fitting athletic apparel that allowed for the automated IMAGE analysis program to identify the required anatomical landmarks in each 2D image (8, 9, 11). Participants with long hair were instructed to pull their hair "back" and "up" to accurately measure the neck diameter in the digital image. Participants stood with their weight evenly distributed on both feet, in front of a white photography background, facing away from the digital camera. They were instructed to place their heels together with the feet pointed outward at a 60-degree angle (8, 9, 11). Participants were asked to remain still with arms abducted at a 45-degree angle away from the torso and aligned within the frontal plane, with palms facing away from the camera. Once correctly positioned, a digital image that comprised the head, feet, and arms of the participant was captured from the posterior view, utilizing a 12.9 inch, 64g iPad Air

2 and analyzed with a commercially available application (version 1.1.2, made Health and Fitness, USA. www.mymadeapp.com). The camera was set in a tripod, at approximately participant waist height, at a distance of 3.05 meters away from the participants. %Fat was derived from the IMAGE, using a proprietary algorithm that automatically identified and measured the horizontal linear diameter of various anatomical landmarks (United States Utility Patent 16/841,944) (8, 9, 11).

Posterior and Anterior Body Position 2D Images

After the reference image was taken in the posterior position, participants were asked to "turn around" and face the camera while a second image was taken of them in the same anatomical position (i.e., heels together, feet pointed outward at a 60-degree angle, arms abducted at a 45-degree angle from the torso, aligned within the frontal plane, with palms facing away from the camera) from the anterior view ($\mathrm{ANT}_{\mathrm{Day1}}$). There was no change in camera angle, light, or device.

Different Days

Participants were asked to return to the laboratory at a minimum of 24 hours later, all participants returned within 48 hours of their first testing day. Clothing and specific time of day were not controlled for. One single image from the posterior position was obtained using a 12.9 inch, 64g iPad Air 2 under the same conditions, within the laboratory ($\mathrm{POST}_{\mathrm{Day2}}$). The participants and examiners followed the same protocols outlined above for the reference condition.

Statistical Analyses

Statistical analyses were performed using SPSS for Mac (SPSS 26.0, Chicago, IL) and Microsoft Excel 365 (Microsoft Corp., Redmond, WA). All data were screened for missing or improbable values using descriptive and frequency statistics. Histograms, skewness, and kurtosis plots were used to identify potential outliers as well as normality of the data, a value of ±2 was used to identify outliers. No data was excluded in the analyses.

Paired samples T-tests were used to analyze mean differences between $\mathrm{ANT}_{\mathrm{Day1}}$ vs $\mathrm{POST}_{\mathrm{Day1}}$ and $\mathrm{POST}_{\mathrm{Day1}}$ vs $\mathrm{POST}_{\mathrm{Day2}}$. The constant error (CE) was calculated as “condition” – the reference and the 95% confidence interval (CI) was calculated as ± 1.96 Standard Deviation (SD). Effect size (ES) of the paired differences were assessed using Cohen’s *d* statistic. The ES of the mean comparisons were qualitatively described as follows: trivial (<0.2), small (0.2-0.6), moderate (0.6-1.2), large (1.2-2.0) or very large (>2.0) (12, 13). Intraclass correlation coefficients (ICCs) were also used to determine the level of agreement, with each of the independent conditions being compared to the CR position. An ICC of greater then 0.75 was considered “good” and greater than 0.9 considered “excellent” relative reliability (14). Statistical significance was assessed using an alpha level of $p < 0.05$.

RESULTS

A final sample of 32 participants was used for analysis and descriptive characteristics including M±SD are listed in Table 2.1. The sample was 68.7% female and age ranged from 20 to 52 years with 56.3% being between 20 and 30 years. The sample was 84.4% Caucasian/White, 9.4% multi-racial, 3.1% Hispanic, and 3.1% Other. The sample BMI ranged from 21.6 to 36.8 $\mathrm{kg/m^2}$, with 29.8% of the sample being between 21.6 and 24.9, classified as “Normal Weight”, 58.2% ranged from 25.1 to 29.2 $\mathrm{kg/m^2}$, classified as “Overweight”, and 12% were classified as “Obese” or “Extremely Obese”. Table 2.2 displays the reliability statistics for all comparisons, except for the limits of agreement which are displayed in Figures 2.1 and 2.2 and described below. The following subsections detail the results of each condition being independently compared to the same reference. As explained in the methods, the reference condition utilized participants being placed in the posterior position against a white background.

Different Body Position

The mean %Fat for ANT_{Day1} was considered significantly lower when compared to the criterion (26.44 ± 6.19 %Fat, $p = 0.04$). However, the ES of the difference was considered *trivial,* with *excellent* reliability between measures (ICC=0.992, $p < 0.001$) (Table 2.2). In addition, the CE was negligible, and indicated that the ANT_{Day1} underestimated %Fat (CE=-0.40 %Fat, 95% LOA 1.72 to -2.52 %Fat) (Figure 2.1).

Different Days

The mean %Fat for the $\text{POST}_{\text{Day2}}$ comparison was not considered statistically significantly different when compared to the criterion (26.65 ± 6.17 %Fat, $p = 0.31$) and the ES value was considered *trivial.* The ICC value (ICC=0.993, $p < 0.001$) was qualified as *excellent* (Table 2.2). The 95% LOA (CE ± 1.96 SD of residual scores) for %Fat ranged from 1.79% above to -2.17% below the CE of -0.19 %Fat (Figure 2.2).

DISCUSSION

The purpose of this study was to evaluate the reliability of IMAGE to measure %Fat across two different body positions and two days separated by no more than 48 hours. These results indicated that the IMAGE produced reliable %Fat results from images obtained while in the ANT_{Day1} and for the $\text{POST}_{\text{Day2}}$ conditions due to the trivial mean differences, excellent ICC values of over 0.99, and narrow limits of agreement.

Test-retest reliability is a commonly used evaluation for body composition measures (15, 16) to ensure a technique can consistently reproduce. While field-based measures such as skinfolds and bioelectrical impedance are valid measures of body composition, their reliability is often reported as inconsistent. For instance, Hume & Marfell-Jones (17) reported a greater than 10% difference in fold measurement (i.e., 8.5 to 7.5mm) between two trained technicians at three different anatomical locations (i.e., bicep, calf, and thigh). The researchers also assessed variations in site location, such that a 1-centimeter deviation, in any direction, from an established centralized site may cause up to an 18% difference in skinfold measurement (i.e., 6.8 to 46.9mm) (17). However, another study reported excellent test-retest reliability (ICC=0.99) and good inter-rater reliability (ICC=0.99), although they did note %Fat estimates with skinfold to be inconsistent and poor when compared to the air displacement plethysmography reference (18). Bioelectrical impedance analysis has displayed excellent test-retest reliability (ICC=0.98); however, it has been noted that strict parameters regarding hydration status and exercise must be followed to ensure accuracy and reliability (19-21). Research on a three-dimensional (3D) imaging device reported good test-retest reliability (ICC=0.98), however measurement error associated with image quality has been reported for the 3D scanner (18, 22). Bourgeois et al. (22) noted differences in how three different 3D scanners estimated body volume and reported that these estimates cannot be used to estimate %Fat via 2-compartment model. In addition, the researchers also noted error associated with anatomical landmarking particularly in individuals with high BMI, are very tall, or have poor balance (22). The present results indicate that the IMAGE application displays a similar level of test-retest reliability (ICC=0.99) as other field methods.

Previous research regarding the IMAGE used a posterior facing position (i.e., facing away from the camera), however the manufacturer settings suggested that it can be used with individual users facing the camera. The IMAGE technique showed excellent agreement, as a measure of BV, and when estimating %Fat in a 3C model compared to two different laboratory criterions. However, these were completed with participants in the posterior facing position (8-10). Thus, it was important to evaluate the reliability of the IMAGE with participants facing the camera, as it was intended to be used. The mean difference for the $\mathrm{ANT}_{\mathrm{Day1}}$ condition was considered significantly lower, however this difference was considered *trivial* and the ICC of 0.99 was considered *excellent*. Therefore, the IMAGE does appear to produce reliable %Fat measurements when users are facing the camera, and $\mathrm{ANT}_{\mathrm{Day1}}$ versus $\mathrm{POST}_{\mathrm{Day1}}$ body positions may be used interchangeably.

This study also sought to evaluate the day-to-day reliability of the IMAGE. The IMAGE technique, when used as commercially intended, can be used to determine the efficacy of an intervention aimed at improving %Fat. Research thus far has only evaluated the IMAGE for estimating body volume and subsequent composition metrics via inclusion in a 3-compartment model with BIA measured total body water (8-10); these measures were completed on a single visit to the laboratory. However, the day-to-day consistency of the IMAGE is an important determinant of performance over time. Jackson (23) reported inter-day reliability statistics of 0.96 and 0.97 for BIA as well as 0.97 and 0.98 for SKF. In addition, Anderson (24) reported an ICC of 0.98 when measuring %Fat via air displacement plethysmography between days. Our results agree with previous work regarding the day-to-day reliability of body composition techniques (23, 24). Thus, it appears that the IMAGE displays acceptable levels of reliability when measuring %Fat over a 48-hour time period.

Several limitations to the current study are worth noting. First, previous literature involving body composition techniques such as air displacement plethysmography and BIA have utilized procedures in which %Fat estimates greater than 3% between two trials, required a third measurement be taken (23, 24). While this was not completed in this study, it may be worth investigating in future studies to ensure the highest levels of precision with this technique. Secondly, this study observed the reliability of the technique between two days separated by no more than 48-hours. However, the technique is designed to evaluate the effectiveness of body composition interventions, which would require a more longitudinal study design. Thus, future research should evaluate the validity and reliability between trials separated by at least 8 weeks or more. Thirdly, the commercial use of the IMAGE technique would require an individual user to complete a scan outside of the laboratory environment. Research to this date has not evaluated the accuracy or reliability of the IMAGE outside of the laboratory. Therefore, a study design in which users complete a scan in such a manner (i.e., at home without technician supervision) is warranted. In addition, this study did not control for clothing between day 1 and day 2. Future research should assess the effect of clothing on the IMAGE technique. However, the limited day-to-day variability observed in this study suggests that clothing may not introduce undo error into %Fat measurement from the IMAGE. Lastly, the IMAGE technique's consistency may be affected by other factors that are specific to it and not relevant to other body composition measurement devices. Future research should examine the reliability of the IMAGE across different conditions that may apply outside of the laboratory, such as different megapixel cameras, background colors, and room lighting levels.

In conclusion, given the *trivial* mean differences and the near-perfect correlations, the IMAGE technique appears to be reliable when estimating %Fat between the posterior and anterior positions, as well as across two different days separated by no more than 48-hours. While this technique does need additional research to examine the accuracy and reliability across conditions that may be seen outside of the laboratory as well as in a longitudinal study design, these results add to the growing body of research in support of the effectiveness of the IMAGE technique. The 2D image smartphone application is an inexpensive and portable technique to measure %Fat in field settings, which may be a valuable alternative when traditional assessment techniques are not available.

REFERENCES

1. Burkinshaw L. Measurement of human body composition in vivo. Progress in Medical Radiation Physics: Springer; 1985. p. 113-37.

2. Behnke AR, Feen B, Welham W. The specific gravity of healthy men: body weight÷ volume as an index of obesity. Journal of the American Medical Association. 1942;118(7):495-8.

3. Wilmore JH. Body composition in sport and exercise: directions for future research. Medicine and Science in Sports and Exercise. 1983;15(1):21-31.

4. Kammar A, Hernández-Hernández ME, López-Moreno P, Ortíz-Bueno AM, Martínez-Montaño MDLC. Probability and body composition of metabolic syndrome in young adults: Use of the bayes theorem as diagnostic evidence of the waist-to-height ratio. Stats. 2018;1(1):21-31.

5. Mottillo S, Filion KB, Genest J, Joseph L, Pilote L, Poirier P, et al. The metabolic syndrome and cardiovascular risk: a systematic review and meta-analysis. Journal of the American College of Cardiology. 2010;56(14):1113-32.

6. Ashrafian H, Harling L, Darzi A, Athanasiou T. Neurodegenerative disease and obesity: what is the role of weight loss and bariatric interventions? Metabolic Brain Disease. 2013;28(3):341-53.

7. Barajas-Galindo DE, Arnáiz EG, Vicente PF, Ballesteros-Pomar MD. Effects of physical exercise in sarcopenia. A systematic review. Endocrinología, Diabetes y Nutrición (English ed). 2021.

8. Fedewa MV, Sullivan K, Hornikel B, Holmes CJ, Metoyer CJ, Esco MR. Accuracy of a Mobile 2D Imaging System for Body Volume and Subsequent Composition Estimates in a Three-Compartment Model. Medicine and Science in Sports and Exercise. 2020.

9. Sullivan K, Hornikel B, Holmes CJ, Esco MR, Fedewa MV. Validity of a 3-compartment body composition model using body volume derived from a novel 2-dimensional image analysis program. European Journal of Clinical Nutrition. 2021:1-8.

10. Sullivan K, Metoyer CJ, Hornikel B, Holmes CJ, Nickerson BS, Esco MR, et al. Agreement Between A 2-Dimensional Digital Image-Based 3-Compartment Body Composition Model and Dual Energy X-Ray Absorptiometry for The Estimation of Relative Adiposity. Journal of Clinical Densitometry. 2022;25(2):244-51.

11. Esco MR, Holmes CJ, Sullivan K, Hornikel B, Fedewa MV. Utilizing a Novel 2D Image Processing System for Relating Body Composition Metrics to Performance in Collegiate Female Rowers. International Journal of Environmental Research and Public Health. 2021;18(5):2413.

12. Hopkins W, Marshall S, Batterham A, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Medicine and Science in Sports and Exercise. 2009;41(1):3.

13. Hopkins WG. A scale of magnitudes for effect statistics. A new view of statistics. 2002;502:411.

14. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin. 1979;86(2):420.

15. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine. 2016;15(2):155-63.

16. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. The Journal of Strength & Conditioning Research. 2005;19(1):231-40.

17. Hume P, Marfell-Jones M. The importance of accurate site location for skinfold measurement. Journal of Sports Sciences. 2008;26(12):1333-40.

18. Totosy de Zepetnek JO, Lee JJ, Boateng T, Plastina SE, Cleary S, Huang L, et al. Test-retest Reliability and Validity of Body Composition Methods in Adults. Clinical Physiology and Functional Imaging. 2021.

19. McLester CN, Nickerson BS, Kliszczewicz BM, McLester JR. Reliability and agreement of various InBody body composition analyzers as compared to dual-energy X-ray absorptiometry in healthy men and women. Journal of Clinical Densitometry. 2020;23(3):443-50.

20. Heyward V. ASEP methods recommendation: body composition assessment. Journal of Exercise Physiology Online. 2001;4(4).

21. Kushner RF, Schoeller DA. Estimation of total body water by bioelectrical impedance analysis. The American Journal of Clinical Nutrition. 1986;44(3):417-24.

22. Bourgeois B, Ng BK, Latimer D, Stannard CR, Romeo L, Li X, et al. Clinically applicable optical imaging technology for body size and shape analysis: comparison of systems differing in design. European Journal of Clinical Nutrition. 2017;71(11):1329-35.

23. Jackson A, Pollock ML, Graves JE, Mahar M. Reliability and validity of bioelectrical impedance in determining body composition. Journal of Applied Physiology. 1988;64(2):529-34.

24. Anderson DE. Reliability of air displacement plethysmography. The Journal of Strength & Conditioning Research. 2007;21(1):169-72.

Table 2.1. Participant Characteristics

Characteristics	**Mean ± SD**
Age (y)	32.00 ± 9.39
Height (cm)	170.07 ± 10.05
Body Mass (kg)	77.48 ± 13.95
BMI (kg/m^2)	26.69 ± 3.52

y = years; cm = centimeters; kg = kilograms; BMI = body mass index; Mean ± SD = Mean ± Standard Deviation.

Table 2.2. Reliability statistics for the experimental conditions compared to the criterion.

Condition	M ± SD	p	ES	ICC	Lower	Upper
CR	26.84 ± 6.07					
			Anterior Condition			
ANT	26.44 ± 6.19	0.04	0.07	0.992	0.984	0.996
			Different Day Condition			
DAY	26.65 ± 6.17	0.31	0.03	0.993	0.986	0.997

CR = Criterion Reference; ANT = Anterior Condition; DAY = Different Day Condition; M ± SD = Mean ± Standard Deviation; ES = Effect Size; ICC = Intraclass Correlation Coefficient; CE ± 1.96 SD = Constant Error ± 1.96 Standard Deviation.

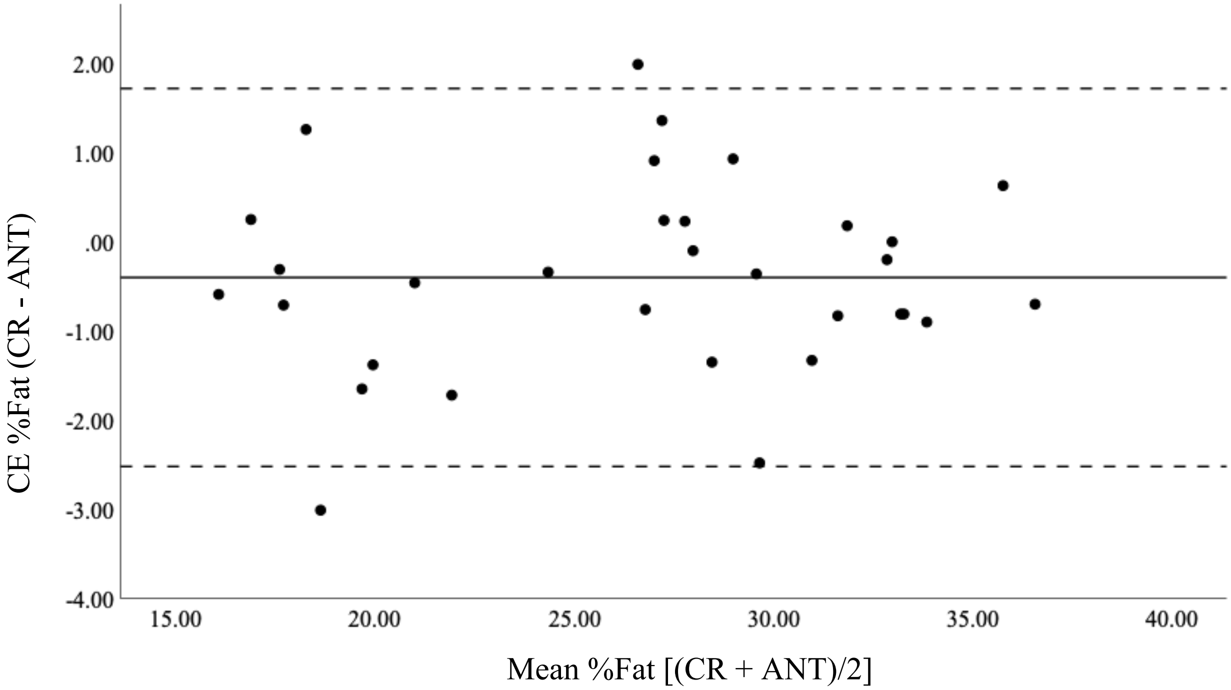


Figure 2.1. Bland-Altman plot showing the agreement between the criterion and the anterior body position measures of %Fat. The solid middle line represents the CE. The two outside dashed lines represent the upper and lower limits of agreement.

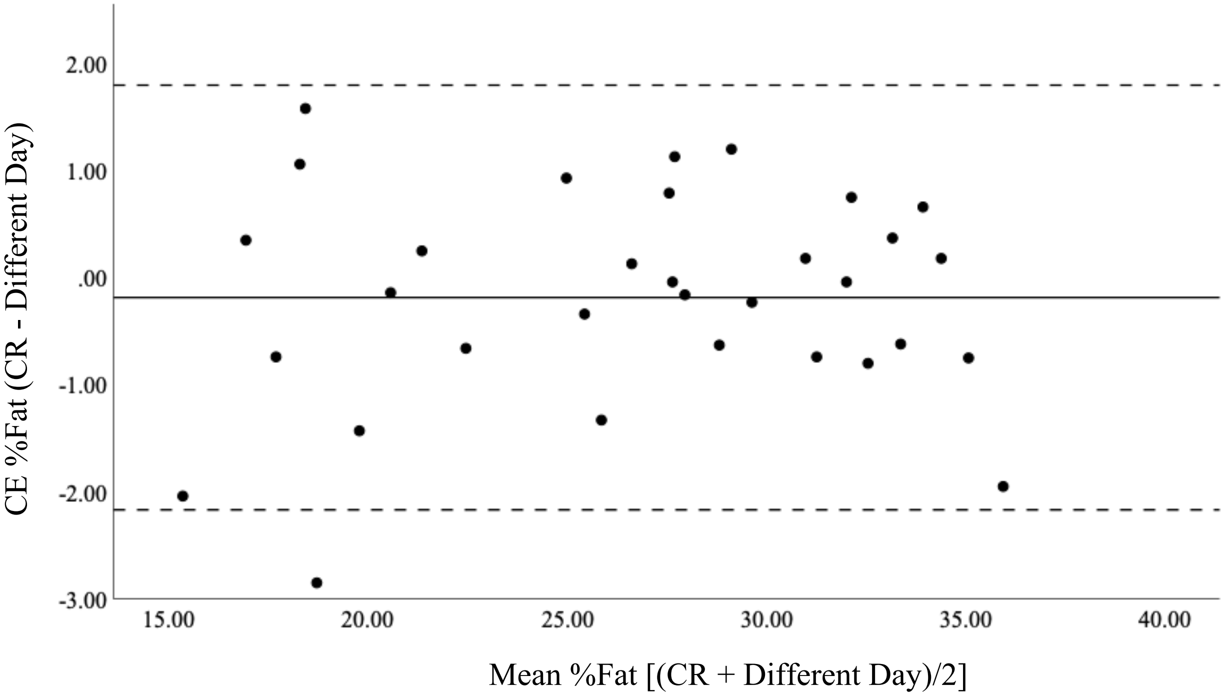


Figure 2.2. Bland-Altman plot showing the agreement between the criterion and the different day condition measures of %Fat. The solid middle line represents the CE. The two outside dashed lines represent the upper and lower limits of agreement.

CHAPTER 3:

RELIABILITY OF A 2D IMAGING SYSTEM ACROSS A VARIETY OF DIFFERENT CONDITIONS.

ABSTRACT

The purpose of this study was to evaluate the reliability of the IMAGE for the assessment of body fat percentage (%Fat) with different megapixel (MP) cameras (0.7MP, 5MP, 8MP, and 12MP). The reliability of the IMAGE was also assessed in Low, Ambient, Moderate, and Bright Light (LL, AL, ML, and BL, respectively). Lastly, the reliability of the IMAGE was also assessed while standing in front of White, Black, Green, Orange, and Gray backgrounds (WB, BB, GnB, OB, and GyB, respectively). To accomplish this, each condition was compared to a single reference condition in which participants standing in a posterior position (i.e., facing away from the camera), in front of the WB with an 8MP camera in ambient lighting (AL). Paired t-tests were used to assess potential mean differences between the experimental conditions and the reference condition. The mean %Fat for the 12MP and 5MP conditions were not different when compared to the 8MP reference condition ($p = 0.16$ and 0.99, respectively). In contrast, %Fat measured via the 0.7MP camera was considered significantly higher ($p < 0.001$). The effect size (ES) of the difference for all three conditions were considered *trivial* and the ICC values (ICC=0.99 for each) were qualitatively *excellent.* The mean %Fat comparison between LL and the AL was significantly higher ($p = 0.02$). However, the mean %Fat for the BL and ML conditions were not different than the AL reference condition ($p = 0.56$ and 0.62, respectively). In addition, the ES of the differences for all three conditions were considered *trivial* and the ICC values (ICC=0.99 for each) were considered *excellent*. The mean %Fat for the BB, GnB, and GyB conditions were all considered significantly higher when compared to the WB reference condition ($p < 0.01$ for each condition). Whereas the mean %Fat comparison between the OB and WB did not display statistical significance ($p = 0.15$). The ES of the difference for the BB condition was considered *small* (ES=0.54) while the other three conditions (GnB, OB, and GyB) were considered *trivial* (0.11, 0.04, and 0.15, respectively). In addition, the statistically significant ICC values were considered *excellent* for OB, GnB, and GyB (ICC=0.99 for each). However, the BB was *moderately* correlated with the WB (ICC=0.71, $p<0.001$). The overall findings indicate %Fat from IMAGE is reliable across a variety of conditions.

KEYWORDS: reliability; lighting; background color; percent fat; 2D image; technology

INTRODUCTION

An accurate and reliable assessment of body composition is necessary for evaluating the effectiveness of healthy lifestyle interventions (1-3). A number of body composition techniques have been shown to provide an acceptable level of accuracy, however most of the research is directed towards studying their accuracies under tightly controlled conditions within a laboratory (4-7). Research has illustrated why standardization of body composition protocols is important to ensure accuracy and reliability (8-10). For instance, participants are often required to undergo 12-hour food fasts, wear specific clothing, and refrain from exercise prior to testing (8-10). Each body composition technique has their own inherent sources of error; therefore, to determine the suitability of any body composition assessment tool, research regarding its ability to provide reproducible measures across multiple conditions is important.

Recently a smartphone application has been developed that estimates body composition from built-in algorithms that are applied to a single two-dimensional image (IMAGE) of a person standing. Once the image is processed by the application, within 30 seconds estimates of body fat percent (%Fat), fat mass (FM), and fat free mass (FFM) are presented on the smartphone screen (4, 5, 11). While the built-in algorithms are proprietary and not available to the public, the conversion of the body composition metrics may first involve prediction of body volume according to previous validation research (4, 5, 11). For example, a previous study demonstrated that the body volume estimates from the image provided an excellent agreement with hydrostatic weighing ($r=0.99$, $p<0.001$) (5). In addition, when combining the body volume estimates with total body water from bioimpedance in a field 3-compartment model, the IMAGE technique also yielded excellent agreement with the laboratory approach ($r=0.96$, $p<0.001$) (11). Therefore, the IMAGE technique shows promise as an accurate field prediction method of body composition.

However, the previous validation research was conducted in tightly controlled laboratory environments in which digital images were obtained while participants stood in front of a white background under ambient room lighting. In addition, the smartphone device and camera was consistent across all images within each study. This technique was commercially designed and is currently available for public use outside of laboratory-based settings. As a result, individual users will likely perform their own measurements and obtain their personal images under a variety of room lighting conditions, with different cameras, and with different colored backgrounds. Therefore, the conditions that can potentially impact the ability of the IMAGE to provide reproducible results are unique compared to other available field techniques. The purpose of this study was to evaluate the reliability of the IMAGE for the assessment of body fat percentage (%Fat) with different megapixel (MP) cameras (0.7MP, 5MP, 8MP, and 12MP). The reliability of the IMAGE was also assessed in Low, Ambient, Moderate, and Bright Light (LL, AL, ML, and BL, respectively). Lastly, the reliability of the IMAGE was also assessed while standing in front of White, Black, Green, Orange, and Gray backgrounds (WB, BB, GnB, OB, and GyB, respectively).

METHODS

Participants

Participants (n = 32, 68.7% female, 84.4% White/Caucasian) were recruited for this study. Descriptive characteristics of the study sample are presented in Table 1. Written informed consent was obtained from each participant before data collection. The institutional review board at the University of Alabama approved the study protocol and data collection procedures.

Procedures

Body mass was measured, to the nearest 0.1 kg, with a calibrated digital scale (Tanita BWB-800, Tanita Corporation, Tokyo, Japan). Barefoot standing height was measured to the nearest 0.1 cm with a manual stadiometer (SECA 213, Seca Ltd., Hamburg, Germany). Height and weight were manually entered into the smartphone application for body composition calculation.

2D Image Processing System / Reference Condition

For image procurement, participants were asked to wear tight-fitting athletic clothing that allows for the automated IMAGE analysis program to identify the necessary anatomical points of interest (4, 5, 11). Participants with long hair were instructed to pull their hair “back” and “up” to allow the digital image to capture the width of the neck. Participants stood with their feet flat in front of a white photography background facing away from the digital camera, with weight evenly distributed on both feet. The heels will be placed together with the feet pointed slightly outward at a 60-degree angle (4, 5, 11). Participants were instructed to remain still with arms abducted at a 45-degree angle away from the torso and aligned within the coronal plane, with palms facing away from the camera. Once correctly positioned, a single digital image that includes the head, feet, and arms of the participant was obtained from the posterior view and analyzed with a commercially available application (version 1.1.2, made Health and Fitness, USA. www.mymadeapp.com). The camera was set in a tripod, at approximately participant waist height, at a distance of 3.05 meters from the participants. %Fat was derived from the IMAGE, using a proprietary algorithm which automatically identifies and measures the horizontal linear diameter of various anatomical landmarks (United States Utility Patent 16/841,944) (4, 5, 11).

Camera Megapixel Conditions

Images were obtained with four different cameras including an iPhone 7, iPad Air 2, iPhone 4, and an iPad 2 (12MP, 8MP, 5MP, 0.7MP). The participants were instructed to stand in the same anatomical position as they did for the reference condition while each picture was taken. All camera megapixel conditions were obtained under AL (300-400 Lux) with participants in front of a WB backdrop.

Room Lighting Conditions

Four digital images of each participant were obtained using the 12.9 inch 64g iPad Air 2 under a series of different room lighting conditions in front of a WB backdrop. Professional photography lighting (AGG856, LimoStudio LLC., Studio City, Ca.) was positioned using a tripod 3.05 meters away from the participant and allowed the researchers to experimentally manipulate the lighting in the room to create the LL, AL, ML, and BL conditions (<50 Lux, 300-400 Lux, 500-800 Lux, and >900 Lux, respectively). The LL condition was created without any overhead fluorescent lighting in the laboratory, and with only one photography bulb illuminated. The AL condition was created using only the overhead fluorescent lighting in the laboratory, and none of the photography bulbs illuminated. The ML condition was created using the overhead fluorescent lighting in the laboratory with two of the photography bulbs illuminated. Lastly, the BL condition was created using the overhead fluorescent lighting in the laboratory with four of the photography bulbs illuminated. Each lighting condition was assessed using a light meter (MT-912, Shenzhen Flus Technology Co., Ltd., Shenzhen, China), which was held at hip-level by the participant, facing the digital camera and photography lighting for 10 seconds to measure illuminance in lux.

Color Background Conditions

Five digital images of each participant were obtained using the 12.9 inch 64g iPad Air 2 under AL conditions in front of a series of different color background. Different solid-colored photography backdrops were used to experimentally create the WB, BB, GnB, OB, and GyB conditions. The participants were instructed to stand in the same anatomical (posterior) position as they did for the reference condition while each picture was taken.

Statistical Analyses

Statistical analyses were performed using SPSS for Mac (SPSS 26.0, Chicago, IL) and Microsoft Excel 365 (Microsoft Corp., Redmond, WA). All data were screened for missing or improbable values using descriptive and frequency statistics. Histograms, skewness, and kurtosis plots were used to identify potential outliers as well as normality of the data, a value of ±2 was used to identify outliers. No data was excluded from analyses.

Pairwise t-tests were used to assess the mean differences between the reference condition and each experimental condition. The participant image obtained using the iPad Air 2 under AL in front of the WB served as the reference condition for each pairwise comparison. The constant error (CE) was calculated as the difference between each experimental condition and the reference, such that a positive CE indicated a higher %Fat measurement in the experimental condition. The 95% confidence interval (CI) was calculated as ± 1.96 Standard Deviation (SD). Effect size (ES) of the paired differences were assessed using Cohen's *d* statistic. The ES of the mean comparisons were qualitatively described as follows: trivial (<0.2), small (0.2-0.6), moderate (0.6-1.2), large (1.2-2.0) or very large (>2.0) (12, 13). Intraclass correlation coefficients (ICCs) were also used to determine the level of agreement, with each of the experimental conditions being compared to the reference condition. An ICC of greater than 0.75 was considered "good" and greater than 0.9 considered "excellent" relative reliability (14). Statistical significance will be assessed using an alpha level of $p < 0.05$.

RESULTS

A final sample of 32 participants was used for analysis and descriptive characteristics including M±SD are listed in Table 3.1. The sample was 68.7% female, 84.4% Caucasian/White, 9.4% multi-racial, 3.1% Hispanic, and 3.1% Other. The sample age ranged from 20 to 52 years with 56.3% being between 20 and 30 years. The sample BMI ranged from 21.6 to 36.8 kg/m^2, with 29.8% of the sample being between 21.6 and 24.9, classified as "Normal Weight", 58.2% ranged from 25.1 to 29.2 kg/m^2, classified as "Overweight", and 12% were classified as "Obese" or "Extremely Obese". In addition, for our lighting conditions the LL condition displayed a range of 9.8 to 25.2 Lux (16.2±3.5 Lux), AL ranged from 346.2 to 394.6 Lux (371.8±13.0 Lux), ML ranged from 593.6 to 788.5 (730.1±49.1 Lux), BL ranged from 984.2 to 1183.0 (1094.2±57.4 Lux). Table 3.2 displays the reliability statistics for all comparisons, except for the limits of agreement which are displayed in Figure 3.2 and described below. The following subsections detail the results of each condition being independently compared to the same reference condition. As explained in the methods, images obtained with the 8MP camera under AL in front of a WB served as the reference condition.

Megapixel Camera Conditions

The mean %Fat for the 12MP and 5MP conditions were not different than the 8MP camera reference condition ($p = 0.16$ and 0.99, respectively). In contrast, the %Fat measured using the 0.7MP camera was considered significantly higher than the 8MP ($p <0.001$). However, despite the higher %Fat observed in the 0.7MP condition, ES of the difference for all three conditions were considered *trivial* and the ICC values (all ICC=0.99) for each were qualitatively *excellent* (all $p < 0.001$) (Table 3.2). The 95% LOA (CE ± 1.96 SD of residual scores) for 12MP was -0.22% ± 1.70%, 5MP was 0.00% ± 1.87%, 0.7MP was 0.96% ± 1.47%. This indicated that %Fat ranged from -1.92% below to 1.48% above the CE for 12MP; -1.87% below to 1.87% above the CE for 5MP; and -0.50% below to 2.43% above the CE for 0.7MP, as illustrated in Figure 2.

Room Lighting Conditions

The mean %Fat measured in LL was significantly higher than the AL reference condition ($p = 0.02$). However, the mean %Fat for the BL and ML conditions were not different than the AL reference condition ($p = 0.56$ and 0.62, respectively; Table 3.2). Despite the higher %Fat observed in the LL condition, the ES of the differences for the LL, ML, and BL conditions were considered *trivial* (Table 3.2). Furthermore, the ICC values, which were all considered *excellent* (ICC=0.99, $p<0.001$) (Table 3.2). The CE and 1.96 SD for the three lighting conditions were as follows: 0.45% ± 1.95 for LL, 0.10% ± 2.10 for ML, and -0.11% ± 2.07 for BL. These values indicate that the 95% LOA ranged from -1.50% below to 2.39% above the CE for LL, -2.01% below to 2.20% above the CE for ML, and -2.18% below to 1.96% above the CE for BL (Figure 2).

Different Background Color

The mean %Fat for the BB, GnB, and GyB conditions were considered significantly higher when compared to the WB reference condition ($p < 0.01$ for each condition). Whereas the mean %Fat comparison between the OB and WHT did not display statistical significance ($p = 0.15$). The observed difference between the BB and WB was considered *small* (ES=0.54), while the GnB, OB, and GyB conditions were considered *trivial* (ES=0.11, 0.04, and 0.15, respectively) (Table 3.2). In addition, the ICC values were considered *excellent* for OB, GnB, and GyB (all ICC=0.99, $p < 0.001$) (Table 3.2). However, the BB condition showed the lowest ICC value (ICC=0.71) which was qualified as *moderate* (Table 3.2). The CE ± 1.96 SD for BB was 5.14% ± 17.59%, which was the largest of all conditions. Thus, its 95% LOA ranged from -12.45% to 22.74%. For the other conditions, the CE ± 1.96 SD for GnB was 0.68% ± 2.15%, for OB was 0.27% ± 2.00%, and for GyB was 0.94% ± 2.38%. This indicated that %Fat ranged from -1.46% below to 2.83% above the CE for GnB; -1.73% below to 2.27% above the CE for OB, and -1.45% below to 3.32% above the CE for GyB, as illustrated in Figure 2.

DISCUSSION

The purpose of this study was to evaluate the reliability of IMAGE to measure %Fat across several independent conditions. To accomplish this, each condition was compared to a single reference condition of an image obtained with an 8MP camera under AL in front of a WB background. The overall findings of the study indicated that the IMAGE produced reliable %Fat results with the use of three different cameras and under three different levels of room lighting, due to the *trivial* mean differences, *excellent* ICC values, and small limits of agreement. Interestingly, different color backgrounds may influence the IMAGE %Fat results and decrease reliability. However, between the various backgrounds that were compared to the WB reference condition, the BB showed the largest mean difference, lowest ICC value, and widest limits of agreement. The findings of GnB, OB, and GyB indicated similar reliability results as the other previously mentioned conditions. Users and practitioners who measure body composition with a smartphone application should avoid using black color backgrounds in the digital images.

Reliability testing is a necessary tool in the evaluation of measurement techniques to ensure their consistency in producing a particular data point across various conditions (15, 16). Several studies have previously been conducted to examine the validity of the IMAGE technique. For example, when comparing body volume, predicted via IMAGE, to underwater weighing, a statistically significant r value of 0.99 was observed (5). In addition, when IMAGE predicted BV was combined with total body water, measured via bioimpedance, in a 3C model, the %Fat estimates were considered equivalent to the underwater weighing criterion (5). The IMAGE technique has also been used to relate fat mass and fat free mass to rowing performance (4). However, no research has been conducted on the reliability of the IMAGE technique across various conditions. While reliability is concerned with consistency, it does not ensure accuracy (i.e., it does not evaluate a valid test). A technique can be considered reliable and not accurate, thus separate evaluation of these two measurements is warranted. This has important practical implications for several reasons, but specifically, the aforementioned studies used a tightly controlled environment for IMAGE procurement, similar to the reference condition used in this study (4, 5, 11). For commercial purposes, the IMAGE was designed with the intention of having public access to perform analyses at their own discretion, which will likely involve an array of different smartphone cameras being utilized for assessments. Thus, based on the current findings, different smartphone cameras can be used interchangeably.

Reliability conceptually evaluates sources of error and their effect on the reproducibility of a test (15, 16). Such sources may include biological variability, instrumentation, subject and technician error (8, 15, 17). The reliability of the IMAGE device has not been previously determined so it is hard to compare results to previous research. However, for other predictions methods, hydration status has been cited as a source of error for bioimpedance and technician error as a source for skinfolds (17-19). While these sources of error may impact reliability across traditional body composition prediction methods, based on what is known regarding the IMAGE technique they may not apply to it (4, 5, 11, 20). Instead, concerns of reliability relate to the ability to take a picture with a smartphone and the ability of the automated, built-in algorithm to distinguish a person within an image and acquire the anatomical landmarks (4, 5, 11, 20). One of our primary findings was that different colored backgrounds (i.e., black, orange, green, and gray) can affect the reliability of the device. While the black background condition displayed a lower ICC and a greater CE than all other conditions, the color and patterns of the participant clothing were not controlled, and the results may indicate that certain clothing colors did not create enough contrast between the participant and the background. Despite the lower reliability with the BB condition, camera megapixel and lighting conditions independently displayed excellent reliability. Thus, the findings support the use of the IMAGE technique under most conditions, however LL (i.e., dark hallways or closets), low megapixel (i.e., < 5MP), and low contrast backgrounds (i.e., black clothing against a black background) should be avoided.

Several limitations should be considered in relation to this study. First, we did not compare the IMAGE technique to any other body composition methods. Future research should directly compare the reliability of the IMAGE to other common techniques. Third, we did not control for clothing in this study, which based on the results, may have contributed to the lower

ICCs and wide LOAs seen in the BB condition. Further exploration of clothing color, in conjunction with background color should be assessed to determine specific methods for the most reliable results. In addition, we did not assess skin pigmentation or control for race/ethnicity in this study design. Future research should assess IMAGE accuracy and reliability across different skin pigments as well as race/ethnicities. Lastly, to our knowledge the IMAGE technique has only been assessed in a laboratory setting with a trained technician. However, the IMAGE was designed to be utilized by individual users for self-assessment, utilizing the front-facing camera and timer setting within the application. Future research should explore the reliability of this technique for self-assessment, to predict %Fat by participants without the aid of a trained technician.

In conclusion, estimating %Fat from a 2D image application is reliable when measured using different camera resolutions, as well as across a variety of lighting conditions and background colors. While similar clothing and background color could increase measurement error, this should be standardized for improved reliability. Although, for the most reliable results, it is recommended that users perform the measurements under similar environmental conditions, it appears that images taken with different megapixel cameras, under different lighting conditions, and in front of various background colors are comparable given the acceptable reliability.

REFERENCES

1. Wilmore JH. Body composition in sport and exercise: directions for future research. Medicine and Science in Sports and Exercise. 1983;15(1):21-31.

2. Högström GM, Pietilä T, Nordström P, Nordström A. Body composition and performance: influence of sport and gender among adolescents. The Journal of Strength & Conditioning Research. 2012;26(7):1799-804.

3. Jones BH, Hauret KG, Dye SK, Hauschild VD, Rossi SP, Richardson MD, et al. Impact of physical fitness and body composition on injury risk among active young adults: a study of army trainees. Journal of Science and Medicine in Sport. 2017;20:S17-S22.

4. Esco MR, Holmes CJ, Sullivan K, Hornikel B, Fedewa MV. Utilizing a Novel 2D Image Processing System for Relating Body Composition Metrics to Performance in Collegiate Female Rowers. International Journal of Environmental Research and Public Health. 2021;18(5):2413.

5. Fedewa MV, Sullivan K, Hornikel B, Holmes CJ, Metoyer CJ, Esco MR. Accuracy of a Mobile 2D Imaging System for Body Volume and Subsequent Composition Estimates in a Three-Compartment Model. Medicine and Science in Sports and Exercise. 2020.

6. Nickerson BS, Tinsley GM, Esco MR. Validity of Field and Laboratory Three-Compartment Models in Healthy Adults. Medicine and Science in Sports and Exercise. 2019;51(5):1032-9.

7. Esco MR, Nickerson BS, Fedewa MV, Moon JR, Snarr RL. A novel method of utilizing skinfolds and bioimpedance for determining body fat percentage via a field-based three-compartment model. European Journal of Clinical Nutrition. 2018;72(10):1431-8.

8. Fosbøl MØ, Zerahn B. Contemporary methods of body composition measurement. Clinical Physiology and Functional Imaging. 2015;35(2):81-97.

9. Nana A, Slater GJ, Hopkins WG, Halson SL, Martin DT, West NP, et al. Importance of standardized DXA protocol for assessing physique changes in athletes. International Journal of Sport Nutrition and Exercise Metabolism. 2016;26(3):259-67.

10. Nana A, Slater GJ, Stewart AD, Burke LM. Methodology review: using dual-energy X-ray absorptiometry (DXA) for the assessment of body composition in athletes and active people. International Journal of Sport Nutrition and Exercise Metabolism. 2015;25(2):198-215.

11. Sullivan K, Hornikel B, Holmes CJ, Esco MR, Fedewa MV. Validity of a 3-compartment body composition model using body volume derived from a novel 2-dimensional image analysis program. European Journal of Clinical Nutrition. 2021:1-8.

12. Hopkins W, Marshall S, Batterham A, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Medicine and Science in Sports and Exercise. 2009;41(1):3.

13. Hopkins WG. A scale of magnitudes for effect statistics. A new view of statistics. 2002;502:411.

14. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin. 1979;86(2):420.

15. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. The Journal of Strength & Conditioning Research. 2005;19(1):231-40.

16. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine. 2016;15(2):155-63.

17. Heyward V. ASEP methods recommendation: body composition assessment. Journal of Exercise Physiology Online. 2001;4(4).

18. Kushner RF, Schoeller DA. Estimation of total body water by bioelectrical impedance analysis. The American Journal of Clinical Nutrition. 1986;44(3):417-24.

19. Medicine ACoS. ACSM's guidelines for exercise testing and prescription: Lippincott williams & wilkins; 2013.

20. Sullivan K, Metoyer CJ, Hornikel B, Holmes CJ, Nickerson BS, Esco MR, et al. Agreement Between A 2-Dimensional Digital Image-Based 3-Compartment Body Composition Model and Dual Energy X-Ray Absorptiometry for The Estimation of Relative Adiposity. Journal of Clinical Densitometry. 2022;25(2):244-51.

Table 3.1. Participant Characteristics

Characteristics	**Mean ± SD**
Age (y)	32.00 ± 9.39
Height (cm)	170.07 ± 10.05
Body Mass (kg)	77.48 ± 13.95
BMI (kg/m^2)	26.69 ± 3.52

y = years; cm = centimeters; kg = kilograms; BMI = body mass index; Mean ± SD = Mean ± Standard Deviation.

Table 3.2. Reliability statistics for all conditions when compared to the criterion.

Condition	M ± SD	p	ES	ICC	Lower	Upper
Criterion	26.84 ± 6.07					
			Different Camera Megapixel Condition			
12MP	26.62 ± 5.88	0.16	0.04	0.995	0.989	0.997
5MP	26.84 ± 6.16	0.99	< 0.00	0.994	0.987	0.997
0.7MP	27.80 ± 6.18	< 0.01	0.16	0.996	0.992	0.997
			Different Room Lighting Condition			
LL	27.29 ± 6.23	0.02	0.07	0.993	0.987	0.997
ML	26.94 ± 6.07	0.62	0.02	0.992	0.984	0.996
BL	26.73 ± 6.10	0.56	0.02	0.992	0.984	0.996
			Different Background Color Condition			
BB	31.98 ± 11.99	< 0.01	0.54	0.714	0.413	0.860
GnB	27.52 ± 6.31	< 0.01	0.11	0.992	0.984	0.996
OB	27.11 ± 6.47	0.15	0.04	0.993	0.986	0.997
GyB	27.77 ± 6.26	< 0.01	0.15	0.990	0.980	0.995

12MP =12 Megapixel; 5MP = 5 Megapixel; 0.7MP = 0.7 Megapixel; LL = Low Light; ML = Medium Light; BL = Bright Light; BB = Black Background; GnB = Green Background; OB = Orange Background; GyB = Grey Background; M ± SD = Mean ± Standard Deviation; ES = Effect Size; ICC = Intraclass Correlation Coefficient; CE ± 1.96 SD = Constant Error ± 1.96 Standard Deviation.

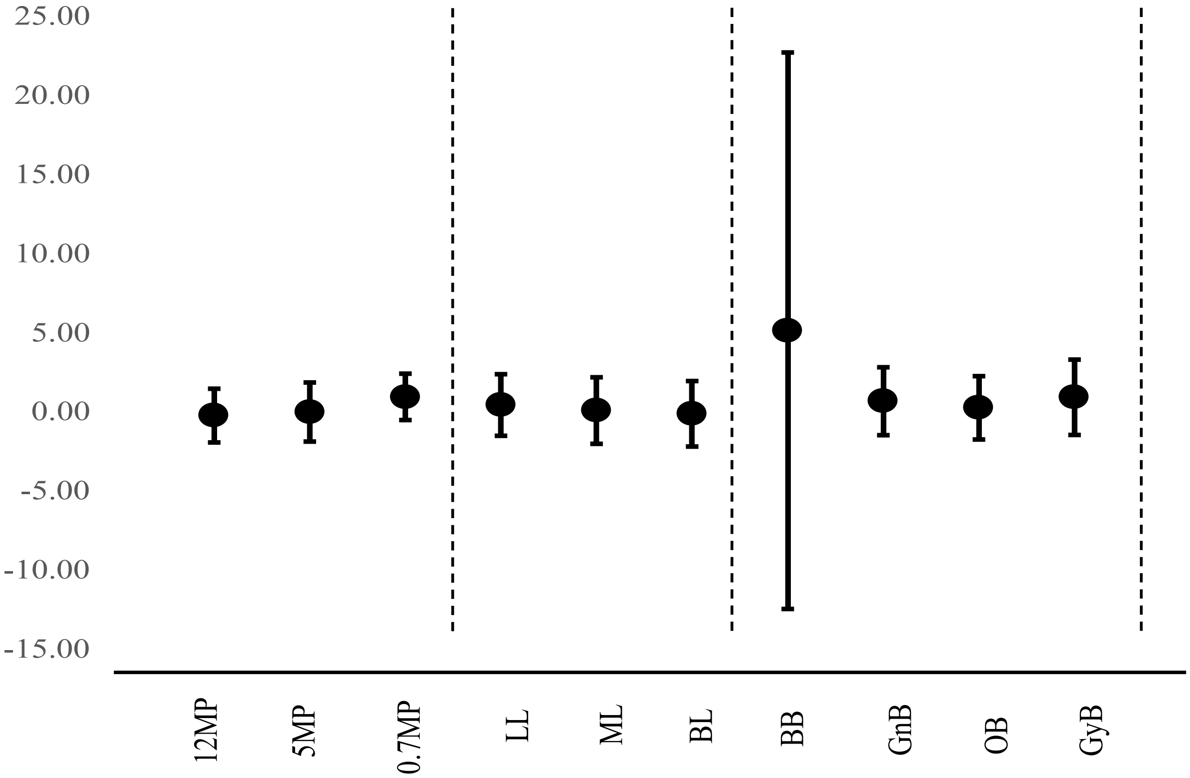


Figure 3.1. The 95% limits of agreement for each condition compared to the criterion. The solid dots represent the mean constant error and the vertical error bars represent the standard deviation*1.96 ± mean.

CHAPTER 4:

THE INFLUENCE OF SELECTED METRICS OF PHYSICAL ACTIVITY AND FITNESS ON THE ERROR ASSOCIATED WITH PREDICTION OF BODY COMPOSITION.

ABSTRACT

The purpose of this study was to determine the extent to which individual error associated with bioimpedance analysis (BIA), skinfolds (SKF), and 2-dimensional image analysis (IMAGE) relates to selected markers of physical activity (PA) and muscular fitness (MF). Markers of physical activity and muscular fitness included: the International physical activity questionnaire, which was used to calculate total, vigorous, and moderate levels of physical activity per week (IPAQT, IPAQV, IPAQM, respectively); Perceived functional ability questionnaire, which was used to determine 1-mile run submaximal pace ($1mile_{SUB}$) and 3-mile run submaximal pace ($3mile_{SUB}$); Physical Activity Rating questionnaire (PA-R): handgrip test (HG); and push-up test (PU).The direction and strength of the association between the PA and MF markers and the individual error between %Fat estimation methods and a 4-compartment criterion (i.e, SKF_{ERROR}, BIA_{ERROR}, and $IMAGE_{ERROR}$) was assessed using Pearson's r. Follow-up stepwise regression procedures were used to determine which PA and/or MF marker accounts for the greatest variation in observed error via each prediction method (SKF_{ERROR}, BIA_{ERROR}, and $IMAGE_{ERROR}$). SKF_{ERROR} was significantly correlated with PU (r=0.17, p=0.038), while none of the independent variables correlated with $BIA_{ERROR.}$ $IMAGE_{ERROR}$ was positively associated with IPAQV (r=0.24, p=0.003), $1mile_{SUB}$ (r=0.24, p=0.003), $3mile_{SUB}$ (r=0.17, p=0.045), PA-R (r=0.24, p=0.003), and PU (r=0.41, p<0.001) indicating that higher physical activity and muscular fitness was associated with greater error. However, only PU was related to both $\text{SKF}_{\text{ERROR}}$ and $\text{IMAGE}_{\text{ERROR}}$ in the regression models, PU independently accounted for 2.9% and 16.9% of the variability ($p < 0.001$), respectively. Based on the results of the study, $\text{SKF}_{\text{ERROR}}$ and $\text{IMAGE}_{\text{ERROR}}$ were associated with maximum repetition push-up test as a measure of muscular fitness such that higher PU was associated with greater error. In addition, none of the physical activity and fitness measures evaluated in this study were associated with $\text{BIA}_{\text{ERROR}}$.

KEYWORDS: percent fat; prediction error; physical activity; 2D image; bioimpedance; skinfolds

INTRODUCTION

The assessment of body composition provides important information regarding disease risk, health status, and performance optimization which can influence decisions of individuals, clinicians, and practitioners. Traditional models for estimating percent body fat (%Fat) are based on two compartments (2C), dividing the body into fat (FM) and fat-free (FFM) components (1-4). Estimation of %Fat in 2C models assumes a constant density of FM and FFM (0.9 and 1.1 g/ml, respectively) (3, 5). However, the degree of error for most 2C models has been associated with the individual variation of several FFM constituents such as total body water, bone mineralization, and protein fractions (1, 5, 6). Because of this, multi-compartment models that account for the variation of FFM components, rather than assume a constant density for all individuals, have emerged as the preferred criterion methods in body composition research (7-9). However, because a multi-compartment approach involves multiple techniques of measurement, 2C models are considered a more feasible approach. Thus, research designed to enhance the accuracy and decrease the range of error associated with 2C approaches is necessary.

Previous research has shown that varying degrees of physical fitness and activity level influence the constituents of FFM in which the error associated with 2C approaches may be related (1, 5, 10, 11). For example, Modlesky et al. (5) reported that in weight trained individuals %Fat may be consistently overestimated by up to 4.2% with 2C models. Such a difference was explained by the weight trained participants displaying a lower density of FFM due to a higher water, lower bone mineral, and lower protein fractions than the assumed constants (5). In addition, Withers et al. (1) reported that the 2C method of hydrostatic densitometry provided an underestimation of %Fat when compared to a laboratory 4-compartment (4C) model in sedentary individuals and aerobically trained participants (2.3-2.8 %Fat). The inaccuracy was associated with a lower hydration status than the assumed constant from hydrostatic densitometry (1). Because the individual discrepancies in the tissues that comprise FFM may lead to error in 2C approaches, proportional bias exists as many techniques overestimate %Fat in lean individuals but underestimate in obese individuals (1, 6). The existing error may be best explained by accounting for individual characteristics that are related to the varying levels of FFM compartments, such as physical activity and muscular fitness.

There are several techniques available for measuring body composition in practical settings. For example, bioelectrical impedance analysis (BIA) & the skinfold method (SKF) are two of the most commonly performed prediction methods (8). Because of their simplicity and low cost, BIA and SKF provide greater accessibility than laboratory based methods (8). In addition, a 2-dimensional image technique (IMAGE) was recently developed using proprietary algorithms that predict %Fat based on diameter measurements of various anatomical landmarks of a person standing within a picture (12, 13). Though the technique further simplifies body composition estimation, the validity of the IMAGE has yet to be thoroughly examined, especially regarding the factors related to its range of error. Typically, the acceptable standard error of estimation for field-based prediction methods is approximately ±3.5 %Fat (8). However, the degree to which aspects of physical fitness and activity contribute to the individual range of error is not well understood (8, 12). Research into this area will help practitioners discern which techniques are most appropriate for the populations they work with, as well as direct future research that is designed to improve the accuracies of these techniques. Therefore, the purpose of this study was to determine the extent to which individual error associated with BIA, SKF, and IMAGE relates to selected markers of physical activity and muscular fitness. The associated error of the three field prediction methods was determined by comparison to a laboratory 4-compartment model. Physical activity and muscular fitness are multi-dimensional paradigms that can be challenging to assess in field settings. Thus, the current study used several questionnaires and basic muscular fitness performance tests to evaluate their relationship with the individual error associated with body composition estimation. The subjective markers of physical activity were measured with the International Physical Activity Questionnaire (IPAQ), the Physical Activity Rating (PA-R) and Perceived Functional Ability (PFA) questionnaire (14-16). The markers of muscular fitness were assessed with the Push-up (PU) and Hand-grip (HG) tests. It was hypothesized that the error for each body composition field predictor would be significantly correlated with each marker of physical activity and fitness.

METHODS

Participants

Participants (n = 148, 50.7% female, 85.1% White/Caucasian) were recruited for this study. Descriptive characteristics of the study sample are presented in Table 1. Written informed consent was obtained from each participant before data collection. The Institutional Review Board at the University of Alabama approved the study protocol and data collection procedures.

All data were collected in the Exercise Physiology Laboratory at the University of Alabama during a single 120-minute visit. Prior to arriving at the laboratory, participants were instructed to abstain from exercise and the ingestion of food and drink, other than water, for a minimum of 12 hours and 4 hours, respectively (13). Adherence to the pre-test study protocol was self-reported and confirmed for each participant prior to data collection using a 24-hour history questionnaire (Appendix A).

Procedures

Each participant provided demographic information including participants self-reported age, biological sex (male/female), and race/ethnicity. Race/ethnicity was recorded on an 8-point scale with the following options (Caucasian/White, Native Hawaiian, Black, Hispanic, American Indian, Asian, Other, Multiple)(17). Hydration status was assessed using a refractometer (Atago SUR-NE, Atago Corp Ltd., Tokyo, Japan) via a urine specific gravity (USG) value of < 1.030 (18). Body mass was measured to the nearest 0.1 kg, with a calibrated digital scale (Tanita BWB-800, Tanita Corporation, Tokyo, Japan). Standing height was measured without shoes to the nearest 0.1 cm with a manual stadiometer (SECA 213, Seca Ltd., Hamburg, Germany).

Laboratory Body Composition Procedures

Dual-Energy X-Ray Absorptiometry

DXA was used to measure BMC for the criterion 4C model. Prior to each whole-body scan, DXA was calibrated according to the instructions provided by the manufacturer using a standard calibration block (GE Lunar Prodigy, software version 14.10.022; GE Lunar Corporation, Madison, WI). Participants were instructed to remove any jewelry as well as bulky clothing and shoes prior to the scan. During each scan, participants laid supine with arms at their sides and palms against their legs, remaining stationary. Velcro straps were placed around the ankles and knees of each participant to reduce lower limb movement. Scans lasted approximately 6-12 minutes. Previous research has shown *excellent* reliability and validity for BMC (ICC = 0.99, SEM = 0.04-0.05kg) (19, 20). The BMC measures from DXA were converted to total body bone mineral (Mo) for entry into the 4C model (9, 21).

$$\text{Mo} = (\text{BMC} \times 1.0436)$$

Underwater Weighing

Residual lung volume was determined using the oxygen dilution technique via nitrogen analysis (ParvoMedics True Max 2400; ParvoMedics, Sandy, UT) while in a seated position before entering the underwater weighing (UWW) tank. Participants completed a minimum of two trials and the average of the nearest two trials, within 5%, was used for residual lung volume. UWW took place in a custom-built, temperature regulated tank with participants wearing tight-fitting clothing or a bathing suit. Participants were instructed to enter the tank and sit in a nylon sling seat suspended from a calibrated Chatillon® 15-kg scale (Model #1315DD-H, Largo, FL). For each participant, the straps of the sling seat were adjusted to ensure the water level rests at approximately chin height. For testing, participants were instructed to perform a maximum expiration and submerge all body parts below water. While participants are fully submerged, their underwater weight was recorded to the closest 0.025 kg. The mean of the three highest values (5-10 trials) was used as the representative value of underwater weight and to derive BV (BV_{UWW}) for inclusion in the criterion 4C model calculation(s).

Bioimpedance Spectroscopy

Hand-to-foot BIS (Imp^{TM} SFB7, ImpediMed Limited, Queensland, Australia) was used to determine TBW (TBW_{BIS}), via algorithms within the BIS device, for the criterion 4C model. Prior to electrode placement, excess hair was removed with a razor from the right hand and foot, these sites were cleaned with alcohol pads. In accordance with manufacturer's specifications, electrodes were placed on the right hand and right foot with participants in a supine position. Participants were instructed to abduct the arms ≥30° away from the body and separate the legs. The measurement was taken after the necessary descriptive characteristics (height, weight, age, and sex) were entered into the BIS device. Previous research has shown a strong correlation between TBW_{BIS} and deuterium oxide ($r = 0.98$, SEE = 2.12L,) (22). TBW_{BIS} was used in the 4C model.

Four-compartment model calculation

Body composition values were calculated utilizing a 4C model as the criterion for %Fat assessment. The 4C model uses BV_{UWW}, TBW_{BIS}, Mo_{DXA}, and body mass to calculate FM. Body composition metrics derived from the 4C model were calculated as described by Wang et al. (9).

$$\text{Fat Mass} = 2.748(BV_{UWW}) - 0.699(TBW_{BIS}) + 1.129(Mo_{DXA}) - 2.051(\text{Body Mass})$$

$$FFM = BM - FM$$

$$\%Fat = (FM / BM) \times 100$$

Field Body Composition Procedures

2D Image Processing System

For image capture, participants were asked to wear form-fitting athletic clothing that allows for the automated 2D image analysis program to identify the necessary anatomical landmarks (12). Participants with long hair were instructed to pull their hair "back" and "up" to allow the diameter of the neck to be viewable in the digital image. Participants stood in front of a white photography background with feet flat on the floor and weight evenly distributed between them, facing away from the digital camera. The heels were placed together with the feet angled slightly outward at approximately 60-degrees. Participants were instructed to remain stationary with arms abducted from the torso, at a 45-degree angle away and aligned within the frontal plane, with palms facing away from the camera. Once correctly positioned, a single digital image that includes the individual's entire body (i.e., head, feet, and arms) was obtained from the rear/posterior view using a 12.9 inch, 64g iPad Pro and analyzed using a commercially available application (version 1.1.2, made Health and Fitness, USA. www.mymadeapp.com) (12, 13). The camera was set in a tripod, at approximately participant waist height, at a distance of 3.05 meters from the participants. %Fat was calculated from the 2D digital image, using a proprietary algorithm which automatically identifies and measures the horizontal linear diameter of various anatomical landmarks (United States Utility Patent 16/841,944) (12, 13, 23). Previous research has shown the accuracy of the IMAGE for deriving BV ($r = 0.99$, $SEE = 0.68$ L) (12). The development of the IMAGE proprietary algorithm was completed using participants that were not included in the current study.

Bioelectrical Impedance Analysis

For the measurement of BIA, participants were instructed to remove all body jewelry. Electrode sites were cleaned with alcohol swabs and excess hair removed with a razor. Participants rested in a supine position on a gurney for approximately 5 minutes with arms abducted to 30° from the lateral aspect of the torso and legs not touching. Once participants were fully prepared, electrodes were placed on the right hand and foot according to manufacturer guidelines. Resistance (R) and reactance (X_c) values were then measured using a single frequency (50 kHz) BIA device (Quantum iV, RJL Systems Inc., Clinton, MI) that uses proprietary software and algorithms to derive %Fat for each participant.

Skinfolds

Skinfold thickness was measured on the right side of the body with a calibrated Lange caliper (Beta Technology, Ann Arbor, MI), utilizing the 7-site skinfold protocol outlined by Jackson & Pollock (24). Skinfolds were measured from 7 sites as follows: 1) a diagonal fold on the chest half way between the nipple and front of the shoulder for men, or 1/3 of the distance between the nipple and shoulder for women; 2) a vertical fold on the triceps halfway between the back of the shoulder joint and elbow, with the arm held freely to this side of the body; 3) a diagonal fold on the shoulder blade; 4) a vertical fold directly on the side the body in between the arm pit and hip bone; 5) a vertical fold on the abdomen 2 cm to the right of the belly button; 6) a diagonal fold 2 cm above the hip bone; 7) a vertical fold on the anterior midline of the thigh, midway between the proximal border of the patella and the inguinal crease (hip). Each site was measured once, and then all measurements were repeated. Each site was measured at least two times and the average was calculated. All measurements were within 2 mm, however if they exceeded that limit a third measurement was taken after the skin had time to regain normal texture and thickness (25). %Fat was calculated utilizing the Brozek equation (4).

Markers of Physical Activity and Muscular Fitness

Participants completed the International Physical Activity Questionnaire – Short Form (IPAQ-SF) which assesses vigorous, moderate, walking, and sitting physical activities, in hours and minutes, over the past 7 days (Appendix B). Information obtained was quantified as a continuous variable in total physical activity minutes per week (IPAQT), total vigorous physical activity minutes per week (IPAQV), total moderate physical activity minutes per week (IPAQM)

(14). Previous research has shown acceptable correlations (>0.50) between the IPAQ-SF and objective techniques for measurement of vigorous physical activity (14). In addition, it has been reported that its proven reliability shows it can be used in repeated measures studies (14).

Next, each participant completed the Perceived Functional Ability (PFA) questionnaire, which asks two questions based on a 13-point scale with minutes per mile labels (i.e., “7 minutes per mile or less”) in addition to qualitive effort descriptions (i.e., “Running at a *fast* pace”); “Suppose you were going to exercise continuously on an indoor track for 1-mile ($1mile_{SUB}$). Which exercise pace is just right for you—not too easy and not too hard?” and “How fast could you cover a distance of 3-miles and NOT become breathless or overly fatigued? Be realistic” ($3mile_{SUB}$; Appendix C) (15). Previous research has shown a strong correlation (r=0.72) with criterion maximal aerobic testing (15).

Next, participants completed the Physical Activity Rating (PA-R) questionnaire, in which physical activity was assessed with one question presented on a 0-10 point scale which describes different types and amounts of PA (i.e., “10 = vigorous activity: run over 25 miles per week or spend over 8 hours per week in a comparable physical activity described above”); “Select the number that best describes your overall level of physical activity for the previous 6 MONTHS.”(15, 16) (Appendix D). Previous research has reported correlations 0.59 and 0.35 with criterion maximal aerobic testing (15, 16).

Participants also performed two objective muscular fitness tests, first of which is a maximum effort hand grip test. Each participant was instructed to perform the handgrip test standing with both feet flat on the floor. The dynamometer was adjusted such that the second joint of the finger is bent at 90°. The dynamometer was reset to zero by the researcher prior to each handgrip test. Participants were instructed to hold the dynamometer with the elbow flexed at 90°. The participants were then instructed to squeeze the dynamometer as hard as possible. The handgrip strength was recorded in kilograms (kgs). The same procedure was then repeated with the opposite hand. This procedure was repeated two additional times for each hand, for a total of 6 tests. The highest value of the three readings for each hand was added together and recorded as the combined sum (HG) (26). Previous research has reported error of less than 1% for this test and measurement tool, given standard operating procedures are followed (27).

For the push-up test, each participant placed their hands shoulder width apart on the ground and maintained a horizontal spinal position. Women were instructed to keep their knees bent at approximately 90° and feet plantar flexed, performing a modified push-up using the knees as a pivot point. Men were instructed to perform standard push-ups with the toes as a pivot point. All participants were instructed to bend their elbows until they achieved 90° of flexion at which point, they were instructed to "push-up" and extend the elbows to return to their starting position. A complete push-up was counted when participants perform this motion correctly without breaking proper from. The test was untimed, with no metronome or cadence required, participants were allowed to rest in the fully extended positioned ad libitum. The test was terminated if the subject was unable to maintain the proper technique or at volitional failure. The maximum number of push-ups performed correctly was counted, recorded by an examiner, and analyzed as a continuous variable. Previous research has noted test-retest correlations of 0.93 for both males and females indicating strong reliability. The push-up test has also been compared to a maximum repetition bench press test with correlations ranging from 0.28 to 0.87 for males and females, for this study PU was used as a standalone measure of muscular fitness.

Individual error

Individual error was calculated by subtracting the %Fat value calculated by each field predictor from the 4C model with the following equation.

%Fat error = predicted %Fat from each method – observed %Fat from the 4C model

%Fat error from SKF, BIA, and IMAGE will be labeled as SKF_{ERROR}, BIA_{ERROR}, and $IMAGE_{ERROR}$, respectively.

Statistical Analyses

Statistical analyses were performed using SPSS version 26.0 (IBM Corporation, Chicago, IL). All data were screened for missing or improbable values using descriptive and frequency statistics. Histograms, skewness, and kurtosis plots were used to identify potential outliers as well as normality of the data, a value of ±2 was used to identify outliers.

Bivariate correlations (Pearson's r) were used to determine the direction and strength of the relationship between the physical activity and fitness markers and SKF_{ERROR}, BIA_{ERROR}, and $IMAGE_{ERROR}$. For all correlation procedures the strength of each r value was qualitatively described as follows: 0-0.30, small; 0.30-0.50, medium; 0.50-0.70, large; and 0.70-1.00, very large (28).

In addition, stepwise regression procedures were used to determine the independent association between measures of physical activity and muscular fitness and the observed error of the %Fat prediction methods. In addition, the percent of observed error that can be explained by the measures of physical activity and muscular fitness was evaluated via the R^2 statistic. Results expressed as mean (M) ± standard deviation (SD), unless otherwise indicated. Statistical significance for all procedures was determined as $p < 0.05$.

RESULTS

A final sample of 148 participants was used for analysis. The M±SD as well as lower and upper ranges of the studied variables are presented in Table 4.2. The sample was 50.7% female, 85.1% Caucasian/White, 6.8% multi-racial, and the other 8% consisting of other, Asian, Black, Hispanic, or Native American. The sample age ranged from 18 to 64 years with 86% being between 18 and 30 years. The sample BMI ranged from 17.8 to 41.1 kg/m^2, with 58% of the sample being between 18.5 and 24.9, classified as "Normal Weight", 2% classified as "Underweight", and 25.8% classified as "Overweight". Table 4.3 displays the correlation coefficients for each pairwise comparison. SKF$_{\text{ERROR}}$ was positively correlated with PU (*small*, p = 0.038). None of the studied variables were significantly correlated with BIA$_{\text{ERROR}}$. However, IMAGE$_{\text{ERROR}}$ was positively associated with IPAQV (*small*, p = 0.003), 1mile$_{\text{SUB}}$ (*small*, p = 0.003), 3mile$_{\text{SUB}}$ (*small*, p = 0.045), PA-R (*small*, p = 0.003), and PU (*medium*, p < 0.001), indicating that higher physical activity or muscular fitness was associated with greater error (Table 4.3).

Stepwise linear regression analysis of the independent variables and SKF$_{\text{ERROR}}$ indicated that only PU independently contributed to the model, such that 2.9% of the error associated with %Fat estimation via SKF was explained via the PU test (p = 0.038). There were no significant contributors to the regression model for BIA$_{\text{ERROR}}$, which may be due to no significant correlations existing between the independent variables and BIA$_{\text{ERROR}}$. Whereas stepwise linear regression analysis of the independent variables and IMAGE$_{\text{ERROR}}$ indicated that only PU independently contributed to the model, indicating that 16.9% of the error associated with %Fat estimation via the IMAGE technique was explained via the PU test (p < 0.001 (Figure 4.1).

DISCUSSION

Based on the results of this study, SKF$_{\text{ERROR}}$ displayed a *small* but significant correlation with PU, however the measures of PA and muscular fitness were not associated with BIA$_{\text{ERROR}}$. Whereas, PU, IPAQV, 3mileSUB, 1mileSUB, and PA-R were associated with IMAGE$_{\text{ERROR}}$. Follow-up stepwise regression analyses revealed that only the PU test independently accounted for the variability of both SKF$_{\text{ERROR}}$ and IMAGE$_{\text{ERROR}}$. Thus, it appears that upper-body muscular endurance independently accounts, to some extent, for the error associated with SKF and IMAGE techniques.

It is well understood that field estimates of %Fat display some level of error (5). There are several previous studies that examined the potential variables that may account for the observed error (8, 29, 30). Sources of error associated with the SKF technique include both biological (i.e., age, sex, race, etc.), as well as technical error (8, 25, 30, 31). Technical error with the SKF technique is associated with variation in measurement site location and interrater reliability testing (8, 25, 30, 32). While these error related co-factors are common with SKF measurement, BIA displays relatively different measurement error and sources. Biologically, BIA is heavily influenced by fluctuations in hydration status (8, 31, 33). Whereas variations in device, frequency, and equations have been cited as technical sources (8, 33).

The focus of the current study was to determine if features of physical activity and muscular fitness could also contribute to the error associated with SKF and BIA. As mentioned previously, a *small* correlation with SKF$_{\text{ERROR}}$ and PU (r=0.17) was found, but no correlations between the independent variables and BIA$_{\text{ERROR}}$. Esco et al. (34), reported strong negative correlations between three skinfold sites (thigh, sub-scapular, and abdomen) and one minute push-up test (r = -0.53, -0.43, and -0.40, respectively). While this study did not look at individual site correlations, results do indicate there is a relationship between body composition measured via skinfolds and muscular endurance measured via push-up test. The specific reasons for insignificant findings for $\text{BIA}_{\text{ERROR}}$ are difficult to determine but previous research has observed significant negative correlations between maximum repetition push-up test and %Fat (r=-0.46) measured via bioimpedance in a large sample of young military men (35). However, in the current study, the sample comprised 50.7% women and mean %Fat values were 5% greater (22.97±8.98 vs. 17.8±7.2 %Fat). While these attributes do not fully explain the non-significant findings between all independent variables and $\text{BIA}_{\text{ERROR}}$, sample characteristics may have contributed to it. Further, previous research has noted potential error associated with body composition measurement in age-matched and/or weight-matched resistance trained athletes in both within group comparisons as well as between non-athlete group comparisons, which may be relevant to the IMAGE technique (36-38).

The IMAGE technique is a novel measure of body composition and thus the sources of potential error are not as well understood. The results of this study indicate that factors related to muscle fitness introduce error into the ability of the IMAGE technique to estimate %Fat. As mentioned previously, maximal effort push-up test displayed the highest correlation with $\text{IMAGE}_{\text{ERROR}}$. While the IMAGE technique does not directly measure muscular performance, the algorithm used for %Fat prediction uses anatomical landmarks and horizontal diameter measurements as part of the process. Previous research has shown that more muscular individuals may display different anatomical widths when compared to the average population. For instance, Katch et al. (36) reported that there were only significant differences in a group of bodybuilders, when compared to power lifters and Olympic weightlifters, in the girths of the shoulders, chest, biceps, and forearms. In addition, Kanehisa et al. (37) reported that muscle cross-sectional areas, in collegiate Olympic weightlifters, were 17 to 56% larger than age-matched non-athletes. Therefore, $\text{IMAGE}_{\text{ERROR}}$ associated with muscular performance may be due to variation in anatomical widths in physically fit individuals. Future research should further investigate $\text{IMAGE}_{\text{ERROR}}$ associated with muscular fitness (i.e., hypertrophy, strength, and endurance) and more specifically muscle cross-sectional area.

There are a few limitations to this study that should be noted. The sample that participated in this study was 85% Caucasian or White and therefore the results cannot be generalized to other races. Future research should examine the error associated with these field-based devices and physical activity in a more racially diverse sample. While the physical activity measures used in this study were selected for their accessibility and evidence-based practice, they are not exhaustive and other measures may yield different results. Future research should examine other measures of physical activity and/or muscular fitness, such as maximal aerobic consumption or maximal strength assessment. Lastly, while the IMAGE technique was previously validated as a measure of body volume, research has yet to thoroughly examine the validity of this technique as a measure of body composition. Thus, future research should continue to explore the validity of field-based measure of body composition using laboratory-based criterion measures, such as 4 compartment models, dual x-ray absorptiometry, and air displacement plethysmography.

In conclusion, roughly 17% of the observed error associated with %Fat prediction by the IMAGE was associated with maximum repetition push-up test as a measure of muscular fitness. In addition, physical activity and muscular fitness were not correlated with SKF and BIA %Fat error in the current sample. Based on these results, it appears that the IMAGE technique is influenced by muscular fitness, which is likely a surrogate for anatomical diameter measurements. While more research should explore the influence of varying degrees of muscle mass and diameter on the validity of $\mathrm{IMAGE}_{\mathrm{ERROR}}$ as a standalone measure of body composition.

REFERENCES

1. Withers RT, LaForgia J, Pillans R, Shipp N, Chatterton B, Schultz C, et al. Comparisons of two-, three-, and four-compartment models of body composition analysis in men and women. Journal of Applied Physiology. 1998;85(1):238-45.

2. Lohman TG. Research progress in validation of laboratory methods of assessing body composition. Medicine and Science in Sports and Exercise. 1984;16(6):596-605.

3. Siri WE. Body composition from fluid spaces and density: analysis of methods. Techniques for Measuring Body Composition. 1961;61:223-44.

4. Brožek J, Grande F, Anderson JT, Keys A. Densitometric analysis of body composition: revision of some quantitative assumptions. Annals of the New York Academy of Sciences. 1963;110(1):113-40.

5. Modlesky CM, Cureton KJ, Lewis RD, Prior BM, Sloniger MA, Rowe DA. Density of the fat-free mass and estimates of body composition in male weight trainers. Journal of Applied Physiology. 1996;80(6):2085-96.

6. Fedewa MV, Nickerson BS, Tinsley GM, Esco MR, Dunbar EG, Boucher AG, et al. Examining race-related error in two-compartment models of body composition assessment: a systematic review and meta-analysis. Journal of Clinical Densitometry. 2021;24(1):156-68.

7. Heymsfield SB, Ebbeling CB, Zheng J, Pietrobelli A, Strauss BJ, Silva AM, et al. Multi-component molecular-level body composition reference methods: evolving concepts and future directions. Obesity Reviews. 2015;16(4):282-94.

8. Heyward V. ASEP methods recommendation: body composition assessment. Journal of Exercise Physiology Online. 2001;4(4).

9. Wang ZS, W,; Withers, R; Heymsfield, SB. Multicomponent molecular-level models of body composition analysis. . In: Heymsfield SB LT, Wang ZM, Going SB, editor. Human Body Composition. Champaign, IL: Human Kinetics; 2005. p. 163-76.

10. Prior BM, Modlesky CM, Evans EM, Sloniger MA, Saunders MJ, Lewis RD, et al. Muscularity and the density of the fat-free mass in athletes. Journal of Applied Physiology. 2001;90(4):1523-31.

11. Withers RT, Laforgia J, Heymsfield S. Critical appraisal of the estimation of body composition via two-, three-, and four-compartment models. American Journal of Human Biology: The Official Journal of the Human Biology Association. 1999;11(2):175-85.

12. Fedewa MV, Sullivan K, Hornikel B, Holmes CJ, Metoyer CJ, Esco MR. Accuracy of a Mobile 2D Imaging System for Body Volume and Subsequent Composition Estimates in a Three-Compartment Model. Medicine and Science in Sports and Exercise. 2020.

13. Sullivan K, Hornikel B, Holmes CJ, Esco MR, Fedewa MV. Validity of a 3-compartment body composition model using body volume derived from a novel 2-dimensional image analysis program. European Journal of Clinical Nutrition. 2021:1-8.

14. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the international physical activity questionnaire short form (IPAQ-SF): A systematic review. International Journal of Behavioral Nutrition and Physical Activity. 2011;8(1):1-11.

15. George JD, Stone WJ, Burkett LN. Non-exercise VO2max estimation for physically active college students. Medicine and Science in Sports and Exercise. 1997;29(3):415-23.

16. Jackson AS, Blair SN, Mahar MT, Wier LT, Ross RM, Stuteville JE. Prediction of functional aerobic capacity without exercise testing. Medicine and Science in Sports and Exercise. 1990;22(6):863-70.

17. Wang Y, Beydoun MA. The obesity epidemic in the United States—gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. Epidemiologic Reviews. 2007;29(1):6-28.

18. Armstrong L, Soto J, Hacker F. Urinary indices during dehydration, exercise, and rehydration. Occupational Health and Industrial Medicine. 1999;2(40):97.

19. Stone TM, Wingo JE, Nickerson BS, Esco MR. Comparison of bioelectrical impedance analysis and dual-energy x-ray absorptiometry for estimating bone mineral content. International Journal of Sport Nutrition and Exercise Metabolism. 2018;28(5):542-6.

20. Moon JR, Eckerson JM, Tobkin SE, Smith AE, Lockwood CM, Walter AA, et al. Estimating body fat in NCAA Division I female athletes: a five-compartment model validation of laboratory methods. European Journal of Applied Physiology. 2009;105(1):119-30.

21. Heymsfield SB, Wang J, Heshka S, Kehayias JJ, Pierson RN. Dual-photon absorptiometry: comparison of bone mineral and soft tissue mass measurements in vivo with established methods. The American Journal of Clinical Nutrition. 1989;49(6):1283-9.

22. Moon JR, Tobkin SE, Roberts MD, Dalbo VJ, Kerksick CM, Bemben MG, et al. Total body water estimations in healthy men and women using bioimpedance spectroscopy: a deuterium oxide comparison. Nutrition & Metabolism. 2008;5(1):7.

23. Esco MR, Holmes CJ, Sullivan K, Hornikel B, Fedewa MV. Utilizing a Novel 2D Image Processing System for Relating Body Composition Metrics to Performance in Collegiate Female Rowers. International Journal of Environmental Research and Public Health. 2021;18(5):2413.

24. Jackson AS, Pollock ML. Practical assessment of body composition. The Physician and Sportsmedicine. 1985;13(5):76-90.

25. Medicine ACoS. ACSM's guidelines for exercise testing and prescription: Lippincott williams & wilkins; 2013.

26. Medicine ACoS. ACSM's health-related physical fitness assessment manual: Lippincott Williams & Wilkins; 2013.

27. Cronin J, Lawton T, Harris N, Kilding A, McMaster DT. A Brief Review of Handgrip Strength and Sport Performance. The Journal of Strength & Conditioning Research. 2017;31(11):3187-217.

28. Rosenthal R, Rosnow RL. Essentials of behavioral research: Methods and Data Analysis, 2008.

29. Heyward VH, Wagner DR. Applied body composition assessment: Human Kinetics; 2004.

30. Lohman T, Pollock M, Slaughter M, Brandon L, Boileau R. Methodological factors and the prediction of body fat in female athletes. Medicine and Science in Sports and Exercise. 1984;16(1):92-6.

31. Lohman T. Biological variation in body composition. Journal of Animal Science. 1971;32(4):647-53.

32. Lohman TG. Skinfolds and body density and their relation to body fatness: a review. Human Biology. 1981:181-225.

33. Kushner RF, Schoeller DA. Estimation of total body water by bioelectrical impedance analysis. The American Journal of Clinical Nutrition. 1986;44(3):417-24.

34. Esco MR, Olson MS, Williford H. Relationship of push-ups and sit-ups tests to selected anthropometric variables and performance results: A multiple regression study. The Journal of Strength & Conditioning Research. 2008;22(6):1862-8.

35. Vaara JP, Kyröläinen H, Niemi J, Ohrankämmen O, Häkkinen A, Kocay S, et al. Associations of maximal strength and muscular endurance test scores with cardiorespiratory fitness and body composition. The Journal of Strength & Conditioning Research. 2012;26(8):2078-86.

36. Katch VL, Katch FI, Moffatt R, Gittleson M. Muscular development and lean body weight in body builders and weight lifters. Medicine and Science in Sports and Exercise. 1980;12(5):340-4.

37. Kanehisa H, Ikegawa S, Fukunaga T. Body composition and cross-sectional areas of limb lean tissues in Olympic weight lifters. Scandinavian Journal of Medicine and Science in Sports. 1998;8(5):271-8.

38. Spenst LF, Martin AD, Drinkwater DT. Muscle mass of competitive male athletes. Journal of Sports Sciences. 1993;11(1):3-8.

Table 4.1. Participant characteristics

Characteristics	**Mean ± SD**
Age (y)	24.06 ± 7.82
Height (cm)	172.26 ± 9.87
Body Mass (kg)	74.08 ± 16.05
BMI (kg/m^2)	24.82 ± 4.09

y = years; cm = centimeters; kg = kilograms; BMI = body mass index; Mean ± SD = Mean ± Standard Deviation.

Table 4.2. Mean, standard deviation, and range of the %Fat and error related to each body composition measure.

			Range	
Variables	**Mean**	**SD**	**Lower**	**Upper**
$4C_{\%Fat}$	21.14	9.37	2.78	39.50
$SKF_{\%Fat}$	18.59	7.82	3.27	33.92
$BIA_{\%Fat}$	22.97	8.98	5.37	40.57
$IMAGE_{\%Fat}$	24.40	6.45	11.79	37.04
SKF_{Error}	-2.55	5.46	-13.25	8.15
BIA_{Error}	2.12	5.52	-8.70	12.94
$IMAGE_{Error}$	3.25	7.04	-10.54	17.05

$4C_{\%Fat}$ = 4-Compartment model percent body fat ; $SKF_{\%Fat}$ = Skinfold percent body fat; $BIA_{\%Fat}$ = Bioimpedance analysis percent body fat; $IMAGE_{\%Fat}$ = 2D image percent body fat; SKF_{Error} = Skinfold individual error; BIA_{Error} = Bioimpedance analysis individual error; $IMAGE_{Error}$ = 2D image individual error.

Table 4.3. Correlation coefficients of each pairwise comparison.

Variable	SKF_{ERROR}	BIA_{ERROR}	$IMAGE_{ERROR}$
IPAQT	-0.13	-0.06	-0.07
IPAQV	0.10	-0.03	0.24*
IPAQM	0.04	0.13	0.04
$1mile_{SUB}$	0.11	-0.10	0.24*
$3mile_{SUB}$	0.13	-0.04	0.17*
PA-R	0.14	-0.05	0.24*
HG	-0.11	-0.10	-0.01
PU	0.17*	0.04	0.41*

SKF_{ERROR} = Skinfold individual error; BIA_{ERROR} = Bioimpedance analysis individual error; $IMAGE_{ERROR}$ = IMAGE individual error; IPAQT = IPAQ total active minutes; IPAQV = IPAQ total vigorous active minutes; IPAQM = IPAQ total moderate active minutes; $1mile_{SUB}$ = PA-R question 1; $3mile_{SUB}$ = PA-R question 2; HG = Handgrip test; PU = Push-up test; *indicates $p < 0.05$

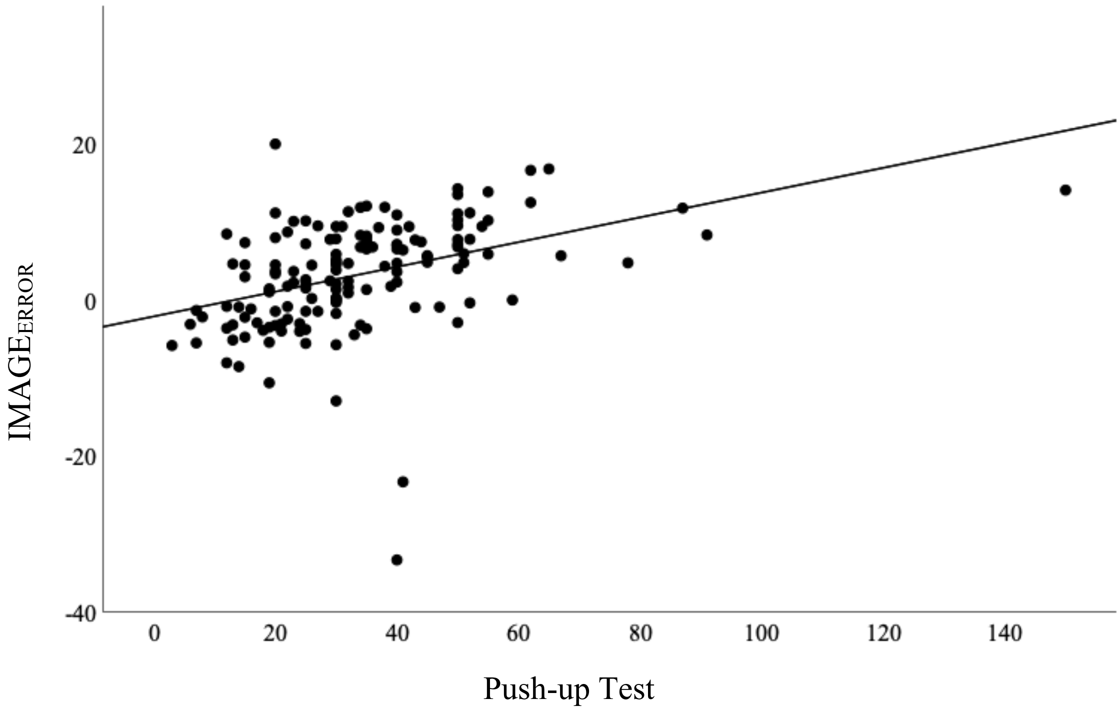


Figure 4.1. Scatterplot representing the relationship between Push-up test and constant error of the 2D image processing system (IMAGE). The solid line represents the regression line based off the stepwise linear regression analysis.

CHAPTER 5:

CONCLUSION

Obesity is a rapidly developing syndrome that has been deemed by the World Health Organization as a growing problem and has only been exacerbated by the recent COVID-19 pandemic (1-4). Traditional approaches to evaluating the outcomes of medical and lifestyle interventions on body composition, such as body mass index and weight-to-hip ratio, do not fully explain the underlying culprit of obesity, excess body fat accumulation (5, 6). While sophisticated equipment, such as dual-energy x-ray absorptiometry, provides a high level of accuracy for evaluating relative adiposity (%Fat), access to this technique can be quite limited (7). Therefore, there remains the need to develop an accurate and reliable, yet portable %Fat technique that can be integrated in both traditional and virtual settings.

With recent technological developments, a novel 2-dimensional image smartphone application (IMAGE) was developed for measuring %Fat. Previous research has evaluated its accuracy as a measure of body volume and subsequently incorporated into a 3-compartment and 4-compartment model for estimation of %Fat against a laboratory criterion (8-10). It has also been used as a standalone measure to relate rowing performance to body composition (11). However, no research to date has evaluated the reliability of the IMAGE across varying conditions that may affect the IMAGE's performance outside of the laboratory setting. Furthermore, the error related co-factors associated with %Fat prediction have yet to be investigated as well. Therefore, the purpose of this dissertation was to critically examine the reliability of a 2D image smartphone application for the measurement of %Fat and evaluate error related co-factors that contribute to field-based measures of body composition.

The first study sought to assess the reliability of the IMAGE across two different body positions and days for the estimation of %Fat. The IMAGE technique produced reliable results for both the anterior camera angle as well as across different days. While the mean difference for the anterior condition was considered significantly greater than the posterior, the effect size of the difference was *trivial* and the intraclass correlation coefficient (ICC = 0.99) was *excellent*. The mean difference for the different day condition was not considered statistically significantly and the effect size was also *trivial* (ES=0.03) with an *excellent* ICC value (ICC=0.99). Results of this study indicate that for the estimation of %Fat be the IMAGE technique the anterior versus posterior body positions may be interchangeable. In addition, the IMAGE appears to be reliable when estimating %Fat between two days separated by no more than 48 hours.

Study 2 aimed at further exploring the reliability of the IMAGE technique for %Fat estimation across different conditions that may affect its performance outside of a laboratory setting. The results of the study indicate that the IMAGE produces reliable %Fat results with the use of three different megapixel cameras and under three different levels of room lighting, due to *trivial* mean differences, *excellent* ICC values of over 0.99, and small limits of agreement. However, different color backgrounds may affect the IMAGE's reliability. The black background condition displayed the largest mean difference, lowest ICC values, and widest limits of agreement, while the other colored backgrounds produced similar reliability results as the other previously mentioned conditions. For estimating %Fat, the IMAGE was reliable across a variety of conditions. It is recommended, for the most reliable results, that users perform measurements under standardized environmental conditions, using the same camera, lighting, and background.

The purpose of study 3 was to determine the extent to which individual error associated with BIA, SKF, and IMAGE relates to selected markers of physical activity and muscular fitness. The results of the study indicated that SKF_{ERROR} was positively associated with PU ($r = 0.17$, $p = 0.04$), while BIA_{ERROR} was not significantly correlated with any of the independent variables. In addition, $IMAGE_{ERROR}$ was positively correlated with the following variables: PU ($r = 0.41$, $p < 0.01$), IPAQV ($r = 0.24$, $p < 0.01$), 3mileSUB ($r = 0.17$, $p = 0.04$), 1mileSUB ($r = 0.24$, $p < 0.01$), and PA-R ($r = 0.24$, $p < 0.01$), indicating that greater error was observed for with higher levels of physical activity or muscular fitness. Follow-up stepwise linear regression analyses between the independent variables, SKF_{ERROR}, BIA_{ERROR}, and $IMAGE_{ERROR}$ revealed that only the PU test was independently associated with SKF_{ERROR} ($R^2 = 0.02$, $p < 0.001$) and $IMAGE_{ERROR}$ ($R^2 = 0.17$, $p < 0.01$), while no association was indicated for BIA_{ERROR}. Although, none of the physical activity and muscular fitness measures evaluated in this study correlated BIA_{ERROR}, it does appear that SKF and the IMAGE technique is influenced by muscular fitness. For the IMAGE technique muscular fitness as measured via PU may be a surrogate for anatomical width measurements.

Overall, this dissertation expands upon the current body of research regarding the IMAGE technique, and specifically its reliability across a variety of different conditions that may affect its ability to estimate %Fat. Additionally, it explored the relationship of select markers of physical activity and muscular fitness with field-based body composition measures. The collective findings indicate that the IMAGE produced acceptable levels of reliability across different conditions, although for the most reliable results the manufacturer guidelines should be followed, with consistent conditions for each measurement. Future research should continue to investigate the reliability and validity of the IMAGE technique across different conditions that may affect the accuracy outside of the laboratory, as well as further explore the effect of muscular fitness on the error observed with %Fat estimation in this technique.

REFERENCES

1. Lee JJ, Pedley A, Hoffmann U, Massaro JM, Fox CS. Association of changes in abdominal fat quantity and quality with incident cardiovascular disease risk factors. Journal of the American College of Cardiology. 2016;68(14):1509-21.

2. Garanty-Bogacka B, Syrenicz M, Goral J, Krupa B, Syrenicz J, Walczak M, et al. Changes in inflammatory biomarkers after successful lifestyle intervention in obese children. Endokrynologia Polska. 2011;62(6):499-505.

3. Jastreboff AM, Kotz CM, Kahan S, Kelly AS, Heymsfield SB. Obesity as a disease: the obesity society 2018 position statement. Obesity. 2019;27(1):7-9.

4. Lohnberg JA, Salcido L, Frayne S, Mahtani N, Bates C, Hauser ME, et al. Rapid conversion to virtual obesity care in COVID-19: Impact on patient care, interdisciplinary collaboration, and training. Obesity Science & Practice. 2022;8(1):131-6.

5. Scheuer SH, Færch K, Philipsen A, Jørgensen ME, Johansen NB, Carstensen B, et al. Abdominal fat distribution and cardiovascular risk in men and women with different levels of glucose tolerance. The Journal of Clinical Endocrinology & Metabolism. 2015;100(9):3340-7.

6. Goodpaster BH, Krishnaswami S, Harris TB, Katsiaras A, Kritchevsky SB, Simonsick EM, et al. Obesity, regional body fat distribution, and the metabolic syndrome in older men and women. Archives of Internal Medicine. 2005;165(7):777-83.

7. Heyward VH, Wagner DR. Applied body composition assessment: Human Kinetics; 2004.

8. Fedewa MV, Sullivan K, Hornikel B, Holmes CJ, Metoyer CJ, Esco MR. Accuracy of a Mobile 2D Imaging System for Body Volume and Subsequent Composition Estimates in a Three-Compartment Model. Medicine and Science in Sports and Exercise. 2020.

9. Sullivan K, Hornikel B, Holmes CJ, Esco MR, Fedewa MV. Validity of a 3-compartment body composition model using body volume derived from a novel 2-dimensional image analysis program. European Journal of Clinical Nutrition. 2021:1-8.

10. Sullivan K, Metoyer CJ, Hornikel B, Holmes CJ, Nickerson BS, Esco MR, et al. Agreement Between A 2-Dimensional Digital Image-Based 3-Compartment Body Composition Model and Dual Energy X-Ray Absorptiometry for The Estimation of Relative Adiposity. Journal of Clinical Densitometry. 2022;25(2):244-51.

11. Esco MR, Holmes CJ, Sullivan K, Hornikel B, Fedewa MV. Utilizing a Novel 2D Image Processing System for Relating Body Composition Metrics to Performance in Collegiate Female Rowers. International Journal of Environmental Research and Public Health. 2021;18(5):2413.

APPENDIX A:

24-HOUR HISTORY QUESTIONNAIRE

24-Hour History

ID ______________
Date ______________
Time ______________

1. How many hours of sleep did you get last night? (please circle one)
   1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 8.5 9 9.5 10 10.5 11 11.5 12 (hrs)

2. How many hours of sleep do you normally get? (please circle one)
   1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 8.5 9 9.5 10 10.5 11 11.5 12 (hrs)

3. How many hours has it been since your last meal or snack? (please circle one)
   1 **1.5** 2 **2.5** 3 **3.5** 4 **4.5** 5 **5.5** 6 **6.5** 7 **7.5** 8 **8.5** 9 **9.5** 10 **10.5** 11 **11.5** 12 **12.5** 13 **13.5** 14 **14.5** 15 (hrs)
   List the items below:

4. When did you last have:
   · a cup of coffee or tea?

   · cigarettes?

   · drugs (including aspirin)?

   · alcohol?

   · herbal or dietary supplements?

5. How many glasses of water or other beverages have you consumed in the last 24 hours?
   1 2 3 4 5 6 7 8 9 10 11 12 13 14

6. When did you last consume water or another beverage? _____________How much? ________ (glasses)

7. What sort of physical activity did you perform yesterday?

8. What sort of physical activity have you performed today?

9. Describe your general feelings by checking one of the following:

______excellent
______very, very good
______very good
______good
______neither good nor bad
______bad
______very bad
______very, very bad
______terrible

APPENDIX B:

INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE

INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE
(August 2002)

SHORT LAST 7 DAYS SELF-ADMINISTERED FORMAT

FOR USE WITH YOUNG AND MIDDLE-AGED ADULTS (15-69 years)

The International Physical Activity Questionnaires (IPAQ) comprises a set of 4 questionnaires. Long (5 activity domains asked independently) and short (4 generic items) versions for use by either telephone or self-administered methods are available. The purpose of the questionnaires is to provide common instruments that can be used to obtain internationally comparable data on health–related physical activity.

Background on IPAQ
The development of an international measure for physical activity commenced in Geneva in 1998 and was followed by extensive reliability and validity testing undertaken across 12 countries (14 sites) during 2000. The final results suggest that these measures have acceptable measurement properties for use in many settings and in different languages, and are suitable for national population-based prevalence studies of participation in physical activity.

Using IPAQ
Use of the IPAQ instruments for monitoring and research purposes is encouraged. It is recommended that no changes be made to the order or wording of the questions as this will affect the psychometric properties of the instruments.

Translation from English and Cultural Adaptation
Translation from English is supported to facilitate worldwide use of IPAQ. Information on the availability of IPAQ in different languages can be obtained at www.ipaq.ki.se. If a new translation is undertaken we highly recommend using the prescribed back translation methods available on the IPAQ website. If possible please consider making your translated version of IPAQ available to others by contributing it to the IPAQ website. Further details on translation and cultural adaptation can be downloaded from the website.

Further Developments of IPAQ
International collaboration on IPAQ is on-going and an ***International Physical Activity Prevalence Study*** is in progress. For further information see the IPAQ website.

More Information
More detailed information on the IPAQ process and the research methods used in the development of IPAQ instruments is available at www.ipaq.ki.se and Booth, M.L. (2000). *Assessment of Physical Activity: An International Perspective*. Research Quarterly for Exercise and Sport, 71 (2): s114-20. Other scientific publications and presentations on the use of IPAQ are summarized on the website.

INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE

We are interested in finding out about the kinds of physical activities that people do as part of their everyday lives. The questions will ask you about the time you spent being physically active in the **last 7 days**. Please answer each question even if you do not consider yourself to be an active person. Please think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

Think about all the **vigorous** activities that you did in the **last 7 days**. **Vigorous** physical activities refer to activities that take hard physical effort and make you breathe much harder than normal. Think *only* about those physical activities that you did for at least 10 minutes at a time.

1. During the **last 7 days**, on how many days did you do **vigorous** physical activities like heavy lifting, digging, aerobics, or fast bicycling?

_____ **days per week**

☐ No vigorous physical activities → ***Skip to question 3***

2. How much time did you usually spend doing **vigorous** physical activities on one of those days?

_____ **hours per day**

_____ **minutes per day**

☐ Don’t know/Not sure

Think about all the **moderate** activities that you did in the **last 7 days**. **Moderate** activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

3. During the **last 7 days**, on how many days did you do **moderate** physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking.

_____ **days per week**

☐ No moderate physical activities → ***Skip to question 5***

4. How much time did you usually spend doing **moderate** physical activities on one of those days?

_____ **hours per day**

_____ **minutes per day**

☐ Don't know/Not sure

Think about the time you spent **walking** in the **last 7 days**. This includes at work and at home, walking to travel from place to place, and any other walking that you have done solely for recreation, sport, exercise, or leisure.

5. During the **last 7 days**, on how many days did you **walk** for at least 10 minutes at a time?

_____ **days per week**

☐ No walking → ***Skip to question 7***

6. How much time did you usually spend **walking** on one of those days?

_____ **hours per day**

_____ **minutes per day**

☐ Don't know/Not sure

The last question is about the time you spent **sitting** on weekdays during the **last 7 days**. Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television.

7. During the **last 7 days**, how much time did you spend **sitting** on a **week day**?

_____ **hours per day**

_____ **minutes per day**

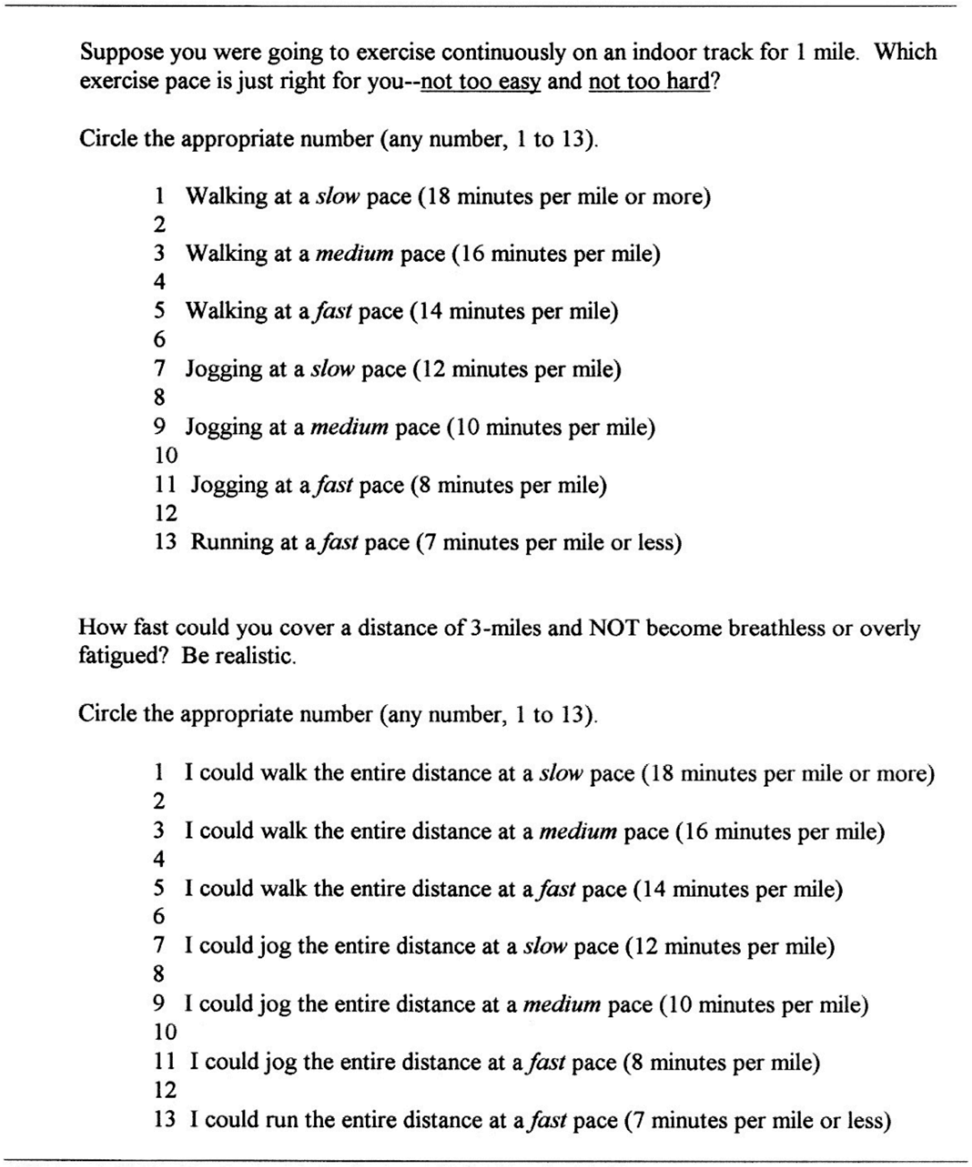
☐ Don't know/Not sure

This is the end of the questionnaire, thank you for participating.

SHORT LAST 7 DAYS SELF-ADMINISTERED version of the IPAQ. Revised August 2002.

APPENDIX C:

PERCEIVED FUNCTIONAL ABILITY QUESTIONNAIRE



Suppose you were going to exercise continuously on an indoor track for 1 mile. Which exercise pace is just right for you--not too easy and not too hard?

Circle the appropriate number (any number, 1 to 13).

1 Walking at a *slow* pace (18 minutes per mile or more)
2
3 Walking at a *medium* pace (16 minutes per mile)
4
5 Walking at a *fast* pace (14 minutes per mile)
6
7 Jogging at a *slow* pace (12 minutes per mile)
8
9 Jogging at a *medium* pace (10 minutes per mile)
10
11 Jogging at a *fast* pace (8 minutes per mile)
12
13 Running at a *fast* pace (7 minutes per mile or less)

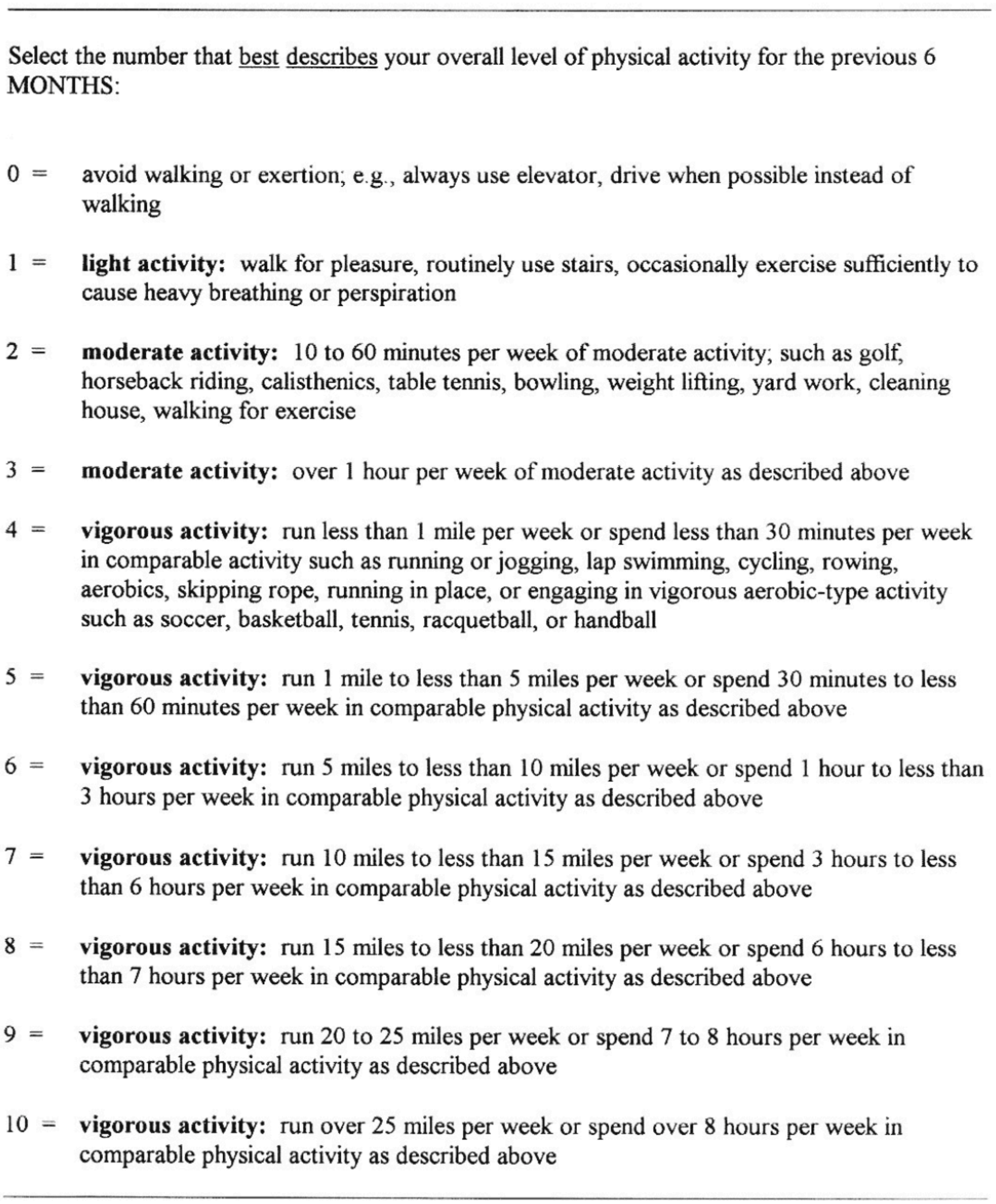
How fast could you cover a distance of 3-miles and NOT become breathless or overly fatigued? Be realistic.

Circle the appropriate number (any number, 1 to 13).

1 I could walk the entire distance at a *slow* pace (18 minutes per mile or more)
2
3 I could walk the entire distance at a *medium* pace (16 minutes per mile)
4
5 I could walk the entire distance at a *fast* pace (14 minutes per mile)
6
7 I could jog the entire distance at a *slow* pace (12 minutes per mile)
8
9 I could jog the entire distance at a *medium* pace (10 minutes per mile)
10
11 I could jog the entire distance at a *fast* pace (8 minutes per mile)
12
13 I could run the entire distance at a *fast* pace (7 minutes per mile or less)

APPENDIX D:

PHYSICAL ACTIVITY RATING QUESTIONNAIRE



Select the number that <u>best</u> <u>describes</u> your overall level of physical activity for the previous 6 MONTHS:

0 = avoid walking or exertion; e.g., always use elevator, drive when possible instead of walking

1 = **light activity:** walk for pleasure, routinely use stairs, occasionally exercise sufficiently to cause heavy breathing or perspiration

2 = **moderate activity:** 10 to 60 minutes per week of moderate activity; such as golf, horseback riding, calisthenics, table tennis, bowling, weight lifting, yard work, cleaning house, walking for exercise

3 = **moderate activity:** over 1 hour per week of moderate activity as described above

4 = **vigorous activity:** run less than 1 mile per week or spend less than 30 minutes per week in comparable activity such as running or jogging, lap swimming, cycling, rowing, aerobics, skipping rope, running in place, or engaging in vigorous aerobic-type activity such as soccer, basketball, tennis, racquetball, or handball

5 = **vigorous activity:** run 1 mile to less than 5 miles per week or spend 30 minutes to less than 60 minutes per week in comparable physical activity as described above

6 = **vigorous activity:** run 5 miles to less than 10 miles per week or spend 1 hour to less than 3 hours per week in comparable physical activity as described above

7 = **vigorous activity:** run 10 miles to less than 15 miles per week or spend 3 hours to less than 6 hours per week in comparable physical activity as described above

8 = **vigorous activity:** run 15 miles to less than 20 miles per week or spend 6 hours to less than 7 hours per week in comparable physical activity as described above

9 = **vigorous activity:** run 20 to 25 miles per week or spend 7 to 8 hours per week in comparable physical activity as described above

10 = **vigorous activity:** run over 25 miles per week or spend over 8 hours per week in comparable physical activity as described above

APPENDIX E:

APPROVED IRB CONSENT FORM

THE UNIVERSITY OF ALABAMA®
Office of the Vice President for
Research & Economic Development
Office for Research Compliance

September 1, 2021

Michael Fedewa, Ph.D.
Department of Kinesiology
College of Education
Box 870312

Re: IRB # 21-08-4874, "Reliability of body composition estimates from a single digital image"

Dear Dr. Fedewa:

The University of Alabama Institutional Review Board has granted approval for your proposed research. Your application has been given expedited approval according to 45 CFR part 46. Approval has been given under expedited review category 4 as outlined below:

(4) Collection of data through noninvasive procedures (not involving general anesthesia or sedation) routinely employed in clinical practice, excluding procedures involving x-rays or microwaves.

The approval for your application will lapse on August 31, 2022. If your research will continue beyond this date, please submit the continuing review to the IRB as required by University policy before the lapse. Please note, any modifications made in research design, methodology, or procedures must be submitted to and approved by the IRB before implementation. Please submit a final report form when the study is complete.

Please use reproductions of the IRB approved stamped consent forms and recruitment flyer.

Good luck with your research.

Sincerely,

[redacted]

Carpantato T. Myles, MSM, CIM, CIP, EXCS™
Director & Research Compliance Officer

Jessup Building | Box 870127 | Tuscaloosa, AL 35487-0127
205-348-8461 | Fax 205-348-7189 | Toll Free 1-877-820-3066

THE UNIVERSITY OF
ALABAMA®
Office of the Vice President for
Research & Economic Development
Office for Research Compliance

July 19, 2021

Michael Fedewa, Ph.D.
Assistant Professor
Department of Kinesiology
College of Education
The University of Alabama
Box 870312

Re: IRB # 19-OR-214-ME-R2 "Accuracy of Body Composition Estimates from a Single Digital Image in Normal Weight and Obese Subjects"

Dear Dr. Fedewa:

The University of Alabama Institutional Review Board has granted approval for your renewal application. Your renewal application has been given expedited approval according to 45 CFR part 46. Approval has been given under expedited review categories 4 & 7 as outlined below:

(4) Collection of data through noninvasive procedures (not involving general anesthesia or sedation) routinely employed in clinical practice, excluding procedures involving x-rays or microwaves.

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

The approval for your application will lapse on July 18, 2022. If your research will continue beyond this date, please submit a continuing review to the IRB as required by University policy before the lapse. Please note, any modifications made in research design, methodology, or procedures must be submitted to and approved by the IRB before implementation. Please submit a final report form when the study is complete.

Please use reproductions of the IRB approved informed consent forms to obtain consent from your participants.

Good luck with your research.

Sincerely,

[redacted]

Carpantato T. Myles, MSM, CIM, CIP
Director & Research Compliance Officer