INFLUENCE OF DIFFERENT ASSESSMENTS OF FAT PERCENTAGE ON NON-EXERCISE VO_{2MAX} ESTIMATION: A VALIDATION STUDY

Konsta Hyttinen

Sports and Exercise Medicine
Faculty of Sport and Health Sciences
University of Jyväskylä, Spring 2020

TIIVISTELMÄ

Hyttinen, K. 2020. Erilaisten kehon rasvaprosenttimittausten vaikutus maksimaalisen hapenoottokyvyn ”non-exercise” -arviointiin: validointitutkimus. Liikuntatieteellinen tiedekunta, Jyväskylän Yliopisto, Liikuntalääketieteen Pro Gradu -tutkielma, 61 s. 3 liitettä.

Rasvaprosenttimittauksen sisällyttäminen VO_{2max}:yn “non-exercise” -arviointiin on todettu tuottavan luotettavia tuloksia. Tutkielman tavoitteena oli arvioida erilaisten rasvaprosentti (rasva%) -mittausten vaikutusta VO_{2max}:yn arviointiin ilman hapenottokyvyn testausta. GE Lunar Prodigy DXA:n, InBody 720:n ja Huawei AH100:n rasva%-mittausten yhteneväisyyttä ja ekvivalenssia testattiin ja näiden eri laitteiden tuottamien rasva%:ien yhteyksiä maksimaalisen hapenottokyvyn kanssa arvioitiin Pearsonin korrelaatiokertoimilla. Lisäksi luotiin ennustemallit maksimihapenottokyvylle askeltavaa backward poistomenetelmää käyttäen jokaista käytettyä kehonkoostumusmittaria kohden. Tutkimukseen rekrytoitiin terveitä ja vaihtelevan kuntoisia miehiä ja naisia (ikä 18-54, n=278). Tutkimus oli poikkileikkausasetelmalla toteutettu ja mittaukset suoritettiin maantieteellisesti kahdessa paikassa: Jyväskylän Yliopiston ja KIHU:n liikunta- ja terveyslaboratoriossa Jyväskylässä sekä Shanghai Jiao Tong -yliopiston liikuntalaboratoriossa Minhangissa. Tutkimuksen analyysit tehtiin kuudessa eri analyysiryhmässä riippuen siitä, mihin mittauksiin tutkittavat olivat osallistuneet (GE Lunar Prodigy & InBody 720 n=146, GE Lunar Prodigy & Huawei AH100 n=95, InBody 720 & Huawei AH100 n=130, VO_{2max} & GE Lunar Prodigy n=146, VO_{2max} & InBody 720 n=278, VO_{2max} & Huawei AH100 n=125).

Huawei AH100 yliarvioi rasva%:ia DXA:aan verrattuna 2,5% (95% LOA -8,7-13,6, MAPE 15,37) ja havaittiin negatiivinen proportionaalinen harha mittausten välillä (p<0.001). Vastaavasti InBody 720 aliarvioi rasva%:n GE Lunar Prodigy DXA:iin nähden -3,8% (95% LOA -10,3-2,7, MAPE 16,75) ja jälleen proportionaalinen harha havaittiin (p<0.001). Huawei AH100 ja InBody 720 erosivat systemaattisesti toisistaan 4,6% (95% LOA -6,7-15,9, MAPE 25,69) eikä proportionaalista harhaa havaittu (p=0.634). Pearsonin korrelaatio VO_{2max}:yn ja GE Lunar Prodigy-rasva%:n välillä oli -0,81 (p<0.001), VO_{2max}:yn ja InBody 720-rasva%:n välillä vastaavasti -0,62 (p<0.001) ja VO_{2max}:yn ja Huawei AH100-rasva%:n välilllä oli -0,60 (p<0.001). Kolme luotua lopullista VO_{2max}:yn ennustemallia luotiin askeltavalla backward poistomenetelmällä ja mallit selittivät VO_{2max}:yn vaihtelusta 71%, 45% ja 40%, ja selkeästi eri analyysiryhmät ja eri laitteilla suoritetut kehonkoostumusmittaukset vaikuttivat malleihin. Ensimmäiseen malliin (SEE 3,72) ennustemuuttujiksi valikoituivat GE Lunar Prodigy DXA:n rasva% (p<0.001), BMI (p<0.001) ja ikä (p<0.001). Toiseen malliin (SEE 5,69) vastaavasti InBody 720:n rasva% (p<0.001), rasvaton pehmytkudosmassa (LBM) (p<0.001), BMI (p<0.001) ja ikä (p<0.001). Kolmanteen malliin (SEE 5,94) taas Huawei AH100:n rasva% (p=0.001), ikä (p=0.009) ja sukupuoli (p=0.044). Ensimmäinen malli osoittautui tarkimmaksi ja kolmas malli heikoimmaksi. Näissä kolmessa ennustemallissa VO_{2max}:n vaihtelun selittäjiksi osoittautuivat siis rasva%:n lisäksi BMI, ikä, LBM ja sukupuoli.

Eri laitteiden rasva%:n arvioinnit olivat toisistaan eroavia ja tilastollisesti mittaukset eivät olleet ekvivalentteja. Tämän vuoksi olisi suositeltavaa tulkita BIA-mittareiden antamia tuloksia varauksella. Toisekseen laitteiden kehonkoostumuksen arvioinnit ja analyysiryhmien erot vaikuttivat maksimaalisen hapenottokyvyn ja rasva% korrelaatioihin. Kehonkoostumuksen arviointimenetelmien ja analyysiryhmien erot vaikuttivat paljon myös luotujen ennustemallien lopullisiin ennustemuuttujiin ja mallien tarkkuuteen. Tutkielman tuloksista voidaan todeta, että tarkasti arvioitu kehonkoostumus voi parantaa maksimihapenottokyvyn ennustettavuutta. Toisaalta ”non-exercise” -menetelmiin liittyy myös muita epätarkkuutta aiheuttavia tekijöitä, joita tulee myös huomioida.

Asiasanat: maksimaalinen hapenottokyky, kehon rasvaprosentti, kehonkoostumus, validointi, yhtäläisyys, ekvivalenssi, ennustus

ABSTRACT

Hyttinen, K. 2020. Influence of different assessments of fat percentage on non-exercise VO_{2max} estimation: A validation study. Faculty of sport and health sciences, University of Jyväskylä, Sports and Exercise medicine Pro Gradu thesis, 61 p. 3 appendices.

Estimated fat% in non-exercise estimations of VO_{2max} can yield reliable predictions. The aim of the thesis was to assess the influence of different assessments of body fat percentage (fat%) on the estimation of VO_{2max}. The level of agreement and equivalence of the assessed fat% were tested between different devices: GE Lunar Prodigy DXA, InBody 720, and Huawei AH100. Their associations with VO_{2max} were evaluated by Pearson correlation coefficients and prediction models on VO_{2max} were created by stepwise backward elimination method regarding each body composition assessment device. Healthy men and women in diverse range in age (18-54), fitness, and body composition were recruited (n=278). The study was cross-sectional, and the data was collected at two centres: the sports laboratories of the University of Jyväskylä and the Research Institute for Olympic Sports in Jyväskylä and Shanghai Jiao Tong University in Minhang. The analyses conducted in six separate groups depending on the completed measurements by the subjects (GE Lunar Prodigy & InBody 720 n=146, GE Lunar Prodigy & Huawei AH100 n=95, InBody 720 & Huawei AH100 n=130, VO_{2max} & GE Lunar Prodigy n=146, VO_{2max} & InBody 720 n=278, VO_{2max} & Huawei AH100 n=125).

Huawei AH100 overestimated fat% compared to GE Lunar Prodigy by 2.5% (95% LOA -8.7-13.6, MAPE 15.37) and a negative proportional bias was found ($p<0.001$). The mean bias of InBody 720 to GE Lunar Prodigy was -3.8% (95% LOA -10.3-2.7, MAPE 16.75) and a negative proportional bias was found as well ($p<0.001$). Huawei AH100 and InBody 720 were systematically different from each other (4.6%, 95% LOA -6.7-15.9, MAPE 25.69) with no proportional bias ($p=0.634$). Pearson correlation between VO_{2max} and GE Lunar Prodigy-estimated fat% was -0.81 ($p<0.001$), between VO_{2max} and InBody 720-estimated fat% -0.62 ($p<0.001$), and between VO_{2max} and Huawei AH100-estimated fat% -0.60 ($p<0.001$). When using Lunar Prodigy-estimated fat% to predict VO_{2max}, the final predictors in the model were fat% ($p<0.001$), BMI ($p<0.001$), and age ($p<0.001$) which explained 71% (SEE 3.72) of the variance in VO_{2max}, indicating that, in addition to fat%, BMI and age contributed to the VO_{2max} variance. When using fat% assessed by InBody 720 to predict VO_{2max}, the final predictors in the model were fat% ($p<0.001$), lean body mass (LBM) ($p<0.001$), BMI ($p<0.001$), and age ($p<0.001$) which explained 45% (SEE 5.69) of the variance of VO_{2max}, showing that LBM, BMI and age are contributed to the VO_{2max} variance. When using fat% assessed by Huawei AH100, the final predictors in the model (SEE 5.94) were fat% ($p=0.001$), age ($p=0.009$), and gender ($p=0.044$) which explained 40% (SEE 5.94) of the variance of VO_{2max} indicating age and also gender contributed to the VO_{2max} variance. The first model turned out to be the most accurate and the third model the least accurate model to predict VO_{2max}.

The estimations of fat% between DXA GE Lunar Prodigy, InBody 720, and Huawei AH100 differed from each other. Because of the lack of statistical equivalence, it would be recommended to interpret the estimations of BIA devices with caution. Second, the use of different body composition assessment methods and study groups notably affected the inverse association between VO_{2max} and fat%. Due to these two affecting factors, the prediction models differed greatly from each other. From these results, it can be stated that the accurate assessment of body composition may lead to better predictions on VO_{2max}, although there are also other factors to be considered when estimating VO_{2max} without exercise.

Keywords: maximal oxygen consumption, body fat percentage, body composition, validation, agreement, equivalence, prediction

ABBREVIATIONS

AT	Adipose tissue
BIA	Bioelectric impedance analysis
BMI	Body mass index
B&A	Bland Altman method
CRF	Cardiorespiratory fitness
CI	Confidence interval
CV	Coefficient of variation
DLW	Doubly labelled water
DXA	Dual-energy x-ray absorptiometry
Fat%	Body fat percentage
FM	Body fat mass
FFM	Body fat-free mass
HR	Heart rate
HW	Hydrostatic weighing
LBM	Lean body mass
LM	Lean muscle mass
LOA	Limits of agreement
MAPE	Mean absolute percentage error
ME/MD	Mean error or mean difference
MRI	Magnetic resonance imaging
PA	Physical activity
SAT	Subcutaneous adipose tissue
SD	Standard deviation
SEE	Standard error of estimate
SEM	Standard error of measurement
TBW	Total body water
TEE	Total energy expenditure
VAT	Visceral adipose tissue
VO_{2max}	Maximal O_2 consumption

CONTENTS

ABSTRACT
1 INTRODUCTION 1
2 BODY COMPOSITION 3
2.1 Body composition assessment 3
2.2 Validity and reliability of the body composition assessment 7
2.2.1 Dual-energy x-ray absorptiometry 8
2.2.2 Bioelectrical impedance 9
2.3 Associations of adiposity with cardiometabolic health 10
3 CARDIORESPIRATORY FITNESS 12
3.1 Direct test of maximal oxygen consumption 13
3.2 Indirect assessment of maximal oxygen consumption 14
3.3 Predicting maximal oxygen consumption 15
4 BODY COMPOSITION AND MAXIMAL OXYGEN CONSUMPTION 17
5 RESEARCH QUESTIONS 19
6 METHODS 20
6.1 Study design and subjects 20
6.2 Measurements and protocols 22
6.2.1 Body composition assessment 23
6.2.2 Direct maximal oxygen consumption test 26
6.3 Analysis 26
6.3.1 Section 1. Agreement in body fat percentage between devices 26
6.3.2 Section 2. Prediction models of maximal oxygen consumption 28
7 RESULTS 29
7.1 Section 1. Agreement in body fat percentage between devices 29

7.2 Section 2: Prediction models of maximal oxygen consumption 35

8 DISCUSSION 37

8.1 Previous research 39

8.1.1 Body fat assessment 39

8.1.2 Non-exercise prediction of VO_{2max} 42

8.2 Strength and limitations 43

8.3 Ethical issues 45

8.4 Future studies 46

8.5 Conclusions 48

REFERENCES 49

APPENDICES

Appendix 1. Bland Altman analysis of agreement in fat% between devices. p. 62-70.

Appendix 2. Descriptive characteristics of subjects in *Section 2* of the study. p. 71-73.

Appendix 3. Linear relationships between measured and predicted VO_{2max}. p. 74-76.

1 INTRODUCTION

There has been a significant effort in research creating valid non-exercise methods to estimate VO_{2max} ($ml \cdot kg^{-1} \cdot min^{-1}$) and many independent predictors have been used in the equations. The most utilised are age, physical activity (PA) status, body mass index (BMI), gender, resting heart rate (HR), weight, height, waist circumference, smoking status, and body fat percentage (fat%) (Wang et al. 2019). This is reasonable as the direct VO_{2max} test requires a significant amount of time and effort, nor it is executable for all population groups, and it is a relatively expensive procedure. The most accurate prediction equations have included variables such as gender, age, objectively assessed PA, moderate- and vigorous PA, perceived functional ability, step counts, waist circumference, and BMI (George et al. 1997; Bradshaw et al. 2005; Cao et al. 2010a; Cao et al. 2010b, Wang et al. 2019). Besides, there is evidence that using estimated fat% in predictions can yield in reliable non-exercise estimations on VO_{2max} (Jackson et al. 1990; Heil et al. 1995; Whaley et al. 1995; Wier et al. 2006; Jackson et al. 2012). Whether the assessment method of body composition influences the prediction of VO_{2max}, has not yet been scientifically evaluated.

While dual-energy x-ray absorptiometry (DXA) as a three-compartment criterion method has been used as the reference commonly in research, the most recent studies have concluded a two-compartment laboratory-based single-frequency bioelectrical impedance analysis (BIA) even more accurate method in estimating body fat mass (FM) and fat-free mass (FFM) (Nickerson & Tinsley 2018). Many studies have indicated underestimation in FM and fat% and overestimation in FFM by commercial single- and multi-frequency BIA devices compared to DXA (Volgyi et al. 2008; Sillanpää et al. 2014; McLester et al. 2018; Moore et al. 2019). However, some other authors have found agreement or contradictory biases between devices, e.g. Ling et al. (2011) and Burns, Fu and Constantino (2019).

The first section of this study tested the agreement between the devices, and after that, the evaluation of the influence of these differences on the correlations with VO_{2max} and non-exercise models on VO_{2max} was possible. At baseline, the null hypothesis was that there are no differences between the devices and the assessments of fat% are equivalent to each other. Therefore, the influence of the different assessments is the same on the correlations with VO_{2max} and non-exercise models as well. This hypothesis was tested in this thesis. To the knowledge of the author, there are no previous studies that have particularly focused on the influence of different body composition assessments on the non-exercise estimation of VO_{2max} and apparently only a few studies are taking advantage of the equivalence testing and the use of the mean absolute percentage error (MAPE) in the analysis of agreement between body composition assessment methods.

Disclosure of the conflicts of interest. Huawei Technologies Oy (Finland) Co., Ltd. partly supported financially the data collection. As in this thesis, Huawei's device was evaluated, it is important to declare that Huawei Technologies Oy Co., Ltd. did not take part in the analysis, data collection, or the resulting conclusions in any way regarding this thesis. There are no other conflicts of interest to declare.

As I was involved in the whole research project, it has provided me with better insight into research and useful learning opportunities. I consider myself lucky that I was hired to conduct the measurements of this research in Jyväskylä and partly organise them. I wish to express my deepest gratitude to my research project supervisor prof. Sulin Cheng and thesis supervisor prof. Ina Tarkka, as well as the project team including Petri Wiklund PhD., Moritz Schumann PhD., Na Wu MSc., and Chris Chen BSc. for the invaluable help and support during this process. I also wish special thanks to Yanxiang Yang MSc., Katja Waller PhD., Heikki Peltonen PhD., and other staff members of sports laboratory of University of Jyväskylä and KIHU, and all others who have taught, supported or assisted me during this project and process.

2 BODY COMPOSITION

2.1 Body composition assessment

As widely known, there are many methodologies developed for measuring or estimating body composition and it is not straightforward to select the most appropriate one for a given purpose. Various factors, such as validity and reliability of measurement, availability, and costs of the measurement technique, safety, and subject cooperability, should be considered before selecting the body composition assessment method (Fosbøl & Zerahn 2015). Besides, the methods for body composition assessment are not always accurate for non-reference groups, i.e. there have been differences observed between methods in different age, ethnicity, sex, health status, etc. (Fosbøl & Zerahn 2015). This aspect, along with individual variation in the estimation (Altman 1990), needs to keep in mind when considering the method for assessment as methods can provide sufficiently accurate results for larger groups of subjects but not at the individual level (Fosbøl & Zerahn 2015).

Body composition can be quantified at several levels, such as atomic, molecular, cellular, and tissue (Roche 1996). According to Roche (1996) and Duren et al. (2008), it can be measured or estimated “at the atomic level with carbon, calcium, potassium, and hydrogen; at the molecular level by amounts of water, protein, and fat; at the cellular level with extracellular fluid and body cell mass; and at the tissue level for amounts and distributions of adipose, skeletal, and muscle tissues”. Accurate assessment from the atomic, molecular and cellular levels can be conducted with direct body composition methods such as neutron activation, isotope dilution, and total body counting (Roche 1996; Duren et al. 2008). For understanding the body composition assessment methods at the tissue level, it is necessary to be introduced to the multi-compartment models of body composition. A summary of the advantages and disadvantages of common body composition assessment methods are presented in TABLE 1. Indirect estimation is based on assumptions regarding the physical or chemical properties of the components that are not directly measured (Fosbøl & Zerahn 2015). These assumptions are based on the known average densities of body compartments measured by direct methods

(Fosbøl & Zerahn 2015). This is the reason the estimate of body composition can be inaccurate, especially at the individual level. The body composition assessment can be improved by using multicompartment models, where methods are combined to minimise the influence of FFM-assumptions (Baumgartner et al. 1991; Wang et al. 1998; Fosbøl & Zerahn 2015).

Currently, the most applied model in studies of body composition is a two-compartment model that divides body mass in FM and FFM, where FFM comprises water, protein, carbohydrates, and mineral (Fosbøl & Zerahn 2015). Adipose tissue does not equal to FM but consists of adipocytes, nerves, blood vessels, and extracellular fluid (Shen et al. 2005). Analogously, FM includes fat mass in adipose tissue and other tissues such as muscle and liver (Shen et al. 2005). FFM includes bone, skeletal muscle, organs, and connective tissue whereas lean body mass (LBM) is simply calculated as total BM minus FM and bone (Prado & Heymsfield 2014). It comprises total body water (TBW), total body protein, carbohydrates, non-fat lipids, and soft tissue mineral. The two-compartment models, such as BIA, cannot itself differentiate between FM, bone mass, and LBM (Dengel, Raymond & Bosch 2017, 27-28). However, the three-compartment model, such as dual-energy x-ray absorptiometry (DXA), can differentiate between bone mass and LBM (Dengel, Raymond & Bosch 2017, 27-28).

The accuracy of the two-compartment model, that assumes constant hydration status, can be improved by measuring TBW. The assumption of stable hydration status has limited validity as hydration varies with age (Lohman, 1986; Hewitt et al. 1993), gender, nutritional status (Waki et al. 1991) and diseases (Fosbøl & Zerahn 2015). A four-compartment model divides body composition in and assumes constant densities for four compartments, FM, water, residuals, and bone mineral at body temperature (Fosbøl & Zerahn 2015). A five-compartment model, further, considers the mineral content of soft tissue, i.e. soluble minerals and electrolytes in both intracellular and extracellular compartments of soft tissue, and a six-compartment model includes also glycogen to the equation (Fosbøl & Zerahn 2015).

TABLE 1. Summary of the advantages and disadvantages of common body composition assessment methods (applied from Dengel, Raymond & Bosch 2017, 27-37).

Compartments	Method	Advantages	Disadvantages
2	BMI	Simple, free, non-invasive, no assessor training needed	Too inaccurate at the individual level
(FM & FFM)	Circumferences, skinfold callipers	Simple, inexpensive, portable, minimal assessor training needed	Too inaccurate, assumptions, interobserver error
	BIA	Portable, simple, non-invasive, reproducible, relatively inexpensive, minimal investigator skills needed, requires little patient cooperation, can be considered reference field method after isotope dilution	Assumptions, sensitive to hydration status, PA, body temperature, and menstrual cycle. No regional estimates, validity is reference group-dependent
	Hydro-densitometry	Valid and reliable (body density, FM, fat%)	Assumptions, cost, patient burden, laboratory method, no regional estimates, investigator expertise needed, laboratory method
	Air displacement plethysmography	Precise and accurate, non-invasive, comfortable for patient, quick, safe, automated, suitable for various subjects (e.g. children, elderly)	Assumptions, cost, no regional estimates, lower validity compared to hydro-densitometry, laboratory method
	Dilution with isotopes	Precise and accurate	Cost, laboratory method

3 (FM, lean soft tissue, bone)	DXA	Cost and time-efficient, low radiation, safe, non-invasive, precise and accurate	Cost, measurement differences between and within manufacturers and software versions, sensitive to hydration status and tissue depth, investigator expertise needed, laboratory method
	Ultrasound	Non-invasive, quick, accurate and precise, capable of assessing total body and regional SAT and VAT, no radiation	Cost, investigator expertise needed, no standard measurement procedures, inherent artefacts may be present on the image produced
4 (FM, bone, muscle, other)	CT	Non-invasive, capable of assessing total and regional body composition, and quantifying SAT, AT, LM, and bone	Cost, accessibility, scan time, investigator expertise needed, high ionising radiation, laboratory method
	MRI	Non-invasive, image resolution, no radiation, capable of assessing total body and regional AT, LM, and bone, automatic and manual segmentation of tissues	Cost, accessibility, scan time, investigator expertise needed, laboratory method
4-6 (FM, water, protein, other)	Multi-compartment models	Precise and accurate	Cost, laboratory setting, use of multiple methods of body composition analysis, time-consuming, little further advantage compared to 3-compartment models

PA= physical activity, FM= body fat mass, Fat%= Body fat percent, FFM= fat-free mass, SAT= subcutaneous adipose tissue, VAT= visceral adipose tissue, AT= adipose tissue, LM= lean muscle mass.

2.2 Validity and reliability of the body composition assessment

The accuracy and reproducibility of methods might be often interpreted higher than they are because of additional causes of measurement error have not been considered when reporting or interpreting results (Nana et al. 2015). Reliability of measurement, e.g. DXA, is affected by technical diversity such as technical inconsistencies, changes between equipment, software or mode of analysis, subject positioning, clothing, preparation, technician-related errors, and biological variation such as nourishment, hydration, the effect of previous physical activity,. (Nana et al. 2015).

The validity and reliability depend on which analytical methods are used to test differences, agreement, and equivalence between assessment methods. More appropriate analytical methods are needed in the assessment of the measurement properties or methods of fitness assessment and physical activity (Staudenmayer, Zhu & Catellier 2012; Hopkins et al. 2009). Likewise, this concerns the indirect methods of body composition. Although Dixon et al. (2018) are discussing fitness- and energy expenditure assessment, precision is particularly important for answering to new research problems as studies in, e.g. fitness assessment, but also body composition, depend on the valid and reliable measures for accurate outcomes (Dixon et al. 2018). The lack of precision, validity, and reliability, concurrently, limits statistical power and truth of results (Dixon et al. 2018). The needs for cost savings and accurate field measurements by developing valid alternatives, such as indirect body composition assessment methods, are also important aims of the research (Dixon et al. 2018).

Accordingly, to test the agreement between two methods, the common analysis method has been the Bland Altman (B&A) analysis which enables error and bias to be visualised across the range of scores (Zaki et al. 2012). However, some note that the B&A method does not enable the degree of agreement to be directly quantified whereas tests of mean differences are routinely used to test systematic differences at the group level, but not equivalence, between the methods (Dixon et al. 2018). As the sample size determines the result of tests of mean differences, there is a risk of concluding erroneously tested methods as valid or not valid. Thus, despite the tests of mean differences are common in measurement agreement research, it does not make them correct (Dixon et al. 2018). Moreover, systematic differences do not detect errors at the individual level. Therefore, it is very informative to report the mean absolute percentage error (MAPE) or the root mean square error (RMSE) (Welk et al. 2019). MAPE presents the error as the average of the unsigned percent errors. It uses the absolute value of the difference before dividing by the criterion and therefore avoids the cancellation of under- and overestimation errors (Welk et al. 2019).

2.2.1 Dual-energy x-ray absorptiometry

The criterion methods of body composition measure the density of tissues and bones or describe amounts and distributions of skeletal muscle, and adipose tissues via x-ray or magnetic imaging techniques. Those methods include densitometry, computed x-ray tomography (CT), magnetic resonance imaging (MRI), and dual-energy x-ray absorptiometry (DXA) (Roche 1996; Duren et al. 2008).

DXA is a popular method in the assessment of body composition and it is based on discriminating three compartments; quantifying fat, lean, and bone tissues (Duren et al. 2008; Fosbøl & Zerahn 2015). Its measurement is fast, it can assess regional body composition, uses very low ionising radiation (~0.5 μSv), and it is non-intrusive (Nana et al. 2015). On the other hand, the equipment is expensive and non-portable, a trained technician is required, scanning bed is too small for larger or taller individuals, body composition estimation algorithms are developed for normal population or some specific groups, and algorithms might differ between device models or software versions (Nana et al. 2015). DXA technique is based on low-dose x-rays of two different energies which it passes through the body (Earthman 2015). As a result of it, an image is created as the photon detector measures the differential attenuation of the low and high x-ray energy by the soft tissue and bone (Earthman 2015).

2.2.2 Bioelectrical impedance

Indirect methods, including anthropometry and bioelectrical impedance analysis (BIA), provide estimates or indices of body composition based on results from direct or criterion methods (Roche 1996; Duren et al. 2008). The measurement technique of BIA is based on small electrical current with one or multiple frequencies which are conducted through the body via electrodes to measure electric impedance, resistance and reactance of the body tissues (Forbes 1994; Lukaski et al. 1985). The advantages of BIA are safety, observer-independency, low costs, and convenience (Fosbøl & Zerahn 2015). Today, there is a wide range of different BIA devices available. The most common commercial devices are bipolar (two-electrodes), which are conducting the small electrical current leg-to-leg or hand-to-hand. The traditional tetrapolar (four-electrodes) and more recently developed octopolar (eight-electrodes) BIA devices measure the impedance throughout the whole body, and, in theory, should be the most accurate of the BIA techniques (Carrion et al. 2019).

BIA technology is estimating TBW and uses equations to estimate body composition based on biological relationships for a specific population and reference data (Duren et al. 2008). Therefore, the mentioned equations are useful only for individuals close to the reference population in body size and shape (Duren et al. 2008). According to Fosbøl and Zerahn (2015), results from the BIA equation in each population should be cross-validated in a random sub-sample against estimates from a criterion method due to the mentioned fact (Fosbøl & Zerahn 2015). It is necessary to select a suitable BIA equation for a given sub-group because the relation between the ratio of the length of the cylinder (L2) and the resistance (R) (measure or BIA) and body composition varies with age, ethnicity, hydration, health status, etc. (Buchholz et al. 2004; Fosbøl & Zerahn 2015).

The BIA uses TBW and isotope-dilution in its estimations of body composition as water is the most abundant molecule in the body (Duren et al. 2008) and it has a stable relationship with FFM (Siri 1961; Chumlea et al. 2002; Chumlea et al. 2007). BIA technology utilises this relationship and uses TBW and isotope dilution as reference for FFM and body composition estimation (Siri 1961; Chumlea et al. 2002; Chumlea et al. 2007). The average proportion of TBW in FFM is 73 % and about 15-30 % of TBW is in adipose tissue as extracellular fluid and it increases with adiposity (Siri 1961; Chumlea et al. 2002; Chumlea et al. 2007).

The agreement of BIA compared to four-compartment methods is moderate. In estimating fat%, the bias of BIA has been -10-5% (LOA ±8%) in a healthy population. The mean bias in FFM,, has been 1.7-6.9 kg (Fuller et al. 1992; Wang et al. 1998; Jebb et al. 2000; Chouinard et al. 2007; Moon et al. 2013). In various populations, limited accuracy at the individual level in longitudinal changes in body composition has been observed, e.g. in obese adults (Evans et al. 1999; Minderico et al. 2008; Johnstone et al. 2014), athletes (Matias et al. 2012) and elderly healthy subjects (Moon et al. 2013). Recent studies have indicated underestimation in FM and fat% and overestimation in FFM by single- and multi-frequency BIA compared to DXA (Volgyi et al. 2008; Sillanpää et al. 2014; McLester et al. 2018; Moore et al. 2019) but not e.g. Ling et al. (2011). It has been stated that laboratory-based bioelectrical impedance spectroscopy might be a more accurate method for body composition than DXA when compared to a five-compartment model and that the assessment of TBW increases accuracy (Nickerson & Tinsley 2018). It has been, consequently, encouraged to use the two-compartment models with estimations of TBW over DXA as a reference method (Nickerson & Tinsley 2018).

2.3 Associations of adiposity with cardiometabolic health

Adiposity has been widely known to be associated with mortality and many causes of morbidity. The global trend in the prevalence of obesity is especially alarming. There has been a steady increase in the prevalence of high abdominal circumference in the general population from 10 to 20% in the 1960s to 40-60% in the year 2000 (Okosun et al. 2004).

In adults, BMI levels above 25 are associated with an increased risk of morbidity and mortality with BMI levels of 30 and greater indicating obesity (WHO 1998, 7-16; Chumlea & Guo 2000). The high abdominal circumference is linked with increased risk for morbidity, specifically, type 2 diabetes and the metabolic syndrome and mortality (Pouliot et al. 1994; Despres et al. 1991;

Nicklas et al. 2004). Central adiposity has also been associated with dementia in a large longitudinal study (Whitmer et al. 2008). Most men with the abdomen-to-hip ratio greater than 1.0 and women with a ratio greater than 0.85 are at increased risk for cardiovascular disease, diabetes, and cancers (Seidell et al. 1987; Fujimoto et al. 1991).

Abdominal adiposity includes both visceral and subcutaneous areas of adipose tissue and they can only be divided from each other by more sophisticated assessment methods, e.g. CT or MRI. Regional accumulation of abdominal subcutaneous and visceral adipose tissue are the major contributors to cardio-metabolic dysfunction (Bjorntorp et al. 1990). They are important clinical targets for reducing the risk of metabolic diseases and can be primarily affected by physical activity (de Lannoy & Ross 2019, 245).

Especially visceral adipose tissue has been linked to a higher risk for type 2 diabetes (Neeland et al. 2012), hypertension (Chandra et al. 2014), cardiovascular disease, dementia, and mortality (Katzmarzyk, Mire & Bouchard 2012; Neeland et al. 2015). Additionally, visceral adipose tissue seems to be associated with higher intrahepatic fat (i.e. liver fat) and insulin resistance, although it was found that only the liver fat explained the variation in serum insulin concentrations (Westerbacka et al. 2004). Thus, visceral adipose tissue seems to be a more significant determinant of cardiometabolic health than total body fat or subcutaneous adipose tissue.

3 CARDIORESPIRATORY FITNESS

Aerobic capacity (VO_{2max}) is defined as the ability of the cardiorespiratory system to supply oxygen for the working muscle tissues and the ability of the muscle tissues to utilise the oxygen as an essential part of energy production (McArdle et al. 2015, 249-453). Aerobic capacity is affected by physiological systems such as respiratory, circulatory, energy metabolic, endocrinologic and nervous regulatory systems (McArdle et al. 2015, 249-453). The determinants of VO_{2max} are stroke volume (SV), muscle tissue mass, haemoglobin mass, blood flow to the active muscles, arteriovenous oxygen difference (a-v O_2-diff) and the efficiency of oxygen utilisation by the tissues (Joyner & Coyle 2008, 35-44). A poor condition of the cardiorespiratory system, aerobic fitness, and low physical activity are associated with mortality (Myers et al. 2002; Sui et al. 2007) and many types of morbidity (Booth et al. 2012).

PA is the primary modifiable determinant of VO_{2max}. Increasing the PA leads eventually to improvements in oxygen delivery and uptake of the exercising muscles (de Lannoy & Ross 2019, 230-231). Thus, according to de Lannoy and Ross (2019), VO_{2max} can be considered as an alternative measure of the amount of PA and an objective measure of the effect of PA in the long-term (de Lannoy & Ross 2019, 230-231). The association between PA and VO_{2max} is, however, only moderate (Williams 2001; Myers et al. 2004) which might be related to subjective questionnaires of PA their poor correlation with direct measures of PA (Prince et al. 2008). VO_{2max} is affected by the quantity and quality of PA, and it can be improved by endurance training (Jones & Carter 2000; Milanovic et al. 2015). Although aerobic capacity is a characteristic that can be improved, according to evidence, the improvements are rather intensity-dependent than simply dose-dependent (Gormley et al. 2008).

The aerobic fitness of an individual is also determined by body composition. Body mass and body fat greatly determines the relative VO_{2max} ($ml \cdot kg^{-1} \cdot min^{-1}$) whereas the amount of activated muscle tissue, especially lower body muscle ($r=0.95$; $p<0.001$), determines the absolute VO_{2max}

($L \cdot min^{-1}$) (Sanada et al. 2005). Also, aerobic fitness is a highly heritable characteristic (Bouchard et al. 1999; Kujala 2016; Joyner 2017).

3.1 Direct test of maximal oxygen consumption

For measurement of aerobic capacity, generally, the direct test of VO_{2max} has been used, and it has been proved to be a valid measure of assessing aerobic capacity and endurance performance (Keskinen et al. 2007, 78-117). The measurement of VO_{2max} is accurate in controlled conditions, and different device manufacturers report 1-2% accuracies of their direct gas analysers. It is also known that the test design, and the length of the test, affect VO_{2max} results due to sport specificity and fatigue accumulation (Keskinen, Häkkinen & Kallinen 2018, 80-100). It has been stated, understandably, that the criteria used to determine VO_{2max} must be the same between tests (Howley, Bassett & Welch 1995). The coefficient of variation for VO_{2max} measurements has been considered around 4% (Katch et al. 1982).

VO_{2max} can be determined by progressively increasing the level of the aerobic demand for energy production until the individual reaches the level of exhaustion or by increasing the speed of a treadmill or the resistance of a bicycle ergometer progressively until the exhaustion when the body's ability to utilise oxygen reaches its peak value (Keskinen 2016, 110-113). VO_{2max} means the highest possible measured oxygen consumption at the maximal load, where the performance continues beyond total exhaustion. The measurement of maximal oxygen consumption results the so-called absolute VO_{2max}, i.e. how much oxygen the body has consumed in $L \cdot min^{-1}$ and often the result of the test is also calculated as relative VO_{2max}, when the result is expressed in relation to the individual's weight, $ml \cdot kg^{-1} \cdot min^{-1}$ (Kenney et al. 2012, 249). The oxygen consumption and the exhalation of the carbon dioxide are measured by the respiratory gas analyser (Keskinen 2007, 78-117).

According to Shvartz and Reibold (1990), the average values of men in absolute VO_{2max} in the North American and European population at age 18 is 3.4 $L \cdot min^{-1}$ and at 30 is 3.2 $L \cdot min^{-1}$. Correspondingly, women’s average values at age 18 and 30 are 2.2 and 1.8 $L \cdot min^{-1}$. In the relative VO_{2max}, gender differences are less pronounced. Men have, on average, values of 50 and 48 ml·kg^{-1}·min^{-1} at age 18 and 30. Similarly, women have values of 44 and 41 ml·kg^{-1}·min^{-1} at the same age (Shvartz & Reibold 1990).

3.2 Indirect assessment of maximal oxygen consumption

Indirect VO_{2max} assessment methods are also sufficiently accurate and safe for many purposes for normal populations (Keskinen et al. 2007, 78-117; Suni & Vasankari 2011, 32-35). Sub-maximal VO_2 tests, evaluate the individual absolute or relative VO_{2max}, but the difference compared to the direct method is that the result will be estimated according to an equation or an algorithm (Suni & Vasankari 2011, 32-35). Many indirect methods are suitable for evaluating the change in aerobic fitness when the used protocol and procedures are the same between tests but there should not be a comparison between the results of different test protocols (Suni & Vasankari 2011, 32-35). The most accurate and reproducible methods of indirect assessment are those based on maximum tests. These tests achieve oxygen consumption as close as possible to individuals own maximal performance (e.g. Ross & Jackson 1991). In the sub-maximal tests carried out on bicycle ergometers, the VO_{2max} has been estimated to be between 7% and 27% of the actual VO_{2max} (Margaria et al. 1965; Davies 1968; Fox 1973; Terry 1977; Greive et al. 1995).

3.3 Predicting maximal oxygen consumption

Direct and indirect VO_{2max} measurements are time-consuming, require expensive equipment, and motivated test participants. Therefore, it has been considerable interest of many researchers in finding more convenient ways of predicting VO_{2max} based on prediction equations derived from exercise- and non-exercise variables (Stahl et al. 2006; Rexhepi & Brestovci 2014). Sanada et al. (2007) found in young Japanese men that the use of thigh muscle mass (r=0.55) and cardiac dimensions (r=0.72-0.74) is a valid method to predict absolute VO_{2max}. Their model reached R^2 of 83% in cross-validation ($p<0.01$) and the standard error of estimate (SEE) was 0,39 $L \cdot min^{-1}$ (Sanada et al. 2007). Their prediction was not significantly different from the measured VO_{2max} in the validation group (Sanada et al. 2007).

Stahn et al. (2006), on the other hand, concluded that using BIA might be an accurate, rapid, and convenient method for predicting absolute VO_{2max} in young and fit men and women. They created a predictive model including impedance index (H^2/Z), age, gender, and self-reported physical activity which accounted for 88% of the variance in VO_{2max} (SEE=258 $ml \cdot min^{-1}$) (Stahn et al. 2006). Afterwards, Moon et al. (2011) validated Stahl et al. (2006) predicting equation. They stated that using the equation to prescribe exercise, will yield underestimated exercise intensities and that the predictive equations are highly specific to the sample where they derive from because the validation group differed significantly from the sample of the study of Stahl et al. (2006) (Moon et al. 2011). Same has been stated on specificity in previous studies on prediction equations of cardiorespiratory fitness (Malek et al. 2004).

Rexhepi and Brestovci (2014) studied active football players aged 16–35 years. Their predictive regression equation could explain 26% of the variance in VO_{2max} ($L \cdot min^{-1}$) and was no different from the actual VO_{2max} (t=-0.28, p=0.78). The model included age, body mass, and resting heart rate as independent variables (Rexhepi & Brestovci 2014). Wier et al. (2006) compared regression models of waist girth (r=0.81, SEE=4.80), fat% (r=0.82, SEE=4.72), and BMI (r=0.80, SEE=4.90) to predict VO_{2max} ($ml \cdot kg^{-1} \cdot min^{-1}$). With these variables, they included gender, age, and PA into the model and concluded them accurate in estimating VO_{2max} but not in individuals at the extremes of fitness (Wier et al. 2006). Previous to that, e.g. Jackson et al. (1990) also predicted VO_{2peak} using fat%, gender, age, and PA (r=0.81, SEE=5.35) as well as with BMI, gender, age, and PA (r=0.78, SEE=5.70) (Jackson et al. 1990).

It has been also found that self-reported PA along with age, gender, BMI and resting HR, are good predictors of maximal energy expenditure measured by maximal or submaximal exercise test (r^2=0.58-0.65, SEE=5.08-6.90) (Jurca et al. 2005). The best non-exercise equations have also utilised objectively assessed total-, moderate- and vigorous PA, perceived functional ability, and step counts (George et al. 1997; Bradshaw et al. 2005; Cao et al. 2010a; Cao et al. 2010b). Similarly, the total energy expenditure (TEE) measured by doubly labelled water method (DLW) can be predicted by accelerometer-derived PA counts and HR relatively well (Plasqui & Westerterp 2007; Zhusheng et al. 2012). In conclusion, maximal or total energy expenditure can be estimated indirectly relatively well in a specific reference population.

4 BODY COMPOSITION AND MAXIMAL OXYGEN CONSUMPTION

Maximal oxygen consumption is associated with body composition. Research has found both positive and negative associations between cardiorespiratory fitness and body composition. Firstly, it is necessary to differentiate absolute VO_{2max} ($L \cdot min^{-1}$) and VO_{2max} relative to BM ($ml \cdot kg^{-1} \cdot min^{-1}$) because associations of body composition with both have been studied.

Correlation between BMI and VO_{2max} ($ml \cdot kg^{-1} \cdot min^{-1}$) in men and women is from weak to moderate (r=-0.32 to -0.41) (Jackson et al. 1990; Wier et al. 2006; Pribis et al. 2010; Mondal and Mishra 2017). Besides that, some studies have also found very high correlations, e.g. in young healthy Indian males (r=-0.88, $p<0.05$) (Prabha et al. 2014). Absolute VO_{2max} does not seem to correlate with BMI or correlates weakly (Maciejczyk et al. 2014).

The association between FFM and relative VO_{2max} has been weak in both genders (r=0.37, $p<0.01$) (Mondal & Mishra 2017). Between absolute VO_{2max} and FFM or skeletal muscle mass, very strong correlations have been observed in young adult males (r=0.78-0.85) (Buskirk & Taylor 1957; Sanada et al. 2007). However, Maciejczyk et al. (2014) concluded only a moderate correlation between these variables in college-aged men (r=0.38, $p<0.05$) (Maciejczyk et al. 2014)

Buskirk and Taylor (1957) concluded long-ago that as the inverse relationship between fat% and VO_{2max} relative to BM is evident, there is no significant difference in absolute VO_{2max} in groups of different body fat. They also proved in 1957 that excess fat itself does not impact the capacity of the cardiorespiratory system to deliver oxygen to muscles under maximal performance conditions (Buskirk & Taylor 1957). Nevertheless, excess body weight and fat mass are still related to decreased running performance in distance e.g. in a 12-minute-running test (Mattila et al. 2007). Interestingly, concerning the above, Farrell et al. (1985) found that in obese women, absolute VO_{2max} ($L \cdot min^{-1}$) and the VO_{2max} relative to FFM ($ml \cdot FFM^{-1} \cdot min^{-1}$) on each incline of the submaximal treadmill test was higher than in lean counterparts ($p<0.05$).

The VO_{2max} relative to BM, however, was significantly lower in obese women due to their higher body weight ($p<0.05$) (Farrell et al. 1985).

Indeed, the inverse correlation between fat% and VO_{2max} ($ml \cdot kg^{-1} \cdot min^{-1}$) is strong in adults ($r=-0.65--0.93$) (Jackson et al. 1990; Wier et al. 2006; Sharma et al. 2016; Mondal & Mishra 2017) but moderate inverse correlations have been concluded as well (Amani et al. 2010; Pribis et al. 2010). However, one study did not find a relationship between fat% and VO_{2max} ($ml \cdot kg^{-1} \cdot min^{-1}$) in young female athletes (Shete et al. 2014). Goran et al. (2000), similarly, found a negative (-3.7%, $p<0.05$) difference in absolute VO_{2max} after weight loss in obese women and a significant increase in VO_{2max} relative to BM (+15%, $p<0.05$) whereas the change in VO_{2max} relative to FFM was not significant (Goran et al. 2000). They (2000) stated that the VO_{2max} relative to FFM could be a valid indicator for comparing VO_{2max} in adults of different body size and body fat composition as the outcome does not vary with the change in body composition (Goran et al. 2000).

5 RESEARCH QUESTIONS

The objective of this thesis is to assess the influence of estimations of fat% on the non-exercise estimation of VO_{2max}. Specifically, the level of agreement and equivalence of the measured fat% between DXA (GE Lunar Prodigy Advance, Madison, WI, USA with software version 9.3.), InBody 720 (Biospace Co., Ltd, Korea), and Huawei AH100 (Huawei Technologies Co., Ltd.) are tested. Secondly, the associations of the fat% measurements with VO_{2max} are evaluated and prediction models created.

The research questions:

1. Does the measured fat% by different devices provide a similar estimate?

   H0: There are no differences between the devices. The level of agreement is determined by 95% limits of agreement (±1.96 SD). The criterion for equivalence is set to ±10%. Individual-level error is determined by the Mean Absolute Percentage Error.

2. Are the non-exercise estimations of VO_{2max} different from each other created by the stepwise backward elimination method? The multiple linear regression models are mainly based on estimations of fat% of the different body composition assessment devices.

   H0: The Pearson correlation coefficients with VO_{2max} are similar between the three different fat% assessments. The models consist of the same independent variables chosen according to the stepwise backward elimination method and provide similar non-exercise estimates of VO_{2max}.

6 METHODS

6.1 Study design and subjects

The present study was a cross-sectional and part of a validation research project of body composition assessment-, wearable fitness- and wellness -devices. It was approved by the Ethical Review Board of the University of Jyväskylä and Shanghai Jiao Tong University Bio-X Ethics Board. The measurements were conducted partly at sports laboratory of the University of Jyväskylä, Research Institute for Olympic Sports (KIHU), and partly at sports laboratory of Shanghai Jiao Tong University during winter 2018-2019, from November to April. Only relatively healthy male and female subjects were recruited for the study, with varying age, levels of physical activity, aerobic fitness, and body composition. The inclusion and exclusion criteria are presented in TABLE 2.

TABLE 2. The inclusion and exclusion criteria for participation.

Inclusion criteria
Age 20 – 45 (in *SJTU, 18 – 45)
Ability to participate in high-intensity exercise and the treadmill running test
Ability to participate in body composition measurements
Exclusion criteria
BMI > 38
Cardio-metabolic diseases, high blood pressure, joint- or skeletal muscle problems
Cardio-metabolic symptoms such as arrhythmias, palpitations, dizziness, chest pain, dyspnoea, etc.
Acute illness or infections
Pregnancy
Issues with bones, joints or ligaments

* SJTU = Shanghai Jiao Tong University

Recruitment was conducted from multiple sources as a convenience sample. Email lists of the University of Jyväskylä, JAMK University of Applied Sciences, and GRADIA Jyväskylä Educational Consortium were used as well as social media, flyers, word of mouth, the local newspaper, and website of the city of Jyväskylä were used to ensure enough quantity and variety in subjects. In Shanghai, the recruitment was conducted via word of mouth and advertisement in Shanghai Jiao Tong University and the local community.

Total group sizes and participation in measurements are presented in TABLE 3. The total number of subjects included in the analyses was n=146 from Jyväskylä and n=138 from Shanghai. Prior to any measurements, subjects signed a written informed consent form. The study was divided into Section 1 and Section 2. In Section 1, the aim was to determine the level of agreement and equivalence in fat% measurement between the three devices. Section 1 included three groups consisting of subjects who completed both, DXA (GE Lunar Prodigy)- and Huawei AH100-, DXA (GE Lunar Prodigy)- and InBody 720-, and Huawei AH100- and InBody 720 -body composition analyses (TABLE 3). The same study subjects are in the three groups. These three groups were used in the analysis. In Section 2, the objective was to determine predictive multiple linear regression equations for VO_{2max} for each body composition assessment device. Section 2 included subjects who completed both VO_{2max}- and DXA (GE Lunar Prodigy) -measurement, VO_{2max}- and InBody 720 -measurement, as well as VO_{2max}- and Huawei AH100 -measurement. The characteristics of the subjects of Section 1 and Section 2 of the study are presented in TABLE 4 and APPENDIX 2.

TABLE 3. The description of which measurements were done in the study groups, the group sizes, and a description of which ethnical groups are represented in the study groups.

Section 1. Agreement in fat%			Section 2. Multiple linear regression models for VO_{2max}		
Measurements	n	Population	Measurements	n	Population
GE Lunar Prodigy InBody 720	146 M 74 F 72	Finnish	Direct VO_{2max} test GE Lunar Prodigy	140 M 69 F 71	Finnish
GE Lunar Prodigy Huawei AH100	95 M 50 F 45	Finnish	Direct VO_{2max} test InBody 720	278 M 164 F 114	Finnish and Chinese
InBody 720 Huawei AH100	130 M 73 F 53	Finnish and Chinese	Direct VO_{2max} test Huawei AH100	125 M 71 F 54	Finnish and Chinese

M= male, F= female.

6.2 Measurements and protocols

The measurements consisted of two laboratory visits and the first visit included a risk assessment, basic information collection by a questionnaire, and a body composition assessment. The second visit included a direct VO_{2max} test on a treadmill.

For the first visit, subjects were instructed to arrive in a fasted state, and they were allowed only to consume water, but not caffeine, alcohol, or smoking. For the second visit, subjects were instructed to avoid strenuous physical activity two days before the VO_{2max} test, and they were instructed to hydrate and eat properly 1-2 hours before the test. Additionally, the researchers ensured that the subjects were symptomless of any infection or illness in the last seven days prior to the VO_{2max} test or had any other contraindications regarding the measurements.

6.2.1 Body composition assessment

Body composition measurements were performed after overnight fasting (12 h). Measurement was done in lightweight clothing and all metal items were removed from the subjects to ensure the accuracy of the measurements. Written instructions were given to subjects before measurement. The instructions guided the subjects for the overnight fasting, hydration and encouraged avoidance of alcohol and caffeine on the previous and the measurement day and avoidance of intense exercise in 24 h before the measurement. Menstrual cycle was not controlled for women. Height was determined using a fixed wall-scale measuring device to the nearest 0.1 cm. Weight was determined within 0.1 kg for each subject using an electronic scale (InBody 720, Biospace Co., Ltd, Korea). BMI was calculated as weight (kg) per height $(m)^2$.

Dual-energy X-ray absorptiometry (DXA) (GE Lunar Prodigy Advance, Madison, WI, USA with software version 9.3.) can measure whole-body FM, fat%, FFM, LBM, skeletal mineral content, as well as other variables (FIGURE 1). DXA quantifies the amount of bone and soft tissue by using the differences in the absorption of high energy and low energy X-ray photons of the body tissues. It uses algorithms to determine three compartments of the body composition: the amount of bone, lean, and fat mass. The K-edge filter of DXA separates the spectrum into two energy peaks, low 38 and high 70 keV, and by this technique, differentiates bone and soft tissues from each other.

FIGURE 1. GE Lunar Prodigy Advance DXA (adapted from Absolute Medical).

DXA measurement was done only at the University of Jyväskylä. The Prodigy software automatically set the Finland Total Body Composition Reference Population (v113) for the measurements. The subjects were positioned in a supine position, arms at the side, on the centre of the table for each scan and four positioning aids were placed for separating arms from the body and two positioning straps were used to keep legs in a consistent position. The researcher assured the proper centring of the subject. In tall participants who exceeded the range of the device, exceeding feet area was left out of measurement. Participants were scanned using the default scan mode automatically selected by the software. Quality control procedures recommended by the manufacturer were conducted by the University of Jyväskylä prior to each day and on a weekly basis. Data were electronically imported to Excel using the Prodigy software.

InBody 720 (Biospace Co., Ltd, Korea with Lookin'Body software) (FIGURE 2) is a multi-frequency bioelectrical impedance plethysmograph body composition, and it emits a multitude of frequencies including 1kHz, 5kHz, 50kHz, 250kHz, 500kHz, and 1MHz. It uses an 8-point tactile electrodes method (FIGURE 3) and BIA technique, measuring impedance at six and reactance at three frequencies. TBW is calculated first with a measured impedance value and the value of FFM using the TBW is calculable. FM is then determined by deducting the FFM from the measured weight. The multi-frequency technology can also separate the intracellular water from the extracellular water. During the measurement, the researcher assured the proper positioning of the subject on the device. Data were electronically imported to Excel using Lookin’Body software. Quality control procedures recommended by the manufacturer were conducted by the University of Jyväskylä, and Exercise, Health, and Technology centre of Shanghai Jiao Tong University. The similar device was used in both measurement centres (InBody 720).

Huawei AH100 (Smart Scale, Huawei Technologies Co., Ltd. with Huawei Body Fat Scale mobile application) (FIGURE 2) is a single-frequency bioelectrical impedance body composition analyser, and it uses a 2-point tactile electrode method (FIGURE 3). In addition to weight, AH100 can estimate fat%, protein mass, visceral fat, muscle mass, TBW, and bone mass using the single-frequency method. Data was automatically synchronised from the Smart Scale to the mobile application and from the application manually entered to Excel. The similar device was used in both measurement sites (Huawei AH100).

FIGURE 2. The used BIA devices in this study. On the left InBody 720 (adapted from InBody) and on the right Huawei AH100 Smart Scale (adapted from Huawei).

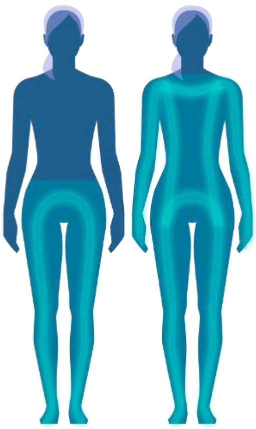


FIGURE 3. Four (two-point)- and eight (four-point)-electrode measurement methods. Huawei AH100 measures foot-to-foot by four electrodes and InBody 720 feet-to-hands by eight electrodes (adapted from Huawei).

6.2.2 Direct maximal oxygen consumption test

The direct maximal oxygen consumption test was performed on a treadmill. The test started from a speed of four kilometres per hour and continued until volitional exhaustion of maximal performance. Increments of the test were four minutes long, including the collection of lactate sample from the subject's fingertip. Oxygen consumption was measured breath-by-breath throughout the test using a portable gas analyser Oxycon Mobile, Jaeger (Germany) in Jyväskylä and K4b2, COSMED (Italy) in Shanghai.

The test protocol proceeded as it started with 5 minutes of data collection at rest (sitting), followed by the test, starting from speed 4 km/h and increased by 1 km/h on each increment, continuing thereafter until maximum. The gradient of the treadmill maintained at 1% throughout the test. Blood lactate was measured at baseline and the end of each increment. After the end of the test, the recovery data of 5 minutes was collected. The VO_{2max} was determined as the best 60-second average of VO_2.

6.3 Analysis

Statistical analysis of data was conducted by Statistical Package for the Social Sciences (SPSS version 24.0; IBM Inc., Chicago, IL) and MedCalc Statistical Software version 19.0.3 (MedCalc Software bvba, Ostend, Belgium). Normality of distribution for the variables was determined by the Kolmogorov-Smirnov (n>50) test, analysis of histograms, Q-Q -plots as well as outliers. Significance was set at $p<0.05$.

6.3.1 Section 1. Agreement in body fat percentage between devices

Paired samples t-test was conducted to assess the difference in fat% measurements between Huawei AH100 and GE Lunar Prodigy, InBody 720 and GE Lunar Prodigy as well as Huawei AH100 and InBody 720. Non-parametric Related samples Wilcoxon Signed-Ranks test was conducted between Huawei AH100 and InBody 720 where assumptions of the parametric testing were not fulfilled in females (normality of distribution). Associations and linearity between the devices were defined as Pearson correlation (r) and Spearman's rank correlation (r_s) coefficients. Spearman's rank correlation was used for the analysis between Huawei AH100 and InBody 720. All the subjects completing the measurements were included in the analysis.

Bland Altman method (B&A) was used to examine the difference and agreement between the body composition assessment devices with 95% limits of agreement (±1.96 SD) (APPENDIX 1), to estimate the error at the group level, and visualise proportional errors. Differences were approximately normally distributed allowing B&A analysis. The mean absolute percent error (MAPE) was calculated to estimate the expected error at the individual level. Equivalence between devices was evaluated with a ±10% equivalence interval of the mean of the reference method GE Lunar Prodigy DXA, using the confidence interval method at the level of $p<0.05$. The equivalence interval was set to be constant throughout all analyses due to the interval is narrower in males because of lower mean fat%. Therefore, the equivalence interval derived from the total sample average (n=146, DXA mean= 23.91, ±10% EqInt= ± 2.4).

Equivalence testing was employed to test the null hypothesis that there was no equivalence between the reference (DXA) and surrogate assessments of fat%. As described by Dixon et al. (2018), a 90% Confidence Interval needs to be calculated for the difference in means. If a calculated mean difference 90% CI falls entirely within the equivalence interval, the null hypothesis will be rejected and it will be concluded that the two observed assessments of body composition are statistically equivalent (Dixon et al. 2018).

The MAPE was calculated using the comparison of multiple methods in MedCalc software, GE Lunar Prodigy as the reference method. In comparison between Huawei AH100 and InBody 720, InBody 720 was used as the reference. MAPE presents the error as the average of the unsigned percent errors (Welk et al. 2019). In addition to B&A plots, fat% tertile-differences determined by Paired samples T-test between devices are presented in histograms (FIGURE 4 and 5) to illustrate the proportional errors in low, middle-range, and high scores of fat%. The tertile-group cut-offs were defined from the total sample measured by GE Lunar Prodigy (n=146) and were defined for males (n=74) and females (n=72) separately.

6.3.2 Section 2. Prediction models of maximal oxygen consumption

Independent samples T- or Mann-Whitney U -test was used to assess differences in descriptive variables between genders. Associations between VO_{2max} and independent variables were determined as Pearson correlation and Spearman's rank correlation coefficients. Mann-Whitney U test and Spearman's rank correlation was used if the assumptions of the related parametric test were not met. All the subjects completing the measurements were included in the analysis. For each body composition assessment device separately, linear multiple regression models predicting VO_{2max} were created using fat% and LBM of the device. The mentioned or the other variables included in the model if they statistically contributed to the model at the alfa-level of < 0.05. The selection process was done according to the step-by-step backward elimination method. The validation of the models is outside of the scope of this study.

7 RESULTS

7.1 Section 1. Agreement in body fat percentage between devices

The characteristics of the subjects are presented in TABLE 4. Men was significantly heavier (MD 12.0, t=8.0, p<0.001), had higher LBM (MD 16.0, t=15.7, p<0.001), and BMI (MD 0.9, t=2.1, p=0.036) compared to women. Females had significantly higher fat% (GE Lunar Prodigy DXA: MD 10.2, t=8.2, p<0.001) whereas the age did not differ between males and females (p=0.241). The correlations in fat% between the different body composition assessment devices are presented in TABLE 5. Assessments of fat% between Huawei AH100 and GE Lunar Prodigy were strongly correlated in females but moderately in males. Between InBody 720 and GE Lunar Prodigy, the correlation was very strong in both genders. Huawei AH100 and InBody 720 were also highly correlated in females and moderately in males. Comparison of differences, equivalence, and correlations in fat% between GE Lunar Prodigy, InBody 720, and Huawei AH100 are presented in TABLE 5 and the B&A analysis of agreement is illustrated in APPENDIX 1.

The mean bias in fat% was 2.5% (95% LOA -8.7-13.6) between Huawei AH100 and GE Lunar Prodigy, -3.8% (95% LOA -10.3-2.7) between InBody 720 and GE Lunar Prodigy, and 4.6% (95% LOA -6.7-15.9) between Huawei AH100 and InBody 720. Systematic differences, therefore, indicate that Huawei AH100 estimates the highest values and InBody 720 the lowest values in comparison of these devices. Agreement at the individual level (MAPE) ranged from 13.53 to 19.83% when GE Lunar Prodigy was used as the reference. In males, Huawei AH100 had lower MAPE than InBody 720, and in females vice versa. The MAPE of Huawei AH100 to InBody 720 ranged from 23.24 to 28.39%. No equivalence was found between the devices (TABLE 5).

TABLE 4. Characteristics of subjects in *Section 1* of the study.

	Total		Males		Females	
	Mean±SD (n = 146)	Range	Mean±SD (n = 74)	Range	Mean±SD (n = 72)	Range
Age	31.5±7.1	20.4-45.1	32.2±6.7	22.6-44.8	30.8±7.4	20.4-45.1
Weight (kg)	72.0±10.9	49.9-101.1	77.9±9.7	55.0-101.1	65.8±8.4	49.9-85.2
BMI	23.9±2.7	17.0-31.5	24.4±2.5	17.0-30.0	23.5±2.8	18.1-31.5
LBM_{LP} (kg)	52.6±10.1	34.5-78.0	60.5±70.8	45.0-78.0	44.4±51.2	34.5-61.0
$Fat\%_{720}$	20.1±7.4	3.0-41.6	16.0±4.9	4.6-26.6	24.3±7.3	3.0-41.6
$Fat\%_{LP}$	23.9±9.0	5.5-45.7	18.9±7.1	5.5-35.6	29.1±7.8	9.8-45.7
$Fat\%_{AH100}$*	24.9 ±7.2	5.7-39.6	20.7 ±4.4	7.4-33.8	30.3 ±6.3	5.7-39.6

* $Fat\%_{AH100}$ total n=130, males n=74, females n=57. SD= standard deviation, BMI= body mass index, LBM= lean body mass, Fat%= body fat percentage, LP= GE Lunar Prodigy, 720= InBody 720, AH100= Huawei AH100.

Significant proportional errors were found between all the devices. Males and females combined, Huawei AH100 (intercept 11.87, p<0.001, slope -0.37, p<0.001) and InBody 720 (intercept 0.60, p=0.395, slope -0.20, p<0.001) were found to have significant negative slopes in comparison to GE Lunar Prodigy. The error by Huawei AH100 to InBody 720 was rather systematic than proportional (intercept 5.40, p=0.004, slope -0.04 p=0.634).

In males, Huawei AH100 (intercept 20.15, p<0.001, slope -0.86, p<0.001) and InBody 720 (intercept 3.85, p<0.001, slope -0.38, p<0.001) had negative slopes compared to GE Lunar Prodigy. Compared to InBody 720, there was a significant negative proportional bias by Huawei AH100 (intercept 9.37, p=0.001, slope -0.29, p=0.046). In females, finally, the only proportional bias was found between Huawei AH100 and GE Lunar Prodigy (intercept 18.73, p<0.001, slope -0.54, p<0.001) whereas InBody 720 underestimated fat% systematically (intercept -2.75, p=0.038, slope -0.08, p=0.110). Huawei AH100 overestimated systematically compared to InBody 720 (intercept 8.42, p=0.057, slope -0.11, p=0.462).

TABLE 5. Comparison of differences, equivalence and correlations in fat% between the different body composition assessment devices.

	Mean Difference (MD, p)	t	95% LOA	90% CI*	MAPE (%)	MAPE (fat%)	Correlation (r, p)
Comparison between Huawei AH100 and GE Lunar Prodigy DXA							
Total (n = 95)	2.46, <0.001	4.22	-8.7 – 13.6	1.49 – 3.43	15.37	3.67	0.79, <0.001
Males (n = 50)	2.78, 0.001	3.39	-8.8 – 14.2	1.41 – 4.16	13.53	2.54	0.59, <0.001
Females (n = 45)	2.09, 0.015	2.52	-8.8 – 13.0	0.70 – 3.49	19.83	5.87	0.70, <0.001
Comparison between InBody 720 and GE Lunar Prodigy DXA							
Total (n = 146)	-3.80, <0.001	-13.75	-10.3 – 2.7	-4.26 – -3.34	16.75	4.00	0.94, <0.001
Males (n = 74)	-2.86, <0.001	-7.16	-9.6 – 3.9	-3.52 – -2.19	17.08	3.23	0.90, <0.001
Females (n = 72)	-4.78, <0.001	-13.67	-10.6 – 1.0	-5.36 – -4.19	16.62	4.84	0.93, <0.001
Comparison between Huawei AH100 and InBody 720							
Total (n = 130)	4.57, <0.001	8.98	-6.7 – 15.9	3.73 – 5.41	25.69	5.22	0.68, <0.001
Males (n = 73)	4.02, <0.001	6.64	-6.1 – 14.1	3.01 – 5.03	28.39	4.71	0.47, <0.001
Females (n = 53)	5.28, <0.001	z = 5.45	-7.5 – 18.0	3.83 – 6.72	23.24	5.83	0.60 (r_s), <0.001

* Equivalence was evaluated by 90% CI and equivalence interval of ±2.4 was used as the criterion: no equivalence was found. LOA= limits of agreement, CI= confidence interval, MAPE = mean absolute percentage error to reference, p= significance, r= Pearson's correlation coefficient, r_s= Spearman's rank correlation coefficient, z= Wilcoxon Signed-Ranks test Z.

The fat% tertile differences between devices are presented in histograms (FIGURE 4 and 5). It seems that in lean subjects, Huawei AH100 overestimates fat% and in fatter individuals, it underestimates it in both genders. In the middle range of fat%, the overestimation is small but significant. InBody 720, in females, systematically underestimates fat% compared to GE Lunar Prodigy although the bias seems to increase towards the higher scores of FM. In males, it seems to agree well in lean individuals but likewise, increasingly underestimates in the 2nd and the 3rd tertiles of fat%.

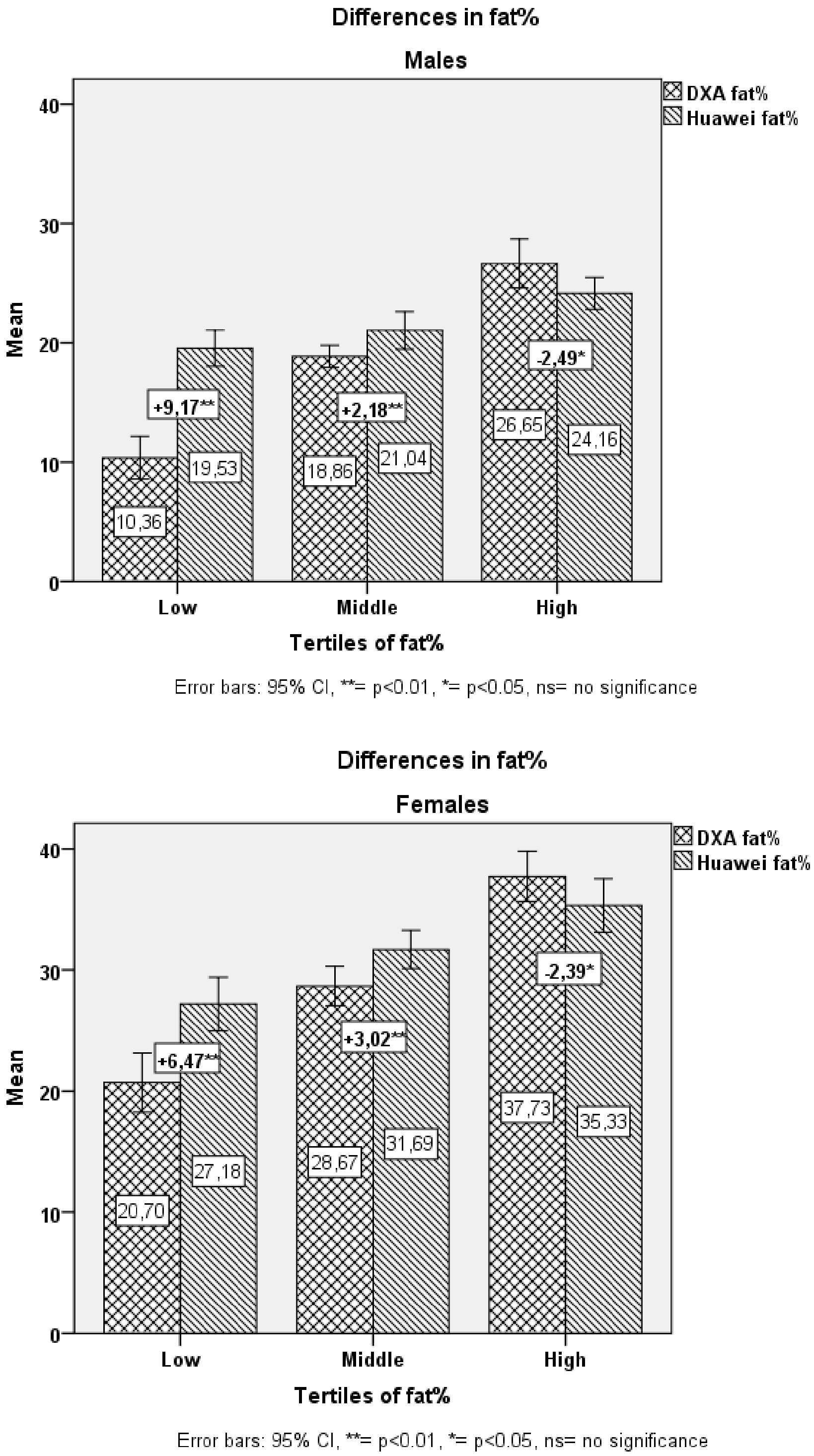
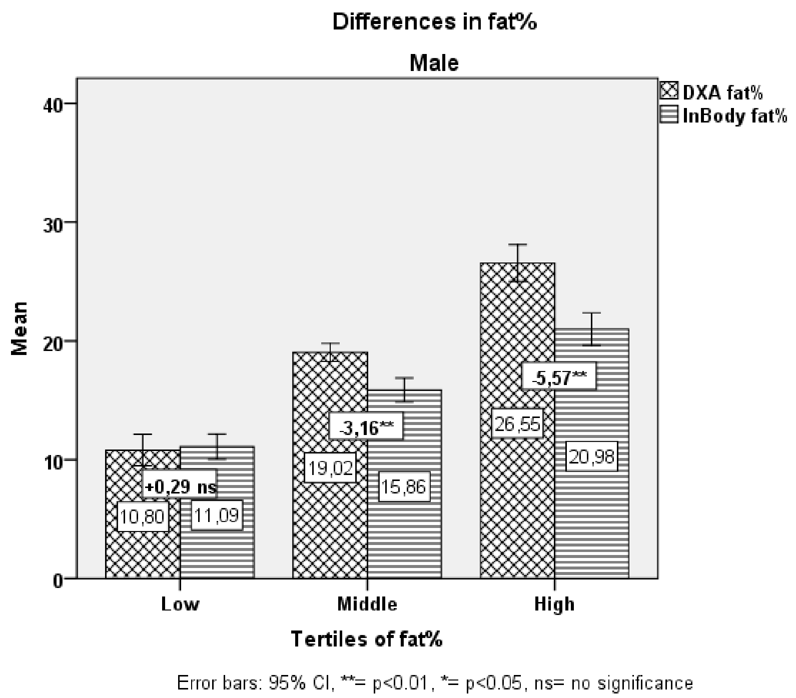


FIGURE 4. Differences in body fat percentage between GE Lunar Prodigy DXA and Huawei AH100 in low, middle, and high tertiles of fat%. Paired samples T-test was used to assess differences between the devices.



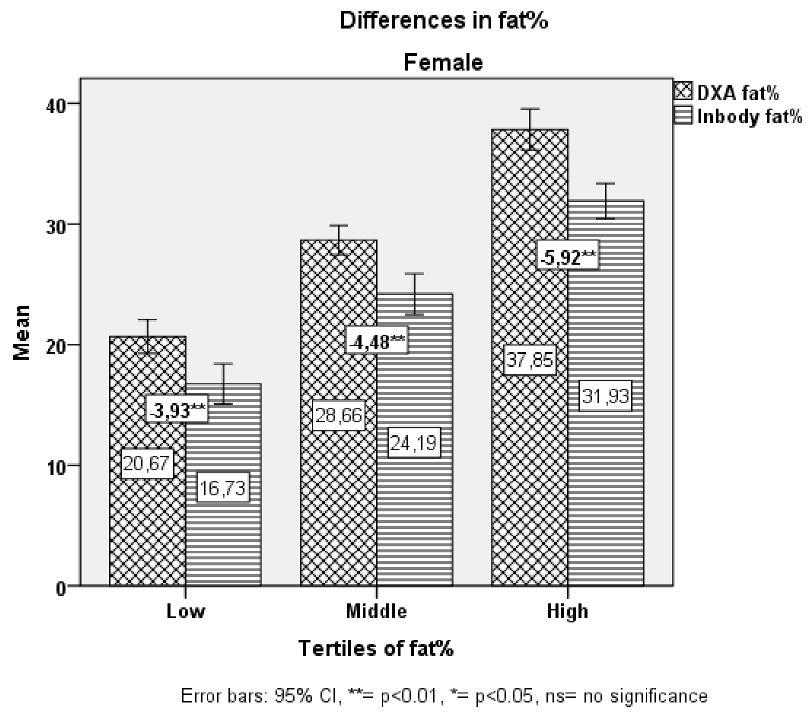


FIGURE 5. Differences in body fat percentage between GE Lunar Prodigy DXA and InBody 720 in low, middle, and high tertiles of fat%. Paired samples T-test was used to assess differences between the devices.

7.2 Section 2: Prediction models of maximal oxygen consumption

Characteristics of subjects and differences between genders are presented in APPENDIX 2. No difference in age between males and females were found whereas males had significantly higher VO_{2max} (MD 6.96, p<0.001) and LBM (MD 15.68, p<0.001) compared to females. Males had also scarcely but significantly higher BMI in the largest group (n=278, MD 0.75, p=0.024). In groups n=140 and n=125, the difference in BMI was non-significant. Females had a higher fat% (MD 10.28, p<0.01). Associations in fat%, LBM, BMI, age, and gender with VO_{2max} are presented in TABLE 6. Body fat percent highly correlated with VO_{2max} and was therefore considered as being the major contributor to the linear multiple regression models. Age, gender (male), BMI, and LBM were also significantly correlated with VO_{2max}. To remark, age did not correlate with VO_{2max} in the group n=140.

TABLE 6. Pearson's correlation coefficients of VO_{2max} with fat%, LBM, BMI, age, and gender.

		Fat%	LBM	BMI	Age	Gender (male)
VO_{2max} and InBody 720						
(n= 278)	r	-0.62	0.26	-0.26	-0.27	0.38
	p	<0.001	<0.001	<0.001	<0.001	<0.001
VO_{2max} and GE Lunar Prodigy DXA						
(n= 140)	r	-0.81	0.57	-0.22	-0.15	0.54
	p	<0.001	<0.001	0.008	0.078	<0.001
VO_{2max} and Huawei AH100						
(n= 125)	r	-0.60	0.31	-0.33	-0.33	0.51
	p	<0.001	0.001	<0.001	<0.001	<0.001

BMI= body mass index, Fat%= body fat percentage, LBM= lean body mass, p= significance, r = Pearson's correlation coefficient.

For the linear multiple regression models, the predicting variables were fat% and LBM of the different devices, as well as age, BMI, and gender. It was found that fat% assessed by different devices was the major contributor (p<0.001) in the models. The lower the fat% was the higher the VO_{2max} was. TABLE 7 represents the influence of the independent predictors on VO_{2max}. In addition to fat%, age contributed significantly to all the prediction models whereas BMI to two of the models, and LBM and gender to only one model. The R^2 value of the models ranged from 40 to 71% and SEE from 3.72 to 5.94 $ml \cdot kg^{-1} \cdot min^{-1}$. The models are used, e.g. model 1: VO_{2max}= 53.57 + ($fat\%_{GE\ Lunar\ Prodigy}$) * -66.34 + (BMI) * 0.50 + (age) * -0.18. The linear relationships between predicted and measured VO_{2max} are presented in APPENDIX 3.

TABLE 7. The final multiple linear regression models of the relationships of VO_{2max} and the predictors.

	Beta	95 % CI	β	t	p
Model 1 (DXA GE Prodigy)					
(Constant)	53.57	47.65 – 59.48		17.92	<0.001
$Fat\%_{DXA}$	-66.34	-74.03 – -58.66	-0.89	-17.08	<0.001
BMI	0.50	0.23 – 0.78	0.19	3.57	<0.001
Age	-0.18	-0.28 – -0.09	-0.19	-3.92	<0.001
R= 0.84, R^2= 0.71, Adjusted R^2= 0.70 – F (3,136) = 109.29, p<0.001, SEE= 3.72					
Model 2 (InBody 720)					
(Constant)	62.95	56.79 – 69.10		20.13	<0.001
$Fat\%_{InBody}$	-0.99	-1.20 – -0.77	-0.96	-9.12	<0.001
LBM_{InBody} (kg)	-0.32	-0.49 – -0.15	-0.41	-3.78	<0.001
BMI	1.03	0.47 – 1.59	0.37	3.62	<0.001
Age	-0.23	-0.31 – -0.14	-0.24	-5.26	<0.001
R= 0.67, R^2= 0.45, Adjusted R^2= 0.45 – F (4,273) = 56.60, p<0.001, SEE= 5.69					
Model 3 (Huawei AH100)					
(Constant)	58.16	51.15 – 65.16		16.43	<0.001
$Fat\%_{Huawei}$	-0.40	-0.63 – -0.18	-0.39	-3.53	0.001
Age	-0.18	-0.32 – -0.05	-0.20	-2.65	0.009
Gender (male)	3.25	0.09 – 6.40	0.21	2.03	0.044
R= 0.64, R^2= 0.40, Adjusted R^2= 0.39 – F (3,121) = 27.23, p<0.001, SEE= 5.94					

R= correlation coefficient, R^2= coefficient of determination of the estimated model, Adjusted R^2= coefficient of determination adjusted by the number of predictors and the sample size, SEE= standard error of the estimate, β= standardised regression coefficient, Beta= unstandardised regression coefficient, CI= confidence interval.

8 DISCUSSION

The objective of this thesis was to assess the influence of different device-based estimations of fat% on the non-exercise estimation of VO_{2max}. The level of agreement and equivalence of the measured fat% was evaluated between GE Lunar Prodigy DXA, InBody 720, and Huawei AH100. Correlations of Lunar Prodigy, 720, and AH100 with VO_{2max} were observed as well and predicting models on VO_{2max} were created and their uniformity was evaluated. It was found that Huawei AH100 estimated the fat% highest and InBody 720 the lowest among the three devices. Overall, the agreement between the devices was relatively good or acceptable but there are some challenges for interpretation.

The bias of Huawei AH100 to GE Lunar Prodigy was relatively small in the middle scores of fat% but it seemed to overestimate at low scores and underestimate at high scores of fat%, therefore introducing a challenge for reliable interpretation if monitoring change in fat mass. InBody 720, on the other hand, seemed to underestimate fat% and the underestimation increased simultaneously with the increase in fat% although in low-fat individuals it appeared to agree well with GE Lunar Prodigy. The slope was not significant in females although the same trend was observed, see FIGURE 5. The bias of Huawei AH100 to InBody 720 was systematic overall, but the B&A analysis revealed a significant slope in males or lower-fat individuals, although the linear relationship in the scatter plot was very weak (APPENDIX 1).

The equivalence testing showed no equivalence between the devices within the ±10% interval (TABLE 5). Huawei AH100 reached an equivalence interval of ±14.3% to GE Lunar Prodigy, InBody 720 ±17.75% to GE Lunar Prodigy, and Huawei AH100 ±22.5% to InBody 720. Despite the smallest equivalence interval between Huawei AH100 and GE Lunar Prodigy, the narrowest 95% LOA was observed between InBody 720 and GE Lunar Prodigy (-10.3 – 2.7) indicating that InBody 720 underestimated fat% consistently and with less variation in the estimation than Huawei AH100. Inversely, Huawei AH100 estimated fat% closer to the estimation of GE Lunar Prodigy but with more variation in the estimation than InBody 720 (APPENDIX 1).

The individual-level agreement (MAPE) with GE Lunar Prodigy was the lowest in Huawei AH100 (15.37%) (TABLE 5). InBody 720 agreed better in individuals higher in FM (i.e. females) (InBody 720 16.62% vs. Huawei AH100 19.83%), and Huawei AH100 agreed better in individuals lower in FM (i.e. males) (Huawei AH100 13.53% vs. InBody 720 17.08%). This seems to be conflicting as the agreement of InBody 720 appears to be better in lower scores of fat% than Huawei AH100 in relation to GE Lunar Prodigy. In TABLE 5, the MAPE has been converted in fat% instead of percent of the mean and the differences appeared not very notable. Nevertheless, it can be argued that the average of the unsigned percentage errors compared to GE Lunar Prodigy DXA was better in Huawei AH100 because its estimations were closer to the average of GE Lunar Prodigy whereas InBody 720 underestimated nearly systematically throughout all scores of fat%.

The second research question concentrated on the similarity of the different non-exercise estimations of VO_{2max}. The models were created using the stepwise backward elimination method and therefore the models were as accurate as possible for each device (GE Lunar Prodigy, InBody 720 & Huawei AH100). It should be noted that the validity of the fat% assessments cannot be evaluated by these models as they include different predictors and derive from different groups of subjects.

Nevertheless, the major finding was that the correlations to VO_{2max} were different between the different fat% assessment methods and the created models considerably differed from each other by the variables included, and the different estimates of fat% seemed to influence the models. Assessed fat% was negatively associated with VO_{2max} and it was the major contributor in all created models which explained the variation in VO_{2max} by 71, 45, and 40%. It was observed that the less the estimation of fat% acquired with the device correlated with VO_{2max}, the less predictive power the model had on VO_{2max}. Indeed, the created model based on variables estimated by GE Lunar Prodigy DXA reached the highest R^2, smallest SEE, and the assessed fat% and FFM highest correlations with VO_{2max}.

8.1 Previous research

8.1.1 Body fat assessment

The differences in linear relationships and R^2's might also be linked with the validity of the estimations of the body composition assessment devices. In this study, DXA was assumed to provide the reference estimates for FM and FFM although this can be criticised. Similar to this study, Burns, Fu, and Constantino (2019) used the B&A method, equivalence testing, and MAPE in evaluating agreement and equivalence in estimates of fat% between Hologic Discovery W DXA and other commonly used body composition assessment methods. They found in college-aged subjects, that BodPod (COSMED, Concord, CA, USA), Valhalla BIA (Valhalla Scientific Model 1990B; Clinton Twp., MI, USA), hydrostatic weighing, and the skinfolds thickness method were equivalent within ±10% interval with DXA in fat%. On the other hand, Tanita BIA (foot-to-foot, BF-556, Arlington Heights, IL, USA), Omron BIA (hand-to-hand, HBF-306, Lake Forest, IL, USA), and near-infrared reactance were not equivalent with DXA and all the methods overestimated fat% compared to DXA. MAPE, however, was lowest in skinfolds thickness (11.7%), hydrostatic weighting (13.4%), and BodPod (14.5%) whereas the others were between 17.0 and 21.9% (Burns, Fu and Constantino 2019). This thesis complements the previous findings on BIA devices, and no equivalence was found between InBody 720, Huawei AH100, and GE Lunar Prodigy DXA within ±10% interval. However, a different DXA model was used as a reference. InBody 720 and Huawei AH100 had slightly smaller MAPE's to DXA compared to the BIA devices used in the study by Burns, Fu and Constantino (2019), except in females Huawei had MAPE of 19.8%.

While DXA as a three-compartment criterion method has been used as the reference commonly in research, the most recent studies have concluded a two-compartment laboratory-based single-frequency BIA even more accurate method in estimating FM and FFM (Nickerson & Tinsley 2018). Indeed, when compared to the five-compartment model, the systematic error by GE Lunar Prodigy DXA has been from +3.7 to +4.1% (LOA ±5.9 to ±6.4%) and larger than the bias of laboratory-based two-compartment models (Moon et al. 2009; Nickerson & Tinsley 2018). When compared to four-compartment models, i.e. CT and MRI, the mean bias for fat%

has ranged from -3.8% to +2.8% (LOA ±3% to ±10%) (Prior et al. 1997; Clasey et al. 1999; Wang et al. 2010; Fosbøl and Zerahn 2015). Finally, when monitoring change in fat% and FFM compared to a four-component model after a strength training program, there has been a mean bias by DXA (DPX-L, Lunar Corp.) of approximately -0.2% and 0.2 kg with relatively large errors (LOA ±3.8% and ±3.1 kg) compared to 4-compartment and 3-compartment (Siri, HW + deuterium dilution) (LOA ±0.69% and ±0.53 kg) (Van Marken Lichtenbelt et al. 2004). Similar findings have been found in a weight-loss setting in athletes where changes of 0.8% and -0.5 kg for fat% and FFM was reported by DXA (Hologic QDR-4500A) (LOA -3.7-5.3% and -3.7-2.7 kg) (Santos et al. 2010). Similar studies and systematic literature reviews on the reliability in detecting changes in body composition should be encouraged in future validation studies, especially for field measurements and BIA devices.

There have been many validation studies, e.g. on the agreement between InBody devices and DXA (GE Lunar Prodigy). Sillanpää et al. (2014) validated segmental multi-frequency BIA (InBody 720) against DXA in a large group of Finnish women and men aged 18-88. They found very similar differences between InBody and DXA compared to this study although no proportional bias was found (women MD 4.7%, LOA -0.9-10.3; men MD 3.1%, LOA -3.7-9.9). In another study by Volgyi et al. (2008), it was also detected that InBody 720 overestimated FFM and underestimated FM and fat% in normal weight and obese Finnish population compared to GE Lunar Prodigy. In obese women, however, the estimates of FFM and FM were similar between the two devices (Volgyi et al. 2008).

Further, McLester et al. (2018) support the findings of the previous research demonstrating that InBody multi-frequency BIA analysers, regardless of the model (InBody 230, InBody 720, and InBody 770), have a small individual error, but a tendency to produce large systematic bias (McLester et al. 2018). As discussed above, a similar trend was found in this study except for an increasing bias towards higher scores of fat%. McLester et al. (2018) also found that the agreement of InBody devices was found to be reliable between measurements compared to DXA (GE Lunar iDXA). However, the underestimation was systematic in fat% by InBody devices: -3.1--3.5 fat% in males (LOA -9.4-3.3) and -2.7--3.3 fat% in females (LOA -8.4-2.4). Positive proportional bias was found in the measurement of InBody 770 in females. FFM was overestimated by 5.7-6.0 kg in males (LOA 0.0-11.3) and by 4.5-4.8 kg in females (LOA 0.7-

8.4) compared to DXA, and in males, significant positive proportional bias was observed in all three devices (McLester et al. 2018).

Before, the validity of InBody 720 compared to Hologic QDR-4500 DXA has been studied (Ling et al. 2011). In middle-aged men and women with 15.9-24.7% prevalence of comorbidities, they found increased underestimation in LBM and overestimation in FM along with the increase in scores of BMI. The B&A analysis did not indicate the same trend completely. Instead, the small underestimation bias in LBM was rather systematic (female -0.8 kg, male -1.5 kg) than proportional. Whereas in this study, an increased underestimation towards higher scores of fat% was found, Ling et al. (2011) found an overestimation bias (female 1.2%, male 2.6%) with higher overestimation towards the higher scores of fat% (Ling et al. 2011). Probably the differences in findings with this thesis are related to the different DXA device used and the different study population. Accordingly, the findings of the validation studies are very dependent on the reference method and the specific study populations.

The direct segmental 8-point multi-frequency BIA has been superior to other BIA devices, i.e. 4- and 8-point single-frequency BIA devices. The difference seems to be related to the ability of the multi-frequency BIA to measure segmental impedances of the body tissues and subsequently estimating TBW more accurately than single-frequency BIA (Demura, Sato & Kitabayashi 2004). For instance, in the study by Volgyi et al. (2008), Tanita (BC 418 MA) 8-point single-frequency BIA failed to detect differences in FFM and FM between individuals of low and high physical activity unlike DXA (GE Lunar Prodigy) and InBody 720 (Volgyi et al. 2008). This may be related to the mentioned fact about the inability of single-frequency BIA to estimate segmental impedances and TBW accurately (Demura, Sato & Kitabayashi 2004) or the algorithms the device is using in its estimation (Volgyi et al. 2008). According to Moore et al. (2019), a laboratory-based single-frequency BIA (RJL Systems, Quantum V) underestimated FM by -2.2 kg (LOA ±5.3) compared to DXA (GE Lunar Prodigy) in a wide range of American adults. The bias was systematic across the scores of FM, except in regional estimates. Simultaneously in lean soft tissue mass, the bias was relatively small and systematic (1.2 kg, LOA ±5.2) although the proportional bias, similarly to FM, was significant in regional estimates of lean soft tissue (Moore et al. 2019).

Unfortunately, the agreement of Huawei AH100 body composition analyser to other methods have not been studied previously. It was found that Huawei AH100 rather overestimated than underestimated the amount of FM and there was a significant negative proportional bias between Huawei and DXA in fat%. More research and systematic literature reviews are needed on the validity of newly developed single- or multi-frequency BIA body scales. In conclusion, the systematic bias and the proportional bias of the bioelectric impedance devices in estimating fat%, FM, and FFM should be taken into consideration when interpreting the results of validation studies (McLester et al. 2018).

8.1.2 Non-exercise prediction of VO_{2max}

Probably the most famous non-exercise model using fat% is created by Jackson et al. (1990). In previously created prediction models using fat%, the assessment has been done mainly by skinfold thickness method (Wang et al. 2019). In that sense, this thesis is very innovative utilising three different body composition assessment methods. Previously created models have used variables such as fat%, smoking status, height, weight, waist circumference, resting HR, gender, BMI, PA, and age (Wang et al. 2019).

The previous studies have found conflicting results on the accuracy of the non-exercise prediction models on VO_{2max} relative to body weight (Wang et al. 2019) and only three studies have reached lower SEE than the model 1 of this study (SEE 3.72) (George et al. 1997; Bradshaw et al. 2005; Cao et al. 2010a). Fifteen of the studies evaluated by Wang et al. (2019), however, have found a lower SEE than the model 3 (SEE 5.94) (Jackson et al. 1990; Ainsworth et al. 1992; Heil et al. 1995; Whaley et al. 1995; Matthews et al. 1999; Jurca et al. 2005; Wier et al. 2006; Cao et al. 2009; Cao et al. 2010b; Nes et al. 2011; Shembre & Riebe 2011; Jackson et al. 2012). The coefficients of determination of the created models have ranged from 0.46 to 0.86 (Wang et al. 2019).

In the cross-validation studies, the correlation coefficients have been ranging from 0.24 to 0.91 demonstrating that non-exercise models can relatively well estimate CRF in other groups than the reference, but the fit must be verified before utilising the models (Wang et al. 2019). The best models have used gender, age, objectively assessed PA, moderate- and vigorous PA, perceived functional ability, step counts, waist circumference, and BMI in their estimations (George et al. 1997; Bradshaw et al. 2005; Cao et al. 2010a; Cao et al. 2010b). This study, along with others, proves that using estimated fat% can yield in reliable non-exercise estimations on VO_{2max} (Jackson et al. 1990; Heil et al. 1995; Whaley et al. 1995; Wier et al. 2006; Jackson et al. 2012), and it very important which method is used to assess body composition (TABLE 6 & 7).

8.2 Strength and limitations

This thesis examines an up-to-date topic as new commercial single- and multi-frequency BIA body composition assessment devices are constantly increasing. The strength of this study is a comprehensive analysis of agreement between the body composition assessment devices in fat%. Furthermore, the maximal direct VO_{2max} test, DXA as the reference method in body composition assessment, and relatively diverse study sample were used in this study. To the knowledge of the author, there are no previous studies on the influence of different body composition estimations on the non-exercise estimation of VO_{2max} and only a few studies are taking advantage of the equivalence testing and the use of MAPE in the analysis of agreement between methods.

There are also many limitations to this study. Perhaps the most problematic question relates to the fact that the body composition assessment methods and prediction models of VO_{2max} are not always accurate for non-reference groups as the estimations are based on certain assumptions and algorithms from the reference population, i.e. there are differences observed in accuracy in different groups of age, ethnicity, sex, health status, etc. (Fosbøl & Zerahn 2015). This can lead to misinterpretations that a certain method is not accurate whereas in the reference population it would be completely valid and vice versa.

The specificity of the prediction models to the reference population is a critical factor in the utilisation of them in practice (Moon et al. 2011; Malek et al. 2004). The same issue limits the validity and usage of the created prediction equations of this thesis. Specifically, the model 1 presented in TABLE 7, derives from a group of healthy Finnish men and women aged 20-45 and the model 2 and the model 3 from healthy Finnish and Chinese men and women aged 18-54. Thus, in addition to the validity of the body composition assessment devices, the differences, age-range and ethnic diversity of the study groups of the models 2 and 3 could have affected to the correlations with VO_{2max} as well as the linear relationships and r^2's of these prediction models. This should be noted when evaluating the validity of these equations and the results of this study. Besides, the measurements were not done in the same place or with the same equipment. Additionally, technical and environmental factors might have affected the results as well.

The prediction models were not cross-validated, and the predicting accuracy of the equations is yet to be confirmed. It seems that the predictive equation using GE Lunar Prodigy DXA as body composition assessment method (Model 1), can be a potentially accurate non-exercise method to assess VO_{2max} in relatively healthy Finnish working-aged adults or a similar population. The validity of the prediction models using InBody 720 and Huawei AH100 (Models 2 & 3) are more questionable as the study groups comprised both Finnish and Chinese subjects.

Although the study sample consisted of young and old working-age adults and from low- to high-fit individuals, the models created in this thesis may not be generalisable to some specific populations such as very low- and high-fit individuals, children and old people, and other specific groups not similar to our study sample. Many studies have reported overestimation of the prediction in the low-fit subjects and conversely, underestimation in the high-fit individuals (e.g. Nes et al. 2011; Jurca et al. 2005) and a similar trend was observed in the predictions of this study (APPENDIX 3). Especially, the model 2 and 3 seem to yield more overestimations than underestimations in the low scores of VO_{2max} and vice versa in the high scores of VO_{2max}. Model 1, on the other hand, seemed to yield relatively accurate predictions even at the extremes of fitness.

Up to the present time, probably the most sophisticated non-exercise prediction models for CRF have been created by the longitudinal prediction model method by Jackson et al. (2012). Unlike linear prediction models, this method can account for the non-linear relationship between age and CRF resulting in unbiased predictions in the older population (Jackson et al. 2012). They could predict VO_{2max} by SEE of 4.9-5.4 $ml \cdot kg^{-1} \cdot min^{-1}$ using fat%, age, waist circumference, resting HR, PA, and smoking status. Although in this study, SEE of even 3.7 $ml \cdot kg^{-1} \cdot min^{-1}$ was reached, it does not account for the non-linear relationships and might be biased towards the extremes of age and fitness.

Criticism can also be placed on the ability of the prediction equations to distinguish between individuals that have lower cardiorespiratory fitness but simultaneously low fat% and vice versa as we found them to be inversely related to each other and the relationship may not be linear. The models of this study or probably any non-exercise model cannot detect such differences between individuals. It is, therefore, possible that accurately assessed PA and physiological variables as predictors can increase the validity but will also add complexity to the models.

Overall, the sample and groups sizes were large in this study, and the validation of the body fat assessment methods is likely to illustrate the true errors, i.e. the Section 1 of the study can be stated valid internally and externally. Speaking of Section 2 of the study, the created non-exercise models may not be generalisable to larger or different populations. Thus, they should be interpreted as not utilisable until further cross-validation.

8.3 Ethical issues

The study protocol was approved by the Ethical Review Board of the University of Jyväskylä as well as the Shanghai Jiao Tong University Bio-X Ethics Board. TRIPOD (Collins et al. 2015) and GRRAS (Kottner et al. 2011) checklists were followed as well for the transparency of reporting.

The used data was not pseudonymised or anonymised for the analysis. However, research subjects cannot be identified from the results or publications. In this study, other data security procedures were done, such as internal actions of the controller and the processor to prevent unauthorised access to personal data and data secured working environments or IT-systems was used. Besides, all the direct identifiers of the subjects were removed in the analysis phase. We followed the EU General Data Protection Regulations in Jyväskylä.

In reward of participation at Jyväskylä, a movie ticket was given to the subjects after the measurements whether they completed the measurements or not. Given the small monetary value of the reward, it is seen as an acceptable way to motivate subjects to participate in measurements of very small health-risk, i.e. dual-energy x-ray absorptiometry and direct VO_{2max} test. Furthermore, this may have enlarged the study sample to individuals that were not otherwise motivated to participate, reducing the selection bias to some extent.

In addition, the risk assessment was done to ensure eligibility of subjects for participation and the subjects were insured against accidents, damages, and injuries during the measurements. The sources of funding of the data collection were the University of Jyväskylä, Shanghai Jiao Tong University, and Huawei Technologies Oy (Finland) Co., Ltd. As in this thesis, Huawei's device was evaluated, it is important to declare that Huawei Technologies Oy did not participate in the analysis, the data collection, or in interpretations in any way regarding this thesis. The Faculty of Sport and Health Sciences had signed an agreement with Huawei. The results of this study are owned by the University of Jyväskylä and they are open to the public.

8.4 Future studies

As the supply of commercial body composition assessment devices to the public is increasing, it raises a concern of the accuracy of their measurements. Consequently, it would be important to assess the validity and reliability of these new devices as well as their predictive potential on clinical indicators, such as cardiorespiratory fitness. If the agreement of estimate with the measured value is not satisfactory, the measurement might mislead the customers. Contrariwise, if the estimate is accurate, the validity serves as a potential competitive advantage against competitors of the manufacturer or the service provider. Furthermore, the more valid the measurement is the more potential the device has predicting another valuable health indicator.

Secondly, research should focus on the reliability of the estimations in the follow-up. As was discussed of both sections of this study, the concerning issue was the inaccuracy at the low and high scores of the estimated variable, whether fat% or VO_{2max}. It is not practically useful to create predictions if we cannot monitor change accurately. One very good approach is suggested by Jackson et al. (2012) utilising longitudinal algorithms instead of cross-sectional linear regression models.

As described very recently (2019), the cross-validation of prediction equations in multiple populations including ethnic groups and with morbidities should be encouraged (Wang et al. 2019). In this study, although the study groups included two ethnic groups, it is not reasonable to conclude any findings on the effect of that until further cross-validation. Furthermore, Wang et al. (2019) recommended additional studies on the associations between predicted cardiorespiratory fitness and different morbidities for further validation of non-exercise models in clinical settings (Wang et al. 2019). It is also acknowledged that self-reported PA in the prediction of VO_{2max} has limitations and therefore objective methods of PA in the models should be utilised more for better outcomes and estimates (Wang et al. 2019).

8.5 Conclusions

In this study, by comprehensively evaluating the agreement and equivalence between GE Lunar Prodigy DXA, InBody 720, and Huawei AH100, it was found that the estimations of fat% differed from each other at the group- and individual-level. Due to the lack of statistical equivalence, it would be recommended to interpret the estimations of BIA devices with caution. The second important finding was that the estimations of fat% and LBM of different body composition assessment devices as well as age, BMI, and gender correlated differently with the directly assessed VO_{2max} in the different study groups. Accordingly, the included predictors and the R^2's of the created prediction equations were notably affected and the prediction models were very different from each other.

The majority of the studies that have predicted maximal energy expenditure have included PA along with other predictors and much fewer studies have exploited the assessment of fat% (Wang et al. 2019). In this study, it was found that even by using only fat% assessment of GE Lunar Prodigy DXA, BMI, and age, it is possible to predict VO_{2max} with low error, although it is acknowledged that some limitations are evident. It can be stated from these results that accurately estimated body composition may lead to better predictions on VO_{2max}.

REFERENCES

Absolute Medical. Medical equipments. Bone densitometers: GE Lunar Prodigy Advance. Cited. 18.4.2020. http://www.absolutemed.com/ge-lunar-prodigy-advance-bone-densitometer.htm.

Ainsworth, B., Richardson, M., Jacobs, D. & Leon, A. 1992. Prediction of cardiorespiratory fitness using physical-activity questionnaire data. Research Quarterly for Exercise and Sport. 63, 75–82.

Altman, D. G. 1990. Practical Statistics for Medical Research. Taylor & Francis Ltd., London.

Amani, A. R., Somchit, M. N., Konting, M. M. B. & Kok, L. Y. 2010. Relationship between body fat percent and maximal oxygen uptake among young adults. Journal of American Science. 6, 1–4.

Baumgartner, R. N., Heymsfield, S. B., Lichtman, S., Wang, J. & Pierson, R. N. Jr. 1991. Body composition in elderly people: effect of criterion estimates on predictive equations. American Journal of Clinical Nutrition. 53, 1345–1353.

Bjorntorp, P. 1990. “Portal” adipose tissue as a generator of risk factors for cardiovascular disease and diabetes. Arteriosclerosis. 10 (4), 493–6.

Bouchard, C., An, P., Rice, T. & Skinner, J. 1999. Familial aggregation of VO2max response to exercise training: Results from the HERITAGE Family Study. Journal of Applied Physiology, 87 (3), 1003-1008.

Booth, F. W., Roberts, C. K. & Laye, M. J. 2012. Lack of exercise is a major cause of chronic diseases. Comprehensive Physiology 2 (2), 1143-1211.

Bradshaw, D. I., George, J. D., Hyde, A., LaMonte, M. J., Vehrs, P. R., Hager, R. L. & Yanowitz, F. G. 2005. An accurate VO2max non exercise regression model for 18–65-year-old adults. Research Quarterly for Exercise and Sport. 76, 426–32.

Buchholz, A. C., Bartok, C. & Schoeller, D. A. 2004. The validity of bioelectrical impedance models in clinical populations. Nutrition in Clinical Practice. 19, 433–446.

Burns, R. D., Fu, Y. & Constantino, N. 2019. Measurement agreement in percent body fat estimates among laboratory and field assessments in college students: Use of equivalence testing. PLoS ONE 14 (3), e0214029.

Buskirk, E., & Taylor, H. L. 1957. Maximal Oxygen Intake and Its Relation to Body Composition, With Special Reference to Chronic Physical Activity and Obesity. Journal of Applied Physiology, 11 (1), 72–78.

Cao, Z. B., Miyatake, N., Higuchi, M., Ishikawa-Takata, K., Miyachi, M. & Tabata, I. 2009. Prediction of VO2max with daily step counts for Japanese adult women. European Journal of Applied Physiology. 105, 289–96.

Cao, Z. B., Miyatake, N., Higuchi, M., Miyachi, M., Ishikawa-Takata, K. & Tabata, I. 2010a. Predicting VO2max with an objectively measured physical activity in Japanese women. Medicine and Science in Sports and Exercise. 42, 179–86.

Cao, Z. B., Miyatake, N., Higuchi, M., Miyachi, M. & Tabata, I. 2010b. Predicting VO2max with an objectively measured physical activity in Japanese men. European Journal of Applied Physiology. 109, 465–72.

Carrion, B. M., Wells, A., Mayhew, J. L. & Koch, A. J. 2019. Concordance Among Bioelectrical Impedance Analysis Measures of Percent Body Fat in Athletic Young Adults. International Journal of Exercise Science, 12 (4), 324–331.

Chandra, A., Neeland, I. J., Berry, J. D., Ayers, C. R., Rohatgi, A., Das, S. R., Khera, A., McGuire, D. K., de Lemos, J. A. & Turer, A. T. 2014. The relationship of body mass and fat distribution with incident hypertension: observations from the Dallas Heart Study. Journal of the American College of Cardiology. 64 (10), 997-1002.

Chumlea, W. M. & Guo, S. S. 2000. Assessment and prevalence of obesity: application of new methods to a major problem. Endocrine. Oct. 13 (2), 135-42.

Chumlea, W. C., Guo, S. S., Kuczmarski, R. J., Flegal, K. M., Johnson, C. L., Heymsfield, S. B., Lukaski, H. C., Friedl, K. & Hubbard, V. S. 2002. Body composition estimates from NHANES III bioelectrical impedance data. International Journal of Obesity Related Metabolic Disorders. Dec 26 (12), 1596-609.

Chouinard, L. E., Schoeller, D. A., Watras, A. C., Clark, R. R., Close, R. N. & Buchholz, A. C. 2007. Bioelectrical impedance vs. four-compartment model to assess body fat change in overweight adults. Obesity. 15, 85–92.

Chumlea, W. C., Schubert, C. M., Sun, S. S., Demerath, E., Towne, B. & Siervogel, R. M. 2007. A review of body water status and the effects of age and body fatness in children and adults. Journal of Nutrition, Health & Aging. Mar-Apr 11 (2), 111-8.

Clasey, J. L., Kanaley, J. A., Wideman, L., et al. 1999. Validity of methods of body composition assessment in young and older men and women. Journal of Applied Physiology. 86, 1728–38.

Collins, G. S., Reitsma, J. B., Altman, D. G., Moons & K. G. 2015. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. Circulation. Jan 13: 131 (2), 211-9.

Davies, C. T. M. 1968. Limitations to the prediction of maximum oxygen uptake from cardiac frequency measurements. Journal of Applied Physiology. 24, 700-706.

de Lannoy L. & Ross R. 2019. Physical Activity, Cardiorespiratory Fitness, and Obesity. In: Kokkinos P., Narayan P. (eds.) Cardiorespiratory Fitness in Cardiometabolic Diseases. Springer, Champaign. 230-231, 245.

Demura, S., Sato, S. & Kitabayashi, T. 2004. Percentage of total body fat as estimated by three automatic bioelectrical impedance analyzers. Journal of Physiological Anthropology and Applied Human Science. May 23 (3), 93-9.

Dengel, D. R., Raymond, C. J. & Bosch, T. A. 2017. Assessment of muscle mass. In: H. C. Lukaski (ed.) Body composition: Health and performance in exercise and sport. Boca Raton: Taylor & Francis. 27-28, 27-37.

Després, J. P., Prud'homme, D., Pouliot, M. C., Tremblay A. & Bouchard, C. 1991. Estimation of deep abdominal adipose-tissue accumulation from simple anthropometric measurements in men. American Journal of Clinical Nutrition. Sep, 54 (3), 471-7.

Dixon, P. M., Saint-Maurice, P. F., Kim, Y., Hibbing, P., Bai, Y. & Welk, G. J. 2018. A Primer on the Use of Equivalence Testing for Evaluating Measurement Agreement. Medicine and Science in Sports and Exercise. 50 (4), 837-845.

Duren, D. L., Sherwood, R. J., Czerwinski, S. A., Lee, M., Choh, A. C., Siervogel, R. M., & Cameron Chumlea, W. 2008. Body composition methods: comparisons and interpretation. Journal of Diabetes Science and Technology, 2 (6), 1139–1146.

Earthman, C. P. 2015. Body composition tools for assessment of adult malnutrition at the bedside: A tutorial on research considerations and clinical applications. Journal of Parenteral and Enteral Nutrition. 39, 787–822.

Evans, E. M., Saunders, M. J., Spano, M. A., Arngrimsson, S. A., Lewis, R. D. & Cureton, K. J. 1999. Body composition changes with diet and exercise in obese women: a comparison of estimates from clinical methods and a 4-component model. American Journal of Clinical Nutrition. 70, 5–12.

Farrell, P. A., Gustafson, A. B. & Kalkhoff, R. K. 1985. Assessment of methods for assigning treadmill exercise workloads for lean and obese women. International Journal of Obesity. 9 (1), 49-58.

Forbes, G. B. 1994. Bioelectrical impedance and body composition: present status and future direction. Nutrition Reviews. Sep 52 (9), 323-5.

Fox, E. L. 1973. A simple, accurate technique for predicting maximal aerobic power. Journal of Applied Physiology. 35, 914-916.

Fujimoto, W. Y., Newell-Morris, L. L., Grote, M., Bergstrom, R. W. & Shuman, W. P. 1991. Visceral fat obesity and morbidity: NIDDM and atherogenic risk in Japanese American men and women. International Journal of Obesity. Sep, 15 (2), 41-4.

Fuller, N. J., Jebb, S. A., Laskey, M. A., Coward, W. A. & Elia, M. 1992. Four-component model for the assessment of body composition in humans: comparison with alternative methods, and evaluation of the density and hydration of fat-free mass. Clinical Science (London). 82, 687–693.

Fosbøl, M. & Zerahn, B. 2015. Contemporary methods of body composition measurement. Clinical Physiology and Functional Imaging. 35, 81–97.

George, J. D., Stone, W. J. & Burkett, L. N. 1997. Non-exercise VO2max estimation for physically active college students. Medicine and Science of Sports and Exercise. 29, 415–23.

Goran, M., Fields, D. A., Hunter, G. R., Herd, S. L. & Weinsier, R. L. 2000. Total body fat does not influence maximal aerobic capacity. International Journal of Obesity Related Metabolic Disorders. 24 (7), 841-8.

Gormley, S. E., Swain, D. P., High, R., Spina, R. J., Dowling, E. A., Kotipalli, U. S. & Gandrakota, R. 2008. Effect of intensity of aerobic training on VO2max. Medicine and Science in Sports and Exercise 40 (7), 1336-1343.

Greiwe, J. S. Kaminsky, L. A. Whaley, M. H. Dwyer, G. B. 1995. Evaluation of the ACSM submaximal ergometer test for estimating VO2max. Medicine and Science in Sports and Exercise. 27 (9), 1315-1320.

Heil, D. P., Freedson, P. S., Ahlquist, L. E., Price, J. & Rippe, J. M. 1995. Non exercise regression models to estimate peak oxygen consumption. Medicine and Science in Sports and Exercise. 27, 599–606.

Hewitt, M. J., Going, S. B., Williams, D. P. & Lohman, T. G. 1993. Hydration of the fat-free body mass in children and adults: implications for body composition assessment. American Journal of Physiology. 265, E88–E95.

Hopkins, W. G., Marshall, S. W., Batterham, A. M. & Hanin, J. 2009. Progressive statistics for studies in sports medicine and exercise science. Medicine and Science in Sports and Exercise. 41 (1), 3-13.

Howley, E. T., Bassett, D. R. & Welch, H. G. 1995. Criteria for maximal oxygen uptake: review and commentary. Medicine and Science in Sports and Exercise. 27 (9), 1292-301.

Huawei. Accessories: HUAWEI Smart Scale. Cited 18.4.2020. https://consumer.huawei.com/en /accessories/smart-scale.

Huawei. Accessories: HUAWEI Smart Scale 2 Pro. Cited. 18.4.2020. https://consumer.huawei.com/cn/accessories/scale-2-pro.

InBody. Tuotteet: InBody 720. Cited. 18.4.2020. https://www.inbody.fi/tuotteet/inbody720.

Jackson, A. S., Blair, S. N., Mahar, M. T., Wier, L. T., Ross, R. M. & Stuteville, J. E. 1990. Prediction of functional aerobic capacity without exercise testing. Medicine and Science in Sports and Exercise. 22 (6), 863–70.

Jackson, A. S., Sui, X., O’Connor, D. P., Church, T. S., Lee, D. C., Artero, E. G. & Blair, S. N. 2012. Longitudinal cardiorespiratory fitness algorithms for clinical settings. American Journal of Preventive Medicine. 43, 512–9.

Jebb, S. A., Cole, T. J., Doman, D., Murgatroyd, P. R. & Prentice, A. M. 2000. Evaluation of the novel Tanita body-fat analyser to measure body composition by comparison with a four-compartment model. British Journal of Nutrition. 83, 115–122.

Johnstone, A. M., Faber, P., Gibney, E. R., Lobley, G. E., Stubbs, R. J. & Siervo, M. 2014. Measurement of body composition changes during weight-loss in obese men using multi-frequency bioelectrical impedance analysis and multi-compartment models. Obesity Research & Clinical Practice. 8 (1) e46–e54.

Jones, A. M. & Carter, H. 2000. The effect of endurance training on parameters of aerobic fitness. Sports medicine (Auckland, N.Z.) 29 (6), 373-386.

Joyner, M. J. & Coyle, E. F. 2008. Endurance exercise performance: the physiology of champions. Journal of Physiology. 586, 35–44.

Joyner, M. J. 2017. Exercise and trainability: contexts and consequences. The Journal of Physiology 595 (11), 3239-3240.

Jurca, R., Jackson, A. S., LaMonte, M. J., Morrow, J. R., Blair, S. N., Wareham, N. J., Haskell, W. L., van Mechelen, W., Church, T. S., Jakicic, J. M. & Laukkanen, R. 2005. Assessing cardiorespiratory fitness without performing exercise testing. American Journal of Preventive Medicine. 29 (3), 185–193.

Katch, V. L., Sady, S. S. & Freedson, P. 1982. Biological variability in maximum aerobic power. Medicine and Science in Sports and Exercise. 14 (1), 21-5.

Katzmarzyk, P. T., Mire, E. & Bouchard, C. 2012. Abdominal obesity and mortality: The Pennington Center Longitudinal Study. Nutrition & Diabetes. 2 (8), e42.

Kenney W, Wilmore J, Costill D. 2012. Physiology of Sport and Exercise. Champaign, IL: Human Kinetics. 249.

Keskinen, K. L. 2016. Fyysinen kunto ja sen testaaminen. In I. Vuori, S. Taimela & U. M. Kujala (eds.) Liikuntalääketiede. 8th ed. Helsinki: Duodecim, 110-113.

Keskinen, O. P. Mänttäri, A. Aunola, S. & Keskinen, K. L. 2007. Kestävyysominaisuuksien mittaaminen. In K. L. Keskinen, K. Häkkinen & M. Kallinen (eds.) Kuntotestauksen käsikirja. 2nd. ed. Helsinki: Liikuntatieteellisen Seuran julkaisu nro 161. 78-117.

Keskinen, K. L., Häkkinen, K. & Kallinen, M. 2018. Fyysisen kunnon mittaaminen – käsi- ja oppikirja kuntotestaajille. Helsinki. Liikuntatieteellisen Seuran julkaisu nro 174. 80-100.

Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajeweski, B. J., Hróbjartsson, A., Robersts, C., Shoukri, M. & Streiner, D. L. 2011. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. Journal of Clinical Epidemiology. 64 (1), 96-106.

Kujala, U. M. 2016. Perintötekijät ja liikunta. In I. Vuori, S. Taimela & U. M. Kujala (eds.) Liikuntalääketiede. 8th. ed. Helsinki: Duodecim, 55-59.

Ling, C. H., A. J. de Craen, P. E. Slagboom, et al. 2011. Accuracy of direct segmental multi-frequency bioimpedance analysis in the assessment of total body and segmental body composition in middle-aged adult population. Clinical Nutrition. 30 (5), 610-5.

Lohman, T. G. 1986. Applicability of body composition techniques and constants for children and youths. Exercise and Sport Sciences Reviews. 14, 325–357.

Lukaski, H. C, Johnson, P. E., Bolonchuk, W. W. & Lykken, G. I. 1985. Assessment of fat-free mass using bioelectrical impedance measurements of the human body. American Journal of Clinical Nutrition. Apr 41 (4), 810-7.

Malek, M. H., Berger, D. E., Housh, T. J., Coburn, J. W. & Beck, T. W. 2004. Validity of VO2max Equations for Aerobically Trained Males and Females. Medicine and Science in Sports and Exercise. 36 (8), 1427-1432.

Maciejczyk, M., Więcek, M., Szymura, J., Szyguła, Z., Wiecha, S. & Cempla, J. 2014. The influence of increased body fat or lean body mass on aerobic performance. PloS ONE, 9 (4), e95797.

Matias, C. N., Santos, D. A., Fields, D. A., Sardinha, L. B. & Silva, A. M. 2012. Is bioelectrical impedance spectroscopy accurate in estimating changes in fat-free mass in judo athletes? Journal of Sports Sciences. 30, 1225–1233.

Matthews, C. E., Heil, D. P., Freedson, P. S. & Pastides, H. 1999. Classification of cardiorespiratory fitness without exercise testing. Medicine and Science in Sports and Exercise. 31, 486–93.

Mattila, M. 2007. Body Composition by DEXA and its association with physical fitness in 140 conscripts. Medicine and Science in Sports and Exercise, 39 (12), 2242-2247.

Margaria, R. Adhemo, P. Rouelli, E. 1965. Indirect determination of maximal oxygen consumption in man. Journal of Applied Physiology. 20, 1070-1073.

McArdle, W. D., Katch, F. I. & Katch, V. L. 2015. Exercise physiology: nutrition, energy and human performance. 8th. ed. Philadelphia (PA): Lippincott Williams & Wilkins. 249-453.

Mclester, C. N., Nickerson, B. S., Kliszczewicz, B. M. & Mclester, J. R. 2018. 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. Nov 3.

Milanovic, Z., Sporis, G. & Weston, M. 2015. Effectiveness of High-Intensity Interval Training (HIT) and Continuous Endurance Training for VO2max Improvements: A Systematic Review and Meta-Analysis of Controlled Trials. Sports medicine (Auckland, N.Z.) 45 (10), 1469-1481.

Minderico, C. S., Silva, A. M., Keller, K., Branco, T. L., Martins, S. S., Palmeira, A. L., Barata, J. T., Carnero, E. A., Rocha, P. M., Teixeira, P. J. & Sardinha, L. B. 2008. Usefulness of different techniques for measuring body composition changes during weight loss in overweight and obese women. British Journal of Nutrition. 99, 432–441.

Mondal, H., & Mishra, S. P. 2017. Effect of BMI, Body Fat Percentage and Fat Free Mass on Maximal Oxygen Consumption in Healthy Young Adults. Journal of Clinical and Diagnostic Research. 11 (6), CC17–CC20.

Moon, J. R., Eckerson, J. M., Tobkin, S. E., et al. 2009. Estimating body fat in NCAA Division I female athletes: a five-compartment model validation of laboratory methods. European Journal of Applied Physiology. 105 (1), 119-30.

Moon, J. R., Dalbo, V. J., Roberts, M. D., Kerksick, Ch. M. & Stout, J. R. 2011. Usefulness of bioelectrical Impedance in the Prediction of VO2max in Healthy Men and Women. Sport Journal. 14 (1).

Moon, J. R., Stout, J. R., Smith-Ryan, A. E., Kendall, K. L., Fukuda, D. H., Cramer, J. T. & Moon, S. E. 2013. Tracking fat-free mass changes in elderly men and women using single-frequency bioimpedance and dual-energy X-ray absorptiometry: a four-compartment model comparison. European Journal of Clinical Nutrition. 67, 40–46.

Moore, M. L., Benavides, M. L., Dellinger, J. R., Adamson, B. T., & Tinsley, G. M. 2019. Segmental body composition evaluation by bioelectrical impedance analysis and dual-energy X-Ray absorptiometry: Quantifying agreement between methods. Clinical Nutrition. Dec. 13.

Myers, J., Kaykha, A., George, S., Abella, J., Zaheer, N., Lear, S., et al. 2004. Fitness versus physical activity patterns in predicting mortality in men. American Journal of Medicine. 117 (12), 912–8.

Myers, J., Prakash, M., Froelicher, V., Do, D., Partington, S. & Atwood, J. E. 2002. Exercise capacity and mortality among men referred for exercise testing. The New England Journal of Medicine 346 (11), 793-801.

Nana, A., Slater, G. J., Stewart, A. D. & Burke, L. M. 2015. Methodology review: Using dual-energy x-ray absorptiometry (DXA) for the assessment of body composition in athletes and active people. International Journal Sport Nutrition and Exercise Metabolism. 25, 198–215.

Neeland, I. J., Turer, A. T., Ayers, C. R., Berry, J. D., Rohatgi, A., Das, S. R., Khera, A., Vega, G. L., McGuire, D. K., Grundy, S. M. & de Lemos, J. A. 2015. Body fat distribution and incident cardiovascular disease in obese adults. Journal of the American College of Cardiology. 65 (19), 2150-1.

Neeland, I. J., Turer, A. T., Ayers, C. R., Powell-Wiley, T. M., Vega, G. L., Farzaneh-Far, R., Grundy, S. M., Khera, A., McGuire, D. K. & de Lemos, J. A. 2012. Dysfunctional adiposity and the risk of prediabetes and type 2 diabetes in obese adults. JAMA. 308 (11), 1150-9.

Nes, B. M., Janszky, I., Vatten, L. J., Nilsen, T. I. L., Aspenes, S. T. & Wisløff, U. 2011. Estimating VO2peak from a non-exercise prediction model: the HUNT Study, Norway. Medicine and Science in Sports and Exercise. 43, 2024–30.

Nickerson, B. & Tinsley, G. 2018. Utilization of BIA-derived bone mineral estimates exerts minimal impact on body fat estimates via multi-compartment models in physically active adults. Journal of Clinical Densitometry. 21 (4), 541-549.

Nicklas, B. J., Penninx, B. W., Cesari, M., Kritchevsky, S. B., Newman, A. B., Kanaya, A. M., Pahor, M., Jingzhong, D. & Harris, T. B. 2004. Health, Aging and Body Composition Study. Association of visceral adipose tissue with incident myocardial infarction in older men and women: The Health, Aging and Body Composition Study. American Journal of Epidemiology. 160 (8), 741–749.

Okosun, I. S., Chandra, K. M., Boev, A., Boltri, J. M., Choi, S. T., Parish, D. C. & Dever, G. E. 2004. Abdominal adiposity in U.S. adults: prevalence and trends, 1960-2000. Preventive Medicine. 39 (1), 197–206.

Plasqui, G. & Westerterp, K. R. 2007. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. Obesity. 15 (10), 2371-9.

Pouliot, M. C., Després, J. P., Lemieux, S., Moorjani, S., Bouchard, C., Tremblay, A., Nadeau, A. & Lupien, P. J. 1994. Waist circumference and abdominal sagittal diameter: best simple anthropometric indexes of abdominal visceral adipose tissue accumulation and related cardiovascular risk in men and women. American Journal of Cardiology. Mar 1, 73 (7), 460-8.

Prabha, V, Padmanabha, B. V. & Doddamani, B. R. 2014. Correlation between obesity and cardiorespiratory fitness. International Journal of Current Scientific Research. 5 (3), 1193-1196.

Pribis, P, Burtnack, C. A., McKenzie, S. O. &Thayer, J. 2010. Trends in body fat, body mass index and physical fitness among male and female college students. Nutrients. 2 (10), 1075-85.

Prior, B. M., Cureton, K. J., Modlesky, C. M., et al. 1997. In vivo validation of whole-body composition estimates from dual-energy x-ray absorptiometry. Journal of Applied Physiology. 83, 623–30.

Rexhepi, A. & Brestovci, B. 2014. Prediction of vo2max based on age, body mass, and resting heart rate. Human Movement. 15.

Roche A. 1996. Anthropometry and ultrasound. In: Roche A, Heymsfield S, Lohman T. (eds.) Human body composition. Champaign, IL: Human Kinetics Press. 167–189.

Sanada, K. Kearns, C., Kojima, K. & Abe, T. 2005. Peak oxygen uptake during running and arm cranking normalized to total and regional skeletal muscle mass measured by magnetic resonance imaging. European Journal of Applied Physiology. 93, 687-93.

Sanada, K., Midorikawa, T., Yasuda, T., Kearns, C. F. & Abe, T. 2007. Development of non-exercise prediction models of maximal oxygen uptake in healthy Japanese young men. European Journal of Applied Physiology, 99 (2), 143–148.

Santos, D. A., Silva, A. M., Matias, C. N., Fields, D. A., Heymsfield, S. B. & Sardinha, L. B. 2010. Accuracy of DXA in estimating body composition changes in elite athletes using a four-compartment model as the reference method. Nutrition & Metabolism (London). 7, 22.

Schembre, S. M. & Riebe, D. A. 2011. Non-exercise estimation of VO2max using the international physical activity questionnaire. Measurement in Physical Education and Exercise Science. 15, 168–81.

Seidell, J. C., Oosterlee, A., Thijssen, M. A., Burema, J., Deurenberg, P., Hautvast, J. G. & Ruijs, J. H. 1987. Assessment of intra-abdominal and subcutaneous abdominal fat: relation between anthropometry and computed tomography. American Journal of Clinical Nutrition. Jan, 45 (1), 7-13.

Sharma, M., Kamal, R. B. & Chawla, K. 2016. Correlation of body composition to aerobic capacity; A cross sectional study. International Journal of Applied Research. 2 (1), 38-42

Shen, W., St-Onge, M. P., Wang, Z. & Heymsfield, S. B. 2005. The science of body composition research. In: Heymsfield, S. B., Lohman, T. G., Wang, Z. & Going, S. B. (eds.) Human Body Composition. Human Kinetics, Champaign, IL. 1–14.

Shete, A. N., Bute, S. S., & Deshmukh, P. R. 2014. A Study of VO_{2Max} and Body Fat Percentage in Female Athletes. Journal of Clinical and Diagnostic Research. 8 (12), BC01–BC3.

Shvartz, E. & Reibold, R. C. 1990. Aerobic fitness norms for males and females aged 6 to 75 years: a review. Aviation, Space, and Environmental Medicine 61 (1), 3-11.

Sillanpää, E., Cheng, S., Häkkinen, K., Finni, T., Walker, S., Pesola, A., et al. 2014. Body composition in 18- to 88-year-old adults - comparison of multifrequency bioimpedance and dual-energy X-ray absorptiometry. Obesity. 22 (1), 101-109.

Siri, W. 1961. Body composition from fluid spaces and density analysis of methods. In: Brozek J, Henshcel A, editors. Techniques for measuring body composition. Washington, DC: National Academy Press. 223–244.

Sui, X., Laditka, J. N., Hardin, J. W. & Blair, S. N. 2007. Estimated functional capacity predicts mortality in older adults. Journal of the American Geriatrics Society 55 (12), 1940-1947.

Suni, J. & Vasankari, T. 2011. Terveyskunto ja fyysinen toimintakyky. In M. Fogelholm, I. Vuori & T. Vasankari. Terveysliikunta. 2nd. ed. Helsinki: Duodecim, 32-35.

Stahn, A., Terblanche, E., Grunert, S. & Strobel, G. 2006. Estimation of maximal oxygen uptake by bioelectrical impedance analysis. European Journal of Applied Physiology. 96 (3), 265–273.

Staudenmayer, J., Zhu, W. & Catellier, D. J. 2012. Statistical considerations in the analysis of accelerometry-based activity monitor data. Medicine and Science in Sports and Exercise. 44 (1), 61-7.

Terry, J. W., Tolson, H., Johnson, D. J., Jessup, G. T. 1977. A workload selection procedure for the Astrand-Ryhming test. Journal of Sports Medicine. 17, 361-365.

Van Marken Lichtenbelt, W. D., Hartgens, F., Vollaard, N. B., Ebbing, S. & Kuipers, H. 2004. Body composition changes in bodybuilders: A method comparison. Medicine and Science in Sports and Exercise 36, 490–7.

Volgyi, E, Tylavsky, F. A., Lyytikainen, A., Suominen, H., Alen, M., Cheng, S. 2008. Assessing body composition with DXA and bioimpedance: effects of obesity, physical activity, and age. Obesity. 16, 700-705.

Waki, M., Kral, J. G., Mazariegos, M., Wang, J., Pierson, R. N. Jr. & Heymsfield, S. B. 1991. Relative expansion of extracellular fluid in obese vs. non-obese women. American Journal of Physiology. 261, E199–E203.

Wang, Y., Chen, S., Lavie, C., Zhang, J., Sui, X. 2019. An Overview of Non-exercise Estimated Cardiorespiratory Fitness: Estimation Equations, Cross-Validation and Application. Journal of Science in Sport and Exercise, 1 (1), 38-53.

Wang, Z. M., Deurenberg, P., Guo, S. S., Pietrobelli, A., Wang, J., Pierson, R. N. Jr. & Heymsfield, S. B. 1998. Six-compartment body composition model: inter-method comparisons of total body fat measurement. International Journal of Obesity Related Metabolic Disorders. 22, 329–337.

Welk, G. J., Bai, Y., Lee, J.-M., Godino, J., Saint-Maurice, P. F. & Carr, L. 2019. Standardizing Analytic Methods and Reporting in Activity Monitor Validation Studies. Medicine and Science in Sports and Exercise. 51 (8), 1767–1780.

Westerbacka, J., Cornér, A., Tiikkainen, M. et al. 2004. Women and men have similar amounts of liver and intra-abdominal fat, despite more subcutaneous fat in women: implications for sex differences in markers of cardiovascular risk. Diabetologia. 47, 1360.

Whaley, M. H., Kaminsky, L. A., Dwyer, G. B. & Getchell, L. H. 1995. Failure of predicted VO2peak to discriminate physical fitness in epidemiological studies. Medicine and Science in Sports and Exercise. 27, 85–91.

Whitmer, R. A., Gustafson, D. R., Barrett-Connor, E., Haan, M. N., Gunderson, E. P., & Yaffe, K. 2008. Central obesity and increased risk of dementia more than three decades later. Neurology.71 (14), 1057–1064.

Wier, L. T., Jackson, A. S., Ayers, G. W. & Arenare, B. 2006. Non-exercise models for estimating VO2max with waist girth, percent fat, or BMI. Medicine and Science in Sports and Exercise. 38 (3), 555–561.

Williams, P. T. 2001. Physical fitness and activity as separate heart disease risk factors: a meta-analysis. Medicine and Science in Sports and Exercise. 33 (5), 754–61.

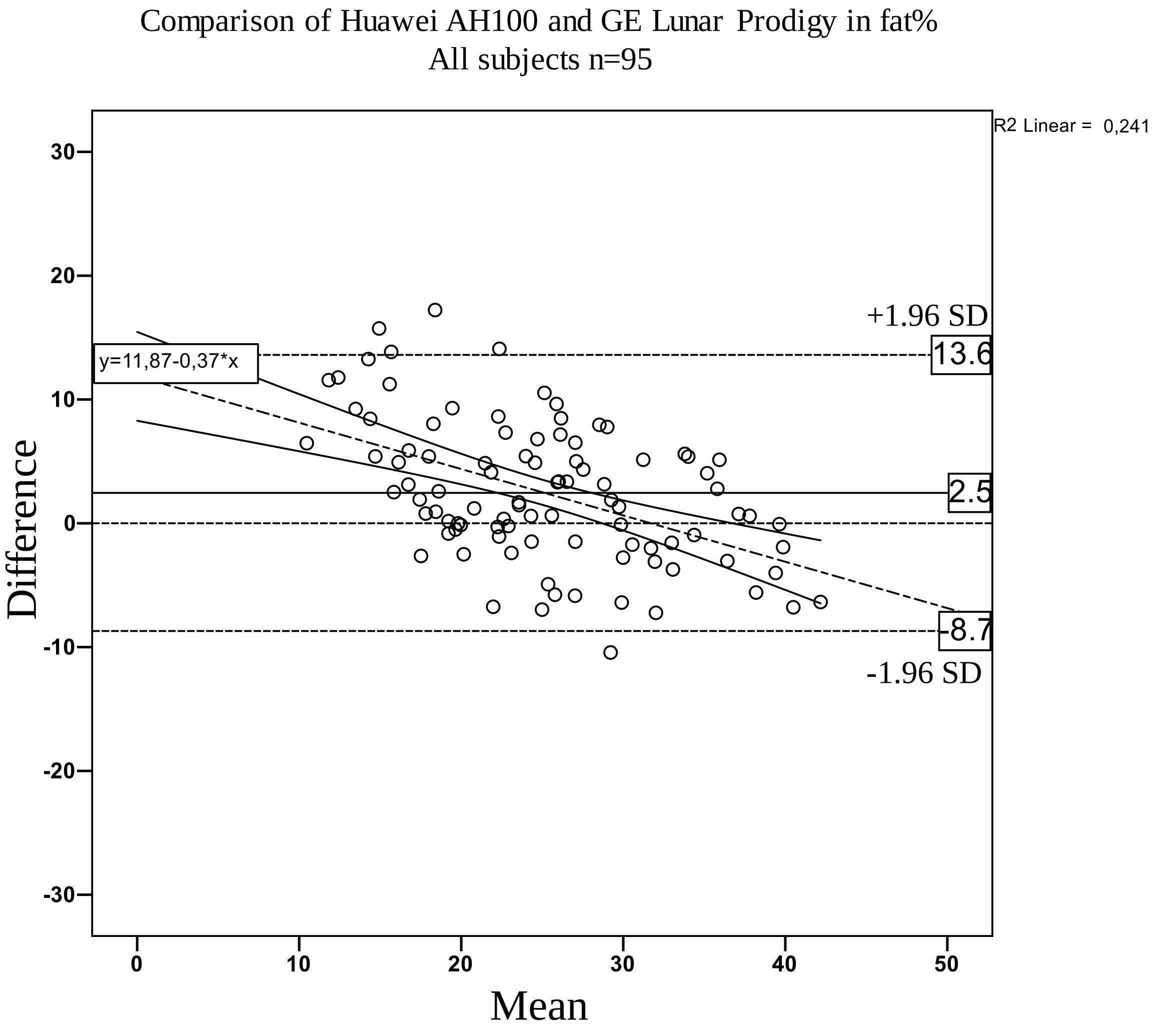
World Health Organization. 1998. Obesity: preventing and managing the global epidemic. Geneva: World Health Organization Programme of Nutrition. 6-3-1997: 7-16.

Zaki, R., Bulgiba, A., Ismail, R. & Ismail, N. A. 2012. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review. PLoS ONE. 7 (5), e37908.

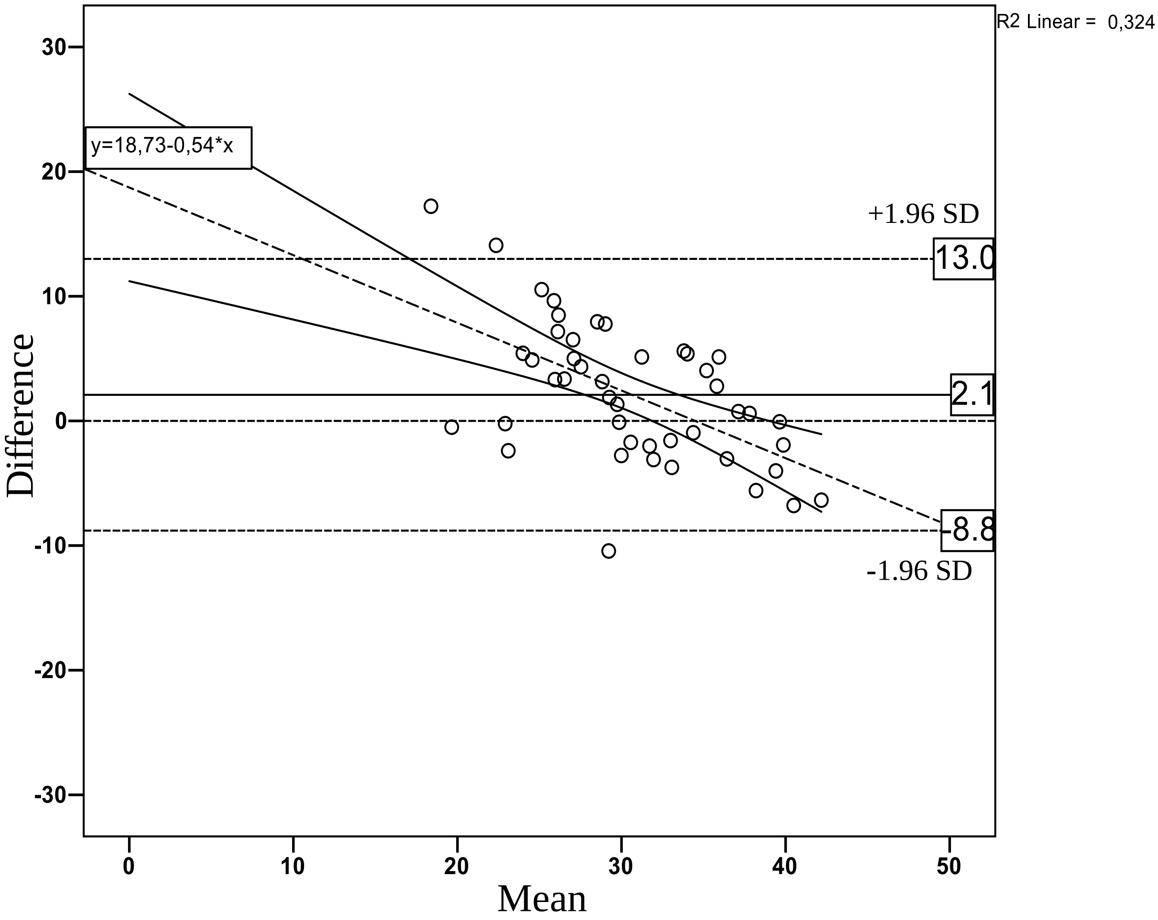
Zhusheng, Y., Volgyi, E., Wang, R., Ember A., Wiklund, P., Alen, M., Tylavsky, F. A. & Cheng S. 2012. Comparison of heart rate monitoring with indirect calorimetry for energy expenditure evaluation. Journal of Sport and Health Science. 1 (3), 178-183.

APPENDICES

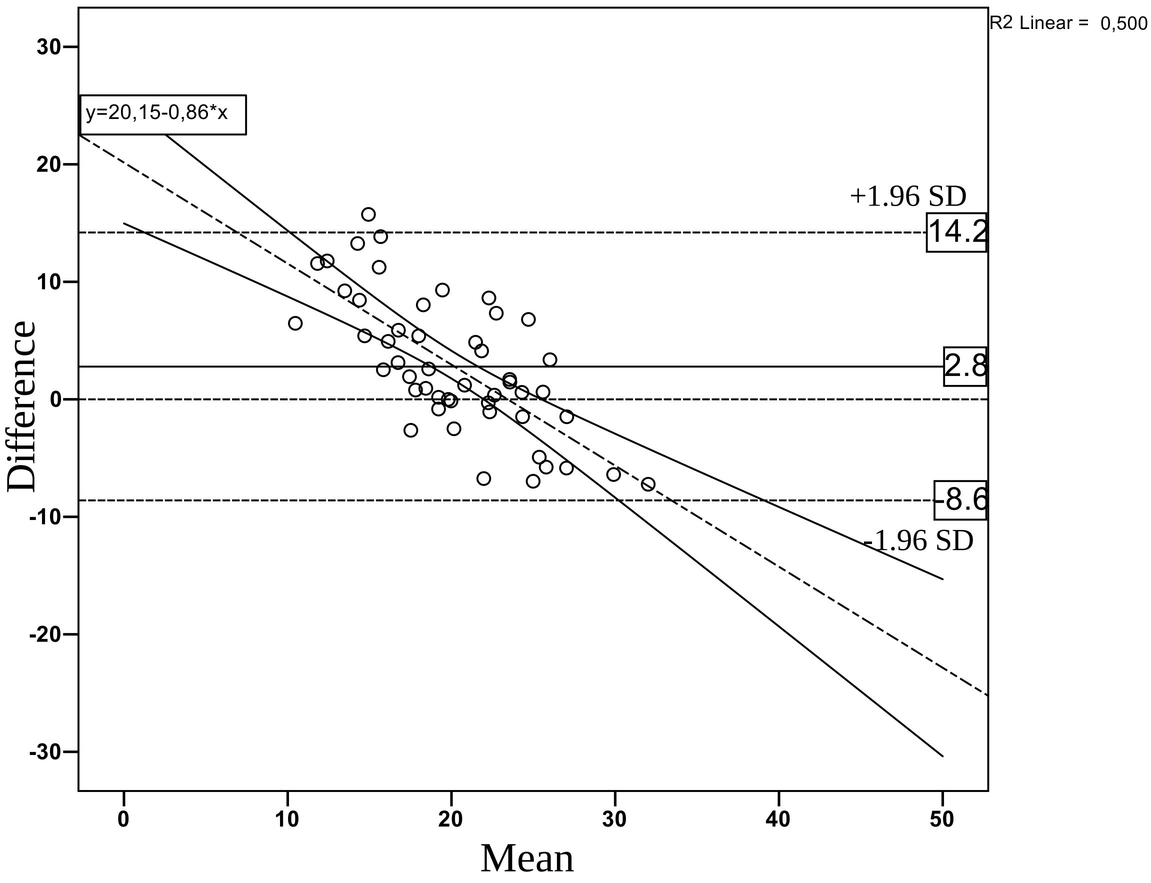
APPENDIX 1. Bland Altman analysis of agreement in fat% between devices.

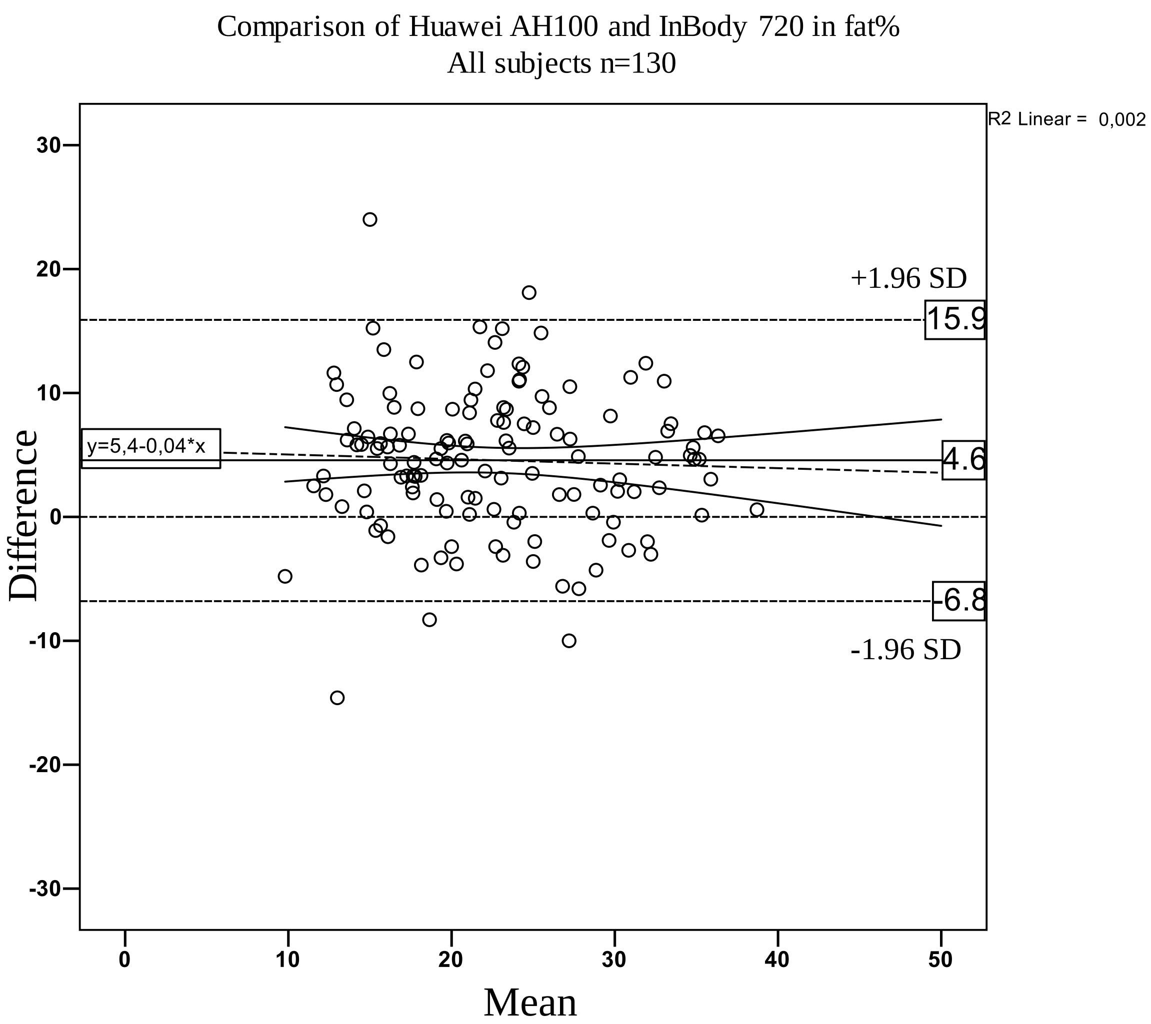


Comparison of Huawei AH100 and GE Lunar Prodigy in fat%
Females n=45

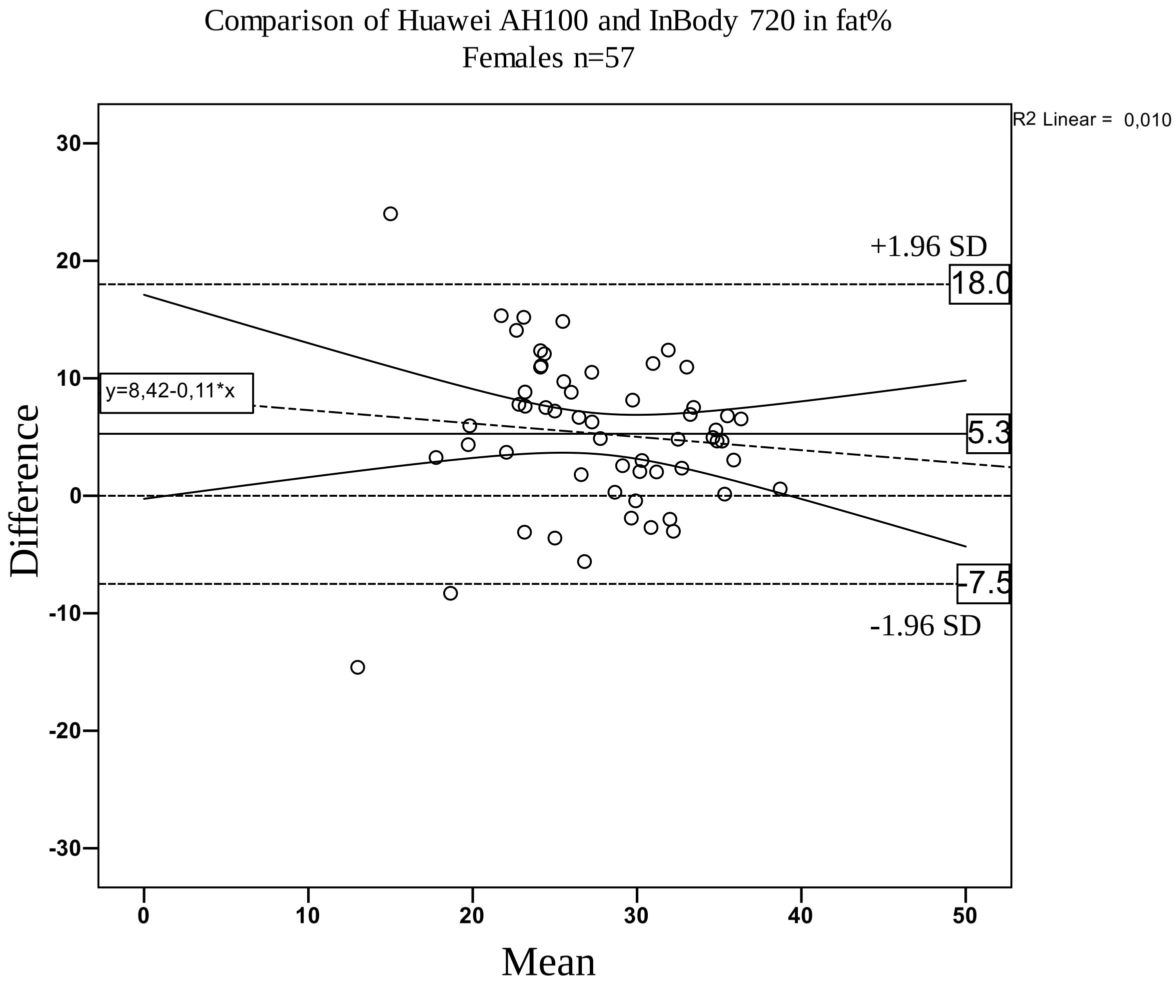


Comparison of Huawei AH100 and GE Lunar Prodigy in fat%
Males n=50

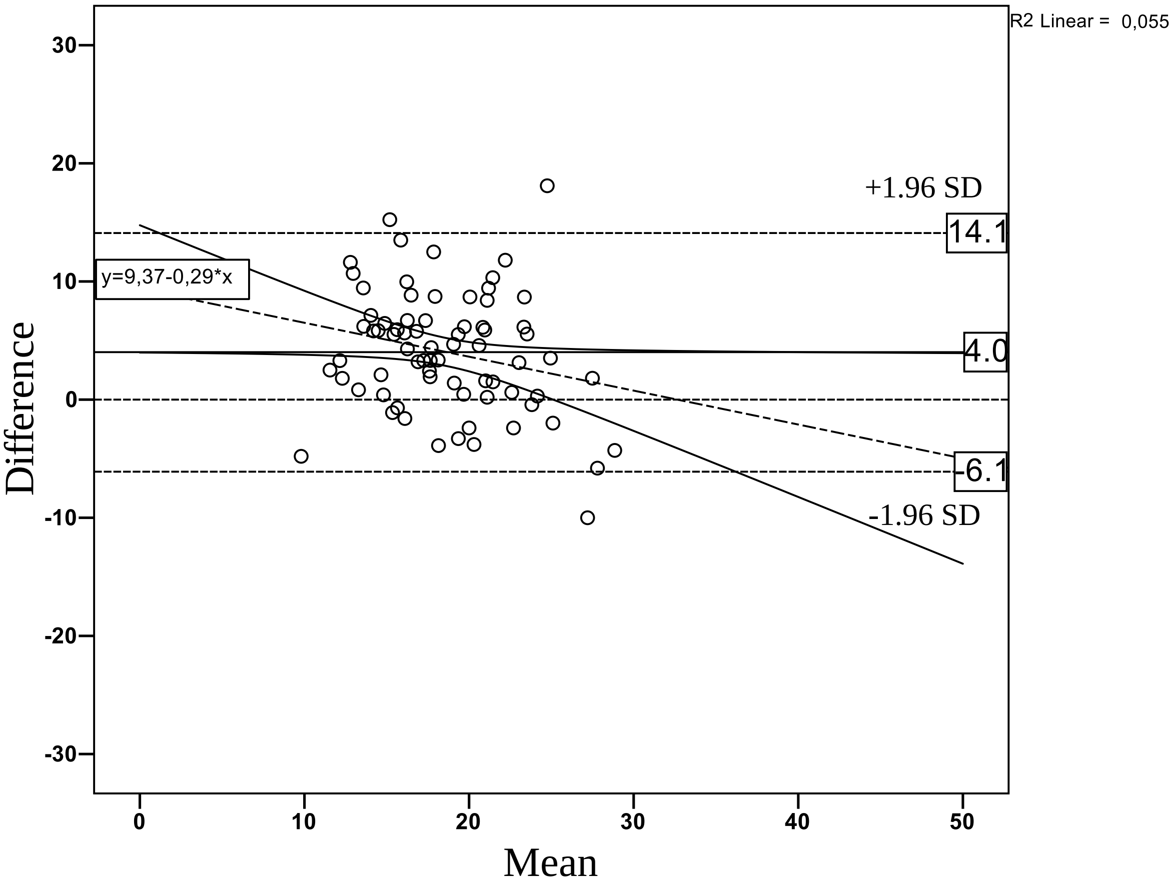




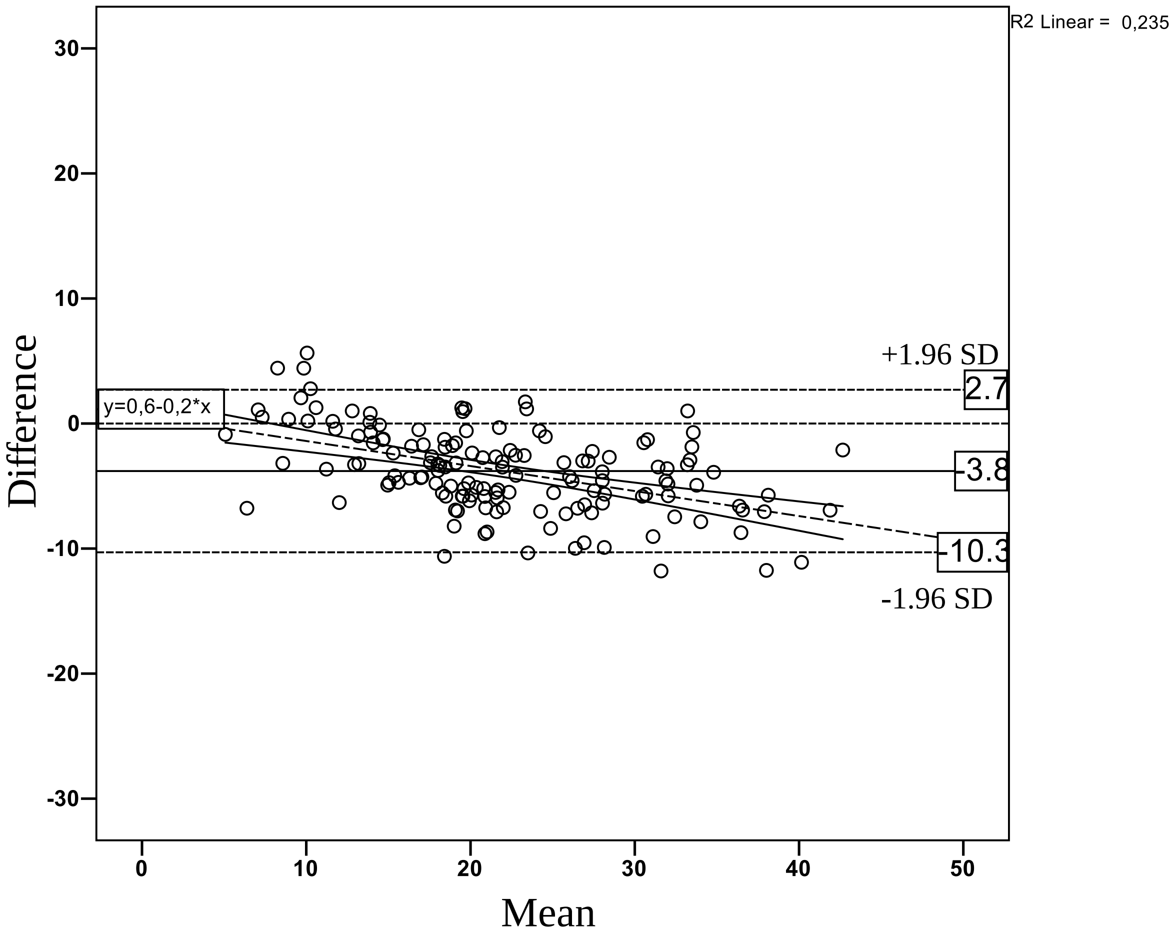
Comparison of Huawei AH100 and InBody 720 in fat%
All subjects n=130
R2 Linear = 0,002
y=5,4-0,04*x
+1.96 SD
15.9
4.6
-6.8
-1.96 SD
Difference
Mean
30
20
10
0
-10
-20
-30
0
10
20
30
40
50

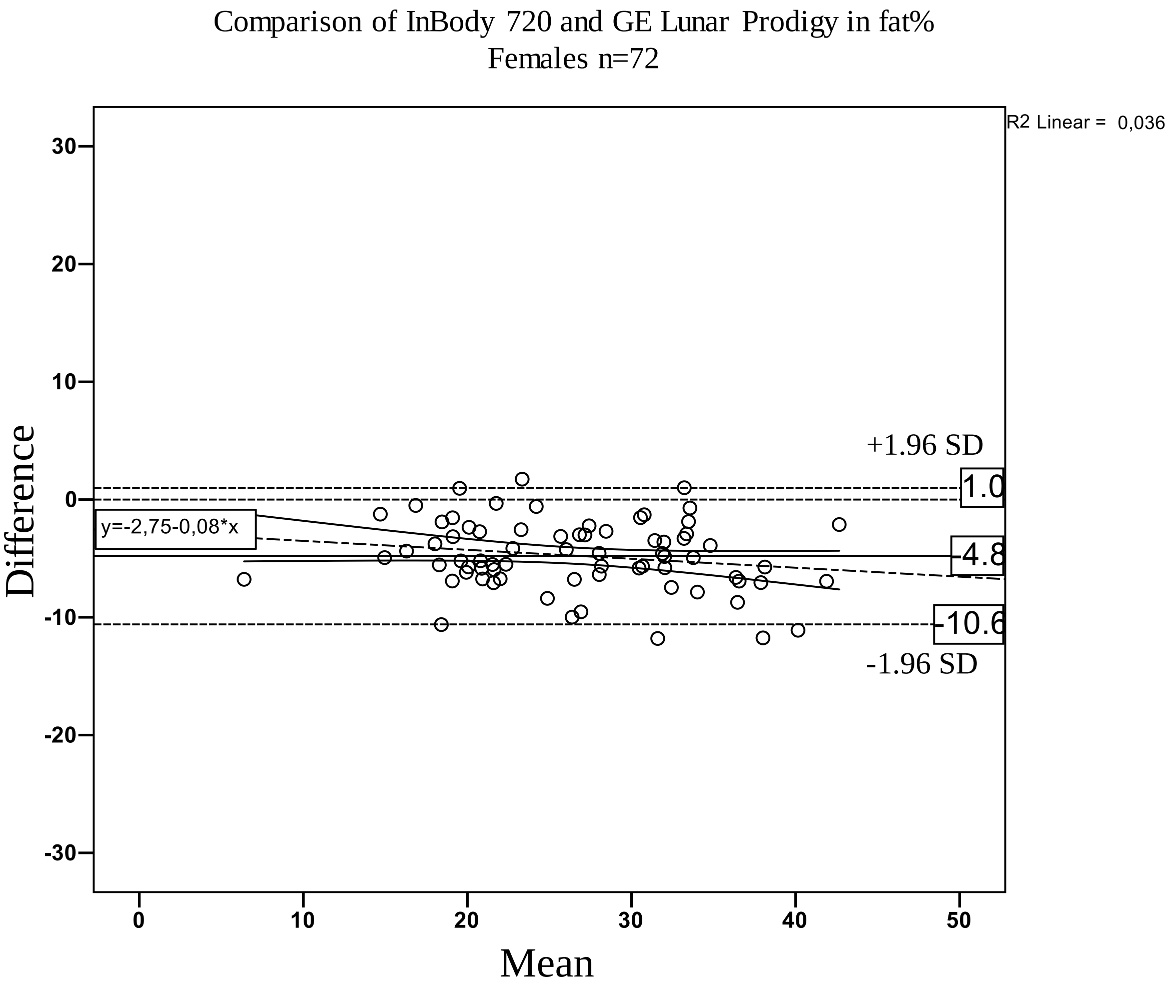

Comparison of Huawei AH100 and InBody 720 in fat%
Females n=57
R2 Linear = 0,010
+1.96 SD
18.0
y=8,42-0,11*x
5.3
-7.5
-1.96 SD
30
20
10
0
-10
-20
-30
0
10
20
30
40
50
Difference
Mean

Comparison of Huawei AH100 and InBody 720 in fat%
Males n=73

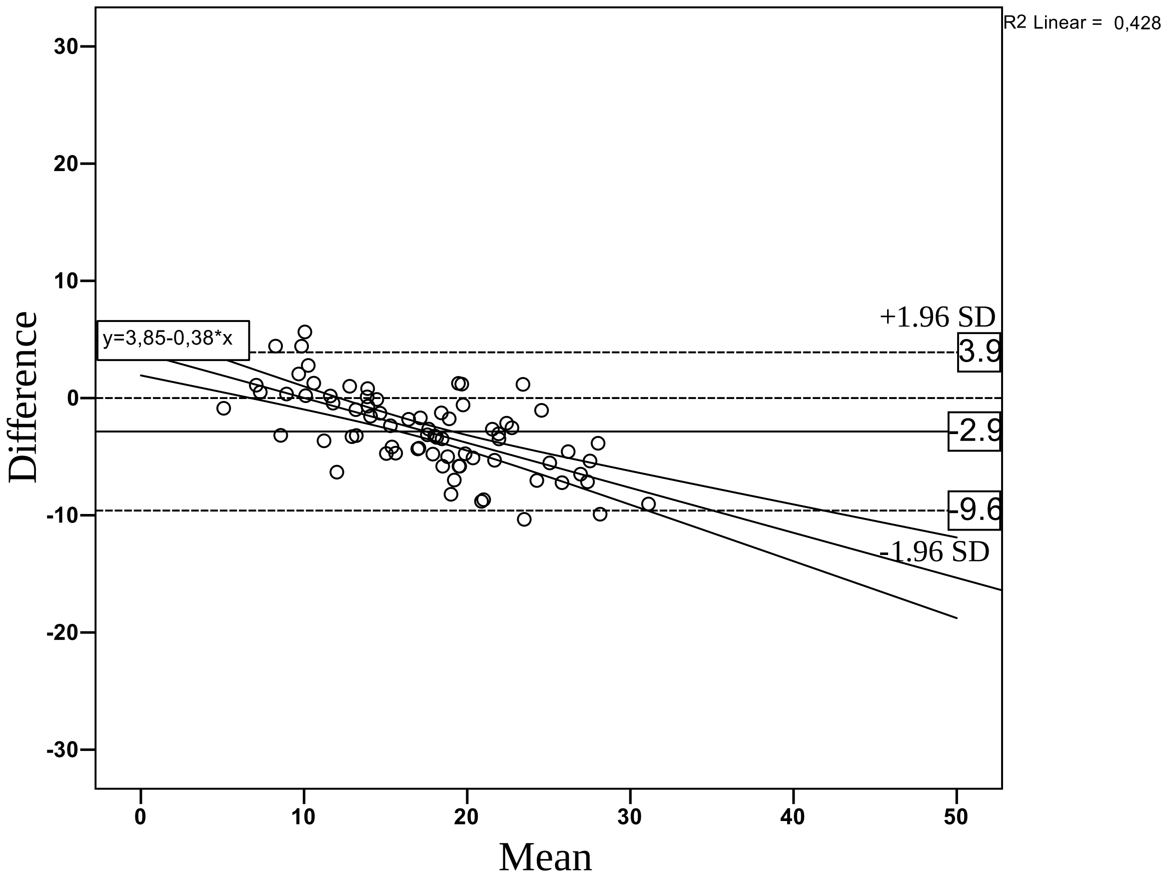


Comparison of InBody 720 and GE Lunar Prodigy in fat%
All subjects n=146




Comparison of InBody 720 and GE Lunar Prodigy in fat%
Females n=72
R2 Linear = 0,036
+1.96 SD
1.0
-4.8
-10.6
-1.96 SD
y=-2,75-0,08*x
Difference
Mean
30
20
10
0
-10
-20
-30
0
10
20
30
40
50

Comparison of InBody 720 and GE Lunar Prodigy in fat%
Males n=74



APPENDIX 2. Descriptive characteristics of subjects in *section 2* of the study.

TABLE 1. Descriptive characteristics and differences between genders in age, BMI, lean body mass, body fat percent, and VO_{2max} in a group where measurements of InBody 720 and VO_{2max} were conducted (n=278).

	Total (n=278)		Male (n=164)		Female (n=114)		Difference	
	Mean (range)	SD	Mean (range)	SD	Mean (range)	SD	MD (t)	p-value[a]
Age	29.22 (18.0-54.0)	8.23	28.46 (18.0-54.0)	7.89	30.30 (18.0-53.0)	8.62	1.84 (1.81)	0.071
BMI	23.39 (16.5-32.7)	2.73	23.70 (16.5-32.7)	2.80	22.95 (18.1-31.5)	2.56	-0.75 (-2.27)	0.024
LBM_{InBody}	52.53 (29.8-78.2)	9.81	58.13 (39.9-78.2)	7.78	44.47 (29.8-63.8)	6.11	-13.66 (-16.38)	<0.001
$Fat\%_{InBody}$	20.53 (3.0-41.6)	7.45	17.04 (4.6-34.1)	5.70	25.56 (3.0-41.6)	6.79	8.52 (10.98)	<0.001
VO_{2max} $L^{-1} \cdot min^{-1}$	3.04 (1.2-4.8)	0.68	3.39 (2.0-4.8)	0.58	2.54 (1.2-4.1)	0.48	-0.85 (-13.28)	<0.001
VO_{2max} $ml \cdot kg^{-1} \cdot min^{-1}$	43.50 (25.1-65.5)	7.64	45.91 (25.4-65.5)	7.42	40.04 (25.1-59.3)	6.59	-5.87 (-6.79)	<0.001

[a] Differences between genders was tested by the Independent samples T-test.

TABLE 2. Descriptive characteristics and differences between genders in age, BMI, lean body mass, body fat percent, and VO_{2max} in a group where measurements of GE Lunar Prodigy DXA and VO_{2max} were conducted (n=140).

	Total (n=140)		Male (n=69)		Female (n=71)		Difference	
	Mean (range)	SD	Mean (range)	SD	Mean (range)	SD	MD (t)	p-value[a]
Age	31.49 (20.4-45.1)	7.07	32.20 (22.6-44.8)	6.61	30.80 (20.4-45.1)	7.47	-1.41 (-1.18)	0.240
BMI	23.85 (17.0-31.5)	2.61	24.25 (17.0-29.9)	2.38	23.45 (18.1-31.5)	2.77	-0.80 (-1.82)	0.070
LBM_{DXA}	52.28 (34.5-78.0)	9.91	60.23 (45.0-78.0)	6.91	44.55 (34.5-61.0)	5.07	-15.68 (-15.26)	<0.001
$Fat\%_{DXA}$	23.91 (5.5-45.7)	9.07	18.69 (5.5-35.6)	7.08	28.98 (9.8-45.7)	7.87	10.28 (8.12)	<0.001
VO_{2max} $L^{-1}\cdot min^{-1}$	3.14 (1.6-4.7)	0.68	3.64 (2.7-4.7)	0.50	2.65 (1.6-4.1)	0.43	-0.99 (-12.56)	<0.001
VO_{2max} $ml\cdot kg^{-1}\cdot min^{-1}$	43.94 (28.2-63.0)	6.79	47.47 (34.6-63.0)	5.46	40.51 (28.2-59.3)	6.20	-6.96 (-7.04)	<0.001

[a] Differences between genders was tested by the Independent samples T-test.

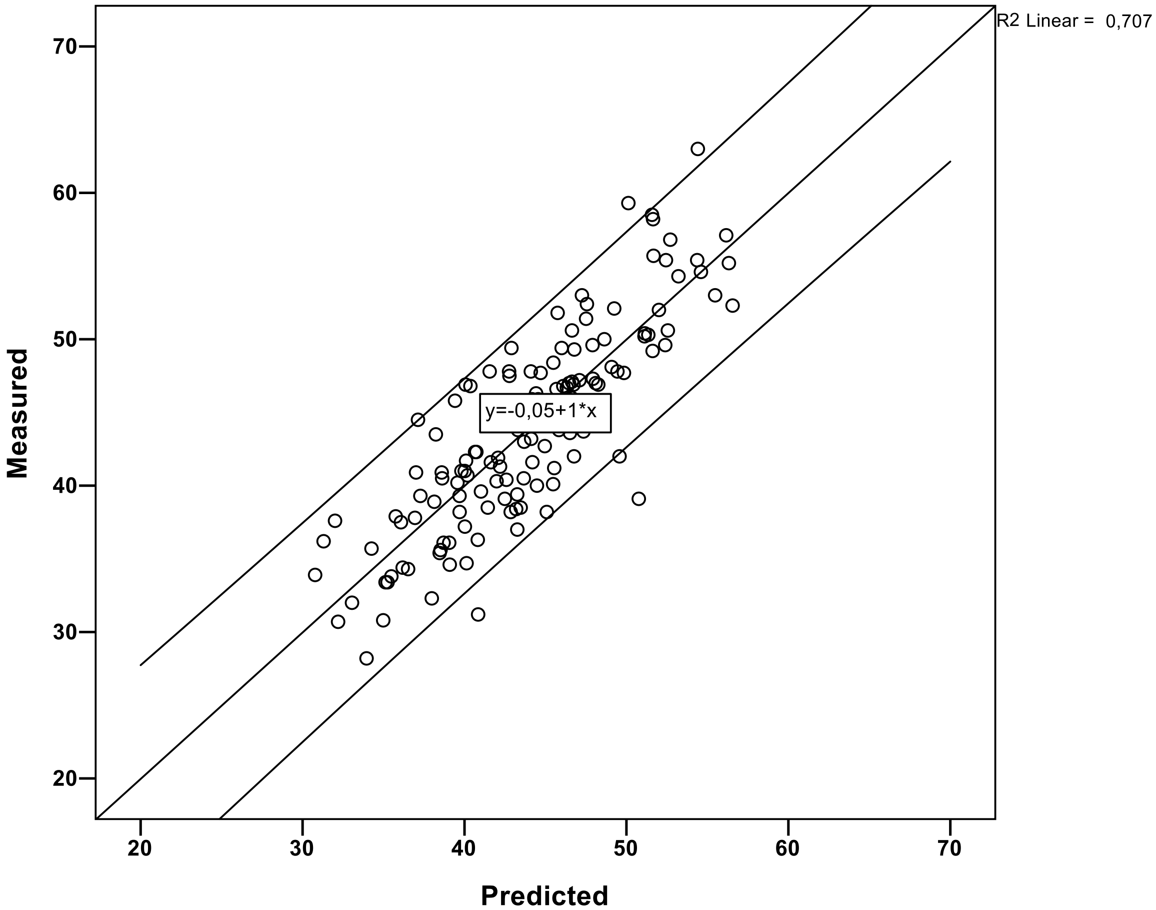
TABLE 3. Descriptive characteristics and differences between genders in age, BMI, lean body mass, body fat percent, and VO_{2max} in a group where measurements of Huawei AH100 and VO_{2max} were conducted (n=125).

	Total (n=125)		Male (n= 71)		Female (n= 54)		Difference	
	Mean (range)	SD	Mean (range)	SD	Mean (range)	SD	MD (t)	p-value[a]
Age	30.68 (18.0-54.0)	8.25	30.12 (18.0-54.0)	8.56	31.42 (19.0-45.0)	7.83	1.30	0.315[u]
						U(123)=1715.5, Z=-1.0 [u]		
BMI	23.85 (17.0-29.9)	2.66	24.02 (17.0-29.9)	2.63	23.64 (18.1-29.6)	2.72	-0.38 (-0.79)	0.431
SMM $_{Huawei}$	50.53 (36.0-69.0)	8.77	56.91 (41.5-69.0)	5.49	42.15 (36.0-56.1)	3.80	-14.75	<0.001[u]
						U(123)=72.0, Z=-9.2 [u]		
Fat% $_{Huawei}$	24.92 (5.0-39.6)	7.25	20.34 (5.0-33.8)	4.74	30.94 (14.5-39.6)	5.32	10.60	<0.001[u]
						U(123)=248.5, Z=-8.3 [u]		
VO_{2max} $L^{-1}\cdot min^{-1}$	3.15 (1.6-4.6)	0.67	3.55 (2.3-4.6)	0.51	2.62 (1.6-4.1)	0.46	-0.94	<0.001[u]
						U(123)=311.0, Z=-8.0 [u]		
VO_{2max} $ml\cdot kg^{-1}\cdot min^{-1}$	44.35 (28.2-63.7)	7.59	47.70 (33.3-63.7)	6.41	39.95 (28.2-59.3)	6.77	-7.75 (-6.54)	<0.001

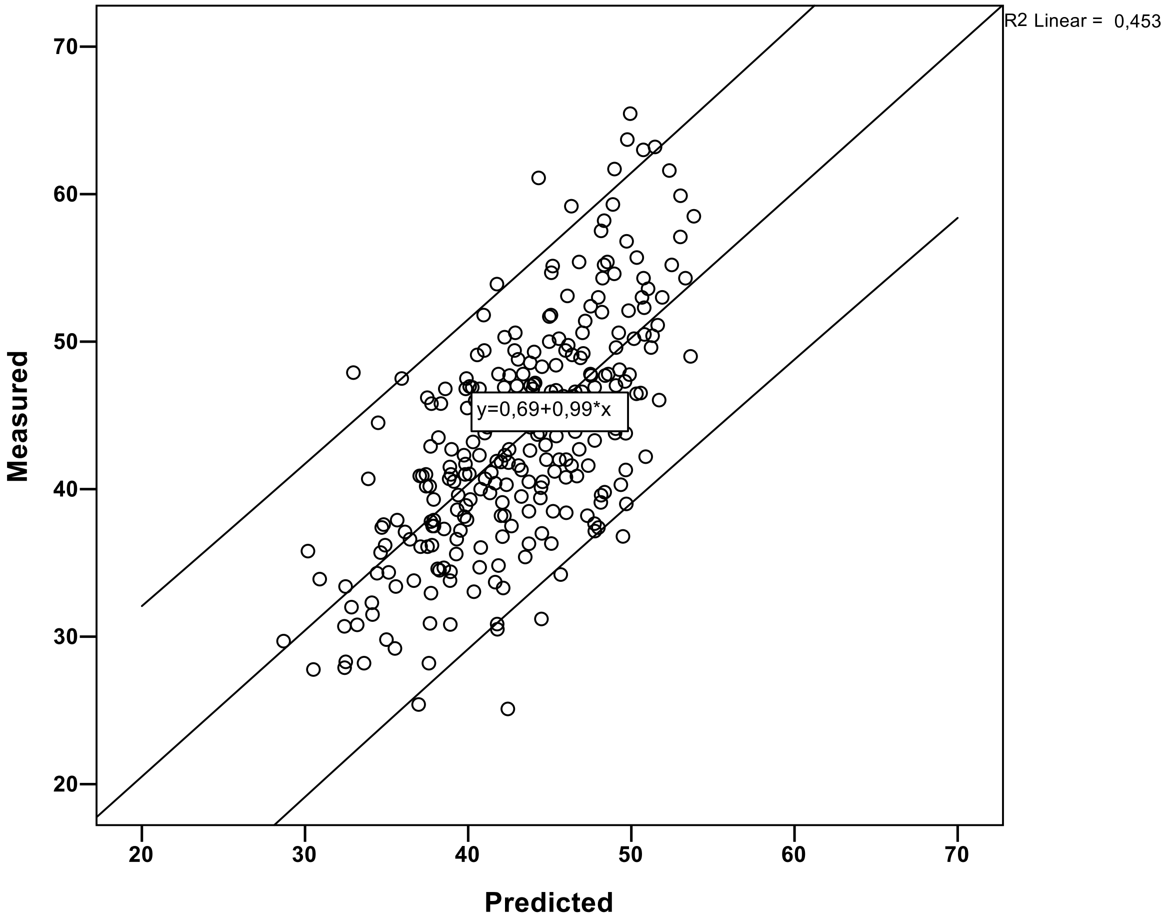
[a] Differences between genders was tested by the Independent samples T or Mann-Whitney U[(u)] test.

APPENDIX 3. Linear relationships between measured and predicted VO_{2max}.

Model 1. VO_{2max}= 53.57 + ($fat\%_{GE\ Lunar\ Prodigy}$) * -66.34 + (BMI) * 0.50 + (age) * -0.18.



Model 2. VO_{2max}= 62.95 + ($fat\%_{InBody\ 720}$) * -0.99 + ($LBM_{InBody\ 720}$) * -0.32 + (BMI) * 1.03 + (age) * -0.23.



Model 3. VO_{2max}= 58.16 + (fat%$_{Huawei\ AH100}$) * -0.40 + (age) * -0.18 + [gender (male=1, female=0)] * 3.25.