Reliability and Validity of Body Fat Determination in Elite Female Athletes and the Implications for Practitioners

By

Daryl M. G. Hurrie

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Faculty of Kinesiology and Recreation Management

University of Manitoba

Winnipeg, Manitoba

ABSTRACT

PURPOSE: To establish the reliability of anthropometric and dual energy X-ray absorptiometry (DXA) techniques used to assess percent body fat (% BF) in female athletes; to establish limits for detecting the smallest real change in % BF associated with anthropometric and DXA testing; to evaluate the validity of commonly used % BF prediction equations recommended by national certification programs along with equations derived from Multicompartment (MC), and DXA, in female athletes; and to create a new DXA based regression equation for elite female athletes. **METHODS:** Female athletes aged 17-31 were recruited into the study and participated in the establishment of anthropometric reliability (N=20), DXA reliability (N=32), and /or skinfold validity (N=95) testing. Anthropometric testing consisted of measurements of skinfolds, circumferences, and breadths. DXA measurements were conducted using a GE Lunar Prodigy DXA which served as the criterion measure (% BF DXA). **RESULTS:** Excellent reliability for both anthropometric sum5 skinfolds (ICC= .997, %TEM=0.9 %) and DXA (ICC =.996, CV =1.13% BF) techniques allows for detection of smallest real differences of 2.2 mm and 721g in summed skinfolds (sum5) and fat mass respectively. The DXA based equation of Ball et al. (2004) displayed the greatest validity of existing equations R=.874, total error (TE) 2.9% BF, and Bland Altman Limits of Agreement -4.7to 6.5 % BF. The newly created regression equation demonstrated a non-linear characteristic and displayed similar predictive ability R= .840, TE 3.0%BF, and Bland Altman Limits of Agreement of -6.1to 6.1 % BF. **CONCLUSIONS:** Anthropometric equations derived from various criteria yielded dissimilar results. Long utilized popular equations advocated in national accreditation schemes (ACSM, CSEP) show considerable bias compared to modern values obtained by current DXA technology. A new

regression equation was created for female Canadian athletes 17-31yrs of age using skinfolds taught in the Canadian national professional certification program (CSEP).

ACKNOWLEDGEMENTS

This thesis has been a labour of love and well.....not so much love. I'm sure that many researchers can share this sentiment, having experienced the simultaneous frustration of trying to decipher many years of literature, while anticipating possible discoveries amidst the book stacks or within their own experiments. Thus, although it has been exciting at times, at times this thesis has been my bane, and I could not have achieved its completion without the notable help of those listed below;

I would first like to thank all the ***participants*** involved in the study. Your time as students and athletes is valuable and I appreciate your donation towards my curiosities. Additionally, I must give recognition to my advisory committee, ***Dr. Todd Duhamel, Dr. Greg Gannon*** and my primary advisor ***Dr. Michelle Porter***. ***Todd***, thank you for your critical eye and objective input, I realize all the tough questions were designed to make me a better researcher. ***Greg***, thank you for helping me start this journey, and giving me the tools to "think outside the box". ***Michelle,*** I am extremely appreciative of your commitment and dedication in helping me and teaching me above all else, how to conduct research. Your exemplary council, guidance and relentless edits were eye opening for my future endeavours into academia. I have learned so much!

I would like to extend my thanks to the staff at the ***Richardson Centre for Functional Foods and Neutraceuticals (RCFFN)*** for allowing me to use their world class equipment, and accommodate my rigorous testing schedule in their most beautiful facility. Additionally I would like to express great gratitude towards my family and home away from home, the staff of the ***Canadian Sport Centre Manitoba (CSCM).*** Thank you all for your support and encouragement. The rigorous debates in the "Hot Stove" allowed me to approach my research with a truly multidisciplinary perspective. I would also like to make special thanks to ***Randy Anderson,*** GM of the CSCM. Your unwavering support over the past couple of years has allowed me to achieve my own "personal excellence".

Lastly, I need to thank all my friends and family. ***Friends***, for consistently asking me, when I will be done? Hopefully, I have silenced you all! ***The Bryant's*** for putting up with my shop talk at all the family dinners and the belief that I could achieve greatness. ***Mom, Dad, Kev,*** thanks for the early years, the middle years, and the final years. It's hard to believe I've been going to school for so long. Thank you for letting me evolve without constraint or pressure. Your unconditional love was enormous. ***Jen***, you have provided me with sympathy when I was down, encouragement when I was broken, accountability when I procrastinated, comfort when it was needed and food when I was too busy writing or reading to eat. Words fail in expressing my thanks, and I'm sure I will spend the next couple years trying to pay back the kindness you have showed me throughout this process. Thank you all, we did it!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1 -INTRODUCTION

1.1 Introduction - Statement of Problem

1.1.1 Rationale for applied body compositional assessment

It is evident that body composition is a significant factor contributing to sport performance (Benardot & Czerwinski, 1991; Sovak *et al.*, 1992; Carter & Ackland, 1998; Anderson *et al.*, 2008). The common observation is that elite athletes are in general leaner and more muscular than their less-elite (Bayios *et al.*, 2006; Kerr *et al.*, 2007; Malousaris *et al.*, 2008) or recreational counterparts (Helmuth, 1980; Sovak & Hawes, 1987; Tsunawake *et al.*, 1995; Bourgois *et al.*, 2001; D'Alessandro *et al.*, 2007). It can be debated whether these are specific adaptations resulting from the demands of the sport, or predisposed genetic traits that allow participation at higher levels, but body composition remains a pertinent factor that must be taken into consideration when investigating sport performance. Attempts to optimize body composition requires objective feedback to assess composition of weight loss or weight gain (Storlie, 1991), as well as create strategies for these changes in a healthy manner, as many athletes could benefit from safe and structured guidance for attaining their compositional goals (AAP, 1996; Rankin, 2002; Hagmar *et al.*, 2008). In addition, coaches and athlete support staff require this information to assess the efficacy of training programs. Two forms of information are conventionally utilized to make decisions about training or performance strategies. Group statistics or values provide objective assessments related to the global success of a program allowing modifications to the training program. However, the ability to identify change within

an individual is of greater importance, as this allows individualization, a key principle to optimizing training stimuli for elite athletes (Bompa, 1999).

1.1.2 Methods of Investigation

Quantification of body composition is undertaken by dividing the body into a series of compositional partitions. Often these theoretical models result in 2 to 5 component models that are used to explain the constitution of the body (Sutcliffe, 1996; Malina, 2007). The quantification of a single partition within a model allows for the extrapolation of the residual component as the sum of the partitions will always yield the composition of the total system. A primary example with the lowest level of complexity is the 2-component (2C) model. The 2C model divides the body into Fat Free Mass (FFM) and Fat Mass (FM) components. The sum of the two represents the composition of the total system, or Total Body Mass (TBM). An investigator only needs to quantify two of the three variables in this example in order to calculate the quantity of the third. Conventionally, the greater the number of partitions that are measured, the fewer assumptions required to "reconstitute" the total system and thus, less error is associated with the method. However, as more partitions are measured, the complexity and accessibility of the method will limit its utility for applied practitioners looking to assess their athletes in an ongoing manner.

1.1.3 Anthropometry

Perhaps the most common and longest standing field based method of body composition analysis is anthropometry (Wang, 2000). Anthropometry utilizes a series of skinfold measurements, circumferences, bone lengths and breadths that are rendered in statistical analysis

to represent %Body Fat (% BF) and/or Fat Free Mass (FFM). The attractiveness of the method among practitioners stems from its portability, low cost, non invasiveness, and a wide array of published norms for comparison (Roche, 1996). The limitations include methods of standardization, as well as the inter and intra tester reliability of the anthropometrist (technician performing the anthropometric measurements). The utility of using common skinfolds to represent relatively small changes in an individual athlete's composition is widely practiced despite a lack of global acknowledgement of the limitations when predicting body fat of individuals.

1.1.4 Hydrodensitometry (HD) Criterion

Early anthropometric prediction equations were commonly created using Hydrodensitometry (HD) as the criterion measure of % BF. The method entails measuring the subject's weight in water. Based on assumptions of densities of Fat and Fat Free Mass (FFM), total body density can be estimated. As these early regression equations are used to predict Body Density (BD), an additional calculation is employed to convert the BD to % BF based on assumptions created by a small series of cadaver dissection studies (Siri, 1956; Brozek *et al.*, 1963). In this regard, these equations are "doubly indirect" in that they represent estimates of fatness, based on an estimate of body density, and are mathematically twice removed from the source of which they attempt to quantify (ISAK). The majority of commonly utilized predictive regression equations utilize hydrodensitometry (HD), and the 2–Component (2C) model as the reference criterion (Sloan *et al.*, 1962; Durnin & Womersley, 1974; Jackson *et al.*, 1980; Thorland *et al.*, 1984). Furthermore, many of these equations are recommended by professional certification agencies in the fields of kinesiology and public health, and remain the primary

pedagogy being taught to new practitioners (e.g., American College of Sports Medicine (ACSM) and Canadian Society of Exercise Physiology (CSEP)). Limitations to these skinfold equations, independent of the error associated with the anthropometrist's reliability, arise from limitations of the HD criterion which may no longer represent the most robust criterion available (Lohman *et al.*, 2000). It has been suggested in some studies that due to the error associated with anthropometric techniques, the ability to detect change within individual subjects is marginal when expressed as % BF from HD regression equations (Piers *et al.*, 2000; Friedl *et al.*, 2001). This has led to exploration of new methodologies for tracking compositional changes within a subject (Slater *et al.*, 2006).

Recently, studies have evaluated the validity of HD equations using multi-compartment methods as the criterion and found that the equations show considerable amounts of error. Magnitudes of % BF underestimation have been confirmed as being 3.0 to 7.6 % BF compared to DXA criterions (Ball *et al.*, 2004; Duz 2009), and up to 6.6 % BF (Peterson *et al.*, 2003) compared to a DXA 4C criterions. This has prompted the investigation of newer, potentially more accurate prediction equations.

1.1.5 Dual Energy X-Ray Absorptiometry (DXA) Criteria

New methods with increased prediction accuracy over older practices have been adopted largely because of precision and speed of implementation. In this regard, the 3 component (3C) model provided by Dual Energy X-ray Absorptiometry (DXA) appears to have promising application in the field of athlete body composition measurement. In the 1990's extensive research was conducted on novel DXA models with promising results. DXA was found to

accurately predict body composition of athletes when compared to multi-component models (Prior *et al.*, 1997) and was shown to accurately detect small changes in body composition (Going *et al.*, 1993). New skinfold regression equations developed using DXA as the criterion have yielded accurate prediction equations with Standard Error of Estimates (SEE) for % BF as low as 2.5% and SEE for FFM as low as 1.1 kg have been reported (Ball *et al.*, 2004; Warner *et al.*, 2004). In roughly the last two decades numerous advancements in DXA technologies have been made, making it necessary to continuously research and establish the validity, and reliability of the machines. Furthermore, since many of the DXA based equations are relatively new, published cross-validation of these equations in different samples is rare.

1.2 Statement of Study Novelty

The reliability of both anthropometric and DXA assessment need review to articulate recommendations for detecting compositional changes within individual subjects, based on magnitudes of reliability. The smallest real difference (SRD) will be calculated to define the magnitude of change necessary between serial measurements to signify change at the 95% confidence level, which is an uncommon statistic in this field despite being advantageous for practitioners. Methodological differences owing to equipment variation, sample selection and statistical methods for regression equation construction will all influence the ability of existing regression equations to accurately predict the body composition of elite female athletes and necessitate the cross-validation of these methods in populations with characteristics that may differ from the original samples (Schmidt & Carter, 1990). Commonly used equations recommended by national certification programs such as the American College of Sports Medicine, (ACSM), and the Canadian Society for Exercise Physiology, (CSEP) need to be

evaluated as they remain fundamental practices taught in post secondary institutes. In addition, a series of equations derived from newer criterion models including Multicompartment (MC) and DXA need to be investigated for their ability to predict % BF in female athletes with a General Electric (GE) Lunar Prodigy DXA as the criterion measure (% BF DXA).

Currently no study has published skinfold regression equations for predicting % BF in female athletes, utilizing measurements obtained with Harpenden skinfold callipers and a GE Lunar Prodigy Fan Beam DXA as the criterion reference. A statistical framework for assessing and interpreting the smallest real change within individual subjects will be investigated along with practical recommendations for utilizing anthropometric methods to interpret and report body composition changes when working with individual or group subjects in applied settings.

1.3 Statement of Purpose

The purpose of this research study is to:

1) Establish confidence limits for detecting within subject smallest real change in % BF associated with anthropometric and DXA testing (GE Lunar Prodigy);
2) Evaluate the validity of existing MC, DXA, and HD based skinfold regression equations in elite female athletes using a GE Lunar Prodigy DXA;
3) Establish a new DXA based regression equation for elite female athletes;
4) Provide recommendations for best practices for practitioners utilizing regression equations for prediction of % BF;
5) Provide suggestions for interpreting within subject change in applied situations.

If you can this page should be removed.

CHAPTER 2 - REVIEW OF LITERATURE

2.1 Theoretical Models of Compositional Analysis

The study of body composition continues to expand as the precision, validity and accessibility of its methodologies improve. What was once thought to be the criterion, the Hydrodensitometry (HD) 2-component model (2C) has been replaced by more sophisticated multi-component models (up to 5C) (Wang *et al.*, 1992). Investigations of these new methods of body composition have been categorized as either descriptive or mechanistic studies (Wang *et al.*, 1995). Regardless of the method used, all compositional research has common limitations owing to the fact that all in vivo methods are to an extent indirect (Malina, 2007). Implementing new methods with increased predictive accuracy over older practices is dependent on financial, temporal, spatial and practical resources. In this regard, the 3C model provided by DXA appears to have promising application in the field of athletic body composition measurement.

To aid understanding of common compositional measurement techniques a taxonomy for systematically classifying the methods of investigation has been proposed (Wang *et al.*, 1992). This system emphasizes the complexity of compositional research but provides a useful framework for identifying the shared assumptions and limitations of existing analytical procedures. The structural framework proposes a 5 level model that allows for distinguishing the five categories of compositional investigation. Progression from one level of investigation to the next is accompanied by increasing levels of complexity but usually increasing level of accuracy.

Briefly, the most basic level is the Atomic level. It is the study of the basic chemical composition. This analysis is only conducted in limited cadaver studies, elemental analysis, and with highly sophisticated machinery such as In Vivo Neutron Activation (IVNA) techniques. The next level is the Molecular level. Five principle components are analysed and include water lipid protein, minerals and carbohydrate. DXA is classically defined at this level as it is said to measure the molecular content or mineral, as well as using aggregation of other elements to quantify lean mass and fat mass from full body scans. The third level is the Cellular which views the body as composed of cells and extracellular substances. Muscle mass estimated from total body water is an example at this level, that uses intracellular water and extracellular water content as a fixed ration of FFM (Sheng & Huggins, 1979). The fourth level is Tissue, and is the analysis of the contributions of mass. There are few methods in vivo, but a well accepted method is Computed Axial Tomography (CAT) which can evaluate skeletal muscle, adipose tissue, bone, viscera, and brain. The final and simplest level is the 5th Whole Body level. The measurements are taken of size, shape, physique, proportions, volume and density. This is common practice in anthropometry, and photographic methods, and is the most widely employed level of investigation for body composition studies.

Further breakdown of the 5-level model of body composition research shows that each individual level can be subdivided into various compartments of investigation (Mattsson & Thomas, 2006). DXA measures molecular quantities of bone mineral, fat mass and residual (calculated as lean mass) which is termed a 3 compartment investigation. HD takes a series of measurements at the whole body level and uses relationships to infer molecular composition

divided into 2 compartments, fat free mass and fat mass. It is common to classify compositional methods by the number of compartment subdivisions being investigated.

In a paper released three years after their 5-Level Model, Wang and colleagues took the taxonomy of body composition further by proposing a classification system for body composition investigation methodologies that proceeds in stepwise fashion allowing identification of similarities and differences between diverse methods (Wang *et al.*, 1995). The first step in the hierarchy is defining whether the methodology is done in vitro, or in vivo. Applied body composition methods are in vivo and thus, only this branch of the classification system will be further articulated (Figure 2.3). The fundamental concept among all in vivo methodologies is that a mathematical function is always necessary to relate a measureable quantity, to the unknown component needing to be quantified.

Eqn 2.1: $C = f(Q)$

Fundamental Mathematical Concept where C represents the unknown component, Q is a measureable quantity, and F is a mathematical function that relates C to Q.

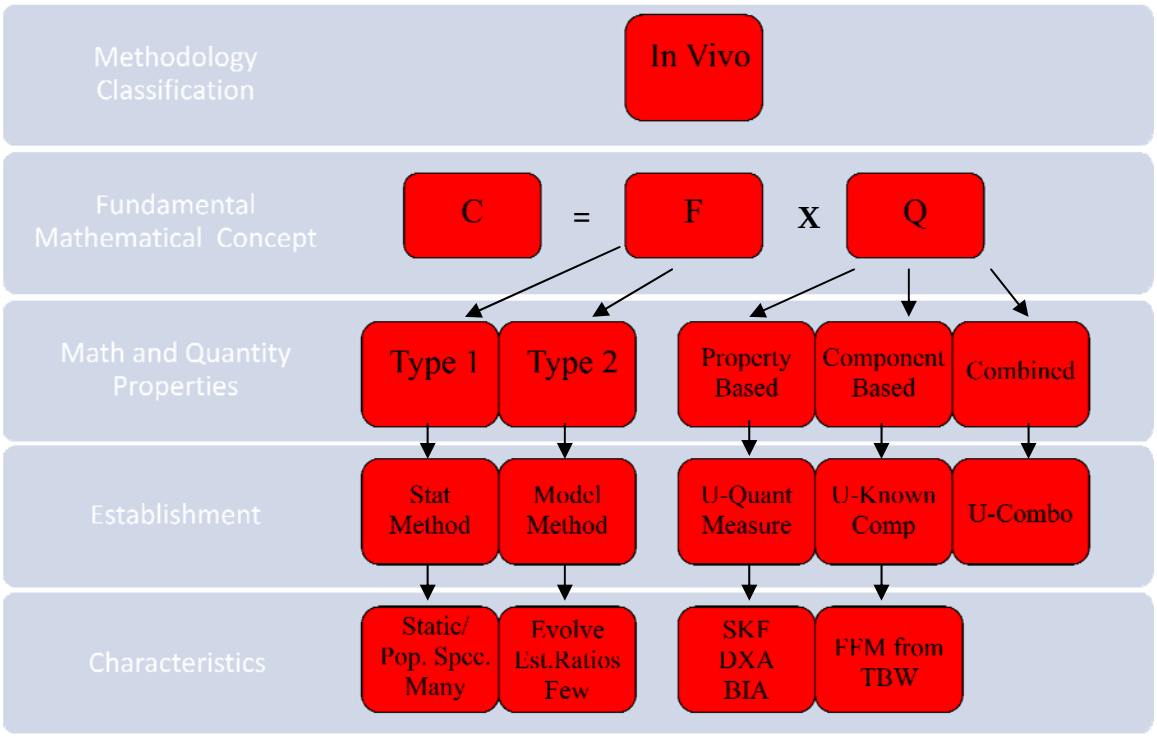


Figure 2.1: A schematic of the classification system for methods of studying body composition proposed by (Wang *et al.*, 1995). A classification of body composition methods proposed by Wang et al, suggests that all in vivo body composition methods may be represented by the fundamental equation where C is the unknown component needing quantification, F is a mathematical function relatingQ a known quantity to C.

Of the variables used to derive the unknown component (C), measureable quantities (Q) are broken down into *property based methods, component based methods,* and a combination of the two. *Property based methods* use measured properties to quantify the unknown component. *Component based methods* use ratios of known components to quantify unknown components. In rare instances there are methods that use combinations of these two to evaluate body composition.

The mathematical function (f) is organized into *Type1* and *Type 2* mathematical characteristics. *Type 1* mathematical functions relating Q to C are derived from statistical approaches such as regression equations, and require a reference point or criterion measure to

quantify the unknown. Thus, the characteristics of a *Type 1* mathematical function are that they are population and condition specific. Any deviations from these parameters require re-evaluation which is why regression equations created from skinfolds must continuously be scrutinized. As well, there are many of these types of equations as they are easily formulated. *Type 2* mathematical functions relating Q to C are created upon assumptions that mathematical constants or ratios inherent to the unknown variables and can be quantified from the measurement of a known variable when the ratio is applied. The characteristics of these mathematical functions are that they are stable between individuals and they evolve slowly over time resulting in fewer models in the literature.

Understanding of the fundamental framework for classification system allows for gap analysis in the literature and study design for the investigation of body composition. This classification system further demonstrates how all methods of *in vivo* body composition analysis are indirect and the varying degrees of separation from the true measurement as well as the mathematical method employed will all contribute to the measuring ability.

2.2 Body Composition and Athletic Performance

In an athlete's quest to excel in their specific sport it has become increasingly evident that body composition is a significant factor contributing to performance (Benardot & Czerwinski, 1991; Sovak *et al.*, 1992; Carter & Ackland, 1998). In a review of athlete evolution, Norton and Olds (2001) reflected on basic measures such as weight, stature, and body mass index (BMI; kg/m2) of the modern elite playing field and stated "sport is Darwinian in that only the 'fittest' reach the highest level of participation". The morphological evolution of today's

athlete is advancing more rapidly than the secular trend of the general population (Norton & Olds, 2001). The concept of morphological optimization signifies that athletes differ from the source population by having either different mean values for particular characteristics, different distribution of characteristics altogether or both (O'Connor *et al.*, 2007). The study of elite athletes has identified this population as being leaner and more muscular than their less-elite (Bayios *et al.*, 2006; Kerr *et al.*, 2007; Malousaris *et al.*, 2008), or recreational counterparts (Helmuth, 1980; Sovak & Hawes, 1987; Tsunawake *et al.*, 1995; Bourgois *et al.*, 2001; D'Alessandro *et al.*, 2007). The increased demand to have specialized physical characteristics for sports competition, as well as the limitations of basic methods such as BMI monitoring to identify sport success (Prentice & Jebb, 2001; Jonnalagadda *et al.*, 2004) has led to the common practice of quantifying individual athlete's body composition so that they may favourably alter their physique through training and diet, for success in their respective sport.

Several studies have shown body composition parameters to have predictive abilities for performance outcomes in heterogeneous samples (Bale *et al.*, 1985) however, as samples become more homogenous, as is the case with elite populations, the ability to differentiate performance outcomes based on compositional indices diminishes (Stoggl *et al.*; Knechtle *et al.*, 2007). However, there still remain instances where the predictive ability is maintained (Claessens *et al.*, 1994; Legaz & Eston, 2005; Arrese & Ostariz, 2006; Legaz Arrese *et al.*, 2006; Anderson *et al.*, 2008) indicating that these relationships may be sport specific. In this regard the greatest value of compositional measurement is in objectively tracking changes of individuals over time. Compositional monitoring with the intent of insuring healthy practices of compositional change in sports has been suggested, where compliance with weight class is part

of the rules, such as rowing and wrestling (Oppliger *et al.*, 1995; Karlson *et al.*, 2001). In sports where aesthetics are inherent to the competition, a great deal of pressure is placed on achieving optimal body composition. An abundance of research on the atypically lean body compositions of female gymnasts, figure skaters, and divers can be found (Claessens *et al.*, 1999; Leone *et al.*, 2002). The focus of body composition in sport has largely evolved from the basic analysis and division of the body into two compartments termed fat free mass (FFM) and fat mass (FM). These two fundamental compartments are the focus of the majority of applied compositional analysis.

2.2.1 Fat Free Mass (FFM)

Conventionally, Fat Free Mass (FFM) is the terminology used to describe all non-fat components and reflects hydration status, lean tissue and bone. The accrual of FFM is thought to be advantageous to nearly all sports but especially in power sports such as volleyball, football, and many athletics events where propelling the bodyweight of oneself or another is integral to the competition. Two main factors dictating expression of neuromuscular power that are affected by increases in lean mass and hence performance are: maximal isometric force, inherently linked to muscle cross sectional area; and the maximal shortening velocity of the muscle fibre, which is contingent upon a series of physiological and anatomical mechanisms such as neural recruitment and fibre pennation (Bamman *et al.*, 2000; Jones *et al.*, 2008). It has been further stipulated that maximal power is related to total muscle volume (O'Brien *et al.*, 2009), while the geometric scaling paradigm indicates that strength is directly linked to muscle cross sectional area (van den Tillaar & Ettema, 2004). When assessing the effects of body size on the throwing performance of high level male and female handball players these authors found that body size

had a strong positive influence on throwing performance, and isometric strength. Throwing velocity showed effects of gender, and the authors concluded that men throw with greater velocity than females because of greater force production resulting from greater amounts of FFM. When normalizing throwing velocity for FFM, the gender effect was strongly reduced (van den Tillaar & Ettema, 2004). Investigations into lower body lean mass associated with performance have also been found in growing children (Temfemo *et al.*, 2009), and university students (Perez-Gomez *et al.*, 2008). When assessing lower limb lean mass via DXA, a linear association between absolute lean mass of lower extremities and peak power during Wingate tests was found in university men and women (Perez-Gomez *et al.*, 2008).

Despite FFM being an important co-contributor to performance, fat mass (FM) may be a better index for tracking compositional changes as the accrual of muscle protein in females may be a slow process relative to their male counterparts. Anabolic sex hormones are cited as a primary difference between genders affecting muscle hypertrophy (Tipton, 2001; Elmlinger *et al.*, 2005). In an investigation of serum concentrations of testosterone from birth to old age it was found that median testosterone levels of healthy adult females (20-99) was only 5% of age matched healthy male counterparts. Stratifying the dataset into smaller age groupings reveals a 17.9 times greater level of testosterone in males (16.1nmol) than females (0.9nmol) aged 21-30 years, indicating a reduced anabolic environment for protein accretion (Elmlinger *et al.*, 2005).

Additionally, hypertrophic capacity across muscle fibre types may differ between sexes despite similar training stimuli. Muscle morphological characteristics were analysed via tissue sample biopsy in a small group of male and female bodybuilders and were compared to non

bodybuilder controls (Bell & Jacobs, 1990). Although both bodybuilding groups displayed larger muscle cross sectional areas than their control counterparts, male body builders displayed greater muscle cross sectional areas than female body builders. Comparing the ratio of fast twitch muscle area to slow twitch muscle area revealed that males had a preferential hypertrophic expression of the fast twitch fibres while females showed no preferential hypertrophy. Reduced anabolic environments as well as a potentially reduced hypertrophic response in the muscle fibre itself make monitoring of FFM and muscle tissue accretion a potentially difficult task for tracking and interpreting compositional changes in the competitive female athletes over short spans of time.

2.2.2 % Body Fat (% BF) & Fat Mass (FM)

In endurance sports, the majority of body composition assessment tends to focus on the fat mass (FM) or the relative fatness (% BF) as it relates to performance. Running performance is influenced by a milieu of factors but is often conceptually thought to be influenced by three primary factors. The maximal aerobic power of an athlete (often represented by their VO_2 max), the running threshold;(represented by lactate steady state or anaerobic threshold), and running economy (the VO_2 consumed for a given workload) are all important factors related to endurance sport performance (Bassett & Howley, 2000). When running economy and relative threshold values are held constant, the addition of non-metabolic mass theoretically causes a detriment to performance by reducing the relative VO_2 and coincidentally the performance expression of the aerobic system, as a significant positive correlation between energy cost of running and % BF was found in non trained females running on treadmills (Bunc, 2000). However, the remaining variables, lactate threshold, as well as running economy, along with a

variety of other variables are not always constant in dynamic systems such as training, nor do the findings in non-trained individuals necessarily hold true for trained athletes. O'Conner et al (2007) performed calculations to account for a 2 kg fat mass increase in a 70kg middle distance runner with a relative VO_2 max of 75ml/kg/min. It was estimated that the increase of approximately 2.8% BF would result in an increased metabolic cost of running by 2.9% when viewed in isolation. Further modelling that incorporated the effects of increased air resistance due to larger body size, but increased kinetic energy return due to the increased elastic recoil of the lower limb musculature resulting from a heavier body mass still resulted in a predicted net decrease in performance represented as approximately 1.8% increase in time over 1500m (O'Connor *et al.*, 2007).

Aside from contributing to non-metabolic weight that must be supported by the aerobic system, extra subcutaneous fat impairs the ability of the body to regulate heat when exercising in environments with thermal strain or when competing under circumstances of non compensable heat stress (Bar-Or *et al.*, 1969; Cheung *et al.*, 2000). In these instances composition has implications not only for performance, but for health as well. Utilizing anthropometric methods to monitor composition when competing in heat stress environments may be beneficial as the practice of simply taking weight is insufficient because the composition of weight lost is unknown and this may be a co-factor for ability to perform in heat stress environments (Selkirk & McLellan, 2001). However, it is acknowledged that a variety factors will contribute to one's ability to perform in these environments (Anderson, 1999).

2.2.3 Composition and Health

As mentioned above the implications for body composition extend beyond that of performance alone. The uses of compositional testing by skinfolds and other methods have been employed to understand, and protect the health of athletes. Attempts to optimize body composition requires objective feedback to assess composition of weight loss or weight gain, as well as create strategies for these changes in a healthy manner, as many athletes could benefit from safe and structured guidance for attaining their compositional goals (AAP, 1996; Rankin, 2002; Hagmar *et al.*, 2008).

In American high school and college wrestling, skinfolds are converted to % BF to identify reasonable body compositional changes throughout a competitive season, ensuring wrestling athletes are not pursuing unhealthy or drastic forms of "weight cutting" in order to compete in a particular "weight class" (Oppliger *et al.*, 1995). In female athletes, drastic reductions in relative fatness (expressed as % BF), is implicated as a co-factor among other physiological and psychological variables that contribute to menstrual disorders (Holschen, 2004; Goodman & Warren, 2005). Perhaps one of the most profound negative effects of menstrual disorders, particularly amenorrhoea, on health, is its deleterious effects on skeletal structure secondary to hypoestrogenism (Burrows & Bird, 2000). In the young female athlete, identification and amelioration of the factors contributing to amenorrhoea is critical as inevitable decay of bone mineral density occurs beyond the 3rd decade of life (Burrows & Bird, 2000).

In terms of acute injury, the quantity of subcutaneous adiposity may dictate how therapy is applied to produce the most efficacious recovery. Otte et al. (2002) found significant

differences in cooling time required to reduce relative muscular temperature in the quadriceps muscle by 7 degrees Celsius in subjects with various skinfold thicknesses. The implications of their study led the group to suggest that perhaps individual therapeutic cooling durations be formulated dependent on adiposity levels of the treated individuals, although they conceded that at the time of publication no ideal cooling threshold for therapeutic purposes were agreed upon (Otte *et al.*, 2002).

2.2.4 Composition and Assessing Training Effects

Sport support teams require information about changes in their athletes' body composition to assess the efficacy of training programs. Two types of information allow for the evaluation of appropriate training stimuli with any given population. Barrages of generalized or sport specific tests are conducted periodically to make objective statements about performance, so that decisions about training or performance strategies can be made. Group statistics provide assessment related to the global success of a program allowing modifications to the training. However, the ability to identify real change within an individual is of greater importance, as this allows individualization of the training stimuli often necessary to elicit very small responses at the level of elite athletes (Hopkins *et al.*, 1999).

2.3 Anthropometrics

2.3.1 Anthropometric Preamble

Anthropometrics is a very old science. Various institutions and organizations have tried to standardize the use of anthropometric measurement and allow the growth of the discipline, but minute differences owing to research backgrounds have led to differing streams of measurement

practices. The two most prominent schools of thought include the interpretation of anthropometric data in its raw form, such as summed skinfolds, or by conversion to % BF through some form of regression equation. The latter is largely embraced by the professional certification system in the United States (American College of Sports Medicine (ACSM)), where the most prominent regression equations were created (Jackson *et al.*, 1980; Jackson & Pollock, 1982; Jackson, 1984). In contrast, European and Australian certification programs propose a separate stream of accreditation where raw anthropometric data is interpreted with additional parameters such as mass to infer compositional changes (Slater *et al.*, 2006). Embracement and perpetuation of this standardized style is championed by the International Society for the Advancement of Kinanthropometry (ISAK).

In Canada the principle professional body responsible for "physical activity, health and fitness research, and personal training" certification is the Canadian Society for Exercise Physiology (CSEP). Briefly, the accreditation scheme has two levels of practitioner accreditation. The first level is the Certified Personal Trainer (CPT) while the second more advanced level is the Certified Exercise Physiologist (CEP) designation. Currently the CPT conducts an anthropometric performa that consists of skinfolds taken at five sites (Subscapular, Biceps, Triceps, Iliac Crest and Calf). The sum of these five skinfolds (sum5 CSEP) are used for interpretation as part of a Healthy Body Composition Module (CSEP – Canadian Physical Activity, Fitness and Lifestyle Appraisal Protocol (CPAFLA)). As the practitioner advances their accreditation to a CEP designation, a series of regression equations are recommended, which include the ACSM`s preferred Jackson and Pollock equation (Jackson *et al.*, 1980). In this instance the practitioner must learn new skinfold sites to employ body composition

interpretation. Therefore, the Canadian accreditation scheme uniquely requires understanding and interpretation of both raw skinfold data as well as use of regression equations to calculate % BF. However, the current certification schemes in the field of anthropometry are flawed in isolation. When using % BF regression equations rarely are there guidelines for interpretation of accuracy of % BF for an individual subject, nor are there directions for assessing change in % BF in individual subjects. On the other hand, ISAK gives meaningful direction on how to assess the magnitude of change using 95% confidence interval about the measures, but this method still does not allow for prescription as there is no conversion to % BF, therefore, no conversions to fat mass and caloric expenditure needed to attain or set compositional goals. Additionally, when defining if true change has occurred, strictly using 95% confidence intervals to categorize change as having occurred or not is too absolute. It does not accurately reflect the dynamic flux or progression of human body composition change, nor does this type of information always provide clarity for the practitioner in assessing small magnitudes of change, suggesting a variety of confidence intervals should be explored.

2.3.2 Equipment

Measurement reliability is a product of both human and equipment error. Identifying the differences and limitations of various testing equipment is essential for the interpretation of results. The primary tools for anthropometric assessment are a basic measuring tape, bone callipers, and skinfold callipers. Circumferences are related to fat free mass, whereas skinfolds are measures of subcutaneous adipose which is reflective of total body fatness (Guo, 1996). The calliper used to measure the subcutaneous adipose tissue will introduce the majority of mechanical difference between results and therefore the model of callipers used needs to be

stated and differences between calliper brands understood to interpret anthropometric based body composition studies.

Skinfold callipers are the primary tool for assessing body composition in the field. A variety of skinfold callipers exist. The most popular calliper used is said to be the American Lange calliper (Wang, 2000). However, the Harpenden calliper has gained popularity due to increased precision on the dial face (0.2mm opposed to the Lange's 1mm) and higher reliability (Cyrino, 2003). Schmidt and colleagues investigated static and dynamic differences among five different types of skinfold callipers. Callipers investigated included Harpenden (HRP), Lange (LNG), Slim Guide (SG), Skyndex (SD), and Lafayette (LF). Results indicated that HRP, SG, and SD callipers yielded the smallest measured values when measuring foam rubber thickness, compared to LNG and LF callipers which displayed the largest measured values (Schmidt & Carter, 1990). The lower values of the HPR, SG, and SD would indicate that for any given skinfold measurement, these callipers will under predict skinfold size compared to the LNG and LF callipers. Similar differences have been reported between the Lange calliper and a Harpenden analog, the Cescorf calliper, used in Brazil, as skinfolds are consistently measured smaller when compared to LNG callipers (Cyrino, 2003).

Despite tests indicating similar spring pressures among the callipers, the combination of small jaw surface areas that pivot and low forces of the LNG and LF design, result in different outcomes than designs utilizing fixed, large surface area jaws with large spring forces, characteristic of the HRP, SG, and SD callipers. Comparing the LNG to the HRP callipers, the LNG was found to yield measurement values that are11to 17% (Gruber *et al.*, 1990), and 40-

110% greater than the HRP (Schmidt & Carter, 1990). The authors suggested that measurements between LNG and LF callipers should not be used interchangeably with HRP, SG and SD callipers. The research supports the notion that regression equations are not only skinfold site specific but calliper specific.

2.3.3 Skinfold Reliability

Regression equations attempting to predict % BF with high levels of validity, require individual skinfold data used to create the regression equation to be measured with very high reliability. Additionally, when the goal is to detect change in an individual over a series of measurements it is equally important, if not more important that reliability of the measurement be excellent. The common unit of reliability in anthropometrics is the Technical Error of Measurement (TEM). The TEM is usually represented in both absolute (mm) and relative (%TEM) forms. Currently, differing magnitudes of acceptable reliability have been reported in the literature in regards to skinfold measurement. Minimal standards for practitioner reliability have been suggested with values ranging from less than 2%TEM for circumferences and less than 10% TEM for skinfolds (Wang, 2000), but more stringent criteria are currently imposed by accreditation schemes where less than 1%TEM and less than 5% TEM respectively, are required (ISAK).

Identification of the exact measurement site is the first necessity for anthropometric measurement. A recent investigation focused on the importance of site location for skinfolds and its relationship to error (Hume & Marfell-Jones, 2008). Two level 4 ISAK (highest level of certification possible) anthropometrists measured eight skinfolds on ten healthy men using

Harpenden callipers. Each site was landmarked according to ISAK protocol. In addition, a 1 cm grid of eight dots were landmarked and measured surrounding each central ISAK skinfold to evaluate the effects of landmark deviations on the measurements. Evaluation of robustness of the measure despite landmark deviations was indicated by calculating the absolute difference from the "true" central measurement, the p-value for significant difference designated by a T-Test, and the Cohen effect size, and standardized Cohen effect sizes. The subscapular was found to be the most robust skinfold because deviation from the "true" central landmark was less crucial for achieving low magnitudes of error. Appendage skinfolds including triceps, biceps, and calf showed significant and non-trivial error when measurements deviate from the central site, indicating that identification of these sites properly is important for high reliability. This was also true of the thigh but only when deviating in the medial and distal directions from the "true" landmark location. For central sites, only the supraspinale showed both a significant and non-trivial deviation indicating that for the majority of central sites, exact location precision is not as crucial as seen in the appendages. The authors concluded, that site identification was important for adequate reliability and that this importance is more pronounced in appendicular skinfolds than central skinfolds.

Aside from the actual measurement itself, the method used for assembling the data can affect the degree of reliability. Common practices of assembling data include taking two or three measurements in cyclic fashion across all skinfolds within a given anthropometric testing barrage. The value that gets placed into the regression equations is commonly the average of two measures, average or median of three measures, and sometimes, the average of the two

closest measures provided they are below a "difference threshold" that would indicate the need to take additional measures.

An investigation into strategies for selecting an anthropometric criterion for assembling data was conducted. The success of data assembly methods was based on its resilience to imposed error investigated on 67 females aged 16-60 years (Ward & Anderson, 1998). Using Harpenden callipers and taking three measurements, Technical Error of Measurement (TEM) values were calculated using five different data assembly strategies reported in the literature. Strategies investigated included taking the first single measurement, the average of the 1st two measurements, the average of all three measurements, the median of three measurements, and the average of the two closest measurements. Introducing two different magnitudes of error (5 and 15% of the relative TEM) the subsequent "pseudo" datasets were analyzed. Of the skinfolds assessed (Tri, SSc, SSp, Abb, Thi, C) Tri, and Thi had the lowest relative TEM (3.6%) while the highest was found in SSp (5.8%) and Abb (4.8%). The median of three strategy of data assembly was the most accurate displaying the lowest TEM values in both the small and large imposed error datasets with its superiority becoming more evident when the introduced error was larger. The authors suggest that the median of three measures be used to derive the most reliable dataset as it does not require calculations that could introduce error, is the best indicator of central tendency, and is a common and easy function to use with any spreadsheet program that can be quickly employed. In instances where two values are to be used , a "difference threshold" for acceptable scores should be used, but this is not the recommended strategy (Ward & Anderson, 1998).

2.4 Predictive Equations

2.4.1 Hydrodensitometry (HD) Criteria

Some of the most widely utilized skinfold regression equations were developed by Jackson & Pollock (Jackson *et al.*, 1980). The adoption of these equations by the American College of Sports Medicine (ACSM, 2000) and the Canadian Society of Exercise Physiology (CSEP) has led to their favourable use among practitioners. Briefly, Jackson and Pollock assessed a large group (n=331, validation sample=249, cross-validation=82) of heterogeneous females aged 18-55 over the span of six years to compose regression equations that would allow prediction of body density (BD) when compared to the criterion measure of Hydrodensitometry (HD). The BD value could then be converted to % BF using a density formula (Siri, 1956). Previous research conducted by the group (Jackson & Pollock, 1976) led to the inclusion of the quadratic sum of skinfolds (sum3), gluteal circumference and age as independent variables used to predict the dependent variable HD BD (Jackson *et al.*, 1980; Jackson, 1984). The series of regression equations created yielded similar correlation coefficients, ranging from R= 0.838 to 0.867 with standard errors of estimates (SEE) of 3.6 to 4.0 % BF. In two subsequent papers, regarding the design and analysis of body composition techniques, Jackson and Pollock speculate that the regression equation has perhaps reached the "limits of accuracy possible with body composition prediction equations", because the SEE of HD methods is 2.5% BF and it is illogical for the accuracy of the predictor to exceed that of its own criterion measure (Jackson & Pollock, 1982; Jackson, 1984).

Another equation that is highly referenced in body composition literature is the seminal study of Durnin and Womersley (1974). The pair investigated the relationship between skinfold

thickness and body density in a large general population of males (n=209) and females (n=272), aged 16 to 72 years of age. The group compiled a generalized skinfold regression equation utilizing skinfolds from the upper body that included the subscapular, triceps, biceps and iliac crest (termed supra-iliac in the original paper) as predictive variables of body density which were further converted to % BF. Recognizing the skewed nature of skinfold frequency distribution, the group were among the first to logarithmically transform the skinfold data in attempts to address this phenomena. Plots of their data revealed a non-linear relationship between skinfolds and body density as relatively large increments in skinfold thickness are associated with only small changes in body density in subjects who were more obese. The authors suggest either logarithmic or quadratic transformation can be utilized to normalize the data. The final equation suggested by the authors included a logarithmic transformation of the sum of four upper body skinfolds.

In addition to the popular equations listed above, a review of 50 of the most frequently used regression equations for predicting % BF or FFM show SEE's ranging from ±3% to 11% (Wang, 2000). As most regression equations originating prior to the 1990's were primarily created using HD criterion, the *Type 1* mathematical functions employed attempt to relate skinfolds to BD which is further converted into % BF using either the Siri or Brozek equation (Siri, 1956; Brozek *et al.*, 1963). As aforementioned (in 2.1), inherent to all *type 1* mathematical relationship are the limitations of sample population and equipment specificity. This notion was confirmed by subsequent publications by the group emphasizing additional limitations of many of the original skinfold regression equations. These limitation included the use of stepwise regression equations, sample sizes that were too small to investigate the number of variables

proposed, and the use of linear regression models for non linear relationships (Jackson & Pollock, 1982).

2.4.2 Evaluation of HD Regression with DXA

As part of a national study, entitled "Cross-validation of generalised body composition equations with a diverse sample of young men and women: the Training Intervention and Genetics of Exercise Response (TIGER) Study" Jackson et al. (2008) revisited the validity of their seminal study on body composition prediction. Using a sample of mixed ethnicity, 706 females aged 17-35 years were tested using the Jackson and Pollock sum3 quadratic equation with age. When compared to a Hologic DXA criterion the authors found the old HD regression yielded a validity coefficient of R=0.84, with a TE of 3.9% BF, but had a CE of -3.0% for women (Jackson *et al.*, 2008). The authors concluded that due to the systematic underestimation the generalized equations created in the 80`s need to be updated as the body composition characteristics of the population evolve. The group created a new equation that accounted for race in addition to their original variables, which showed better fit statistics of R=0.88 and TE= 3.4 % BF. The long standing use of the original Jackson and Pollock HD equations make the revised DXA equation a potential candidate for modern compositional assessment of females as the original Jackson and Pollock equations have been consistently shown to underestimate % BF by as much as 6.6 % BF to 7.6% BF in other research (Peterson *et al.*, 2003; Duz 2009).

2.4.3 Dual Energy X-Ray Absorptiometry (DXA) Criterion

The establishment of new female specific regression equations using DXA as the criterion method has been carried out by a variety of research groups (Ball *et al.*, 2004; Warner

et al., 2004; Kagawa *et al.*, 2007; Jackson *et al.*, 2008; Kohli *et al.*, 2009). Ball et al. (2004) evaluated the accuracy of the three most popular and professionally recommended regressions equations from Jackson et al. (1980) on a population of women (n=150) 18-55yrs, that were characteristically very similar to the population used in the original Jackson et al research. They found the mean difference between DXA % BF and predicted SKF % BF were 3.2 to 5.6% indicating an under prediction of the SKF % BF regression equations. Although each of the SKF equations correlated reasonably highly (r = >0.8) with DXA, a systematic bias for underestimating % BF in each of the skinfold equations was the impetus for the group to create and validate a new non-biased regression equation. The new regression equation was cross-validated in a separate group of 25 women and displayed a correlation coefficient of $R^2 = 0.93$ with a mean difference of only 0.4% BF. The PRESS SEE was 2.1% BF which is lower than the range of 3.4 to 4.6 % BF seen in the literature. Although very good validity was obtained, limitations include the fact that validation was done on a Lunar DXA model whereas the equation was created on a Hologic DXA machine. Lange SKF callipers were utilized as well making inferences about similar validity when using Harpenden callipers questionable. Finally, it is unknown if the accuracy can be extrapolated to athletic populations, as it is not recommended since athletes' FFM differs from the generalized populations (Prior *et al.*, 2001).

Warner et.al (2004) developed an athlete specific regression equation for predicting FFM from skinfolds by testing female NCAA Division 1 athletes (n=101) from a variety of sports. Using a Hologic DXA as the criterion method, 4 skinfold sites (suprailiac, thigh, triceps, and abdominal) were assessed using Lange callipers. Multiple regression using the four skinfolds, height and mass as candidate variables was conducted yielding a regression equation that

contained only mass and the thigh and abdomen skinfold measurements. The equation had a validity coefficient of R =0.98 with a very low SEE of 1.1kg FFM. If transformed, the resulting affect on % BF error is 1.9%.

Although superficially more accurate than the equation of Ball et al. (2004), the study has notable methodological limitations. The first methodological error comes from the inclusion of four independent skinfolds added into the regression equation instead of utilizing the sum of skinfolds suggested by Jackson and Pollock's 1976 factor analysis paper, because it does not address the issues of multi-collinearity when creating a regression equation with skinfold sites that have shared variance (Jackson & Pollock, 1976). The second error is the professed ease of measurement with the sites selected for their regression. The authors state that their equations is advantageous because of the minimal invasiveness and ease of measurement with the abdomen and thigh skinfold measurement when it has been reported that these sites, particularly the thigh are among the most difficult to take despite having high correlations with fatness (Eston *et al.*, 2005). Also, they have been shown to have a higher degree of variability in measurement using callipers (Lohman *et al.*, 1984). In addition, the use of circumferences is more commonly associated with FFM relationships (Guo, 1996). To calculate % BF an additional step is needed where FFM is first computed, it must then be subtracted from total mass to yield fat mass, which is finally divided by total mass to calculate % BF.

A lack of understanding of the theoretical background behind regression equation creation for body composition, and the relationships to the measurements being employed, discredits the results from this research paper. Finally, some of the citations in the article appear

to be overly ambitious, or misprinted as it is stated in the discussion section that the "largest mean difference between methods was 0.4 kg FFM and 0.4 % BF". Upon analyzing the data it is evident that at least two and possibly three groups of athletes (softball, soccer, and cross country) in the study exceed the stated mean difference between methods for predicting % BF. Furthermore, this regression equation has not been cross validated.

Other DXA based regression equations that may have predictive abilities but are unlikely to be as generalizable due to population or equipment variations include the equation by Duz et al. (2009) on young Turkish females aged 18-26, Kohli et al. (2009) who used the CSEP skinfolds on a South Asian population, and Kagawa et al. (2007), who were one of the first groups to use ISAK skinfold performa to develop a predictive regression equation with a DXA criterion on Japanese females.

2.4.4 Multicompartment (MC) Criterion

Recently, the creation of new regression equations derived from multi-compartment criterion models have surfaced in the literature (Peterson *et al.*, 2003; Evans *et al.*, 2005; Moon *et al.*, 2009b). Peterson et al. (2003), investigated the association of skinfolds, circumferences, summed skinfolds and quadratically transformed summed skinfolds, with % BF derived from a four compartment criteria (Peterson *et al.*, 2003). They found that the sum3 (Tri, SSp, Thi), quadratic sum3, along with gluteal circumference and weight used in the original Jackson and Pollock equations (Jackson *et al.*, 1980) did not significantly predict the 4C % BF in the large sample of 230 women aged 18-56 years. The group proposed a new regression equation that included a sum of four skinfolds that added a subscapular skinfold to the aforementioned trio.

The equation still displayed large errors for individual % BF prediction as the limits of agreement were ± 9.5%. In addition, the Bland Altman graph showed a marked slop of the correlation line perhaps indicating the presence of heteroscedasticity, but this was not mentioned by the authors. When compared to older existing HD equations, the new equation did not yield greater predictive accuracy to the criterion as total error values were 5.0% BF, in contrast to 4.9% BF for the HD equations of Jackson et al. (1980) and Durnin and Womeresely (1974).

Similar to the investigation of Peterson et al. (2003), Evans et al. (2005) used a 4C model to investigate prediction equations in a group of collegiate athletes. Harpenden callipers were used to measure several skinfold sites. A new regression equation was created that included the use of sum 3 skinfolds (abdomen, triceps and thigh) in addition to accounting for race and gender. The total error of the equation was 3.76% BF which was better than the existing equations tested by the group and was accompanied by tighter limits of agreement of ± 6.9% BF than those displayed in other MC regression creation studies (Peterson *et al.*, 2003; Moon *et al.*, 2009b). The equation provided appears to be a valid method of predicting % BF in elite collegiate athletes of various sports.

Moon and colleagues validated both the original Jackson and Pollock HD equations along with updated MC equations proposed by Evans et al. (2005), in a population of NCAA Division 1 female athletes (n=29). Using similar equipment to the Jackson and Pollock research (Lange calliper), but slightly different equipment than the Evans equation (Harpenden Calliper) as well as horizontal as opposed to vertical skinfolds at the axilla and abdomen, the group found the HD equations to have equal validity (TE of 2.65% BF and 2.71% BF respectively) as the

new MC equation when assessed to a 4C criterion (Moon *et al.*, 2009b). The group proposed that due to the difference of callipers used, the Evans equation may still prove to be a more valid equation once the considerable amount of the constant error (-1.20% BF) was removed with a correction factor equation that was presented.

2.5 Dual Energy X-Ray Absorptiometry (DXA)

2.5.1 General

A recent review of DXA systems used for the analysis of total body composition traces DXA's origins to the evolution of bone scanning methods (Lohman, 2005). The progression from single to dual photon absorptiometry (DPA), has recently been transformed from single to dual energy X-ray absorptiometry (DXA). Similarly, the refinement and improvement within individual DXA technology necessitates reassessment of the technology's applications and limitations. Currently, three major DXA distributing companies exist. Norland, Hologic, and GE all have unique models of DXA's, each with proprietary hardware and software, making interpretation of research results across models difficult (Lohman, 2005). The authors noted that at the time of publication, minimal research had been conducted on newer fan beam models such as the GE Lunar Prodigy, and that the novelty of these updated machines needed to be explored.

Independent of DXA Brand or model, many of the early generation DXA's operated with the same assumptions and thus, shared similar limitations. At a summit on body composition assessment, DXA validation studies were reviewed and limitations of the technology were identified and published, with recommendations to minimize their impact on result reporting

(Lohman *et al.*, 2000). Briefly, among the validation studies of DXA and MC criterion models, the SEE for % fat was generally 3 ± 1% BF, but this may be an overestimation as some of the error may have been introduced into the MC through the different methods used to assess total body water as one of the compartments. Furthermore, concerns of hydration affecting DXA results seemed negligible as normal variations in hydration (1.0-3.0 % BW) did not affect DXA compositional assessments. Reasonable euhydration should be employed as changes greater than 5% BW were shown to alter DXA estimates by 1 to 2.5% BF (Prior *et al.*, 1997). Lastly, studies on the detection of rapid weight loss through DXA revealed that in some cases DXA sum of tissue weight differed significantly from scale weight. It was hypothesized that this discrepancy could be due to changes in composition of fat free mass and that researchers and practitioners should always check to ensure that scale weight and DXA weight are within 1kg of each other to ensure technical errors are not present (Lohman *et al.*, 2000).

2.5.2 GE Lunar Prodigy DXA

Among the various distributors, the GE Lunar prodigy is one of the newest most advanced DXA units available. The Lunar Prodigy has been validated against a 2 & 3C criterion in healthy men and women aged 18-45 (Norcross & Van Loan, 2004) and offers advantages in terms of its ease of use, quickness, and validity when compared to multi-compartment methods (Prior *et al.*, 1997). Possessing a narrow fan beam (4.5 degrees) which scans trans-axially in a rectilinear fashion, significant changes have been made to newer fan beam technologies that are likely to have remedied some of the imprecision with older models. Advancements of this DXA model over its predecessors have increased its measurement accuracy while decreasing measurement time and radiation exposure. In particular, older fan beam units used A/D detectors

which were replaced by vastly superior Cadmium Zinc Telluride (CZT) digital detectors allowing efficient photon count, at a lower radiation levels. The resulting efficiency has led to faster scanning times (approximately 7 minutes per subject). Conventional fan beans have wide angles (35 degrees) that originate below the table and extends upward causing distortion and penumbra effect when the body breadth is large. In older models, assumptions were made about the distance (10 cm) between measured bone and table top which is not always consistent resulting in measurement error (Mazess *et al.*, 2000). This was remedied by actual measurement of the distance of the bone from the table top in the new GE Lunar models.

Initial GE research compared the Prodigy fan beam model to their flagship model, the DPX pencil beam in phantoms (*in vitro*) as well as *in vivo* (Mazess *et al.*, 2000). The research found that there was minimal effect of simulated tissue thickness as 30 cm of acrylic placed above the bone, decreased BMD by only 1%. The results indicated that, theoretically, BMC can be affected by extremes in tissue mass above the bone, but BMD would not. BMD can be determined in a majority of patients with an uncertainty of less than 1%, while BMC is less robust. The precision of the total body composition parameters on the phantom of the Prodigy fan beam was 0.004 g/cm^3 (BMD), and 100g for fat and lean mass, while in vivo values were slightly greater at 0.01 g/cm^3, and 300g for fat and lean mass.

Additional research from GE scientists further investigated the Lunar Prodigy fan beam precision, and agreement with DPX-IQ pencil beam model (Nord *et al.*, 2000). The precision of 46 subjects scanned in triplicate on the Prodigy yielded CV's of 0.8, 1.0, 1.3,0.9, and 0.3 for bone mineral density (BMD), bone mineral content (BMC), fat mass, lean mass and total mass

respectively. Twenty subjects scanned on both DPX and Prodigy scanning beds revealed high correlation coefficients for all measurements ranging from 0.969-0.999 with SEE of .019g/cm^3, 93g , 959g, 911g, 298g, and 1.41% BF for BMD, BMC, fat mass, lean mass, total mass, and tissue percent fat respectively (Nord *et al.*, 2000). Although the two machines correlate well, the inter-instrument reliability differs and thus caution should be taken when tracking and interpreting longitudinal changes within individuals across machines (Hull *et al.*, 2008).

External investigators found similar results when comparing the reliability of one Lunar DPX and two different Lunar Prodigy DXA systems. Soft tissue phantoms representing obese (50% BF) and athletic (10% BF) populations were scanned at three scanning speeds (Slow, Medium and Fast) on each machine for a total of 60 scans per machine (180 grand total). There was an effect of scan speed on measurement of fat in the athletic phantom at the fast setting but the effects were non-significant at medium and slow speeds on all DXA systems. Similar results were found for the obese phantom (Guo *et al.*, 2004) and for the GE pencil beam DPX units (Vozarova *et al.*, 2001). When adjusting for scan speed, the DPX resulted in significantly higher fat percentage for the obese phantom, and lower fat percentage for the athletic phantom than the Lunar Models. The authors conclude that the differences between models, within models, and scanning methods are likely to be a source of error in published equations limiting universal application (Guo *et al.*, 2004).

Despite different technologies, all DXA's are still subject to basic assumptions. Returning to the "Body Composition Classification System" (Figure 2.3) of Wang et al 2000", DXA technology relies on a *type 2* assumption relating the attenuation of its X-ray signal from

source to detector to the composition of the tissue. Any alterations to these characteristics such as the drastic changes in hydration listed above will violate these assumptions and result in errors in the DXA report. Additionally measurement of soft tissue over bone is extrapolated from the area of tissue adjacent to the bone. Some authors suggest that in larger individuals, the more total mass *estimated* over the bone, the larger potential errors in these values. This was deduced by comparing error results of appendages in results from NCAA athletes to smaller but similar high school aged athletes which would likely have less mass than the NCAA athletes (Silva *et al.*, 2006; Moon *et al.*, 2009a). The ever-changing technology and improvements associated with DXA require continuous research and development for the validity and reliability of the machines particularly when testing specialized populations such as elite athletes.

2.5.3 DXA Least Significant Change (LSC)

In clinical densitometry it is common practice to calculate the Least Significant Change (LSC) to interpret an individual's BMD results over time (ISCD, 2007) . A recent investigation suggested that the Ideal Least Significant Change (ILSC) can be utilized under "ideal" circumstances, which denote that baseline and follow up measurements are conducted on the same system with no systematic changes in calibration or precision of the machine (Shepherd & Lu, 2007). The suggested equation is:

Eqn. 2.2 : ILSC=2.77 X Precision Error

Where precision error is the standard error of measurement (SEM) value from a reliability study. The ILSC represents true difference in the two measures with 95% confidence.

Non-ideal circumstances often involve the interpretation of individual results originating from different DXA units, models of brands, each with unique precision values. In this instance the authors generated and validated a complex "Generalized Least Significant Change" (GLSC) equation that incorporated not only individual precision values of the machines but cross calibration data such as correlation coefficient, sample variances of the cross calibration population, and regression coefficient of the slope (Shepherd & Lu, 2007).

2.6 Novelty of the Study

The reliability of both anthropometric and DXA methods need assessment to articulate recommendations for detecting compositional changes within individual subjects, based on magnitudes of reliability. Additionally, methodological differences owing to equipment variation, sample selection and statistical methods for regression equation construction will all influence the ability of existing regression equations to accurately predict the body composition of elite female athletes and necessitate the cross validation of these methods in populations with characteristics that may differ from the original experiment (Schmidt & Carter, 1990). Commonly used equations recommended by national certification programs such as ACSM and CSEP, need to be evaluated along with a series of equations derived from newer models including Multicompartment (MC) and Dual Energy X-ray Absorptiometry (DXA), for their ability to predict % Body Fat (% BF) in Canadian female athletes using a GE Lunar Prodigy DXA as the criterion measure (% BF DXA). Currently no study has validated new skinfold regression equations created with MC and DXA criterions for predicting % BF in female athletes.

CHAPTER 3 - METHODOLOGY

3.1. Overall Study Design

The study was divided into three distinct assessments. The anthropometric reliability assessment was conducted to establish the reliability of the technician performing the anthropometric measurements. A DXA reliability assessment defined the reliability of the DXA machine and software, while a skinfold validity assessment tested the validity of existing skinfold equations and their ability to predict % BF in a group of female athletes compared to the DXA criterion.

3.2. Subject Recruitment

Athletes recruited into the study were allowed to participate in a minimum of one section and a maximum of all three sections of the study. Twenty subjects (n=20) were recruited for anthropometric reliability assessments, thirty two (n=32) subjects were recruited for the DXA reliability assessments, and ninety-five (n=95) subjects were recruited for the skinfold validity assessments. Potential candidates wishing to participate in the study needed to be apparently healthy females, aged 18-30, currently competing in sport at a varsity or national level. Varsity athletes were defined as being on rosters of University or College teams, while National level athletes were defined as athletes who were on rosters or competing for the Canadian national team, indicating that they were among the best in the country. Individuals who met the inclusion criteria, were excluded if they were pregnant.

All athletes were informed of the study intentions including potential risks, time required for testing, and testing methods. All athletes gave written consent prior to participation

in the study (Appendix A). Ethical approval was obtained from the University of Manitoba Education and Nursing Research Ethics Board (REB) prior to commencement of the study.

3.2.1 Sample Size Rationales

The study focused on female athletes due to the aforementioned existing gaps in the literature. Twenty (n=20) subjects were recruited for establishing anthropometric reliability as suggested by professional accreditation guidelines (ISAK; Anthrpometrica, 1996). Similarly, a minimum of thirty (n=30) subjects has been suggested for assessing DXA reliability (ISCD, 2007). An *a priori* decision was made to recruit 105 female athletes for participation in the skinfold validity component of the study. This number was based on suggestions for the creation of regression equations based on the ratio of subjects to candidate variables being investigated for inclusion in the regression equation (Heyward, 1996). A minimum ratio of 10 to 20 subjects per candidate variable has been suggested by Heyward (1996), but a ratio as low as 4 subjects per candidate variable of investigation has been utilized and would theoretically allow the investigation of 10, 5 or 26 variables respectively. However, difficulty in recruiting individuals satisfying all eligibility criteria as well as time restraints ultimately resulted in the achieved sample of 95 subjects. However, the reduced sample was not thought to be detrimental to study design as the majority of existing regression equations contain 5 or less variables, and in this regard our sample size is of sufficient size to satisfy the subject to variable ratio (Jackson & Pollock, 1982; Jackson, 1984; Guo, 1996).

3.2.2 Process of Recruitment

Subject recruitment began in March 2009 and concluded in December 2009. A subject recruitment flowchart is contained in the appendices (Appendix B). Recruitment emails were sent to all coaches of women's sports at the University of Manitoba and University of Winnipeg. In addition, emails were sent out to athletes on the Canadian Sport Centre Manitoba list serve and Sport Manitoba list serve. This email outlined the opportunity for athletes to participate in the study (Appendix C). One week from the date the email was sent, the primary investigator conducted recruitment presentations/information sessions prior to or after team practices, with no coaches present (Appendix D). After the script was read, the investigator answered questions. Prior to leaving, the investigator allowed the opportunity for immediate scheduling and handed out a "study outline and informed consent" document summarizing the subject's involvement in the study (Appendix A). At this time, if athletes were interested they were allowed to sign up for the study and were given instructions on when to pick up the registration package. If the athlete chose not to sign up for the study immediately, they were left with contact information should they choose to participate at a later time.

In addition, recruitment posters (Appendix E) were distributed on the University of Manitoba and the University of Winnipeg campuses. The posters contained tabs that could be ripped off containing the contact information of the primary investigator. Upon receiving a call, a phone script was read (Appendix F) allowing for questions from the potential subject, as well as providing opportunities to sign up and schedule assessments.

3.3 Setting

3.3.1 Anthropometric Reliability Assessments

Assessments were conducted in the Frank Kennedy High Performance Testing Lab (FKHPTL), RM128 of Frank Kennedy Centre, University of Manitoba. The lab contains a private room that was used for anthropometric assessment and private consult allowing separation from individuals undergoing assessment in the general lab space.

3.3.2 DXA Assessments

DXA assessments were conducted in the DXA room on the 1st floor of the Richardson Centre for Functional Foods and Neutraceuticals (RCFFN), 196 Innovation Drive in the Smart Park, University of Manitoba. The facility is access secured, upon entering the facility subjects waited in the main foyer for the principal investigator to grant them access. The DXA room is a private room that houses the DXA scanner. Subjects were assessed individually in this room.

3.3.3 Skinfold Validity Assessments

Due to the assessment consisting of both anthropometric and DXA methods, the RCFFN facility was used for these assessments in accordance with the information above.

3.4 Study Protocols

3.4.1 General Protocol

Subjects arrived at the designated testing location ten minutes prior to their scheduled assessment time between 7:00 am and 12:15 pm, having followed the pre assessment guidelines

(Appendix G). Informed consent sheets were collected by the primary investigator at this time along with the training history document (Appendix A & H).

3.4.2 Anthropometric Reliability Protocol

Twenty (n=20) subjects were recruited to participate in the examination of the Technical Error of Measurement (TEM) for anthropometric variables. These individuals were not necessarily the same individuals utilized in the DXA precision testing. All of the anthropometric measurements were taken by a single technician (D.H.) who was experienced in performing the anthropometric performa (6 years, > 1000 tests conducted). Following the recommendations of the International Society for the Advancement of Kinanthropometry (ISAK), TEM was established to indicate the Anthropometrist's level of reliability (Anthrpometrica, 1996). The guidelines require that:

- The same anthropometrist takes all measurements
- Subjects used to establish the TEM should be from the population which will be tested in the future, or at least represent a similar population
- Repeated measurements should be carried out on at least 20 subjects
- For ease of analysis, the same number of measurements should be made on each subject
- Data on all subjects should be acquired within a reasonable time frame (1 week) as to minimize biological variability

The subjects followed the same procedural instructions as indicted in the "General Study Protocol" above, arriving at the FKHPL. At this time only the anthropometric assessment was

carried out. The time required for the anthropometric assessment was approximately 20 minutes. Before the subjects left, they scheduled a re-assessment returning to the lab as early as the next day or within 7 days upon which the procedure was repeated.

3.4.3 DXA Reliability

Thirty two (n=32) subjects were recruited to participate in the assessment of DXA reliability. These individuals were not necessarily the same individuals as those in the anthropometric TEM assessment. All measurements were taken by a single investigator (D.H.) trained in conducting total body scans on the DXA system. Following the recommendations of the International Society for Clinical Densitometry (ISCD) DXA reliability was assessed (ISCD, 2007). The guidelines suggest:

- Each DXA facility should determine its precision error and calculate the Least Significant Change (LSC)
- The precision error supplied by the manufacturer should not be used
- If a DXA facility has more than one technologist, an average precision error combining data from all technologists should be used to establish precision error and LSC for the facility, provided the precision error for each technologist is within a pre-established range of acceptable performance
- Every technologist should perform an *in vivo* precision assessment using patient's representative of the clinic's patient population

- Each technologist should do one complete precision assessment after basic scanning skills have been learned (e.g., manufacturer training) and after having performed approximately 100 patient-scans
- Measure 15 patients 3 times, or 30 patients 2 times, repositioning the patient after each scan

The subjects followed the same procedural instructions as indicted in the "General Study Protocol" above but, in addition, were required to undergo repeat DXA scans after repositioning within the same session. As well, due to the low dose of X-ray exposure during the DXA scan, subjects who agreed to participate in this part of the study were instructed to pick up a urine sample container and pregnancy test the day before their scheduled assessment. The subjects were instructed to take the pregnancy test and collect the first urine sample of the day and bring the result stick along with their sample to their assessment. Upon arrival, the primary investigator received the sample and result stick. Subjects were asked to fill out the "Pre-Assessment Medical Background Questionnaire" (Appendix I). Visual confirmation by the investigator was made to ensure the subject was not pregnant prior to the scan. Specific gravity of urine samples were assessed within 24 hours to indicate hydration status of the subject. If the subject was menstruating they were not required to take a pregnancy test or provide a urine sample. Total testing time per subject was approximately 20 minutes. An additional 10 minutes was taken at the end of measurements to answer any questions the subject may have had in regards to their assessment.

3.3.4 Skinfold Validity Protocol

The subjects followed the same procedural instructions as indicted in the "General Study Protocol" as well as the "DXA protocol" but, for this assessment subjects began their assessment with a DXA scan followed by the anthropometric assessment. DXA and anthropometric assessments were completed consecutively. Total testing time per subject was approximately 30 minutes. An additional 10 minutes was taken at the end of measurements to answer any questions the subject may have had in regards to their assessment.

3.5 Specific Experimental Protocols

3.5.1 Standardized Anthropometric Assessment

Procedures

Before testing was initiated it was ensured that all data collection forms were filled out, informed consent was signed, and pre assessment protocols were followed. All anthropometric measurements were carried out according to the ISAK methods unless otherwise specified (Anthrpometrica, 1996).

• **Body Mass (weight)** Weight was taken using a Life Source scale recorded to the nearest 0.1kg. The subject was weighed in sports bra and sport shorts.

• **Free Standing Stature (height)** The measurement was taken using a portable SECA stadiometer. The maximum vertical distance from the floor to the vertex of the head was measured to the nearest mm. The vertex was defined as the highest point on the skull when the head was held in the Frankfort plane. In making the stature measurement, the technician

instructed the barefoot or socked subject to stand erect with heels together beneath the headboard of the stadiometer. The subject was instructed to “look straight ahead” and “take a deep breath”. The headpiece was firmly brought down crushing the hair and making firm contact with the vertex. This measurement is a small deviation from ISAK as ISAK protocols utilize stretch stature.

Skinfold Measurement

Instruments

• *Rosscraft Retractable Metallic Metric measuring tape*

• *Harpenden Skinfold Callipers* - A new Harpenden calliper (< 6months) was used to ensure calibration as all Harpendens are precision calibrated prior to leaving the factory. The instrument is designed to provide a constant pressure of 10.0 g/mm2 at the calliper face at all thicknesses. The dial is calibrated in 0.2 mm increments.

• *Rosscraft Campbell Calliper 10* – Utilized for the measurement of bone breadths

Landmarking

The body can assume a variety of postures, so anthropometric standardization is necessary. The descriptions of landmarking are always in reference to the anatomical position. The subject is oriented to a standing position with head and eyes directed forward, upper limbs hanging by the sides with the palms forward, thumbs pointing away from the sides with fingers pointing directly downward. Landmarking was made with an eyeliner pencil.

General Skinfold Technique

The calliper was used to obtain a skinfold thickness. This includes a double layer of skin and the underlying adipose tissue, but not the muscle. The skinfold was raised by the pinching, and slight rolling action of the left thumb and index finger. The fold is raised perpendicularly to the surface of the body at the measurement site. The long axis of the fold was parallel to the natural cleavage lines of the skin (Langer's lines) in the region of measurement. The grasp was large enough to get a complete double layer. The amount of skin elevated forms a fold with approximately parallel sides. The fold was grasped firmly and held throughout the measurement. The skinfold was raised at the designated site and the calliper applied so that the near edge of the pressure plate was one centimetre from the lateral side of the controlling thumb and index finger. Care was taken to assure the calliper was applied at right angles to the fold at all times. The reading on the dial was taken after permitting full spring pressure of the instrument by a complete release of the calliper trigger. The investigator allowed time for the full pressure of the calliper to take effect, but not so long that the adipose tissue became compressed out of the skinfold. Considerable practice is required to make this judgment for skinfolds of varying sizes and varying degrees of compressibility. The reading was made approximately **two seconds** after application, when the needle slowed. When measuring skinfolds with higher levels of subcutaneous fat, firm pressure of the thumb and index finger helped reduce excessive movement of the indicator. When skinfolds were difficult to raise, the calliper was forced to the muscle level and then slightly withdrawn when the fold was controlled by the grasp. Measurements were taken to the nearest 0.2mm in triplicate. All measurements were taken on the right side of the body with the subject standing in the anatomical position unless otherwise

indicated. All skinfold measurements were taken in a cyclic manner and recorded on the anthropometric performa sheet (Appendix J).

Specific Measurements

All measurements were taken with the calliper applied 1 cm distally from the left thumb and index finger on the raised fold at the sites listed below:

• **Triceps (Tri)** the marked mid-acromiale-radiale line on the posterior surface of the right arm.

• **Biceps (Bi)** the marked mid-acromiale-radiale line on the anterior surface of the right arm.

• **Subscapular (SSc)** a fold oblique to the inferior angle of the scapula in a direction running obliquely downwards in a lateral direction at an angle of about 45° from the horizontal along the natural fold (Langer line)

• **Mid-Axilla (Ax)** a vertical fold is taken at level of xiphoid process along the midaxillary line

• **Iliac Crest (IC)** 3 cm superior to the iliac crest at the mid-axillary line (i.e., above the crest on the mid-line of the body). The fold runs anterior and downwards and is usually progressively smaller as one moves in this direction away from the designated site.

• **Supraspinale (SSp)** the intersection of the border of the ilium (project a horizontal line from the iliac crest mark) and a line from the spinale to the anterior axillary border (armpit). The fold follows the natural fold lines running medially downwards at about a 45° angle from horizontal.

• **Abdominal 1 (Ab1)** taken **vertically 2 cm** lateral to the lateral border of the omphalion (belly button). This measurement is used in the Jackson et al. (1980) regression equations.

• **Abdominal 2 (Ab2)** taken **vertically 5 cm** lateral (approximately in the midline of the belly of rectus abdominis) to the omphalion (navel midpoint). This measurement is used in the ISAK anthropometric performa.

• **Mid Thigh (Thi)** the anterior of the right thigh, along the long axis of the femur, when the leg is flexed to 90° at the knee by placing the foot on a box. The mid-thigh position for this measure is the estimated half-distance between the inguinal crease and proximal border of the patella. In subjects where the fold was difficult to raise, the callipers can be pushed to the muscle level and slightly retracted with the subject assisting by supporting the underside of the leg. In particularly heavy-thighed subjects, the anthropometrist gave further support to the underside of the leg by using his/her own knee and thigh.

• **Medial Calf (C)** The calliper was applied 1 cm distally to the left thumb and index finger, raising a vertical fold on the relaxed medial right calf at the estimated level of the greatest circumference. This was easiest to obtain when the subject's leg was flexed to an angle of 90 degrees at the knee by placing the foot on a box.

Circumferences

General technique

The metal case was held in the right hand throughout all the girth measurements. The girths were measured with the tape at right angles to the long axis of a bone or body segment. The tape was passed around the part and held so the stub end and the scale calibrations are in juxtaposition, i.e., one reads to a scale mark and not across a tape space. The reading edge of the

tape is manipulated to the designated level. When measuring, the tape is pulled out of its housing and around the part by the left hand, which then transfers the stub end to the right hand. The tape is then controlled by the right hand which can pull it slightly to maintain it at the designated level. The left hand resumes control of the stub end and can make any further adjustments to the tape. The so-called "cross-handed" technique was therefore employed. A perimeter distance of the part was taken by contact, but not depressing, the fleshy contour. The development of this light touch requires considerable practice since the pressure on the tape is not constant, but is governed by the compressibility of the fleshy contour which varies among individuals. Careful attention was given to the girth specifications. All measurements were taken with the superior portion of the tape lined up with the inferior portion of the landmarking lines. Circumferences were recorded to the nearest 0.1 cm.

• **Arm girth relaxed (ArmR)** was the perimeter distance of the right arm parallel to the long axis of the humerus when the subject was standing erect and the relaxed arm was hanging by the sides. The level of the tape was at the measured and marked mid-acromiale-radiale distance.

• **Arm girth flexed (ArmF)** was the perimeter distance of the right arm parallel to the long axis of the humerus when the subject was standing erect and the arm was abducted 90 degrees. The subjects were asked to flex the bicep and the measurement was read during the contraction at the measured and marked mid-acromiale-radiale distance.

• **Waist** was the perimeter at the level of the noticeable waist narrowing located approximately half way between the costal border and the iliac crest. In subjects where the waist was not

apparent, an arbitrary waist measurement was made at this level. Measurement was taken at end of normal expiration.

• **Gluteal girth (Glut; hip girth)** was the perimeter at the level of the greatest posterior protuberance and at approximately the symphysis pubis level anteriorally. The subject, during this measure, stood erect, feet together, without volitionally contracting the gluteal muscles.

• **Proximal Thigh Girth (Thi)** was the perimeter of the right thigh which was measured when the subject stood erect, legs slightly parted, weight equally distributed on both feet. The tape was raised to a level just distal to the gluteal fold.

• **Mid Thigh (MThi)** was the perimeter of the right thigh perpendicular to the long axis of the femur at the mid trochanterion-tibiale laterale level. The subject position was the same as for proximal thigh girth.

• **Calf girth (C)** was the subject in the same position as the medial calf skinfold, with the foot placed on an elevated surface, causing the leg to be in a position of 90degrees of hip and knee flexion. As aforementioned, the tape was manoeuvred to obtain the maximum perimeter of the calf. This measure was obtained by manipulation of the tape taking a series of girth measurements to assure the largest value. This was achieved by relaxing and tightening the tape with manipulation to various levels facilitated by the anthropometrist's third digits.

Breadths

General technique

The bone callipers were held in pistol-grip mode. The branches were gripped by the thumb and the index finger and rested on the backs of the hands. The middle finger was used to locate the landmark. Firm pressure was applied to the branches. The scale on the smaller bone callipers was read to the nearest 0.1 cm.

• **Humerus Breadth:** (bi-epicondylar) was the distance between medial and lateral epicondyles of the humerus when the arm was raised forward to the horizontal and the forearm was flexed to a right angle at the elbow. The small bone calliper was applied pointing upwards to bisect the right angle formed at the elbow. The epicondyles were palpated by the third digits starting proximal to the sites. The measured distance is somewhat oblique since the medial epicondyle is lower than the lateral. The measurement was taken with the callipers placed at a 45degree angle to the joint, ensuring the pressure plates were applied firmly to the encompassed sites.

• **Femoral Breadth:** (bi-epicondylar) the distance between medial and lateral epicondyles of the femur when the knee was bent to 90˚ of flexion. The small bone calliper was applied pointing downwards to bisect the right angle formed by the knee. The epicondyles were palpated by the third digits starting proximal to the sites. The measured distance is somewhat oblique since the medial epicondyle is lower than the lateral. The measurement was taken with the callipers placed at a 45˚ angle to the joint, ensuring the pressure plates were applied firmly to the encompassed sites.

3.5.2 Dual Energy X-Ray Absorptiometry Assessment

Prior to DXA scanning, daily Quality Control (QC) was conducted utilizing manufacturer recommendations. QC procedures consisted of scanning a phantom block prior to each testing day (ISCD, 2007). Phantoms are artificial representations of human tissue of differing compositions. Once QC was complete, the subjects were individually assessed. Individuals were asked to remove jewellery, garments containing zippers or buttons made of metal and shoes prior to being placed on the scanning bed**.** Subjects were not required to disrobe providing they followed the pre scan procedure and were wearing comfortable loose fitting clothing. The subjects were measured for height and scale weight. The subject's date of birth and ethnicity were inputted into the software to allow comparison to normative data. The subjects were asked to sit on the edge of the DXA and bring their legs up onto the scanning bed and turn so that they were sitting parallel to the long axis of the bed. Once in position the subjects' upper body was slowly guided down by the DXA technician, aligning the central long axis of the spine with the centre of the DXA scanning bed. Once lying, the technician advised the subject to slowly track their body either upwards or downwards on the scanning bed ensuring that the feet were not outside the scanning window at the bottom of the bed, and that the head was not obstructing the initial scanning signal at the top of the bed. The latter technique is required for the scanner to establish a baseline that does not include tissue prior to the scan. Once satisfied with the position, the technician placed one or two velcro straps around the subjects' legs at approximately the mid shin level to prevent lower limb movement or "splaying" of the legs. The technician instructed the subject to keep their arms and hands tight to the body, and completed one final check to ensure that all body parts were within the scanning bed window. Subjects

were instructed to lie still during the entire process of the scan (5-10 minutes), and refrain from talking while the scanning arm was above the torso.

3.5.3 Assessment of Hydration

The 1st urine sample of the day was collected the morning of the assessment in order to assess hydration as well as screening for pregnancy prior to DXA testing. At home upon waking, or, the first urination after waking, the subject provided a minimum 20 ml sample in a collection container from the study package. The subject then inserted the pregnancy test dipstick into the urine as per the manufacturer's instructions and placed the used pregnancy dipstick in the provided Ziploc bag. Subjects were asked to place their study identification number on both the ziplock bag as well as the sample container and were reminded to NOT write their names on any of this material. Subjects brought this material to their scheduled assessment. Subjects menstruating were not asked to provide a sample, nor were they required to provide the results of a pregnancy test. Urine specific gravity was tested at room temperature using an ATAGO PAL-10S refractometer (Atago USA, Inc. WA). Samples were measured individually by pipetting a small amount of urine (approximately 3 drops) into the refractometer reservoir. The average of 3 measurements was taken to represent the hydration status of the individual.

3.6 Skinfold Regression Validity

A series of female anthropometric regression equations (Tables 3.4, 3.5, 3.6) derived from multicompartment (MC) (n=3), DXA (n=9), and hydrodensitometry (HD) (n=8) criterion models were used to calculate % Body Fat (% BF) and were validated against our % BF DXA

criterion measure. Subjects included in the regression validation underwent both DXA and anthropometric measurements within the same assessment session as outlined in 3.5

3.7 Statistical Analysis

Multiple statistical packages were utilized to perform analyses. SPSS 17.0 (SPSS Inc, Chicago, Illinois), Sigma Plot 11.0 (Systat Software Inc., San Jose, California) and statistical functions in Microsoft Excel 2007 (Microsoft Corporation) were the three statistical packages employed. The primary independent variable under investigation was % Body Fat (% BF). When discussing the ability to predict % BF in the regression equation, this variable will be referred to as the Response Variable (RV) (Guo, 1996). Anthropometric variables investigated in the study included weight, height, circumferences (n=8), bone breadths (n=2), skinfolds (n=10), and summed skinfold combinations (n=9) as they constitute a typical anthropometric performa (testing barrage). DXA variables analysed included % BF DXA, Tissue Total (TT), Fat Total (FT), Lean Total (LT), Total Mass (TM) and total Bone Mineral Density (BMDg/cm^2). All DXA units are in kilograms (kg) with the exception of % BF DXA and BMD which are represented by the units shown above.

3.7.1 Reliability

The ability to detect compositional changes in the field of anthropometry is contingent upon the skill of the technician. The most valued characteristic of a good anthropometric technician is the ability to attain a high level of reliability through meticulous repetition quantified by a test-retest procedure. The test-retest procedure allows the practitioner to quantify reliability on the same group of subjects over a designated time interval. The resulting “within

practitioner" statistics are commonly referred to as intra-rater reliability. Better reproducibility of a test-retest is not a singular quality but constitutes time, equipment, study design, and tries to eliminate or minimize all these sources of variance (Beckerman *et al.*, 2001). A variety of statistical indices were used to quantify reliability in relative and absolute terms. For DXA body composition analysis the reliability score reflects the error of the machine, as technician involvement is minimal, and therefore, sometimes referred to as "precision" as the replication of a machine or instrument is being evaluated instead of a person. For continuity of terminology the term "DXA reliability" will be used in the analysis, although it is synonymous with DXA precision from the literature.

3.7.2 Relative Reliability

Relative reliability is often represented by a unitless correlation coefficient (r). Commonly used correlation coefficients include the parametric Pearsons Moment Correlation Coefficient (Pearsons r), non-parametric Spearmans Correlation Coefficient (Spearmans r) and the Intra class Correlation Coefficient (ICC). The preferred method of quantifying reproducibility uses the ICC as it is the ratio of variance between subjects to total variance using the F value from repeated measures ANOVA (Beckerman *et al.*, 2001). Designed for assessing intra-rater reliability the ICC displays an additional advantage over methods such as Pearsons r, as it is bivariate rather than univariate allowing relative reliability indices to be calculated for multiple rather than single datasets. However, the ICC is subject to the same limitations as Pearsons r in that it is influenced by sample heterogeneity because it includes the variance term for individuals in the calculation, in addition to complying with the assumption of distribution normality (Atkinson & Nevill, 1998). In circumstances where this is not true, or when the

dataset is non-parametric the Spearmans r may be a better statistic for assessing relative reliability as it assesses the maintenance of rank from trial to trial and is not subject to errors in assumption of shape of data distribution while being less affected by outliers (Atkinson & Nevill, 1998).

Relative reliability was calculated using SPSS 17.0 for both anthropometric and DXA reliability data. The single measure, two-way mixed ICC with 95% consistency model, in addition to the 95% confidence intervals of the ICC was calculated for all combinations of summed skinfold and all DXA variables. ICC values indicate a perfect positive relationship when the value is 1, while a perfect negative relationship is represented by -1. Suggested guidelines for basic ICC interpretation have been proposed and are referenced in relation to degrees of acceptable reliability. In this scheme, values of 0.70-0.79 are defined as below average acceptable, 0.80089 average acceptable, 0.90-1.0 above average acceptable (Baumgartner, 2007). The existence of standardized, well agreed upon limits for "reliable enough" data are equivocal, indicating that interpretation of sufficient ICC values are contingent upon the desired scope of practice (Atkinson & Nevill, 1998).

3.7.3 Absolute Reliability

In the field of anthropometry the Technical Error of Measurement (TEM), serves as an index of absolute measurement precision (Gore, 2000). In other fields, this error is referred to as the Standard Error of the Measurement (SEM) (Beckerman *et al.*, 2001), Typical Error (TE) (Hopkins, 2000), and Measurement Error (ME) (Eliasziw *et al.*, 1994). The term TEM is primarily restricted to the field of anthropometry (Gore, 2000) while most other exercise science

disciplines have adopted the Standard Error of Measurement (Baumgartner 1989). Maintaining terminology within the respective fields, TEM's will be calculated for anthropometric data, while the term SEM will be used when handling DXA data.

Similar to the variety of names, TEM or SEM can be calculated from a variety of statistical methods all yielding similar if not identical results. Commonly cited equations are listed in Table 3.1.

Table 3.1: Commonly Cited Calculations for TEM / SEM Values

	The difference method
eqn 1.*	TEM = standard deviation of the differences/√2 (Anthropometrica 1996)
eqn 2.	TEM = √ (sum of differences squared / divided by 2n) (Gore 2000) Note: difference is squared then summed
eqn 3.	TEM =Sample Standard Deviation √1- Intraclass Correlation Coefficient (Atkinson and Nevill 1998)
	ANOVA Method
eqn 4.	TEM = √mean square error (AAP 1996; Stratford and Goldsmith 1997; Atkinson and Nevill 1998) Note: data are derived from MSE in a repeated measures ANOVA

***** Equation used for TEM / SEM calculations in the current study

TEM values were calculated by the difference method using eqn 1 in an excel spreadsheet, where the differences are calculated by taking the first anthropometric test results (AnthroT1) and subtracting the results from the second anthropometric test (AnthroT2). TEM values were multiplied by 1.96. This value corresponds to 95% confidence intervals of the TEM.

The TEM was also expressed as a percentage by the equation

eqn 5. %TEM = [TEM / (meanAnthroT1+meanAnthroT2/2)] x 100

where TEM = the TEM value of the variable, divided by the average of the means (mean T1+meanT2 /2) multiplied by 100 to get a percentage.

The calculation of % TEM is equivalent to the more commonly utilized statistic called the Coefficient of Variation (CV%) which will be used when handling DXA data. The ISAK guidelines for acceptable levels of absolute precision were utilized to evaluate the reliability and level of expertise of the anthropometrist (Table 3.2).

Table 3.2: %TEM Values for Various Levels of ISAK Accreditation

	Variable	Level 1	Level 2	Level 3
Intra-Tester Reliability*	Skinfolds	7.5%	5.0%	5.0%
(Within Tester)	Other	1.5%	1.0%	1.0%

*Guidelines utilized for assessing anthropometrist reliability in the current study. Other = circumferences, bone breadths and segment lengths. Level 1 is the first level of ISAK accreditation while Levels 2 and 3 are increasingly advanced levels requiring greater mastery of larger measurement performa's.

The acceptable levels of reliability are expressed as %TEMs, allowing anthropometrists to evaluate their degree of precision on a variety of individual or aggregate anthropometric measurements (e.g., summed skinfolds). Greater levels of reliability, hence, lower TEM values, allow greater probability of detecting the relatively small changes that can occur in elite athlete populations.

3.7.4 Anthropometric Smallest Real Change

The TEM represents the biological and technical error associated with a measure in the units that the measurement was taken and reflect the magnitude of error a single measure will contain two thirds of the time (Hopkins, 2000). In clinical or applied settings the primary

objective is to assess responsiveness (the ability of a scale to detect clinically relevant changes between serial measurements (Beckerman *et al.*, 2001)). Several authors have suggested that a value be calculated that accounts for the measurement error of a single measurement as well as signifying 95% certainty that the smallest real difference (SRD) (Eliasziw *et al.*, 1994; Clark *et al.*, 2006), least significant change (LSC)(Shepherd & Lu, 2007), or minimal detectable change (MDC) (Beckerman *et al.*, 2001) has occurred between two measurements. SRD values will be the term used when calculated from anthropometric data and is synonymous with the term LSC or more appropriately Ideal LSC, which will be utilized when referring to DXA data (Shepherd & Lu, 2007).

The SRD can be calculated as:

eqn 6. SRD = 1.96 x √2x SEM (Eliasziw *et al.*, 1994)

where SRD = the smallest real difference at the level of 95% certainty, where 1.96 is the T-statistic for 95% confidence used in large sample selection.

This equation can be further simplified by multiplying 1.96x√2 yielding

eqn 7. SRD = 2.77 x TEM (Hopkins, 2000; Shepherd & Lu, 2007)

When assessing serial measurements for genuine change, the threshold for change must exceed the SRD value to ensure that the magnitude of change is unlikely to be a result of measurement error. Some authors believe that the 95% confidence interval and its arbitrary designation as the criterion for critical threshold should be challenged in sport settings where the magnitude of change may be clinically significant although not statistically significant. In these

instances, such as elite sport evaluation, it may be necessary to calculate a range of values representing differing confidence intervals and probabilities that "true" change has occurred (Hopkins, 2000). Utilizing alternative confidence limits and applying qualitative inferences on clinical chances of change allows for a "description of change that occurred" rather than an absolute evaluation (Table 3.3).

Table 3.3: Alternative Confidence Limits and Clinical Chances of Change

Alternative Confidence Limits		**Multiplicative Factor for Limits**		**Clinical Chances of Change**	
Probability	**Odds**	**Single Measurement**	**Change in a Measurement**	**Chance Range**	**Qualitative Outcome**
52%	1 to 1	0.71	1	25-75%	is possible (not), may (not) be
68%	2 to 1	1	1.41		
80%	4 to 1	1.28	1.81	75-95%	is likely to be, is probably
90%	9 to 1	1.65	2.33		
95%	19 to 1	1.96	2.77*	95-99.5%	is very likely to be

Alternative Confidence Limits *Probability;* the probability represented by percentage that the true value lies within the confidence limits when applied to a single measurement, or the probability that real change has occurred between multiple measurements. *Odds*; the odds associated with the given probability level. ***Multiplicative Factor for Limits*** represents the values that are multiplied by the TEM/SEM values to yield the confidence intervals in the case of *single measurements*, or the SRD in the case of *change in measurement* (multiple serial measures); ***Clinical Chances of Change***, a standardization of alternative confidence limits denoted as *Chance Range*, with corresponding standardized *Qualitative Outcomes* for interpreting and reporting the data. * factor generates the commonly utilized 95% limits of agreement

3.7.5 Systematic Change in the Mean

Paired Sample, Two Tailed Student's T-Tests were performed to identify significant systematic bias by comparing the means of the test-retests. A p-value < 0.05 was used to signify statistically significant differences between pairs of data. A p-value of < 0.05 leaves a 5% chance that a type I error can occur, or, stated differently, a 5% chance that we reject the null hypothesis, when in fact a true difference does exist. Criticism of using this method solely to identify bias in paired data is that there is no indication of random variation between tests, so

that significant systematic bias is less likely to be detected if there is a large amount of random error between tests and vice versa (Atkinson & Nevill, 1998).

In addition, Bland Altman plots were created to assess systematic bias and random error by identifying magnitude and direction of scatter about the zero line. The limits of agreement (LOA) set at ± 95% above and below the mean difference were identified during this process (Bland & Altman, 1986). For reliability analysis it represents a 95% chance that a subject's "true" value lies within the limits of agreement. Plots were constructed by plotting the y-axis (independent variable) consisting of the individual differences between the measures from T1 vs. T2 (T1-T2) against the x-axis, values that were calculated as the average of the means from T1 and T2.

eqn 8. y- value = datapoint 1 (T1) – datapoint 1 (T2)

eqn9. x- value = (datapoint 1 (T1) + datapoint 1 (T2)) /2

Limits of agreement were calculated as

eqn10. $LOA_{95} = SD_{diff} \times 2$

Where SD diff is the standard deviation of the differences between T1 & T2 and 2 is the approximate multiplicative factor for establishing 95% confidence limits.

Once calculated, the LOA were graphically plotted as threshold lines, with the mean difference for T1- T2 representing the bias.

eqn11. Bias= Mean (T1) – Mean (T2)

3.7.6 Bias of the Error

Heteroscedasticity is the term used to designate that the error of a measurement is related to the magnitude of the measurement (e.g., as the measurement value increases, so does the error value). One of the cautions of using absolute reliability with anthropometric data, is that skinfolds often display heteroscedasticity as skinfold sites with higher amounts of subcutaneous fat are susceptible to larger TEM values (Gore, 2000). For variables that display this characteristic, it has been suggested that log or quadratic transformation will be necessary for further interpretation of these variables using the limits of agreement analysis (Durnin & Womersley, 1974; Atkinson & Nevill, 1998).

Presence of heteroscedasticity was investigated by examining the Bland Altman plots. A Pearsons correlation coefficient (r) was calculated between the absolute differences and the individual means. Correlations that were non significant at $p < 0.05$, were interpreted as being homoscedastic (the magnitude of error is **not** related to magnitude of measurement) justifying interpretation of the data using limits of agreement without transformation.

3.8 VALIDITY

Concurrent validity was investigated to identify predictive anthropometric regression equations that provide results that give close agreement to the "true" value taken from DXA measurements. The response variable (RV) serving as the validation criterion was the % BF DXA measurement from the GE Lunar Prodigy DXA. The predictive variables (PV) were taken from regression equations from the literature that predicted % BF, and were created using Multi-

compartment (MC) (n=3), DXA (n=9), and Hydrodensitometry (HD) (n=8) models as the criterion (Table 3.4, 3.5, 3.6). Group statistics are presented as the mean and standard deviation.

Table 3.4: Anthropometric Variables used in Various Regression Equations Developed with Multi Compartment as the Criteria

Author	Skinfolds								
Multicompartment Regression	Tri	Bi	SSc	Ax	IC	SSp	Abb	Thi	C
Peterson 2003 % BF = 22.18945 + (age x .06368) + (BMI x .60404) - (height x .14520) + (sum4 x .30919) - (sum4^2 x .00099562)	X		X			X		X	
Evans 2005 % BF= 8.997+ 0.24658 x (sum3) 6.343 x (gender) 1.998 (race)	X						X	X	
Moon 2009 % BF = 7.787 + .24658 x (sum3)	X						X	X	

Tri=triceps, Bi= biceps, SSc= subscapular, Ax= axilla, Abb= abdomen, Thi = thigh, C= Calf *Peterson 2003*; sum4= Tri, SSc, SSp, Thi; *Evans 2005*; sum3 = Tri, Abb, Thi ; *Moon 2009*; sum3= Tri, Abb, Thi

Table 3.5: Anthropometric Variables used in Various Regression Equations Developed with DXA as the Criterion

Author	Skinfolds								
DXA Regression	Tri	Bi	SSc	Ax	IC	SSp	Abb	Thi	C
Ball 2004									
% BF = [-6.40665 + .41946 (sum3) - .00126 $(sum3)^2$+ .12515 (Glut) + .06473 (age)]	X					X		X	
Warner 2004									
FFM = 8.51 + (.809 x mass) - (.178 x Abb) - (.225 x Thi)							X	X	
Kagawa 2007									
% BF 8 Var. = -4.504 + .016 (tri) + .0154 (IC) + .0281 (Bi) + .263 (Glut) + .229 (C) -3.249 (bH) + .517 (ArmR) + .125 (Abb#2)	X	X			X				X
% BF Sum7. = 8.48 + .151 (sum7)	X	X	X		X	X	X	X	X
% BF 3 Var. = 10.091 + .410 (Tri)+ .316 (IC) + .750 (Bi)	X	X			X				
% BF Sum3. = 11.294 + .328 (sum3)	X		X	X					
Jackson 2008									
% BF = (0.4446 x (sum3)) - (0.0012 x $(sum3)^2$) + 4.3387	X					X		X	
Duz 2009									
% BF = 27.988 + 17.477 (log Tri) + 8.467 (log SSp) + 19.854 (Thi)	X					X		X	
Kohli 2009									
BFM = -25.26 - 17.36 (height) + .44 (mass) + .34 (Glut) + 8.87 (log Tri) + .14 (age)	X								

Tri=triceps, Bi= biceps, SSc= subscapular, IC = Iliac Crest, Ax= axilla, Abb= abdomen, Abb#2 = abdomen #2, Thi = thigh, C= calf , Glut = gluteal circumference, bH= humerous breadth, ArmR= arm relaxed *Ball* 2004 sum3; Tri, SSp, Thi ; *Kagawa* 2007 sum7; SSc, Tri, Bi, IC, SSp, Thi, C; Kagawa 2007 sum3; SSc, Tri, Abb #2, *Jackson 2008 sum3*; Tri, SSp, Thi

Table 3.6: Anthropometric Variables used in Regression Equations Developed with Hydrodensitometry as the Criterion

Author	Skinfolds								
Hydrodensitometry Regression	Tri	Bi	SSc	Ax	IC	SSp	Abb	Thi	C
Sloan 1962									
BD = 1.0764 - (.00081 (SSp)) - (.00088 (Tri))						X		X	
Durnin & Womersley 1974									
BD = 1.1567- .0717 (log10 x sum4) 16 - 68 year regression	X	X	X		X				
BD = 1.1599 - .0717 (log10 xsum4) 20 - 29 year regression	X	X	X		X				
Jackson & Pollock 1980									
BD(1) = (1.1470292 - .0009376 (sum3) + .000003 (sum3)^2- .0001156 (age) - .0005839 (Glut)	X					X		X	
BD (2) = [(1.096095 - .0006952 (sum4) + .0000011 (sum4)^2- .0000714 (age)]	X					X	X	X	
*BD (3) = [(1.0994921 - .0009929 (sum3) + .0000023 (sum3)^2 - .0001392 (age)]	X					X		X	
Jackson & Pollock 1985									
*BD = 1.089733 - .0009245 x (sum3) + .0000025 x (sum3)^2- .0000979 x (age)	X					X	X		
Thorland 1984									
BD = (1.0987 - .00122 (sum3) + .00000263 (sum3)^2	X		X		X				

* preferred equation of CSEP CEP and ACSM. Tri=triceps, Bi= biceps, SSc= subscapular, Ax= axilla, Abb= abdomen, Thi = thigh, C= Calf *Durnin & Womersley 1974 sum4;* Tri, Bi, SSc, IC; Jackson & Pollock 1980 BD1*BD3 sum3; Tri, SSp, Thi ; *Jackson & Pollock 1980 BD2 sum4;* Tri, SSp, Abb, Thi; *Jackson & Pollock 1985 sum3* Tri, SSp, Abb; *Thorland 1984 sum3;* Tri, SSc, IC Note: The Siri equation (Siri, 1956) was utilized to convert predicted body densities to % BF

3.8.1 Relative Agreement

Pearson's product moment correlations (r) were conducted to assess the relationship between the criterion and prediction. The resulting Validity Coefficient (R) is representative of how well the predicted variables describe the criterion. The Coefficient of Determination (R^2) represents the proportion of total variance explained in the response variable by the predictive variables (Guo, 1996). Similar principles apply to interpretation of the validity coefficient as when interpreting the reliability coefficient. The larger the value of R and R^2, the greater the validity. It should be noted that the coefficient in validity studies plays a more prominent role in interpretation than in reliability. When used to compare a predictive variable to its criterion, the correlation will reflect both between subject variance, as well as within-subject variance. This contrasts to the coefficients application in reliability studies where between-subject variance is omitted and within-subject variance is the primarily analysed (Hopkins, W. *A New View of Statistics – Precision of a Measurement,* <http://www.sportsci.org/> "accessed February 2010"). When the validity coefficient is near 1 the predicted % BF is similar to the % BF DXA criterion. When the value is close to 0, the ability to predict the criterion score is poor (Baumgartner, 2007). Although a value for success of a prediction equation has not been agreed upon, it is common to see values in the literature of > 0.80 to be defined as having moderately high or acceptable predictive ability (Guo, 1996; Heyward, 1996; Baumgartner, 2007).

3.8.2 Absolute Agreement

Paired sample, two tailed Students T-Tests were performed between the criterion and predicted values with a p-value of < 0.05 indicating significant differences between

% BF values. In addition the Standard Error of Estimate (SEE) was calculated as a measure of the accuracy of the predictions.

eqn.12 $SEE = SD\sqrt{1-r^2}$ (Moon *et al.*, 2009b)

Where SD is the standard deviation of the difference between predictor – criterion, and r is the Pearsons product moment correlation coefficient. The SEE represents the average variance of individual predicted variables around the regression line.

Total Error (TE) or Pure Error (PE) was calculated to represent the standard deviation of the difference between predictor and criterion. The TE was calculated as

eqn.13 $TE = \sqrt{\text{sum (predicted-actual)}^2 / n}$ (Moon *et al.*, 2009b)

The TE is used as the primary indicator of an equation's efficacy. The TE represents the average deviation of individual scores from the line of identity. In this regard, the TE will be larger than the SEE because the SEE only represents the degree of deviation of individual scores from the regression line without accounting for the influences of the constant error (CE) or bias as the TE does. In her book "Applied Body Composition Assessment (1996)", Heyward cites validity standards proposed by Lohman, allowing evaluation of the predictive ability of regression equations (Heyward, 1996) (Table 3.8).

Table 3.7: Lohman's Standards for Evaluating Predictive Ability of Regression Equations

Total Error % BF	Subjective Rating
2.0	Ideal
2.5	Excellent
3.0	Very Good
3.5	Good
4.0	Fairly Good
4.5	Fair
5.0	Poor

Lohman's Proposed Evaluation of Validity using Total Error and Subjective Rating Values as seen in "Applied Body Composition Assessment, 1996" (Heyward, 1996).

3.8.3 Agreement of Predictions for Individuals

Constant Error (CE) was calculated as the mean difference between the predicted minus the criterion. In this manner, if the CE was negative, then the prediction equation underestimated % BF in comparison to DXA, if the CE value was positive, then an over prediction of % BF was evident. Bland Altman plots were created to visually inspect bias in the mean of the two methods. Limits of agreement set at 95% above and below the mean difference were established (Bland & Altman, 1986).

> **eqn 14.** Limits of Agreement (LOA) = Standard deviation of the differences multiplied by ± 1.96

For the analysis of validity, the LOA represent the generalizability of the study to a larger sample. The LOA articulates the range of values for which 95% of subjects values will be predicted in relation to the criterion value. It is therefore indicative of the error in a method's predictive ability. Large LOA represent large error, while small LOA represents small error. In the case of skinfold measurements, when individual skinfold measurements are converted to % BF by a regression equation the predicted % BF for any given individual will be within the range specified LOA from true values obtained by the DXA criterion.

3.9 New Regression Equation using DXA as the Criterion

3.9.1 Identification of Predictive Variables

Sigma Plot 11.0 was used in an attempt to create new regression equations with better reliability and validity than existing equations. The median of the three skinfold values for all skinfold measurements were included as candidate variables along with several variables that have previously been selected as significant predictors of % BF in existing equations (age, height, weight, and gluteal circumference).

Regression equations were developed using two differing statistical approaches. The first approach, forwards and backwards stepwise multiple linear regressions were used to identify the best anthropometric variables for describing the response variable (% BF DXA) avoiding the use of extraneous variables. Stepwise multiple regression utilizes statistical criteria to identify independent variables that share the most variance with the dependent variable. In forward stepwise regression, variables are continually added to the regression equation based upon their ability to significantly increase the magnitude of the predictive power of the equation as indicated by the R^2 value. When addition of new variables is non-significant, the process is complete (Heyward, 1996). Backwards stepwise regression is similar except that all variables are added initially, and variables are sequentially removed from the equation based on their inability to contribute to the prediction based on a specified threshold for their F value. F value thresholds for entry were held at 4.0 while F-values for removal were set at 3.90 as recommended by SigmaPlot (SigmaPlot 11.0) After predictive variables were selected, they were entered into a multiple linear regression to develop a regression equation.

Best Subsets Regression analysis was performed including the same candidate variables to further investigate if there were any additional subsets of independent variables that could explain the response or dependent variable. Subsets with the highest adjusted R^2 were selected for entry into a multiple regression equation, as the adjusted R^2 value represents how well the regression model describes the data but accounts for the number of independent variables.

The second method of regression creation utilized was the hierarchical approach. This approach allows predetermined candidate variables to be entered into the regression by the researcher based on theoretical models or logical considerations (Heyward, 1996). This method has been suggested by various authors in their creation of regression equations (Jackson, 1984). To articulate this, Jackson (1984) conducted a forward stepwise regression on previously published data to show problems occurring with stepwise procedures. Using body density as the dependent variable, and twenty two anthropometric measures as independent variables, the data was randomly split into three groups of 225, 40, and 18 females. Each group entered into stepwise regression yielded different primary predictive variables in the first four steps despite differing only by sample size. Furthermore, the associated multiple correlation at each step was found to be a function of sample size as the smallest group with 18 subjects displayed the highest coefficients. The authors concluded that these pitfalls can be avoided with sufficient sample size and by selection of physiological variables as opposed to statistical edit.

Taking theoretical and logical considerations into hand, the study investigator chose to enter sites that were part of the CSEP CPT performa implemented by the national certification program. The variables included the SSc, Tri, Bi, IC, and Calf

skinfolds. As well, age and gluteal circumference were included as individual measurements. In addition, a correlation matrix between the skinfold variables was created to assess the magnitude of multicollinearity between measurements. To minimize the inherent limitations of covariance between skinfold measures, independent variables were created by summing various combinations of skinfolds. These skinfolds were summed (sum5) and entered into the regression equation to account for multicollinearity. Summed values allow variables to be placed in a regression without violating the rules of regression creation, in addition to providing a more representative indices of total body fatness, and create a more favourable ratio of subjects utilized to create the regression equation (Jackson *et al.*, 1980). Furthermore, to ensure normality, the sum5 was logarithmically transformed and entered into a regression equation. The last investigation utilized polynomial regression creation to see if the sum5 displayed a curvilinear relationship between the independent and dependent variables as suggested in previous literature (Jackson, 1984).

3.9.2 Interpretation of Regression Results

A series of additional tests options were selected from SigmaPlot to ensure all regression equation assumptions were checked. The Shapiro-Wilk test was selected to assess normality with a p-value of 0.05. If a regression failed the normality test, values were logarithmically or quadratically transformed, or a different model was investigated as failure of the test indicated that outliers were present or there was an incorrect model selection. A Constant variance test was calculated with a p-value of 0.05. The Durbin-Watson statistic was used to ensure there was no correlation of the residuals. A value of 2 for this statistic indicates that the residuals are not correlated. SigmaPlot automatically

investigates multicollinearity. Variance Inflation Factors exceeding the threshold value of 4.0 would cause a multicollinearity flag to appear on the output.

Supplementary diagnostic options were selected to automatically detect the presence of influential data points. The diagnostics used included the DFFITS, Leverage, and Cooks distance test as suggested in the software advisor. All additional tests were interpreted as per recommendations in the SigmaPlot software.

3.9.3 Cross Validation

Cross validation of newly created regression equations was conducted utilizing the Predicted Residual Sum of Squares (PRESS) method (Holiday *et al.*, 1995). The PRESS method was used to anticipate the stability of the regression when applied to external populations. An option in the SigmaPlot software was selected to calculate the PRESS residuals in the multiple regression output. The PRESS SEE and PRESS R^2 were then calculated. The calculations are summarized below;

eqn 15. PRESS SEE = $\sqrt{}$PRESS/n

where PRESS is the PRESS residuals from the SigmaPlot multiple regression output and n is the number of subjects in the study.

eqn 16 . PRESS $R^{2 =}$ 1- (PRESS/SStotal)

where PRESS is the PRESS residuals from the SigmaPlot multiple regression output and SS total is the Sum of Squares total from the ANOVA table output.

CHAPTER 4- RESULTS

4.1. Subjects

Twenty (n=20), thirty two (n=32), and ninety five (n=95) female athletes (aged 17-31 yrs) were recruited from the University of Manitoba, University of Winnipeg, and Canadian Sport Centre Manitoba to participate in the anthropometric reliability, DXA reliability, and skinfold validity components of the study respectively. All subjects participating in the skinfold validity component were the same as those in the anthropometric and DXA reliability components. Table 4.1 indicates descriptive characteristics for subjects in the study.

Table 4.1: Subject Descriptive Characteristics for each Component Investigated in Study

	ANTHROPOMETRIC RELIABILITY*		DXA RELIABILITY		SKINFOLD VALIDITY	
Variable	**Mean ± SD**	**Range**	**Mean ± SD**	**Range**	**Mean ± SD**	**Range**
n	**20**		**32**		**95**	
Age (years)	21.0 ± 3.1	18.0-29.0	22.0 ± 3.2	18.0-29.0	22.0 ± 3.4	17.0-31.0
Body mass (kg)	69.6 ± 7.6	60.2-83.9	65.0 ± 9.0	52.5-86.8	67.8 ± 8.8	51.1-86.8
Height (cm)	173.0 ± 7.9	162.5-188.8	169.5 ± 8.7	155.6-185.9	172.0 ± 9.0	154.1-191.0
Body mass index (kg/m^2)	23.2 ± 1.7	20.6-26.4	22.54 ± 1.9	19.0-25.9	22.8 ± 2.2	17.9-28.2

*Descriptive values taken from Anthropometric Time point 1 (AnthroT1), *Anthropometric Reliability*; data collected from subjects participating in the establishment of anthropometric TEM values, *DXA Reliability;* data collected from subjects participating in duplicate DXA assessments to establish the DXA SEM values and *Skinfold validity;* data collected from subjects participating in assessments of both DXA and Anthropometric determination of % BF.

Of the 93% of participants that indicated training status, 59% had competed within four weeks (1 month), while 82% had indicated they had competed within the last nine to twelve weeks (2-3 months). Athletes practiced on average six times per week for approximately 120 minutes with an average intensity that was deemed moderate to high.

Additionally, an average of four aerobic session and three resistance training sessions of 45and 60 minutes were stated as being supplemental to the practices.

A broad variety of sports were represented in the sample (Figure 3.1) with the majority of athletes competing at an elite level, (International or professional 44%, National or Varsity 50%, and Lower than National or Varsity 6%) with those competing below national or varsity levels hoping to qualify for national competition in the near future.

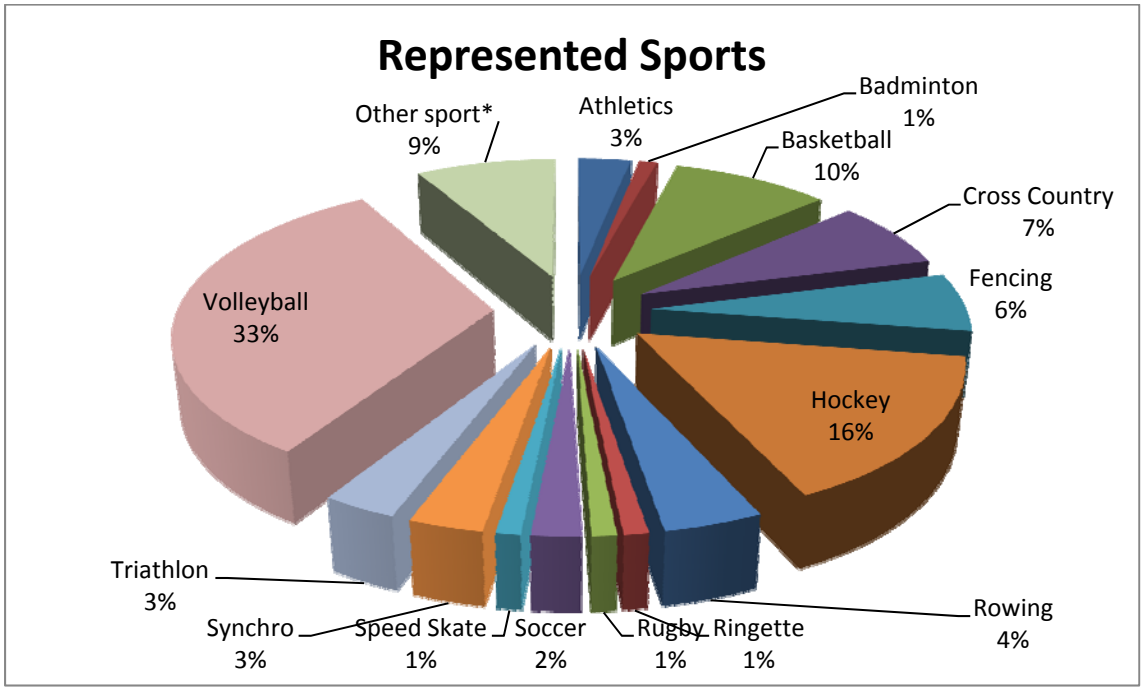


Fig4.1: Representation of sport by athletes participating in *Skinfold Validity* component of study; Synchro = synchronized swimming, Cross Country = cross country running, *Other sport = fitness & figure competitors, ultimate frisbee, and cross-fit competitors

For the DXA reliability component of the study 21.9% of the subjects were menstruating and were assumed to be non pregnant and did not provide a urine sample. All remaining subjects tested negative for pregnancy. Urine specific gravity was 1.020 ± .006 g/cm3 (range 1.010-1.030) indicating relative euhydration (Kavouras, 2002). Of the twenty five samples, urine specific gravity was not significantly related to the difference

in DXA T1 – DXA T2 measurements using a pearson's product moment correlation ($p < 0.05$).

For the skinfold validity component of the study similar statistics were observed. Of the ninety five subjects tested, 21% were menstruating at time of assessment, thus, did not take a pregnancy test or provide a urine sample. All remaining subjects tested negative for pregnancy. Urine specific gravity was 1.019 ± .006 g/cm3 (range 1.007-1.034) indicating relative euhydration. Of the seventy five samples, urine specific gravity was not significantly related to % BF DXA using a pearson's product moment correlation ($p < 0.05$).

4.2 Anthropometric Reliability

Twenty subjects (n=20) were tested twice (AnthroT1 & AnthroT2) within 7 days of each other to establish anthropometric reliability. Anthropometric variables investigated included weight, height, circumferences (n=8), bone breadths (n=2), skinfolds (n=10), and sum of skinfold combinations (n=9). There were no clear advantages of using one method of data compilation over another for calculating individual TEM values. The median of 3 for each variable was used for subsequent calculations as per previous recommendations in the literature (Hume & Marfell-Jones, 2008). All values were analysed by comparing AnthroT1 to Anthro T2, with values presented as means and their SD (Table 4.1).

4.2.1 Relative Reliability

All test–retest ICC's indicated above average acceptable levels of relative reliability (Baumgartner, 2007) (Table 4.1). The anthropometrist displayed an excellent

level of reliability as all relationships between AnthroT1 and AnthroT2 displayed reliability coefficients of 0.984 or greater.

Table 4.2. Relative Reliability of Various Skinfold Sums utilized in Regression Equations

Sum of Skinfolds	Trial #1	Trial #2	ICC[a]	95% CI[b] for ICC
	Mean ± SD	Mean ± SD	3 MED	Low- High
Sum 3 (Tri, Abb, Thi)	48.7 ± 9.5	49.3 ± 9.7	0.984	0.980- .997
Sum 3 (Tri, Ssp, Thi)	42.4 ± 9.5	42.9 ± 9.9	0.987	0.966-.995
Sum 3 (Tri, Ssc, IC)	40.0 ± 9.5	40.2 ± 9.5	0.996	0.989-.998
Sum 3 (Tri, SSp, Abb)	41.2 ± 10.8	41.9 ± 10.8	0.989	0.972-.995
Sum 4 (Tri, SSp, Abb, Thi)	58.6 ± 13.1	59.8 ± 13.1	0.987	0.967-.995
Sum 4 (Tri, Bi, SSc, IC)	45.8 ± 11.2	46 ± 11.3	0.996	0.990-.998
Sum 5 (CSEP SSc, Tri, Bi, IC, C)	57.9 ± 13.8	58.4 ± 13.9	0.997	0.992-.999
Sum 7 (ISAK Level 1 - SSc,Tri,Bi,Abb2,SSp,Thi,C)	86.4 ± 20.3	87.7 ± 20.9	0.995	0.988-.998
Sum 9 (Anthropometrica- SSc,Tri,Bi,Ax,IC,SSp,Abb2,Thi,C)	109.6 ± 27.1	110.9 ± 26.9	0.997	0.993-.999

[a] ICC single measure, two way mixed, consistency, [b]CI = 95% confidence interval SD; standard deviation, SSc; subscapular; Tri; triceps, Bi; biceps; Ax; axilla, IC; iliac crest, SSp; supraspinale, Abb; Abdominal 1, Abb2; Abdominal 2, Thi; thigh, C; calf. Summed skinfolds are combinations of skinfolds used in existing regression equations or professional performas.

4.2.2 Absolute Reliability

All skinfolds showed low absolute TEM values with magnitudes ranging from 0.2 to 0.8 mm across the various skinfold measurements. %TEM values showed excellent reliability and would be classified as having error values that satisfy the ISAK level 3 anthropometrist expertise for all skinfolds (<5%), with the exception of axilla, which would only satisfy the requirements for a level 1 (<7.5%) anthropometrist (Table 4.2). Similarly, all circumferences and breadths yielded %TEM values at or below 0.6% which satisfy the ISAK level 3 requirements of < 1.0%, for all "other" variables.

Table 4.3. %TEM of Individual Skinfolds

SKINFOLDS	Anthro T1	Anthro T2	%TEM
	Mean± SD	Mean± SD	3 MED
Subscapular	9.8 ± 2.5	10 ± 2.5	2.9
Tricep	14.4 ± 3.2	14.4 ± 3.3	2.4
Bicep	5.8 ± 2.0	5.7 ± 1.9	3.0
Axilla	7.5 ± 2.8	7.4 ± 2.1	6.7
Illiac	15.8 ± 4.9	15.8 ± 4.8	2.2
Supraspinale	10.3 ± 3.5	10.6 ± 3.8	3.4
Abdomen2 (5cm)	16.2 ± 5.6	16.7 ± 5.5	2.5
Abdomen1 (2cm)	16.5 ± 4.4	16.9 ± 4.6	2.0
Thigh	17.7 ± 4.7	17.9 ± 4.7	2.3
Calf	12.1 ± 3.2	12.4 ± 3.3	2.8

SD; standard deviation, Anthro T1; Trial 1 values were taken during the first anthropometric assessment; Anthro T2; Trial 2 values were taken at a second assessment that occurred within 7 days from Trial 1. Values from Trial 1 & 2 were calculated using the median of triplicate measurements. % TEM; an index of reliability used to quantify measurement error across various skinfolds.

When skinfold variables were summed in different combinations for their entry into their respective regression equations, the absolute error represented by TEM values was still very low, ranging from 0.6 to 1.4 mm across the combinations. All %TEM values were below 2% (range 0.8 to 1.7%) with sum5 (CSEP), and sum 9 (anthropometrica) skinfold combinations yielding the lowest %TEM at 0.9% and 0.8% respectively (Table 4.4).

Table 4.4: Reliability Values for Various Summed Skinfold Combinations

SKINFOLDS	TEM (mm)	95% CI for TEM (mm)	%TEM (%)
sum 3 (Tri, Abb, Thi)	1.3	2.5	1.7
sum 3 (Tri, Ssp, Thi)	1.1	2.2	1.7
sum 3 (Tri, Ssc, IC)	0.7	1.4	1.1
sum 3 (Tri, SSp, Ab)	1.1	2.2	1.8
sum 4 (Tri, SSp, Ab, Thi)	1.5	2.9	1.6
sum 4 (Tri, Bi, SSc, IC)	0.8	1.6	1.1
sum 5 (CSEP SSc, Tri, Bi, IC, C)	0.8	1.6	0.9
sum 7 (ISAK Level 1 - SSc,Tri,Bi,Ab5,SSp,Thi,C)	1.3	2.5	1.0
sum 9 (Anthropometrica-SSc,Tri,Bi,Ax,IC,SSp,Ab5,Thi,C)	1.4	2.7	0.8

TEM, Technical Error of Measurement; CI, confidence interval, % TEM, Percent Technical Error of Measurement, Summed skinfolds represent combinations of skinfolds used in existing regression equations, or combinations that are commonly utilized to assess the reliability of practitioners.

4.2.3 Smallest Real Difference

The smallest real difference (SRD) was calculated to assess the magnitude of change necessary to identify "true change" independent of measurement error in a clinical setting at the level of 95% confidence. The magnitude of SRD values ranged from 2.2 to 4.2 mm across the summed skinfold combinations (Table 4.5).

Table 4.5 Systematic Bias ± Smallest Real Difference at 95% Certainty Level.

SKINFOLDS	Systematic Bias ± SRD
SUM VARIABLES	
Sum 3 (Tri, Abb, Thi)	-0.6 ± 3.6
Sum 3 (Tri, SSp, Thi)	-0.5 ± 3.0
Sum 3 (Tri, SSc, IC)	-0.2 ± 1.9
Sum 3 (Tri, SSp, Ab)	-0.7 ± 3.0
Sum 4 (Tri, SSp, Ab, Thi)	-0.9 ± 4.2
Sum 4 (Tri, Bi, SSc, IC)	-0.2 ± 2.2
Sum 5 (CSEP SSc, Tri, Bi, IC, C)	-0.4 ± 2.2
Sum 7 (ISAK Level 1 - SSc,Tri,Bi,Ab5,SSp,Thi,C)	-1.3 ± 3.6
Sum 9 (Anthropometrica-SSc,Tri,Bi,Ax,IC,SSp,Ab5,Thi,C)	-1.3 ± 3.9

Systematic Bias is the amount by which the mean of Anthro T2 values differed from the mean of Anthro T1 values. Negative numbers denote a systematic underestimation of Anthro T2, when compared to Anthro

T1. *SRD; Smallest Real Difference* ; this values represents the necessary change in the sum of skinfolds in millimetres to signify change at 95% certainty level.

4.2.4 Bias of the Mean

Bias in the mean was calculated from the mean difference between T1 and T2. All mean differences showed a trend towards negative values indicating that T2 skinfold values were on average larger than T1 values (Table 4.4). Paired Students T-Tests indicated that 2 out of 9 skinfold combinations were significantly different using the median of three values data compilation strategy. Only the sum7 (ISAK) and sum9 (anthropometrica), showed a significant bias at the level of $p < 0.05$ (Table 4.5).

Bland Altman's plots were created to visually investigate the bias. The LOA are also displayed to show the 95% true value confidence interval (Figure 4.1). In this instance Anthro T1 serves as the criterion value. The sum5 contained the smallest mean difference (-0.43mm) represented by the solid line and had the narrowest limits of agreement at +1.89 and – 2.75 mm represented by the dashed lines (Figure 4.2).

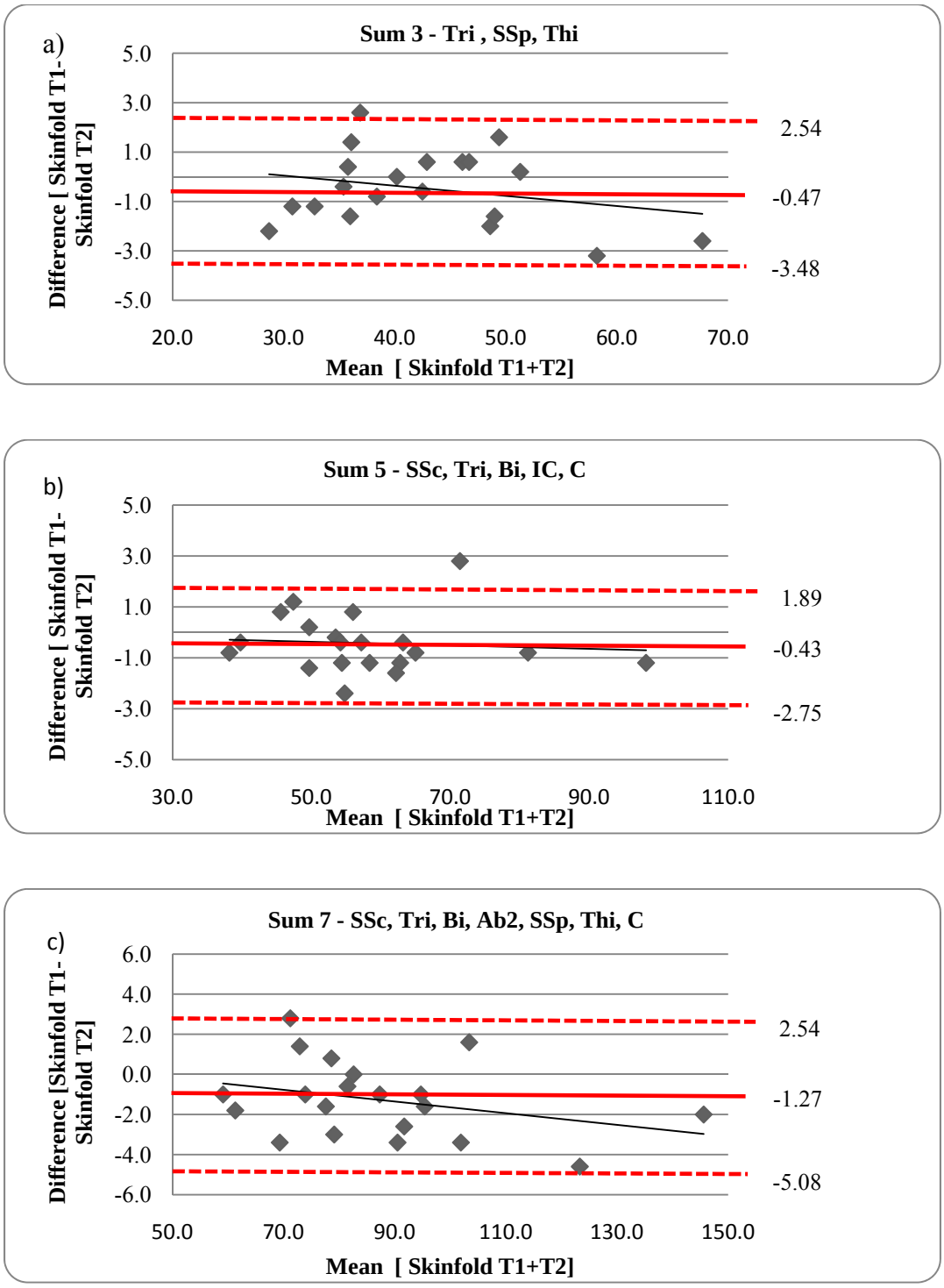


Figure 4.2: Bland Altman plots comparing summed skinfolds (mm) from AnthroT1 and AnthroT2. The dashed lines represent the upper and lower 95% LOA (+- 2 SD). The solid line represents the mean difference or bias. The trend line is a linear line that represents the relationship between the differences of AnthroT1 and AnthroT2 and the mean of AnthroT1 and AnthroT2. The shaded diamonds are individual datapoints while the solid black line is the linear relationship of the datapoints. a) *Sum 3*; skinfold combination used in regression equation by (Jackson *et al.*, 1980; Ball *et al.*, 2004), of which the former is

taught to CSEP-CEP's; b) *Sum5*; skinfolds taught in CSEP–CPT as part of the Healthy Body Composition module; c) *Sum 7* – skinfolds assessed in ISAK Level 1 accreditation.

4.2.5 Bias of the Error

Tests for heteroscedasticity were non-significant for all summed skinfold combinations when using a Pearson's Product Moment Correlation at a level of $p<0.05$. Relationships between the magnitude of error and the magnitude of the measure were not present, indicating that the data was homoscedastic, and did not need logarithmic or quadratic transformation for further analysis.

4.3 DXA RELIABILITY

Thirty Two (n=32) subjects were scanned twice to analyze the reliability of the DXA instrument (ISCD, 2007). Variables analyzed included % Body Fat (% BF DXA), Fat Total (kg), Lean Total (kg) and BMD Total (g/cm^2). All values were analysed with the second DXA scan (DXA T2) compared to the first (DXA T1).

4.3.1 Relative Reliability

All DXA values showed very high relative reliability (range 0.961-1.000). Of greatest importance to this study, % BF DXA reliability was excellent with a reliability coefficient of 0.996 (Table 4.6).

Table 4.6: Relative Reliability of DXA Measurements

Variable	DXA T1 Mean ± SD	DXA T2 Mean ± SD	ICC	95% CI for ICC
% Body Fat DXA	22.61 ± 5.77	22.80 ± 5.86	0.996	.991-.998
Tissue Total (kg)	61.891 ± 8.610	61.798 ± 8.581	1.000	1.000
Fat Total (kg)	14.918 ± 5.289	15.022 ± 5.344	0.998	.995-.999
Lean Total (kg)	46.967 ± 5.434	46.776 ± 5.416	0.998	.995-.999
BMDTotal (g/cm^3)	1.24 ± .09	1.24 ± .09	0.989	.977-.994
Total Mass (kg)	64.77 ± 8.99	64.69 ± 8.96	1.000	1.000

ICC single measure, two way mixed, consistency Intraclass Correlation Coefficient, CI = 95% confidence interval for the ICC.

4.3.2 Absolute Reliability

DXA measurements displayed a high degree of absolute reliability as the SEM value for % BF DXA was 0.38%. Other SEM values of interest in body composition studies show that absolute Fat Total (kg) and Lean Total (kg) were also very reliable as well with SEM values < 300 g (Table 4.7).

Table 4.7Absolute Reliability of DXA measurements

Variable	SEM (units)	95% CI for SEM (units)	CV (%)
%Fat Region (%)	0.380	0.7448	1.13
Tissue Total (kg)	78.930	154.703	0.09
Fat Total (g)	250.030	490.059	1.11
Lean Total (g)	260.00	509.60	0.38
BMD Total (g/cm^2)	0.010	0.0196	0.50
Total Mass (kg)	0.090	0.1764	0.09

SEM; Standard Error of Measurement, *CI;* Confidence Interval, *CV;* Coefficient of Variation, DXA variables and their absolute reliability statistics

Reliability of the DXA indicated by the CV (%) shows that measurements of Fat Total, Lean Total, and BMD Total are all highly reliable with CV's less than 1.5%.

4.3.3 Least Significant Change

The LSC was calculated to assess the magnitude of change necessary to identify "true change" independent of measurement error in a clinical setting at the level of 95% confidence. The primary independent variable under investigation (% BF DXA) requires a change in magnitude of > 1.05% for the change to be considered "true" at the 95% confidence level (Table 4.8). Absolute changes in the mass of lean and fat components need to be in the magnitude of nearly ¾ of a kg, or 1.65lbs to be identified as true changes with 95% confidence.

Table 4.8: Systematic Bias of DXA Variables ± Least Significant Change at 95% Certainty Level

Variables	Systematic Bias ± LSC (units)
% BF DXA	0.19 ± 1.05
FatTotal (g)	103.63 ± 693.05
LeanTotal (g)	-190.84 ± 720.68
BMDTotal (g/cm^2)	0.00 ± 0.03
Total Mass (kg)	-0.08 ± 0.25

Systematic Bias is the amount by which the mean of DXA T1 values differed from the mean of DXA T2 values. Negative numbers denote an underestimation of DXA T1, when compared to DXA T2. Positive numbers denote an overestimation of DXA T1 when compared to DXA T2. *LSC,* Least Significant Change; this value represents the necessary change in the DXA parameter under investigation in units specified to signify change at 95% certainty level.

4.3.4 Bias of the Mean

Bias in the mean was calculated from the mean difference between T2 and T1. Only Lean total (kg) showed statistically significant (p=0.023) systematic bias with lean mass being higher at T1 vs. T2. No other values were significant although BMD, and % BF DXA were approaching significance (Table 4.9).

Table 4.9: Two Tailed Repeat Measures T-Test of DXA Variables

Variable	Mean DXA T2 ± SD	Mean DXA T1 ± SD	Sig. (2-tailed)
%Body Fat DXA	22.8 ± 5.9	22.6 ± 5.8	.065
Fat Total	15.022 ± 5.344	14.918 ± 5.289	.124
Lean Total *	46.776 ± 5.416	46.967 ± 5.434	.023
BMD Total	1.24 ± .09	1.24 ±.09	.056

* Paired T-Test p<.05 *Sig. (2-tailed)*; p-value of a two tailed paired t-test DXA variables represented by their mean and standard deviation.

Bland Altman's plots were created to visually investigate the bias for the DXA data. The LOA are also displayed to show the 95% true value confidence interval of the measurement Figure 4.3.

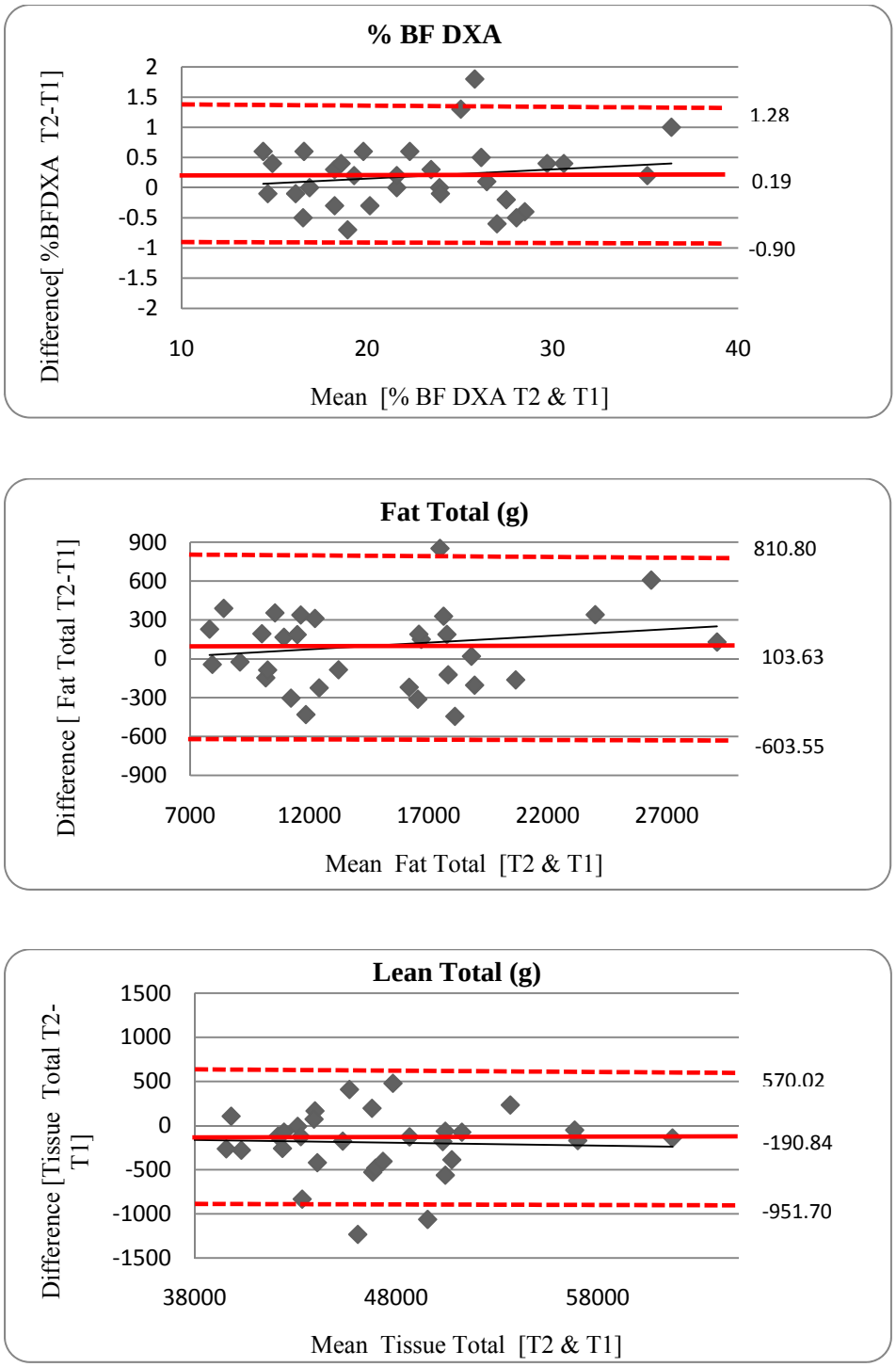


Figure 4.3: Bland Altman plots comparing DXA T2 to DXA T1 for a) %Body Fat DXA, b) Fat Total (g), c) Lean Total (g). The dotted lines represent the upper and lower 95% LOA (± 2 SD). The solid line represents the mean difference or bias. The trendline is a linear line that represents the relationship between the differences of DXA T2 to DXA T1.

4.3.5 Bias of the Error

Heteroscedasticity was assessed by plotting the absolute differences against the mean values and calculating the Pearsons correlation coefficient between these variables. No significant relationships existed between the magnitude of error and the measurement for the primary variables investigated.

4.4 VALIDITY

Ninety-five (n=95) subjects were assessed by DXA and anthropometric methods on the same day. Concurrent validity was investigated between the criterion and predicted % BF which were calculated from existing regression equations that utilized MC (n=3), DXA (n=9), and HD (n=8) for their creation. Group statistics are presented as the mean and standard deviation.

4.4.1Relative Agreement

A validity coefficient was calculated between each regression equation's predictive value and the DXA criterion using a Pearsons Product Moment Correlation (Table 4.9). In general, all equations displayed moderate to high validity coefficients with MC criterion equations ranging from 0.734-0.839, DXA criterion 0.727-0.874, and HD criterion equations 0.801-0.882. The equations displaying the highest validity coefficient are the DXA derived equation by Ball 2004 (r = 0.874) and the HD derived Jackson & Pollock #1 (r = 0.882) (Table 4.10).

Table 4.10: Validity Parameters of Multiple Regression Compared to DXA criterion

					Agreement	
Method	**Mean ± SD**	**Pearsons (r)**	**SEE**	**TE**	**CE ± SD**	**LL - UL**
% BF DXA (criterion)	**24.2 ± 5.7**					
MC						
Peterson 2003	26.25 ± 3.7	0.734	2.6	4.4	2.09 ± 7.7	-5.6 to 9.8
Evans 2005	21.54 ± 3.0	0.839	1.9	4.4	-2.63 ± 7.1	-9.7 to 4.5
Moon 2009	20.33 ± 3.0	0.839	1.9	5.2	-3.80 ± 7.1	-10.9 to 3.2
DXA						
Ball 2004	25.03 ± 4.4	0.874	1.4	2.9	0.87 ± 5.6	-4.7 to 6.5
Warner 2004	20.56 ± 4.1	0.836	1.0	4.0	-3.60 ± 3.6	-7.2 to 0.0
Kagawa 2007 - 8 variables	28.61 ± 5.0	0.879	1.3	5.2	4.45 ± 5.4	-0.9 to 9.8
Kagawa 2007 - sum8	22.67 ± 3.6	0.855	1.6	3.5	-1.50 ± 6.4	-7.9 to 4.9
Kagawa 2007 - 3 variables	26.39 ± 4.7	0.804	2.0	4.0	2.23 ± 6.7	-4.5 to 9.0
Kagawa 2007 - sum3	25.28 ± 3.9	0.796	2.1	3.6	1.11 ± 6.9	-5.8 to 8.1
Jackson 2008	21.65 ± 3.6	0.850	1.7	4.1	-2.52 ± 6.5	-9.0 to 3.9
Duz 2009	25.83 ± 4.7	0.839	1.7	3.5	1.67 ± 6.1	-4.5 to 7.8
Kohli 2009	21.74 ± 5.4	0.727	2.8	4.7	-2.42 ± 8.2	-10.6 to 5.7
HD						
Sloan Burt Blyth (1962)*	21.69 ± 3.3	0.801	2.2	4.3	-2.47 ± 7.2	-9.7 to 4.7
Durnin Womersley (1974) 16-68*	27.46 ± 3.8	0.824	1.9	4.7	3.30 ± 6.6	-3.3 to 9.9
Durnin Womersley (1974) 20-29	25.99 ± 3.8	0.824	1.9	3.8	1.83 ± 6.6	-4.8 to 8.4
Jackson Pollock (1980)*	19.76 ± 4.4	0.848	1.6	5.3	-4.40 ± 6.0	-10.4 to 1.6
Jackson Pollock (1985) (no thigh)*	21.28 ± 4.2	0.832	1.8	4.3	-2.88 ± 6.4	-9.2 to 3.5
Jackson Pollock (1980) #1	22.34 ± 4.8	0.882	1.3	3.2	-1.83 ± 5.3	-7.2 to 3.5
Jackson Pollock (1980) #2	19.47 ± 4.1	0.862	1.5	5.5	- 4.69 ± 5.9	-10.6 to 1.2
Thorland (1984)	20.56 ± 5.1	0.800	2.1	5.0	-3.61 ± 6.9	-10.5 to 3.3

* regressions suggested by CSEP-CEP resource manual, *r* = pearsons correlation coefficient, *SEE* = standard error of estimate, TE = Total Error, CE = Constant Error or Bias, Lower and Upper Limits = CE ± 2SD. The validity statistics are presented for various existing *MC* = Multicompartment, *DXA*, and *HD* = Hydrodensitometry equations.

4.4.2. Absolute Agreement

Paired sample, two tailed Students T-Tests were performed between the criterion and predicted values. All predicted % BF were significantly different from the DXA criterion at a p-value of < 0.05 (Table 4.11).

Table 4.11 Paired T-Test of DXA vs. Regression

Regression equation	Mean ± SD	Pearsons (r)
% BF DXA	**24.16 ± 5.65**	
MC		
Peterson2003	26.25 ± 3.73†	0.735
Evans2005	21.54 ± 2.98†	0.839
Moon2009	20.32 ± 2.98†	0.840
DXA		
Ball2004	25.03 ± 4.39†	0.874
Warner2004	16.96 ± 2.82†	0.836
Jackson2009	21.65 ± 3.55†	0.851
Duz2009	25.83 ± 4.70†	0.839
Kagawa20078 variables	28.61 ± 5.02†	0.879
Kagawa2007sum8	22.67 ± 3.60†	0.855
Kagawa2007three variables	26.39 ± 4.68†	0.803
Kagawa2007sum3	25.28 ± 3.89†	0.796
Kohli2009	21.74 ± 5.37†	0.727
HD		
Jackson Pollock (1980) #1	22.34 ± 4.81†	0.882
Jackson Pollock (1980) #2	19.47 ± 4.09†	0.862
Jackson Pollock (1980)*	19.76 ± 4.41†	0.848
Jackson Pollock (1985) (no thigh)*	21.28 ± 4.17†	0.832
Thorland (1984)	20.56 ± 5.06†	0.800
Sloan Burt Blyth (1962)*	21.69 ± 3.30†	0.801
Durnin Womersley (1974) 16-68*	27.47 ± 3.84†	0.824
Durnin Womersley (1974) 20-29	26.00 ± 3.82†	0.824

* regressions suggested by CSEP-CEP resource manual †p<.05 significant difference from % BF DXA, *r* = pearsons correlation coefficient between the regression equation and the DXA criterion, *MC* = Multicompartment, *DXA*, and *HD* = Hydrodensitometry equations.

Constant error (CE) was calculated as the mean difference between the methods. The lowest CE and range for each method of prediction equation were 2.09 (-3.80 to 2.09 % BF) MC equations, 0.87% BF (-3.60 to 4.45) DXA, and -1.83 and +1.83% BF (-4.69 to 3.30) for HD equations. The Ball2004 equation displayed the lowest CE out of all the regression equations tested at 0.87% BF (Table 4.11).

The Standard Error of Estimate (SEE) was calculated as a measure of the average error of the predictive equation to predict the criterion score of an individual. The ranges for the SEE for each method were 1.9-2.6 MC, 1.0-2.8 DXA, and 1.3-2.1 HD. The Warner (2004) DXA equation displayed the lowest SEE at 1.0% BF.

Total Error (TE) was calculated to represent the variance of the difference between predictor and criterion. TE was highest in non DXA equations. HD equations yielded a range of TE from 3.2 – 5.5, while MC equations had similar amounts of total error ranging from 4.4-5.2. DXA equations showed a wide range of TE values (2.9-5.2) but had the lowest overall TE value of 2.9% BF in the equation by Ball2004.

4.4.3 Agreement Within An Individual

Bland Altman plots were created to visually inspect the agreement between the two methods. LOA were plotted allowing interpretation of how closely individual % BF could be predicted when compared to % BF DXA criterion (Figure 4.4).

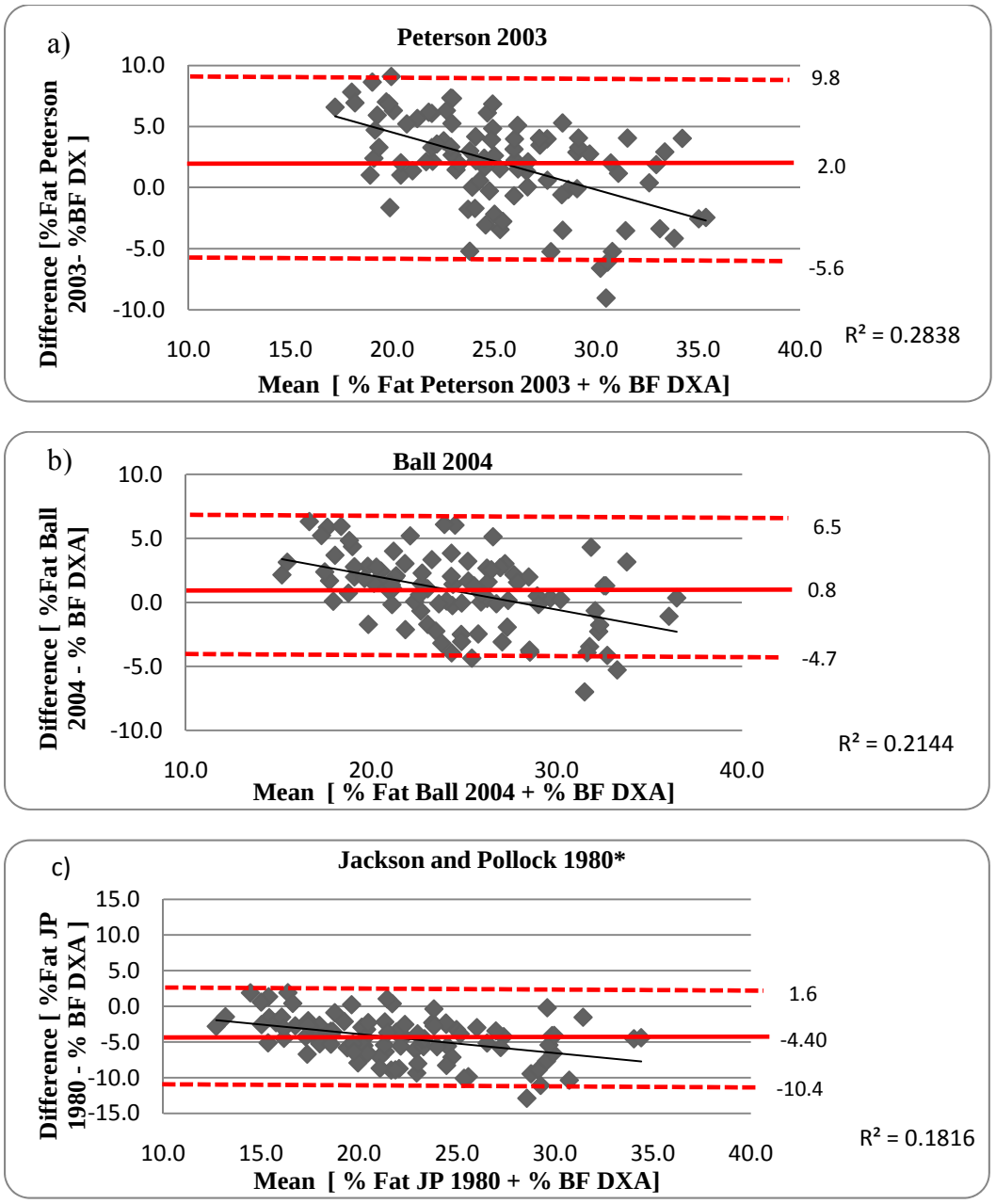


Figure 4.4: Bland Altman Plots of The Best Regression Equations from MC, DXA and HD criteria respectively. Dashed lines represent the upper and lower 95% LOA (+- 2 SD). The solid red line represents the mean difference or bias. The thin black line is the trend line and is a linear line that represents the relationship between the data of the X and Y axis. a) The Peterson 2003 MC equation, b) the Ball 2004 DXA equation, c) The Jackson and Pollock 1980 HD equation. * Equation utilised by CSEP-CEP

4.5 New Regression Equations using DXA as the criterion

New regression equations were developed using both stepwise and hierarchical methods to identify and create regression equations. Forward and backward stepwise linear regressions were conducted with % BF DXA as the response variable (RV), and all skinfolds (SSc, Tri, Bi, Ax, IC, SSp, Ab, Ab2, Thi, C), in addition to age, height, weight, BMI, and gluteal circumference were input as candidate variables (CV). Forward and backward stepwise regression identified 4 variables as significant predictors of % BF DXA. The variables included gluteal circumference, age, and SSp, and Thi skinfolds (Table 4.4). These variables were then input into multiple linear regression to create a predictive equation (Appendix J). The statistical summary data of the forward and backward stepwise linear regressions can be seen in Table 4.12 & 4.13 below.

Table.4.12: Forward Stepwise Regression Summary Output

Variables in Model					
Group	**Coef.**	**Std. Coeff.**	**Std. Error**	**F-to-Remove**	**P**
Constant	-17.277		4.9500		
SSp	0.711	0.487	0.0736	93.350	<0.001
Thi	0.258	0.234	0.0565	20.918	<0.001
Glu	0.340	0.349	0.0499	46.433	<0.001
Age	-0.230	-0.137	0.0767	9.023	0.003

Summary Table					
Step #	**Vars. Entered**	**R**	**R^2**	**Delta R^2**	**Vars in Model**
1	SSp	0.78	0.608	0.608	1
2	Glut	0.88	0.774	0.165	2
3	Thi	0.905	0.819	0.0452	3
4	Age	0.914	0.836	0.0165	4

The forward stepwise regression procedure indicates that only four variables significantly contribute to the prediction of the response variable % BF DXA. The variables include SSp, supraspinale; Glut, gluteal circumference; Thi, thigh; and age. *R* is the correlation coefficient for the regression equation, while R^2 is the coefficient of determination representing the cumulative amount of the response variable explained by

the predictive variables. *Delta* R^2 is the change in the coefficient of determination, allowing the identification of how much each individual variable added to the predictive ability of the regression. In this instance SSp had the greatest ability to predict % BF DXA followed by Glut.

Table 4.13 Backward Stepwise Regression Summary Output

Summary Table					
Step #	**Vars. Removed**	**R**	**R^2**	**Delta R^2**	**Vars in Model**
1	Bi	0.918	0.843	0.843	14
2	Ab	0.918	0.843	-0.00021	13
3	C	0.918	0.842	-0.0005	12
4	Ab2	0.917	0.842	-0.00066	11
5	IC	0.917	0.841	-0.00065	10
6	SSc	0.917	0.84	-0.00066	9
7	Scale Weight (kg)	0.916	0.839	-0.00078	8
8	Height (cm)	0.916	0.839	-0.00016	7
9	BMI	0.916	0.839	-2.9E-05	6
10	Tri	0.916	0.839	-0.00062	5
11	Ax	0.914	0.836	-0.00305	4

The backward stepwise regression procedure indicates that 11 variables do not significantly contribute to the prediction of the response variable % BF DXA. The variables with low predictive ability include Bi, biceps; Ab, abdomen; C, calf; Ab2; abdomen 2; IC; iliac crest; SSc; subscapular; Scale Weight; Height; BMI; body mass index; Tri; triceps and Ax; axilla. The remaining variables SSp; supraspinale; Glut; gluteal circumference; Thi; thigh; and age are the only significant contributors to the creation of the regression equation. *R* is the correlation coefficient for the regression equation, while R^2 is the coefficient of determination representing the cumulative amount of the response variable explained by the predictive variables. *Delta* R^2 is the change in the coefficient of determination, allowing the identification of how much each individual variable added to the predictive ability of the regression. In this instance SSp had the greatest ability to predict % BF DXA followed by Glut.

To further investigate potential candidate variables, a best subset analysis was performed to describe combinations of candidate variables that could predict the response variable % BF DXA. The Best Subsets analysis is summarized in Table 4.14.

Table 4.14: Best Subsets Analysis indicating combinations of variables and their predictive ability of the response variable, % BF DXA.

Mod	Cp	R^2	Adj R^2	MSerr	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	105.985	0.608	0.604	12.619									*						
2	24.755	0.774	0.769	7.366									*					*	
3	4.018	0.819	0.813	5.959									*			*		*	
4*	-2.275	0.836	0.828	5.476									*			*		*	*
5*	-1.809	0.839	0.830	5.435							*		*			*		*	*
6	0.124	0.839	0.828	5.475							*		*			*	*	*	*

Legend: A; Height, B; Scale Weight C; BMI, D; SSc, E; Tri, F; Bi, G; Ax, H; IC, I;, SSp, J; Ab, K; Ab2, L; Thi, M; Calf, N; Glut, O; Age.

Cp; mallows statistic, represents if the appropriate number of values were entered in the equation when these values are minimal. Minimal mallows statistic will indicate maximal R^2 with minimal RMSE and therefore yield subsets with the lowest degree of multi collinearity. R^2 is the *coefficient of determination* representing the cumulative amount of the response variable explained by the predictive variables. Adj R2 is the adjusted coefficient of determination, allowing the identification of the most parsimonious subset of variables. Where the highest number indicates the optimal number of variables. As this number begins to fall it indicates that additional variables added contribute little to the cumulative predictive ability of the equation thus, how much each individual variable added to the predictive ability of the regression. *MSerr*; is the *Error Mean Square* and is an estimate of the variability in the underlying population, computed from the random component of the observations. The lower the MSerr value, the greatest predictive ability as there is less residual error.*Subsets entered into multiple linear regression

The variables selected in model #5 resulted in the best combination for their relationship with the response variable. This was signified by a high R value as well as the highest adjusted R-squared value, indicating that 83.9% of the variance in the response variable was explained by the combination of predictor variables. The Cp or mallows statistic was also low indicating that the appropriate number of variables were in the equations (no unnecessary variables were added) which was also confirmed by the greatest adjusted R^2 value indicating parsimony. The variables in model #5 included gluteal circumference, age, and axilla, supraspinale, and thigh skinfolds. These variables were input into multiple linear regression to create a predictive equation (Appendix K).

For the hierarchical approach, skinfolds containing the five skinfold measurement sites endorsed by CSEP-CPT certification as part of the Healthy Body Composition Module (SSc, Tri, Bi, IC, C), were selected for entry into multiple linear and polynomial regression equations in various combinations. A correlation matrix was created to investigate the relationship between the predictor variables (SKF) and the response variables (% BF DXA) to investigate ways to reduce the degree of multicollinearity (Table 4.15).

Table 4.15: Correlation Matrix of Skinfolds with % BF DXA

	SSc	**Tri**	**Bi**	**Ax**	**IC**	**SSp**	**Ab**	**Ab2**	**Thi**	**C**	**% BFDXA**
SSc		0.568	0.651	0.832	0.771	0.796	0.694	0.743	0.339	0.514	*0.629
Tri			0.765	0.701	0.709	0.698	0.627	0.621	0.624	0.670	*0.741
Bi				0.750	0.725	0.770	0.681	0.701	0.539	0.620	*0.704
Ax					0.862	0.862	0.762	0.822	0.400	0.583	0.700
IC						0.889	0.862	0.872	0.436	0.573	*0.744
SSp							0.858	0.898	0.418	0.644	0.780
Ab								0.937	0.363	0.495	0.703
Ab2									0.316	0.535	0.718
Thi										0.569	0.645
C											*0.679
% BFDXA											

Correlation matrix showing the relationship between various skinfolds as well and % BF DXA. *Skinfolds included in CSEP-CPT performa.

All skinfolds from the sum of 5 displayed moderate correlations (range 0.629-0.744) with the response variable. The matrix indicated instances where strong mulitcollinearity existed between some skinfolds. A correlation matrix exclusively of skinfolds from the sum5 (CSEP-CPT) was created to assess multicollinearity within the subgroup of skinfolds. The majority of the variables showed modest correlation, except for Tri and Bi which showed moderate to strong correlation (0.765) with each other. The

IC showed moderate to strong correlations with all skinfolds (0.709 - 0.771) except calf (0.573) indicating moderate levels of multicollinearity (Table 4.16).

Table 4.16: Correlation Matrix of only CSEP-CPT skinfolds used as a sum5 for regression creation.

	SSc	Tri	Bi	IC	C
SSc		0.568	0.651	0.771	0.514
Tri			0.765	0.709	0.67
Bi				0.725	0.62
IC					0.573
C					

In the anticipation of multicollinearity being flagged during the regression process, the sum of the five individual CSEP skinfolds was selected for entry into the regression equations. A test for normality was performed on the sum5. Upon failing the test, the sum5 was logarithmically and quadratically transformed and retested. The normality test results are listed below:

sum5: W-Statistic = 0.953 P = 0.002 Failed
sum5sqr: W-Statistic = 0.873 P < 0.001 Failed
log10: W-Statistic = 0.988 P = 0.556 Passed

To investigate the relationship between the sum5 and % BF DXA, a scatter plot was created. Fitting trendlines through the data it was found that both a second order polynomial and a logarithmic trend line had greater fit than the linear trendline (Figure 4.5). In this regard, the sum5 was also investigated using 2nd order polynomial regression analysis to create a quadratic equation.

To summarize, variables entered into the regression equations included the individual CSEP skinfold sites, the sum5, logarithmically transformed sum5, quadratically transformed sum5, and a polynomial relationship of the sum5 was explored.

The derived method along with the predictive equations are shown in the figure below (Figure 4.5).

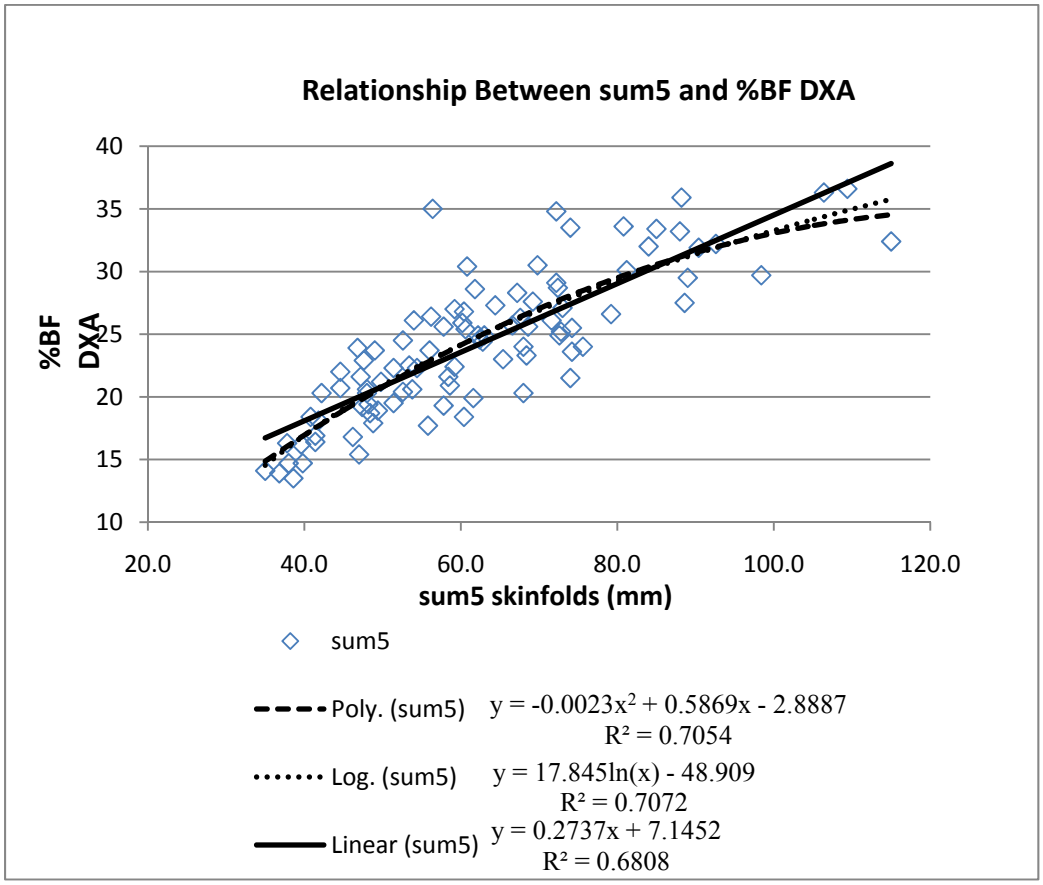


Figure 4.5 Relationship Between sum5 and % BF DXA*; Dashed Line*; represents a 2nd order polynomial relationship of sum5 skinfold data to % BF DXA data. This is also known as a quadratic relationship. *Dotted Line*; represents a logarithmic relationship of sum5 skinfold data to % BF DXA data. *Solid Line*; represents a linear relationship of sum5 skinfold data to % BF DXA data. The equations representing their respective trendline are displayed in addition to the R^2 values which indicate how well the trendline represents the data.

Out of the 10 equations created, only three equations passed the selected assumption tests necessary for confidence in the validity of the statistical output. Appendix L shows the Analysis of Assumption Checking and Outlier Identification for the various new regression equations indicating their validity shortcomings. Outlier

diagnostics were also employed, indicating the datasets that contain values that will have impact on the regression accuracy.

Validation of the three equations that passed all assumption tests were conducted by reintroducing the equations into the dataset and calculating the validity statistics. As these statistics only represent fit statistics in the current sample, cross validation via the PRESS method was utilized to anticipate the stability of the regression to external populations (Table 4.17).

Table 4.17: New Regression Equations with Cross Validation Statistics.

		Fit Statistics					Agreement	
Equation	**Mean ± SD**	**R**	**R^2**	**Adj R^2**	**SEE**	**TE**	**CE ± 2SD**	**LL to UL**
Ball 2004	**25.03 ± 4.39**	**0.874**	**0.764**		**1.4**	**2.9**	**0.87 ± 5.6**	**-4.7 to 6.5**
dh#9 polynomial sum5	24.18 ± 4.75	0.840	0.706	0.699	1.7	3.0	0.02 ± 6.1	-6.1 to 6.1
dh#10 - quadratic sum5 & age	24.18 ± 4.56	0.806	0.650	0.642	2.0	3.3	0.01± 6.7	-6.7 to 6.7
dh#8 - quadratic sum5	24.15 ± 4.45	0.788	0.621	0.617	2.1	3.5	-0.01 ± 6.9	-7.0 to 6.9

dh#9: polynomial regression equation created using sum5 skinfolds, *dh#10;* linear regression equation created using quadratically transformed sum5 skinfolds and age, *dh#8*; linear regression equation created using quadratically transformed sum5 skinfolds. *sum5*; SSc; subscapular, Tri; triceps, Bi; biceps, IC; iliac crest, C; calf. SD; standard deviation, R; correlation coefficient, R2; coefficient of determination, Adj R2; adjusted coefficient of determination, SEE; standard error of estimate, TE; total error, CE; constant error, LL; lower limit of Limit of Agreement, UL; upper limit of Limit of Agreement.

The dh#9 equation that was created with a polynomial model of the second order (quadratic equation) displayed the greatest validity. When applied to the study sample the SEE of 1.7 and TE of 3.0% BF are classified as very good (Heyward, 1996). The CE was eliminated giving symmetrical limits of agreement at ± 6.1% BF. The cross validation's

PRESS statistics verify that the equation has generalizable integrity as the PRESS residuals were lowest in equation #9, the TE was still considered good, and the inflation of the TE was only 1.14% (Table 4.18).

Table 4.18: Cross validation of new regression equations using Predicted Residual Sum of Squares (PRESS) statistics

		Cross Validation		
Equation	**Mean ± SD**	**PRESS Residual**	**PRESS TE**	**% Change in TE**
Ball 2004	**25.03 ± 4.39**			
dh#9 - polynomial sum 5 *	24.18 ± 4.75	932.594	3.13	-1.14
dh#10 - quadratic sum 5 & age*	24.18 ± 4.56	1137.242	3.46	-2.46
dh#8 - quadratic sum 5 *	24.15 ± 4.45	1207.698	3.57	-2.07

SD; standard deviation, PRESS residual, the residual error of the regression line, lower values indicate greater fit of regression line to data, PRESS TE, is the PRESS total error, %Change in TE; represents the change in total error from the TE values of the study vs. the TE values from the PRESS method. A low value indicates good generalization of the equation as the error is thought to be stable when applied to external populations.

CHAPTER 5- DISCUSSION

The practice of body composition assessment is widely employed across a spectrum of fields, utilizing numerous methods. One of the most common assessment tools is the use of population specific regression equations to predict % BF based on anthropometric measures. Methodological differences owing to equipment variation, sample selection and statistical methods for regression equation construction will all influence the ability of existing regression equations to accurately predict the body composition of elite female athletes and necessitate the cross validation of these methods in populations with characteristics that may differ from the original study (Schmidt & Carter, 1990). Few, if any studies, have published skinfold regression equations for predicting % BF in elite female athletes, utilizing measurements obtained with Harpenden skinfold callipers and a GE Lunar Prodigy Fan Beam DXA as the criterion reference. Our study assessed the reliability of both anthropometric and DXA methods, outlining the implications that variations in reliability will have on applied practice. Pertinent existing regression equations from the literature were selected for cross-validation using our subject population and the GE Lunar Prodigy DXA as our criterion measure. A new regression equation was created using skinfold sites commonly taught as part of the Canadian Society of Exercise Physiology (CSEP-CPT) module for Healthy Body Composition, Canada's professional certifying body for physiologists. The new equation showed equivalent levels of validity to the best existing equation currently in the literature.

5.1 Reliability

5.1.1. Anthropometry

Anthropometric reliability was excellent as ICC's values were 0.984 or greater. The values are in line with reliability coefficients found in the literature for sum7 (0.980) (Ball *et al.*, 2004) and individual sites (0.950-0.990)(Warner *et al.*, 2004; Moon *et al.*, 2009b). However, many recent publications fail to state the reliability values of their anthropometrist or methods for establishing their anthropometric reliability despite publishing regression equations(Peterson *et al.*, 2003; Evans *et al.*, 2005; Duz 2009). In lieu of this, it is suggested that regression equations utilized from these studies should be interpreted with caution as insufficient levels of reliability negate validity.

As the ICC is unitless, it can only infer relative reliability, absolute indices of reliability are more clinically relevant when the goal is to interpret change between serial measures, or the likelihood of a true value of a single measure. The absolute reliability of this study's anthropometrist exceeded the expectations of acceptable reliability outlined in the accreditation scheme of the International Society for the Advancement of Kinanthropometry (ISAK). Levels of reliability commensurate with a Level 3 accreditation were evident among all individual skinfold sites as %TEM values were <5% (range 2-3.4%) with the exception of the axilla measurement which had a %TEM value of 6.7%. The reliability is very similar to other publications by professional anthropometrists practicing on elite athlete populations, as %TEM values for similar individual sites have been recorded, ranging from 1.5%-3.1% on 262 Canadian national team athletes(Scholz, 2008). Although the axilla measurement showed the greatest degree

of variability in our study, the %TEM values were still within the limits of the Level 1 ISAK accreditation. All measured circumferences were also reliable enough (<1.5% (range .3-.6%) to satisfy requirements for level 3 accreditation. The high level of reliability of the antrhopometrist in the study is likely due to experience (6 yrs or > 1000 assessments) in conducting the anthropometric performa, as it has been suggested that a minimum of 50-100 subjects should be tested by anthropometrists to develop a sufficient level of skill and proficiency (Heyward, 1996).

5.1.2. DXA

As anticipated, the criterion measure showed very high levels of reliability. ICC`s for all values were greater than 0.960. DXA values for % BF had an ICC value of 0.996, while Fat Total and Lean Total showed equal levels of reliability with ICC values of 0.998 each. Measures of total composition in general displayed greater reliability (range 0.989-1.000) than regional measures of composition (range 0.961-0.991). This is in line with the literature and is likely a result of error due to the extrapolation of tissue composition over areas of large bone as seen on the appendages and trunk (Lohman *et al.*, 2009).

The absolute errors of the compositional measurements were of similar magnitude as reported in the literature. Prodigy reliability was tested on a group of 46 males and females three times with repositioning. The study revealed CV's values for BMD (0.8%), BMC (1.0%), Fat Mass (1.3%), Lean mass (0.9%) and total mass (0.3%) that were all slightly greater than our values, despite a subject pool that was similar in absolute mass (approx 71 kg , but had higher relative proportions of fat than our group (38.9%) (Nord *et al.*, 2000). Although many of the studies utilizing DXA as a criterion or as component of

a MC model do so without drawing reference to the precision of the unit used. In an investigation of *in vivo* and *in vitro* precision of GE Lunar pencil (DPXL) and fan beam (Prodigy) DXA machines, it was found that the fan beam displayed greater short term precision than its Pencil beam counterpart. SEM values and %CV (in brackets) were calculated for 10 subjects scanned twice on the same day with repositioning. CV's of 0.6 % BF (2.7%) ,72g tissue total (0.1%), 409g fat total (2.5%), and 407g lean total (0.8%) were superior in the fan beam articulating differences in reliability within manufacturer (Oldroyd *et al.*, 2003). Citation of the exact system used is a critical step when utilizing DXA technology. The SEM and CV values will be unique to each DXA system as well as the population upon which these values are calculated, dictating the ability to interpret the data in any relevant matter (Guo *et al.*, 2004; ISCD, 2007).

As with the anthropometric data, relative reliability statistics reflect only a small component of the usefulness of reliability measurements to applied practitioners. The use of the Least Significant Change is common terminology in clinical densitometry and is used to assess true change in the density kinetics of bone (Shepherd & Lu, 2007). For true change in % BF to be detected at the 95% certainty level with our DXA, a change in composition must exceed our system's LSC value of 1.05% BF. Alternatively, because testing for heteroscedasticity revealed no relationship between the error of the measure and the magnitude of the measure, allowing for the use of absolute LSC values, the absolute LSC threshold values of 693grams of fat or 720 grams of lean mass need to be surpassed for the 95% confidence of detecting change.

Although the former data suggests that given DXA's high degree of reliability there appears to be minimal need to continually invest in anthropometric measures, in

applied settings anthropometrics will always be utilized and has several practical advantages over DXA. One of the most obvious benefits of anthropometrics is accessibility. Anthropometric measuring kits are inexpensive (approximately $1000 for a full kit), and are portable. Whereas DXA machines are very expensive (approximately $60,000 to $200,000 and need their own housing facility). The cost and portability, will limit accessibility for testing. This is perhaps one of the most important elements for consistently monitoring elite athletes as a large number of athletes spend a great deal of time travelling and / or training in various parts of the world. For this reason, anthropometrics is practically superior to the DXA as it does not require athletes to centralize around a DXA housing facility. Rather, a trained anthropometrist can travel with the athlete or team and take measurements throughout an intensive travelling or competitive phase without requiring the athlete to seek out the limited DXA housing facilities. Taken a step further, anthropometrics can also be utilized to track compositional changes in remote areas such as at altitude camps and or in 2nd and 3rd world countries where resources are sparse.

5.2 Validity

The capacity of a regression equation to predict % BF is indicated by a series of statistics that can be described as “fit statistics” and include the validity coefficient R, the Standard Error of the Estimate (SEE), the Total Error (TE), the Constant Error (CE) or bias, and in some studies, citations of the Limits of Agreement (LOA). These statistics indicate how closely the predictive values agree with the measured values of the criterion. As the correlation indicates the association, not the accuracy of the measures, equation accuracy is better defined by interpreting the level of agreement. Among these statistics

the TE is used as the primary indicator of an equation's efficacy. Briefly, the TE represents the average deviation of individual scores from the line of identity, whereas, the SEE only represents the degree of deviation of individuals scores from the regression line without accounting for the influences of CE (systematic error).

The regression equation by Ball *et al.*, 2004, displayed the greatest accuracy in predicting % BF of female athletes in our study when compared to a DXA criterion. A validity coefficient R= 0.874, was accompanied by the lowest CE of out of all 20 regression equations evaluated. The CE showed a <1% BF bias for overestimation in our study. A low SEE (1.4% BF) and the lowest magnitude of TE (2.9% BF) are qualitatively considered "very good" to "excellent" for predictive equations (Heyward, 1996). The validity statistics in our study confirm the accuracy of the equation, further supporting the original cross validation conducted by the researchers using both PRESS and sample splitting methods. The accuracy of the equation is hypothesized to be largely a function of its development as a generalized equation as the subjects in our study tended to be younger (28.7 vs. 22.9 years), taller (172.0 vs. 165.1cm), heavier (67.8 vs. 63.3kg) and leaner (24.16 vs. 27.9 % BF DXA) than the sample of normal healthy females used in the creation of the equation. It is conceivable that the accuracy is also due in large part to employing sound theoretical methods of regression development originally proposed by Jackson& Pollock (Jackson & Pollock, 1976; Jackson *et al.*, 1980; Jackson & Pollock, 1982; Jackson, 1984). This notion is further supported by the low SEE (1.3-1.8% BF) obtained in our study when the original Jackson and Pollock HD equations were used to predict % BF DXA. However, the Jackson Pollock equations recommended by CSEP and ACSM did show a large tendency to underestimate % BF when compared to DXA as

the CE was -4.40 and -2.88 % BF for the thigh and no thigh equations respectively, with ranges of CE`s from -1.83 to -4.68 across the four different Jackson & Pollock equations investigated. Magnitudes of % BF underestimations in the HD equations have been confirmed in other research as being 3-4% BF (Ball *et al.*, 2004), compared to a Hologic QDR 4500A fan beam DXA system, 7.6% BF (Duz 2009) compared to a Lunar DPX DXA system and 6.6 % BF (Peterson *et al.*, 2003) when compared to 4C criterions.

As part of a national study called TIGER, Jackson et al (2008) revisited the validity of their seminal study on body composition prediction. Using a sample of mixed ethnicity, 706 females aged 17-35years were tested using the Jackson and Pollock (1980) sum3 quadratic equation with age. The authors found the old regression yielded a validity coefficient of R=0.84, SEE 3.9 % BF when compared to a Hologic DXA criterion (Jackson *et al.*, 2008). Once again, this is in line with our findings that indicate nearly equal validity coefficients (r =0.848 vs. 0.84) but our Total error was greater at 5.3% vs. 3.9% BF, likely due to our greater systematic bias of – 4.4 vs. -1.05. This is potentially due to the fact that not only does the Jackson & Pollock equation underestimate % BF in relation to DXA, the Hologic DXA used in the study tends to underestimate body fatness when compared to the Lunar Prodigy DXA that was used in our study by 1.2-1.4 kg in 39 obese females, and 2.24 kg in 21 healthy males and females respectively (Genton *et al.*, 2002; Aasen *et al.*, 2006). This magnitude of underestimation by DXA alone would constitute the systematic underestimation of the equation by 1.7-3.3% BF in our population with a mean mass of 67.8 kg. Jackson et al (2008), concluded that the generalized equations need to be updated as the body composition characteristics of the population evolve. The group created a new equation that accounted for race, which

showed fit statistics of R=0.88 and TE= 3.4 but did not show greater predictive ability than the Ball2004 equation. However, contradicting results are still noted in the literature stating negligible bias of the Jackson and Pollock (1980) equation (r=0.84 SEE = 2.56, TEE = 2.65, and CE = 0.07) when compared to a 4C model (Moon *et al.*, 2009b).

The Ball2004 equation showed further promise as plots generated for detection of bias in the error were non-significant indicating that the magnitude of error is independent of the fatness of the subject (homoscedastic). This is likely a function of utilizing a quadratic equation which better represents the curvilinear relationship between skinfolds and fatness, which is not recognized in many of the other regression equations.

Bland-Altman plots showed that all equations derived from non DXA criteria contained significant bias in their ability to predict % BF in female athletes and that when applied to individual athletes, % BF can be as great as approximately ±10% BF depending on the equation used. Analysis of the Bland Altman plots for the Ball2004 equation indicates that for any given individual the % BF may still be within -4.7 to 6.5 % BF of the true value, and indicates why associations such as ISAK and some authors prefer to use reliability data of raw skinfold for interpreting change in body composition (Slater *et al.*, 2006). Recommendations for application and interpreting of reliability statistics in concert with regression equations can be found in the appendices under the "Applied Composition Analysis – Worksheet" (Appendix L).

It is not too surprising to find that the Durnin & Womersly 1974, (DW) and Jackson Pollock (JP) equations are the most valid out of the group. The DW paper is very similar to the CSEP skinfolds with the exception that the DW excludes lower body

skinfolds (calf). Furthermore the DW 20-29 equation subsample of 100 females, is the more accurate equation tested because it contains fewer subject characteristic deviations from our population. In addition, the original regression was created using a combination of Harpenden and Lange callipers. Although the authors cited earlier work that did not show a significant difference between these two callipers, more recent investigations have found that the Lange measures values significantly higher than Harpenden (Schmidt & Carter, 1990). The creation of an equation with the use of Harpenden callipers hypothetically eliminates much of the error attributed to calliper type used in the creation of other equations. As well, the group found a significant effect of stratified age groups on the intercept of the regression, indicating the justification for the use of age as a separate consideration when considering the relationships. This was likely a result of the unique fat patterning due to increased internal adiposity and decreased proportions of subcutaneous adiposity as age increases, as well as the decrease in skeletal density with age in females over 30 (Durnin & Womersley, 1974).

5.3 New DXA Based Regression Equation

Although it has been suggested that a linear regression represents a "good fit" for using skinfolds to predict body fat with a homogenous sample, the data from our study indicate that the sum5 is better represented by a quadratic relationship. Possible explanations for this may reside in the characteristics of the sample population. The selection criteria were initially intended to lead to a homogenous sample, and thus, a population specific regression equation. However, the diverse range of sports and levels of fatness (13.5 – 36.6 % BF DXA) represented by the sample of athletes tested, have led to sample heterogeneity which is better represented by relationships for the development

of generalized equations (Jackson, 1984). This notion is further supported by the fact that during cross-validation of existing equations the Ball2004 equation showed the greatest agreement with DXA in our sample, despite being developed as a generalized equation, much like the early Jackson Pollock 1980 equations.

It is not surprising that the forward and backward linear stepwise methods yielded the same predictive variables in our study, but this is not always the case (Heyward, 1996). These variables were identical to the variables used in the Ball2004 equation with the exception of the addition of the triceps in the Ball2004 equation, and may once again explain why the Ball2004 equation did so well in predicting % BF DXA. Following the direction of the seminal studies conducted by Jackson and Pollock, Ball et al. chose to enter variables into the regression that had previously shown good relationships to % BF through factor analysis (Jackson & Pollock, 1976). Additionally, since the data from the Jackson and Pollock study of 1976 showed multicollinearity between skinfold sites, summing of several individual skinfolds yields a more stable regression coefficient. This was evident in our analysis as regression equations created out of individual skinfolds were flagged for multicollinearity (eqn. dh#1-dh#3 – Appedix K) whereas the summed equations did not. Furthermore, taking the square of the sums was thought to be more appropriate as the relationship between skinfolds and fatness is curvilinear (Jackson & Pollock, 1982). However, despite statistical associations, the forwards, backwards, best subsets, and selected summed skinfold generated equations (eqn dh#1-dh#7) failed to pass critical assumption tests and as a result must be disregarded (Appendix K).

Perhaps more interesting, but not surprising, is that by selecting biologically similar variables from the CSEP-CPT performa as those identified statistically, we were

able to create an equation with equal predictive abilities as the best existing regression equation. It was hypothesized that with large magnitudes of shared variance it might be possible to exchange variables and maintain a high level relationship with the response variable. For example, the SSp site, which was selected among both stepwise regressions and best subsets analyses, has a strong correlation (0.889) and shared variance of about 79% with the IC skinfold (Table 4.14). The two have moderate to strong correlations with the response variable at 0.780 and 0.744 respectively. It is possible to exchange these two skinfolds and still maintain central representation of adiposity in the predictive regression. Guo *et al.* 1996 has suggested that it is advisable to include a variety of skinfold sites to adequately reflect general fatness through principles of subcutaneous adipose tissue distribution (Guo, 1996). The dh#9 equation was able to represent subcutaneous adipose distribution through use of a sum of 5 skinfolds that contain sites from the upper and lower appendages (Tri, Bi, C) as well as incorporating central representation (SSc, IC) which is critical for accurate reflection of fatness (Guo, 1996).

Several studies have shown conceptual support for the individual skinfolds selected in the sum5. As the Durnin and Womersley 1974 equation utilizes the same skinfolds with the exception of the lower body representation by the calf, research done on this equation and its shortcoming are a learning opportunity. In an investigation of Durnin and Womersley's equation, it was found that the largest correlation between skinfolds and %Fat DXA was found in thigh (0.904), calf (0.887) and biceps (0.841) (Eston *et al.*, 2005). When a MC % BF model was substituted for % BF DXA the greatest predictors were still thigh (0.874), and calf (0.843), while the bicep skinfold was replaced by the iliac crest (0.740). This information stresses the importance of inclusion of these

sites for reflection of body fatness of healthy females. However, the authors acknowledged that the anterior thigh is "perhaps one of the most difficult and least convenient sites to measure in some individuals" (Eston *et al.*, 2005). The authors' recommendations advocate for creating a validity equation using Durnin & Womersley with the addition of a thigh or calf skinfold for healthy young women, which is fulfilled by the new dh#9 equation.

The PRESS statistic has been suggested as a superior method to data splitting for regression cross-validation. The PRESS statistic assesses a prediction equation's ability to perform with independently selected predictor data (Holiday *et al.*, 1995). The PRESS SEE of the dh#9 equation showed good stability as there was only a 1.14% inflation of the SEE (SEE 3.0 to PRESS SEE 3.13), indicating stability of the regression equation. Hence, there is potential for utilizing the dh#9 equation in assessing female athletes with skinfolds common to the national certification program offered by CSEP.

5.4 RESULT APPLICATION

5.4.1. Anthropometry

When interpreting the magnitude of change in raw data, the change in the sum of skinfolds must exceed the SRD value to be defined as a true change with 95% certainty. The magnitudes of SRD values across all summed skinfolds combinations in this study allows a general interpretation for this study's anthropometrist, that a change in summed values that exceeds ± 4.2 mm will be true no matter which sum combination is used to reflect adiposity. This can be deduced as this was the largest SRD among all the summed skinfold combinations. However, one must also consider the percentage that the SRD represents of the summed values. In this instance the sum5 and sum9 yield SRD`s that

represent the smallest percentage of the raw values at 3.7% (2.2 mm) and 3.5% (3.9 mm) respectively. Attempting to utilize these results for all anthropometrists should be done with caution. As calculations using values from this study show (see Table 5.1) that even if an anthropometrist displays proficiency at the minimum reliability criteria level for a Level 1 or Level 3 ISAK accreditation, the resulting SRD for the sum5 and sum9 becomes 8.0 &16.0mm, and 15.2 & 30.4 mm respectively, which are approximately 4-8 times greater than the SRD values of the current anthropometrist. For each increase in %TEM of the anthropometrist there is a 1.6% change in the SRD raw value for sum5 and a 3% change for sum9. In addition, as the sum9 skinfold displayed a significant systemic bias of -1.4 % from Anthro T1 to Anthro T2, it implies that the greater resolution for detecting change and the absence of bias of sum5, identify it as a potential candidate for the creation of a regression for tracking changes in body composition.

Table 5.1 : Raw skinfold values from 95 female subjects and projected SRD values given thresholds of reliability for Level 1 & Level 3 ISAK accreditation.

		ISAK LEVEL 3		ISAK LEVEL 1	
		5%	TEM	10%	TEM
Sum of Skinfold Combinations	Raw SKF	TEM(mm)	SRD (mm)	TEM(mm)	SRD (mm)
Sum 3 (Tri, Abb, Thi)	48.7	2.4	6.7	4.9	13.5
Sum 3 (Tri, SSp, Thi)	42.4	2.1	5.9	4.2	11.8
Sum 3 (Tri, SSc, IC)	40.0	2.0	5.5	4.0	11.1
Sum 3 (Tri, SSp, Ab)	41.2	2.1	5.7	4.1	11.4
Sum 4 (Tri, SSp, Ab, Thi)	58.6	2.9	8.1	5.9	16.2
Sum 4 (Tri, Bi, SSc, IC)	45.8	2.3	6.3	4.6	12.7
Sum 5 (CSEP SSc, Tri, Bi, IC, C)	57.9	2.9	8.0	5.8	16.0
Sum 7 (ISAK Level 1 - SSc, Tri, Bi, Ab5, SSp, Thi, C)	86.4	4.3	12.0	8.6	23.9
Sum 9 (Anthropometrica-SSc, Tri, Bi, Ax, IC, SSp, Ab5, Thi, C)	109.6	5.5	15.2	11.0	30.4

ISAK minimal reliability values for accreditation at level 3 (< 5% skinfold %TEM) and level 1 (< 10% skinfold %TEM) anthropometrist. The effects of differing levels of expertise are reflected in the magnitude of detectable change represented by the SRD, smallest real difference values when applied to summed skinfold combinations from the study.

5.4.2. Dual Energy X-ray Absorptiometry

When interpreting the magnitude of change in raw DXA results it is common to use Least Significant Change (LSC) values to determine whether the observed change is of clinical significance at the level of 95% certainty. The results of this study show that the LSC for fat mass on our Lunar Prodigy DXA is approximately 700 g fat. The findings of this study yield three considerations when assessing individual athletes. First, because the error values are absolute regardless of body mass, the relative percentage of fat needed to be lost or required for identification of change, will differ for individuals of different masses. For example, a 50 kg distance runner will need to change her body composition by 1.39 % for detection of change, while an 85 kg volleyball player will only need to change 0.82 % for change to be detected, because the absolute LSC value for fat mass detected by the DXA is approximately 700 g. It will be easier to detect change in larger individuals as the absolute value of change required will represent a smaller relative change in % BF for that individual (Figure 5.1).

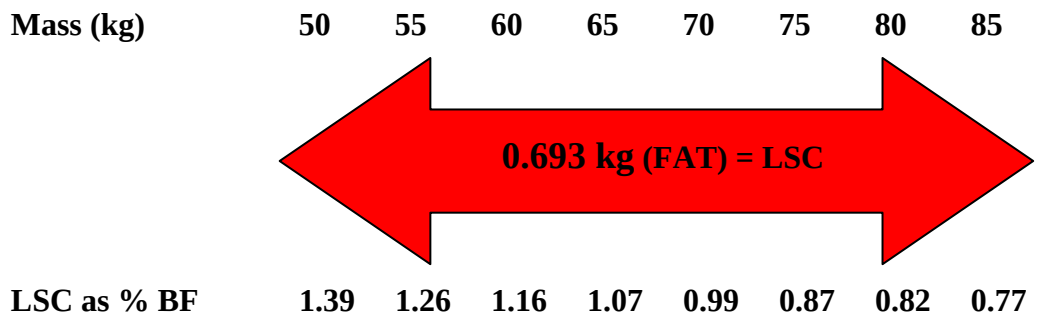


Figure 5.1: A representation of how the LSC value for fat will affect the threshold of detecting % BF change based on the mass of the athletes within the range of the study. As the mass of the athlete increases the LSC of 0.693 kg fat will represent a smaller % BF change.

Second, the magnitude of the LSC has implications for the frequency of anthropometric testing. The ACSM guidelines for healthy weight loss indicate that a negative caloric balance should not exceed 500-1000 kcal/day, that the deficit should be comprised of both nutritional deficit, as well as increased energy expenditure through exercise, and that a maximal weight loss of 1 kg per week should not be exceeded. Given that a deficit of 5336 kcal is the caloric equivalent of 693 g of fat, one could theoretically detect change in as little as 5-6 days or more realistically 10-11days if the guidelines of 1000kcal and 500kcal daily deficits were implemented (ACSM, 2000).

Finally, because the reliability of a prediction equation cannot be less than the criterion in which it was created, anthropometric monitoring should be conducted with greater time intervals between testing than outlined for DXA, or if similar timing is used there would need to be an adjustment in the interpretation of the degree of certainty that the change is not merely due to measurement error.

Although the results for lean mass yielded near equivalent LSC values as those for fat mass (720 g vs. 693 g respectively), there are several small complications

associated with interpreting lean mass changes and these limitations must be acknowledged. When defining whole body composition the DXA defines the three compartments as bone or osseous mineral, fat mass and lean mass. The value of the lean mass measurement comprises of the sum of total protein, non osseous mineral and total body water (Fuller *et al.*, 2001). Of these lean mass constituents, total body water fluctuates due to shifts in fluid homeostasis from sweating, as well as ingestion of fluids. In this regard, small changes in fluid and or levels of hydration could potentially skew interpretation of changes in lean mass. In a study manipulating hydration status, Going et al (1993) found that DXA accurately detected small changes in body mass by inducing a dehydration and then subsequent rehydration of 2% body mass over 3 days. Approximately 48-52% of the changes in body mass induced by dehydration were reflected in lean tissue mass. The authors suggest that a greater percentage was not accounted for likely due to the undefined effects of the ratio of extracellular to intracellular water ration on x-ray attenuation, but acknowledged that DXA is sensitive to shifts in fluid balance and a large component is reflected in the lean mass quantification.

In our study, we assessed hydration status via refractometry and did not find a significant correlation between hydration status and DXA reliability. Similarly DXA was not correlated to DXA %BF values. It was beyond the scope of this study to assess the impact of hydration status on lean tissue measurements, but is something that warrants further investigation. It is recommended that standardized pre assessment protocols be followed to minimize the effect that ingested foodstuffs and liquids may have on lean mass readings. Furthermore, it may be valuable to analyse lean mass values in conjunction with hydration status as this will likely provide more insight into the kinetics

of this compartment change. It may also be beneficial to wait to collect baseline compositional data a few days after training commencement so that the greatest perturbations in fluid homeostasis such as plasma volume expansion (up to 700 ml; Fellmann, 1992) can be normalized, or accounted for, giving a more stable initial reference point when assessing change (Fellmann, 1992). From thereon, the likelihood of identifying changes associated with protein accretion is more probable. Additionally, there may be utility in applying newly created equations in the literature that attempt to predict total body protein from DXA lean tissue results using multiplicative constants from the literature (Fuller *et al.*, 2001).

CHAPTER 6 - SUMMARY

6.1 Conclusions

Anthropometric equations derived from various criteria yielded dissimilar results. Long utilized popular HD equations advocated in many national accreditation schemes (ACSM, CSEP) show considerable bias when compared to modern values obtained by current DXA technology. The results of the study indicate several important points that should be recognized when analyzing body composition in female athletes:

1. The reliability of the assessment method employed should be established prior to attempting to interpret change. The most valuable reliability statistics for practitioners is the smallest real difference (SRD) or least significant change (LSC) value.
2. For anthropometric assessments, ISAK guidelines for establishing reliability should be conducted. Summed skinfolds should be used in combination with the smallest real difference value to assess change. Standards represented by %TEM should not be viewed as thresholds for practicing, but as continuums where increasing levels of reliability will directly affect the magnitude of change that can be detected.
3. DXA is an extremely quick and reliable method of assessing body composition. The ability to measure changes in fat mass in the GE Lunar Prodigy DXA system is independent of body size and thus, fat mass must change by approximately 700 g for true change to be detected.

4. Considerable error exists in current MC, DXA, and HD based regression equations. When assessing female athletes using existing published equations the Ball (2004) equations should be utilized to obtain the most valid representation of % BF indicated by a DXA criterion. When applied to individuals, the values may be up to -4.7 % below to 6.5 % above the predicted value, but may still be valuable when goal setting for compositional change particularly when calculating caloric balance, and the time course for expected change to occur.
5. A new regression equation was created using skinfolds that are familiar to Canadian practitioners as they are taught as part of the national professional certification scheme (CSEP-CPT) as part of the measures for the Healthy Body Composition Module. The dh9 equation can be used to accurately predict % BF in female athletes with equal validity as the Ball 2004 equation. The equation will yield similar results in individual populations with errors potentially as large as ± 6.1% BF while allowing consistent teaching methodology in the measurement procedure.

6.2 Limitations

As with all regression equations, the interpretation of the results can only be extended to populations that share characteristics to that of the present study. This means that the current results should be employed with caution when deviating from the sample characteristics. As well, considerable change in predictive abilities is influenced by relatively small changes in equipment or testing procedures indicating that the utility of the equation may only apply to a small subset of practitioners utilizing the unique combination of equipment and methodology.

Anthropometric reliability is unique to each anthropometrist. In our study it was found that the sum5 was the most reliable summed skinfold combination and thus, a natural candidate for the creation of the new regression equation. This may be a function of the experience and educational background of the anthropometrist in question, and not necessarily a function inherent to the skinfold combination itself. This may not necessarily be a negative limitation as it may allow generalization of the results to the typical professional with similar educational background making the results of this study even more pertinent for professionals nationally.

6.3 Recommendations

Practitioners should follow the steps below and utilize the tools in the appendices to provide interpretation to body composition in female athletes.

6.3.1. Anthropometrics

1. Establish TEM values using desired anthropometric performa as per ISAK guidelines.
2. Calculate or use charts to identify magnitude of change in summed skinfolds necessary to denote “true change” as designated by SRD (95% confidence level).
3. Calculate or use charts to identify magnitude of change in summed skinfolds necessary to denote “true change” has occurred with 90%, or 80% confidence, as these are still important indices of change in applied settings (Appendix M).
4. Convert summed skinfolds into % BF utilizing the equation by Ball2004 for the ACSM performa, or dh#9 equation for CSEP performa.

5. Use the reliability and likelihood of change scores to interpret the % BF results.
6. Only test over time intervals when it is likely that these magnitudes of change have been met.

6.3.2. DXA

1. When assessing change in body composition utilizing a GE Lunar Prodigy DXA, the magnitude of fat loss necessary for detecting "true change" is approximately 700 g.
2. DXA has the potential to theoretically identify fat loss in as little as 5-6 days, but it is more likely that the shortest time course for testing should be 10-11 days based on ACSM healthy weight loss guidelines.

6.4 Future Directions

The evolution of measuring devices will constantly require the reassessment of old methodologies. As seen with DXA, newer, quicker devices that require less technician skill will be developed and prove to be useful, but only to personnel with limited access. Anthropometric measurement has been around for hundreds of years and has developed a legacy due to its ease of use, potential reliability, versatility, and economic accessibility. Although these characteristics are its virtue, the widespread use of anthropometrics have also led to a perpetuation and in some instances deterioration of the field as the necessary theoretical background for proper implementation and interpretation is misunderstood. Many new contributions to the field of research are lacking in execution or articulation of adequate reliability testing prior to regression creation. A more meticulous approach in documentation, reporting, and adhering to

existing standardized protocols is warranted. Future research regarding this particular study should be the cross-validation of the existing dh#9 equation in an independent sample of female athletes to obtain real data in addition to the PRESS statistics that were employed. If successful, a similar equation utilizing variables from the CSEP performa should be developed and implemented with male athletes.

REFERENCES

AAP. (1996). Promotion of healthy weight-control practices in young athletes. American Academy of Pediatrics. Committee on Sports Medicine and Fitness. *Pediatrics* **97,** 752-753.

Aasen G, Fagertun H & Halse J. (2006). Body composition analysis by dual X-ray absorptiometry: in vivo and in vitro comparison of three different fan-beam instruments. *Scand J Clin Lab Invest* **66,** 659-666.

ACSM. (2000). *ACSM's Guidelines For Exercise Testing and Prescription - sixth edition*. Lippincott Williams & Wilkins.

Anderson GS. (1999). Human morphology and temperature regulation. *Int J Biometeorol* **43,** 99-109.

Anderson M, Hopkins W, Roberts A & Pyne D. (2008). Ability of test measures to predict competitive performance in elite swimmers. *J Sports Sci* **26,** 123-130.

Anthrpometrica. (1996). *Anthropometrica*. UNSW Press.

Arrese AL & Ostariz ES. (2006). Skinfold thicknesses associated with distance running performance in highly trained runners. *J Sports Sci* **24,** 69-76.

Atkinson G & Nevill AM. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* **26,** 217-238.

Bale P, Rowell S & Colley E. (1985). Anthropometric and training characteristics of female marathon runners as determinants of distance running performance. *J Sports Sci* **3,** 115-126.

Ball S, Swan PD & DeSimone R. (2004). Comparison of anthropometry to dual energy X-ray absorptiometry: a new prediction equation for women. *Res Q Exerc Sport* **75,** 248-258.

Bamman MM, Newcomer BR, Larson-Meyer DE, Weinsier RL & Hunter GR. (2000). Evaluation of the strength-size relationship in vivo using various muscle size indices. *Med Sci Sports Exerc* **32,** 1307-1313.

Bar-Or O, Lundegren HM & Buskirk ER. (1969). Heat tolerance of exercising obese and lean women. *J Appl Physiol* **26,** 403-409.

Bassett DR, Jr. & Howley ET. (2000). Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med Sci Sports Exerc* **32,** 70-84.

Baumgartner TA, Jackson, Andrew J, Mahar, Matthew T, Rowe, David A. (2007). Measurement for Evaluation: In Physical Education & Exercise Science. **Eight Edition,** 544.

Bayios IA, Bergeles NK, Apostolidis NG, Noutsos KS & Koskolou MD. (2006). Anthropometric, body composition and somatotype differences of Greek elite female basketball, volleyball and handball players. *J Sports Med Phys Fitness* **46,** 271-280.

Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD & Verbeek AL. (2001). Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res* **10,** 571-578.

Bell DG & Jacobs I. (1990). Muscle fibre area, fibre type & capillarization in male and female body builders. *Can J Sport Sci* **15,** 115-119.

Benardot D & Czerwinski C. (1991). Selected body composition and growth measures of junior elite gymnasts. *J Am Diet Assoc* **91,** 29-33.

Bland JM & Altman DG. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1,** 307-310.

Bompa TO. (1999). *Periodization - Theory and Methodology of Training*. Human Kinetics.

Bourgois J, Claessens AL, Janssens M, Van Renterghem B, Loos R, Thomis M, Philippaerts R, Lefevre J & Vrijens J. (2001). Anthropometric characteristics of elite female junior rowers. *J Sports Sci* **19,** 195-202.

Brozek J, Grande F, Anderson JT & Keys A. (1963). Densitometric Analysis of Body Composition: Revision of Some Quantitative Assumptions. *Ann N Y Acad Sci* **110,** 113-140.

Bunc V. (2000). Energy cost of treadmill running in non-trained females differing in body fat. *J Sports Med Phys Fitness* **40,** 290-296.

Burrows M & Bird S. (2000). The physiology of the highly trained female endurance runner. *Sports Med* **30,** 281-300.

Carter JE & Ackland TR. (1998). Sexual dimorphism in the physiques of World Championship divers. *J Sports Sci* **16,** 317-329.

Cheung SS, McLellan TM & Tenaglia S. (2000). The thermophysiology of uncompensable heat stress. Physiological manipulations and individual characteristics. *Sports Med* **29,** 329-359.

Claessens AL, Hlatky S, Lefevre J & Holdhaus H. (1994). The role of anthropometric characteristics in modern pentathlon performance in female athletes. *J Sports Sci* **12,** 391-401.

Claessens AL, Lefevre J, Beunen G & Malina RM. (1999). The contribution of anthropometric characteristics to performance scores in elite female gymnasts. *J Sports Med Phys Fitness* **39,** 355-360.

Clark DJ, Condliffe EG & Patten C. (2006). Reliability of concentric and eccentric torque during isokinetic knee extension in post-stroke hemiparesis. *Clin Biomech (Bristol, Avon)* **21,** 395-404.

Cyrino. (2003). Impact of the use of different skinfold calipers for the analysis of the body composition. *Rev Bras Med Esporte* **9**.

D'Alessandro C, Morelli E, Evangelisti I, Galetta F, Franzoni F, Lazzeri D, Piazza M & Cupisti A. (2007). Profiling the diet and body composition of subelite adolescent rhythmic gymnasts. *Pediatr Exerc Sci* **19,** 215-227.

Durnin JV & Womersley J. (1974). Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged from 16 to 72 years. *Br J Nutr* **32,** 77-97.

Duz S, Kocak,Mehmet & Korkusuz, Feza. (2009). Evaluation of body composition using three different methods compared to dual-energy X-ray absorptiometry. *European Journal of Sport Science* **9**.

Eliasziw M, Young SL, Woodbury MG & Fryday-Field K. (1994). Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. *Phys Ther* **74,** 777-788.

Elmlinger MW, Kuhnel W, Wormstall H & Doller PC. (2005). Reference intervals for testosterone, androstenedione and SHBG levels in healthy females and males from birth until old age. *Clin Lab* **51,** 625-632.

Eston RG, Rowlands AV, Charlesworth S, Davies A & Hoppitt T. (2005). Prediction of DXA-determined whole body fat from skinfolds: importance of including skinfolds from the thigh and calf in young, healthy men and women. *Eur J Clin Nutr* **59,** 695-702.

Evans EM, Rowe DA, Misic MM, Prior BM & Arngrimsson SA. (2005). Skinfold prediction equation for athletes developed using a four-component model. *Med Sci Sports Exerc* **37,** 2006-2011.

Fellmann N. (1992). Hormonal and plasma volume alterations following endurance exercise. A brief review. *Sports Med* **13,** 37-49.

Friedl KE, Westphal KA, Marchitelli LJ, Patton JF, Chumlea WC & Guo SS. (2001). Evaluation of anthropometric equations to assess body-composition changes in young women. *Am J Clin Nutr* **73,** 268-275.

Fuller NJ, Wells JC & Elia M. (2001). Evaluation of a model for total body protein mass based on dual-energy X-ray absorptiometry: comparison with a reference four-component model. *Br J Nutr* **86,** 45-52.

Genton L, Hans D, Kyle UG & Pichard C. (2002). Dual-energy X-ray absorptiometry and body composition: differences between devices and comparison with reference methods. *Nutrition* **18,** 66-70.

Going SB, Massett MP, Hall MC, Bare LA, Root PA, Williams DP & Lohman TG. (1993). Detection of small changes in body composition by dual-energy x-ray absorptiometry. *Am J Clin Nutr* **57,** 845-850.

Goodman LR & Warren MP. (2005). The female athlete and menstrual function. *Curr Opin Obstet Gynecol* **17,** 466-470.

Gore CJ. (2000). *Physiological Tests for Elite Athletes*.

Gruber JJ, Pollock ML, Graves JE, Colvin AB & Braith RW. (1990). Comparison of Harpenden and Lange calipers in predicting body composition. *Res Q Exerc Sport* **61,** 184-190.

Guo SS, Chumela, W.C. (1996). Statistical Mehtods for the Development and Testing of Predictive Equations. In *Human Body Composition*, pp. 191-202. Human Kinetics.

Guo Y, Franks PW, Brookshire T & Antonio Tataranni P. (2004). The intra- and inter-instrument reliability of DXA based on ex vivo soft tissue measurements. *Obes Res* **12,** 1925-1929.

Hagmar M, Hirschberg AL, Berglund L & Berglund B. (2008). Special attention to the weight-control strategies employed by Olympic athletes striving for leanness is required. *Clin J Sport Med* **18,** 5-9.

Helmuth HS. (1980). Anthropometric survey of young swimmers. *Anthropol Anz* **38,** 17-34.

Heyward VH, Stolarczyk, L.M. (1996). Applied Body Composition Assessment, pp. 221. Human Kinetics.

Holiday DB, Ballard JE & McKeown BC. (1995). PRESS-related statistics: regression tools for cross-validation and case diagnostics. *Med Sci Sports Exerc* **27,** 612-620.

Holschen JC. (2004). The female athlete. *South Med J* **97,** 852-858.

Hopkins WG. (2000). Measures of reliability in sports medicine and science. *Sports Med* **30,** 1-15.

Hopkins WG, Hawley JA & Burke LM. (1999). Design and analysis of research on sport performance enhancement. *Med Sci Sports Exerc* **31,** 472-485.

Hull H, He Q, Thornton J, Javed F, Allen L, Wang J, Pierson RN, Jr. & Gallagher D. (2008). iDXA, Prodigy, and DPXL Dual-Energy X-ray Absorptiometry Whole-Body Scans: A Cross-Calibration Study. *J Clin Densitom*.

Hume P & Marfell-Jones M. (2008). The importance of accurate site location for skinfold measurement. *J Sports Sci* **26,** 1333-1340.

ISAK. International Society for the Advancement of Kinanthropometry.

ISCD. (2007). OFFICIAL POSITIONS& PEDIATRIC OFFICIAL POSITIONS of the INTERNATIONAL SOCIETY FOR CLINICAL DENSITOMETRY.

Jackson AS. (1984). Research design and analysis of data procedures for predicting body density. *Med Sci Sports Exerc* **16,** 616-622.

Jackson AS, Ellis KJ, McFarlin BK, Sailors MH & Bray MS. (2008). Cross-validation of generalised body composition equations with diverse young men and women: the Training Intervention and Genetics of Exercise Response (TIGER) Study. *Br J Nutr***,** 1-8.

Jackson AS & Pollock ML. (1976). Factor analysis and multivariate scaling of anthropometric variables for the assessment of body composition. *Med Sci Sports* **8,** 196-203.

Jackson AS & Pollock ML. (1982). Steps toward the development of generalized equations for predicting body composition of adults. *Can J Appl Sport Sci* **7,** 189-196.

Jackson AS, Pollock ML & Ward A. (1980). Generalized equations for predicting body density of women. *Med Sci Sports Exerc* **12,** 175-181.

Jones EJ, Bishop PA, Woods AK & Green JM. (2008). Cross-sectional area and muscular strength: a brief review. *Sports Med* **38,** 987-994.

Jonnalagadda SS, Skinner R & Moore L. (2004). Overweight athlete: fact or fiction? *Curr Sports Med Rep* **3,** 198-205.

Kagawa M, Kuroiwa C, Uenishi K, Mori M, Hills AP & Binns CW. (2007). New percentage body fat prediction equations for Japanese females. *J Physiol Anthropol* **26,** 23-29.

Karlson KA, Becker CB & Merkur A. (2001). Prevalence of eating disordered behavior in collegiate lightweight women rowers and distance runners. *Clin J Sport Med* **11,** 32-37.

Kavouras SA. (2002). Assessing hydration status. *Curr Opin Clin Nutr Metab Care* **5,** 519-524.

Kerr DA, Ross WD, Norton K, Hume P, Kagawa M & Ackland TR. (2007). Olympic lightweight and open-class rowers possess distinctive physical and proportionality characteristics. *J Sports Sci* **25,** 43-53.

Knechtle B, Duff B, Amtmann G & Kohler G. (2007). Cycling and running performance, not anthropometric factors, are associated with race performance in a Triple Iron Triathlon. *Res Sports Med* **15,** 257-269.

Kohli S, Gao M & Lear SA. (2009). Using simple anthropometric measures to predict body fat in South Asians. *Appl Physiol Nutr Metab* **34,** 40-48.

Legaz A & Eston R. (2005). Changes in performance, skinfold thicknesses, and fat patterning after three years of intense athletic conditioning in high level runners. *Br J Sports Med* **39,** 851-856.

Legaz Arrese A, Munguia Izquierdo D & Serveto Galindo JR. (2006). Physiological measures associated with marathon running performance in high-level male and female homogeneous groups. *Int J Sports Med* **27,** 289-295.

Leone M, Lariviere G & Comtois AS. (2002). Discriminant analysis of anthropometric and biomotor variables among elite adolescent female athletes in four sports. *J Sports Sci* **20,** 443-449.

Lohman M, Tallroth K, Kettunen JA & Marttinen MT. (2009). Reproducibility of dual-energy x-ray absorptiometry total and regional body composition measurements using different scanning positions and definitions of regions. *Metabolism* **58,** 1663-1668.

Lohman TG, Harris M, Teixeira PJ & Weiss L. (2000). Assessing body composition and changes in body composition. Another look at dual-energy X-ray absorptiometry. *Ann N Y Acad Sci* **904,** 45-54.

Lohman TG, Pollock ML, Slaughter MH, Brandon LJ & Boileau RA. (1984). Methodological factors and the prediction of body fat in female athletes. *Med Sci Sports Exerc* **16,** 92-96.

Lohman TGaC, Zhao. (2005). Dual Energy X-Ray Absorptiometry. In *Human Body Composition*, pp. 523.

Malina RM. (2007). Body composition in athletes: assessment and estimated fatness. *Clin Sports Med* **26,** 37-68.

Malousaris GG, Bergeles NK, Barzouka KG, Bayios IA, Nassis GP & Koskolou MD. (2008). Somatotype, size and body composition of competitive female volleyball players. *J Sci Med Sport* **11,** 337-344.

Mattsson S & Thomas BJ. (2006). Development of methods for body composition studies. *Phys Med Biol* **51,** R203-228.

Mazess RB, Hanson JA, Payne R, Nord R & Wilson M. (2000). Axial and total-body bone densitometry using a narrow-angle fan-beam. *Osteoporos Int* **11,** 158-166.

Moon JR, Eckerson JM, Tobkin SE, Smith AE, Lockwood CM, Walter AA, Cramer JT, Beck TW & Stout JR. (2009a). Estimating body fat in NCAA Division I female athletes: a five-compartment model validation of laboratory methods. *Eur J Appl Physiol* **105,** 119-130.

Moon JR, Tobkin SE, Smith AE, Lockwood CM, Walter AA, Cramer JT, Beck TW & Stout JR. (2009b). Anthropometric estimations of percent body fat in NCAA Division I female athletes: a 4-compartment model validation. *J Strength Cond Res* **23,** 1068-1076.

Norcross J & Van Loan MD. (2004). Validation of fan beam dual energy x ray absorptiometry for body composition assessment in adults aged 18-45 years. *Br J Sports Med* **38,** 472-476.

Nord RH, Homuth JR, Hanson JA & Mazess RB. (2000). Evaluation of a new DXA fan-beam instrument for measuring body composition. *Ann N Y Acad Sci* **904,** 118-125.

Norton K & Olds T. (2001). Morphological evolution of athletes over the 20th century: causes and consequences. *Sports Med* **31,** 763-783.

O'Brien TD, Reeves ND, Baltzopoulos V, Jones DA & Maganaris CN. (2009). Strong relationships exist between muscle volume, joint power and whole-body external mechanical power in adults and children. *Exp Physiol* **94,** 731-738.

O'Connor H, Olds T & Maughan RJ. (2007). Physique and performance for track and field events. *J Sports Sci* **25 Suppl 1,** S49-60.

Oldroyd B, Smith AH & Truscott JG. (2003). Cross-calibration of GE/Lunar pencil and fan-beam dual energy densitometers--bone mineral density and body composition studies. *Eur J Clin Nutr* **57,** 977-987.

Oppliger RA, Harms RD, Herrmann DE, Streich CM & Clark RR. (1995). The Wisconsin wrestling minimum weight project: a model for weight control among high school wrestlers. *Med Sci Sports Exerc* **27,** 1220-1224.

Otte JW, Merrick MA, Ingersoll CD & Cordova ML. (2002). Subcutaneous adipose tissue thickness alters cooling time during cryotherapy. *Arch Phys Med Rehabil* **83,** 1501-1505.

Perez-Gomez J, Rodriguez GV, Ara I, Olmedillas H, Chavarren J, Gonzalez-Henriquez JJ, Dorado C & Calbet JA. (2008). Role of muscle mass on sprint performance: gender differences? *Eur J Appl Physiol* **102,** 685-694.

Peterson MJ, Czerwinski SA & Siervogel RM. (2003). Development and validation of skinfold-thickness prediction equations with a 4-compartment model. *Am J Clin Nutr* **77,** 1186-1191.

Piers LS, Soares MJ, Frandsen SL & O'Dea K. (2000). Indirect estimates of body composition are useful for groups but unreliable in individuals. *Int J Obes Relat Metab Disord* **24,** 1145-1152.

Prentice AM & Jebb SA. (2001). Beyond body mass index. *Obes Rev* **2,** 141-147.

Prior BM, Cureton KJ, Modlesky CM, Evans EM, Sloniger MA, Saunders M & Lewis RD. (1997). In vivo validation of whole body composition estimates from dual-energy X-ray absorptiometry. *J Appl Physiol* **83,** 623-630.

Prior BM, Modlesky CM, Evans EM, Sloniger MA, Saunders MJ, Lewis RD & Cureton KJ. (2001). Muscularity and the density of the fat-free mass in athletes. *J Appl Physiol* **90,** 1523-1531.

Rankin JW. (2002). Weight loss and gain in athletes. *Curr Sports Med Rep* **1,** 208-213.

Roche AF. (1996). Anthropometry and Ultrasound. In *Human Body Composition*, ed. Roche AF, Heymsfield SB & Lohman T, pp. 167-189; 366. Human kinetics.

Schmidt PK & Carter JE. (1990). Static and dynamic differences among five types of skinfold calipers. *Hum Biol* **62,** 369-388.

Scholz N, Neil, R., Smith, D.J., . (2008). Technical error of measurement (TEM) of skinfolds measured on 262 national team athletes *Applied Physiology Nutrition and Metabolism* **33**.

Selkirk GA & McLellan TM. (2001). Influence of aerobic fitness and body fatness on tolerance to uncompensable heat stress. *J Appl Physiol* **91,** 2055-2063.

Sheng HP & Huggins RA. (1979). A review of body composition studies with emphasis on total body water and fat. *Am J Clin Nutr* **32,** 630-647.

Shepherd JA & Lu Y. (2007). A generalized least significant change for individuals measured on different DXA systems. *J Clin Densitom* **10,** 249-258.

Silva AM, Minderico CS, Teixeira PJ, Pietrobelli A & Sardinha LB. (2006). Body fat measurement in adolescent athletes: multicompartment molecular model comparison. *Eur J Clin Nutr* **60,** 955-964.

Siri WE. (1956). The gross composition of the body. *Adv Biol Med Phys* **4,** 239-280.

Slater GJ, Duthie GM, Pyne DB & Hopkins WG. (2006). Validation of a skinfold based index for tracking proportional changes in lean mass. *Br J Sports Med* **40,** 208-213.

Sloan AW, Burt JJ & Blyth CS. (1962). Estimation of body fat in young women. *J Appl Physiol* **17,** 967-970.

Sovak D & Hawes MR. (1987). Anthropological status of international calibre speed skaters. *J Sports Sci* **5,** 287-304.

Sovak D, Hawes MR & Plant K. (1992). Morphological proportionality in elite age group North American divers. *J Sports Sci* **10,** 451-465.

Stoggl T, Enqvist J, Muller E & Holmberg HC. Relationships between body composition, body dimensions, and peak speed in cross-country sprint skiing. *J Sports Sci* **28,** 161-169.

Storlie J. (1991). Nutrition assessment of athletes: a model for integrating nutrition and physical performance indicators. *Int J Sport Nutr* **1,** 192-204.

Sutcliffe JF. (1996). A review of in vivo experimental methods to determine the composition of the human body. *Phys Med Biol* **41,** 791-833.

Temfemo A, Hugues J, Chardon K, Mandengue SH & Ahmaidi S. (2009). Relationship between vertical jumping performance and anthropometric characteristics during growth in boys and girls. *Eur J Pediatr* **168,** 457-464.

Thorland WG, Johnson GO, Tharp GD, Housh TJ & Cisar CJ. (1984). Estimation of body density in adolescent athletes. *Hum Biol* **56,** 439-448.

Tipton KD. (2001). Gender differences in protein metabolism. *Curr Opin Clin Nutr Metab Care* **4,** 493-498.

Tsunawake N, Tahara Y, Yukawa K, Katsuura T, Harada H & Kikuchi Y. (1995). Characteristics of body shape of female athletes based on factor analysis. *Appl Human Sci* **14,** 55-61.

van den Tillaar R & Ettema G. (2004). Effect of body size and gender in overarm throwing performance. *Eur J Appl Physiol* **91,** 413-418.

Vozarova B, Wang J, Weyer C & Tataranni PA. (2001). Comparison of two software versions for assessment of body-composition analysis by DXA. *Obes Res* **9,** 229-232.

Wang TJC, Kolesnik S, and Pierson Jr, RN. (2000). Anthropometry in Body Composition. *Annals New York Academy of Sciences***,** 317-326.

Wang ZM, Heshka S, Pierson RN, Jr. & Heymsfield SB. (1995). Systematic organization of body-composition methodology: an overview with emphasis on component-based methods. *Am J Clin Nutr* **61,** 457-465.

Wang ZM, Pierson RN, Jr. & Heymsfield SB. (1992). The five-level model: a new approach to organizing body-composition research. *Am J Clin Nutr* **56,** 19-28.

Ward R & Anderson GS. (1998). Resilience of anthropometric data assembly strategies to imposed error. *J Sports Sci* **16,** 755-759.

Warner ER, Fornetti WC, Jallo JJ & Pivarnik JM. (2004). A Skinfold Model to Predict Fat-Free Mass in Female Athletes. *J Athl Train* **39,** 259-262.

APPENDICES

Appendix A

Informed Consent

Research Project Title:
An Applied Approach to Defining Within Subject Body Composition Change
Researcher(s):
Daryl Hurrie
Sponsor:
Canadian Sport Centre Manitoba

CONSENT TO BE A RESEARCH SUBJECT

This consent form, a copy of which will be left with you for your records and reference, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

Assessment Personnel

All assessment procedures and data analysis will be conducted by the primary investigator Daryl Hurrie.

Project Summary

The purpose of my study is to investigate existing body composition techniques focusing on how well methods using skinfolds, girths, breadths and lengths work compared to dual energy x-ray absorptiometry (DXA), the gold standard.

Current ways of calculating body composition variables (%body fat and muscle mass) for female athletes are based on old technologies. Also, existing measurement techniques, from a mathematical point of view, do not provide good information for coaches or sport scientists to really track body composition changes that may occur with training and/or diet in an individual athlete. This study will allow us to use state of the art equipment and statistics to provide recommendations to practitioners working with female athletes to maximize performance related to body composition.

Study Participants will be required to sign up for one or more testing sessions. Depending on which session is chosen, subjects will be instructed to arrive at the testing location indicated below. Before participating in the study subjects will be asked to fill out the

Informed Consent form, as well as take time to read through and fill out necessary questionnaires and forms. These questionnaires will ask questions about your medical and training history, and will take about 10 to 15 minutes to complete.

Pre - Assessment Procedures

Subjects should arrive 10-15 minutes prior to their scheduled assessment time. Prior to testing the subject should follow the pre-assessment instructions listed below in order to maximize the accuracy of the testing procedure.

- have filled out all necessary forms (Informed Consent / Medical Background Questionnaire)
- have fasted for a minimum of 8 hours, (as foodstuff will affect DXA reading)
- have NOT trained for a minimum of 8 hours prior to assessment
- have NOT undergone barium tests/exams, or had a nuclear medicine scan or injections with an x-ray dye within the last two weeks
- wear clothing that permits access to the measurement sites located in the "standardized anthropometric assessment" protocol (i.e., sports bra and athletic shorts)
- remove all metallic objects including zippers (as metal may affect bone density values, thus, affecting body composition values)
- be reasonably hydrated
- have brought sample and pregnancy test results from 1st urine sample of the day as per "Urinalysis" section of methodology*

* if menstruating, please indicate on medical background questionnaire and urine and pregnancy test will not be necessary prior to assessment.*

Methodology

Urinalysis (Urine sample)

The 1st urine sample of the day must be collected in order to assess hydration as well as screening for pregnancy prior to DXA testing. At home upon waking, or, the first urination after waking, the subject will provide a minimum 20ml sample in the provided collection container. The subject will then insert the pregnancy test dipstick into the urine as per the manufacturer's instructions and place the used pregnancy dipstick in the provided Ziploc bag. Subjects will place their study identification number on both the ziplock bag as well as the sample container, along with the time of collection. Subjects are instructed to NOT write their names on any of this material. Subjects will bring this material to their scheduled assessment. The **urine sample is to be collected the morning of assessment**. Subjects need not refrigerate samples as room temperature of

the sample is required for testing. Urinalysis will be carried out prior to anthropometric precision assessments but will not require a pregnancy test.*

subjects menstruating will not be asked to provide a sample, nor will they need to provide the results of the pregnancy test

Anthropometric Precision
(25 mins per anthropometric assessment x 2 assessments = 50 mins total)

The first thing that must be established is the Technical Error of Measurement (TEM) for the anthropometric procedure. This tells us how accurate the person making the measurements is. Twenty subjects volunteering for this part of the study will be asked to undergo two anthropometric assessments. The anthropometric assessments will be conducted in the performance lab, Room 128 Frank Kennedy Centre. **Briefly the anthropometric assessment requires the subject to wear clothing that permits unobstructed access to measurement sites.** (See below). The two assessments will NOT be conducted consecutively, but will require the subject to return for an additional anthropometric assessment at roughly the same time of day, 7 days after their initial assessment. The establishment of the anthropometric precision must be completed prior to commencement of the other components of the study.

Standardized Anthropometric Assessment (Anthrpometrica, 1996)

General Procedures & Techniques

All measurements will be taken by the primary investigator Daryl Hurrie, with the athlete wearing a sports bra and sport shorts, on the right hand side of the body in triplicate at sites landmarked with pen.

Before testing is to commence it should be ensured that all pre assessment protocols have been followed.

Specific Techniques - Measurement Sites (Anthrpometrica, 1996)

General Measurements

Weight – taken wearing sports bra and athletic shorts using a Lifesource precision scale

Height - Maximal standing height is recorded at end of inspiration, using a portable Seca stadiometer.

Seated Height - Portable stadiometer is placed on a box, subjects height is taken at end of inspiration when seated.

Skinfold Measurement

Briefly, skinfolds will be collected using a Harpenden skinfold calliper. The Harpenden calliper is a common calliper used to quantify fat beneath the skin. In appearance the calliper looks like a set of tongs with a dial face that indicates how wide the jaw is opened. The subject will stand upright, relaxed, arms by sides as the technician negotiates around the subject to take measurements. The technician may ask participants to move to provide access to measurement sites. Once located the technician will raise the skinfold with thumb and forefinger in a pinching motion, taking care to not grasp muscle, callipers will be applied in proximity to the raised fold and measurements taken. All skinfold measurements are taken in a cyclic manner. The following locations will be measured: back of the upper arm, front of the upper arm, below the shoulder blade on the back, on the side of the body below the armpit, the side of the body just above the hip, the front of the hip, the stomach just beside the belly button, the front of the thigh, and the side of the calf.

Circumferences

Circumferences will be assessed using a Rosscraft Retractable metallic measuring tape. Care will be taken to make the tape snug, but not compress underlying tissue.

Arm girth relaxed – measured at the midpoint of the upper arm.

Arm girth flexed- measured at the midpoint of the upper arm while the arm is flexed.

Waist/Abdominal girth – measured at end of a normal breath around the waist

Gluteal girth – measured around the torso below the hips while standing

Proximal Thigh Girth – measured just below where the buttocks ends

Mid Thigh – middle of the thigh

Calf girth – approximately the middle of the calf while standing with the foot placed on chair creating a 90 degree angle.

Breadths

Breadths will be assessed using a Rosscraft bone calliper. Bones will be palpated with middle finger prior to applying calliper faces to surface. The following locations will be measured: elbow, wrist, ankle and knee.

Lengths

Lengths will be assessed using a Rosscraft Retractable metallic measuring tape at the following locations: thigh, shin, upper arm, and forearm.

Combined Anthropometric and Dual Energy X-Ray Absorptiometry (DXA) assessment

(25 mins per anthropometric assessment & 15mins per DXA assessments = 40 mins total)

The combined anthropometric – DXA assessment will be conducted at the Richardson Centre for Functional Foods and Neutraceuticals on the Fort Garry campus of the University of Manitoba. This assessment will consist of an anthropometric assessment conducted prior to a DXA scan (see above for a full description of the Anthropometric Protocol). The DXA scan itself takes about 7-10 minutes.

The DXA scan uses a very low dose of two distinct x-ray signals. These signals are passed through your body while you lay still on an open bed. The scanning arm calculates the amount of x-rays absorbed by different tissues allowing quantification of the composition of your body. The dose of x-ray is extremely small and safe ranging from .02 to 1.5mrem. Put into context, you would need 95 DXA scans to reach the energy exposure of a single x-ray. Similarly, you would be exposed to more x-rays by flying from Vancouver to St. Johns.

Combined Anthropometric and Dual Energy X-Ray Absorptiometry (DXA) Precision assessment
(25 mins per anthropometric assessment & 15mins per DXA assessments x 2 assessments = 55 mins total)

The procedures for the DXA precision group will be identical to the previously mentioned protocol which includes both and anthropometric and DXA scan, but will require you to get scanned by the DXA machine twice instead of once. In this situation you will begin your testing session by getting a DXA scan, followed by an anthropometric assessment, and finished with an additional DXA scan. In terms of additional time requirements this will add approximately an additional 15 minutes to your assessment. For this component of the study I will need 30 athletes.

Volunteers may choose to take part in one of 4 combinations of research components for the study. Please indicate with an **X** which component you wish to be a part of;

- ☐ Participate ONLY in establishment of Anthropometric Precision
- ☐ Participate ONLY in the Combined Anthropometric and Dual Energy X-Ray Absorptiometry (DXA) assessment
- ☐ Participate ONLY in the Combined Anthropometric and Dual Energy X-Ray Absorptiometry (DXA) Precision assessment
- ☐ Participate in ALL components including Anthropometric Precision, as well as the Anthropometric – Dual Energy X-Ray Absorptiometry (DXA) Precision assessment

Benefits

Potential benefits for participating in this study are the opportunity to get a free body composition assessment that not many people have access to and in the event they do, it would cost anywhere from $80-$200 depending on where the assessment is done. As well, participation in this study has the potential to contribute to our understanding of how to assess changes in female athletes composition, with the possibility of changing the way in which athletes are assessed in the future, and how assessment results are interpreted to maximize performance.

Risks

There are minimal risks involved in participating in this study. The radiation from dual energy X-ray absorptiometry is very low dosage (equivalent to approximately 1 day of natural background radiation). The dosage is 1000 times less than the limit for trivial exposure. To further safeguard the subjects, pregnancy tests will be taken on the day of assessment; screening questions will be used to identify participation contraindication.

Confidentiality

All data will be coded and names will not be revealed at any time. Identification numbers will be used for all documentation. Data will be stored in a secure hard and digital form. The hard copy will be kept in a locked filing cabinet in a locked room accessible only by the primary researcher and authorized individuals. Digital files will be password protected on a personal computer of the primary researcher, and will only include identification numbers and not names. The key code for subjects' names and identification numbers will be kept in a different filing cabinet. Subject information will be destroyed 7 years after collection.

Feedback

Participants will receive a copy of their DXA body composition results at the testing session. Also, participants indicating that they would like to receive the study results (on general intake questionnaire) or any other time, will receive a written summary of individual and group results by mail after the study has been completed and data has been analyzed.

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the researchers, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from the study at any time, and /or refrain from answering any questions you prefer to omit, without prejudice or consequence. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

Daryl Hurrie, Principal Investigator xxx
Dr. Michelle Porter, Primary Advisor xxx

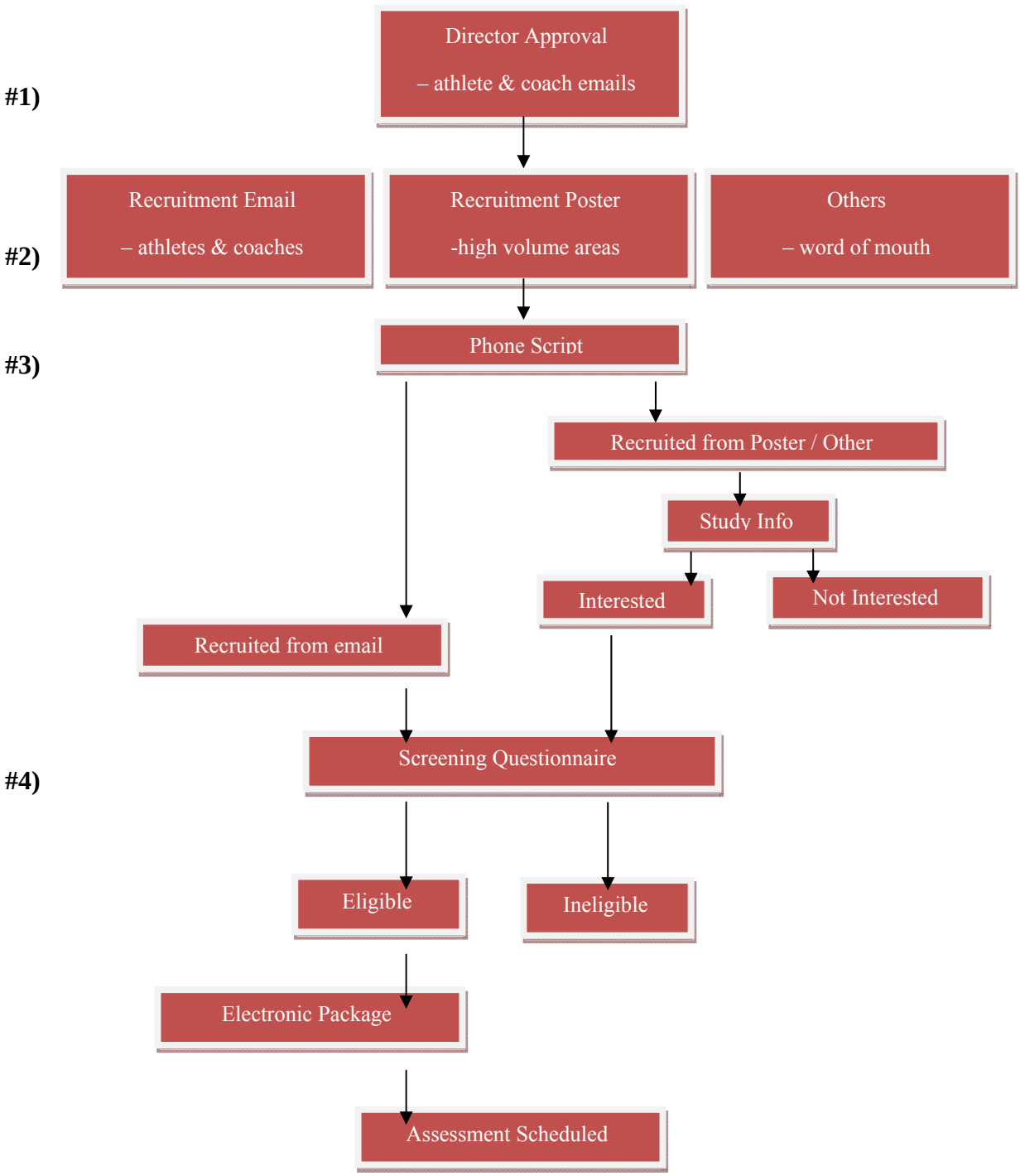
This research has been approved by the Education/Nursing Research Ethics Board. If you have any concerns or complaints about this project you may contact any of the above-named persons or the Human Ethics Secretariat at 474-7122, or e-mail margaret_bowman@umanitoba.ca. A copy of this consent form has been given to you to keep for your records and reference.

_____
Participant's Signature Date

_____
Researcher and/or Delegate's Signature Date

Appendix B

Subject Recruitment Flow Chart



Subject Recruitment Flowchart Explanation

1) Initial Recruitment Approval

University Athletic Directors will be asked to forward the email recruitment script to email lists which include:

- individual or team female athlete emails

CSCM Athlete Services manager will be asked to forward the email recruitment script to their email lists which include:

- female athlete emails

2) Initial Recruitment

Upon approval, "recruitment email" will be forwarded to athletes to make them aware of the study. Recruitment posters will be posted in high traffic areas and on post boards near athlete training and competing facilities on campus.

3) Responses to Initial Recruitment / Initial Inquiries

Phone script will be used to convey study details and determine the eligibility of individuals responding to email, poster, or other source. Accordingly, phone script will be read to screen the individual for participation eligibility, provide more study information, or acquire contact information so that these conversations can be completed at a more appropriate time.

4) Screening for Participant Eligibility

If **ineligible**, reasons for ineligibility will be explained.

If **eligible**, a digital package will be sent out for the potential volunteer to review. (If necessary arrangements can be made to provide a paper package.) The package will include the informed consent document and training history questionnaire. Upon call-back, volunteers will be asked to articulate what component of the study they wish to participate in, and schedule the pickup of their sample container, and pregnancy test along with their assessment time. Consent forms and questionnaires will be collected at the pickup time or at the time of assessment.

Appendix C

Email Recruitment Script

Script:

RE: Thesis Research Subjects Needed - An Applied Approach to Defining Within Subject Body Composition Change

Hi,

My name is Daryl Hurrie, I am a graduate student in the Faculty of Kinesiology and Recreation Management, at the University of Manitoba. Currently I am doing research to fulfill my thesis requirements for my masters' degree.

Your email address was on the ___name of institute____ athlete list serve which means you may be eligible to participate in my study. If you are interested or have time, please read over the study outlined below.

Purpose

The purpose of my study is to investigate existing body composition techniques focusing on how well methods using skinfolds, girths, breadths and lengths work compared to dual energy x-ray absorptiometry (DXA), the gold standard. By assessing existing methods, I will be able to create, validate, and disseminate the best practical methods and recommendations to practitioners trying to track body composition changes in athletes.

Population

The study requires 105 female athletes 18-30 years old, competing at an elite level. I am currently recruiting athletes from University CIS sports teams as well as the Canadian Sport Centre Manitoba.

Study Components

Volunteers may choose to take part in one of 4 combinations of research components for the study. Research assessments will be conducted at the University of Manitoba campus at either the Frank Kennedy Performance lab, or the Richardson Centre for Functional Foods and Neutraceuticals located in the University of Manitoba Smartpark. Below is a list of the study options available and the time required to complete each component

☐ Participate ONLY in establishment of Anthropometric Precision

- *(25 mins per anthropometric assessment x 2 assessments = **50 mins total**)*

☐ Participate ONLY in the Combined Anthropometric and Dual Energy X-Ray Absorptiometry (DXA) assessment

- *(25 mins per anthropometric assessment & 15mins per DXA assessments = **40 mins total**)*

☐ Participate ONLY in the Combined Anthropometric and Dual Energy X-Ray Absorptiometry (DXA) Precision assessment

- *(25 mins per anthropometric assessment & 15mins per DXA assessments x 2 assessments = **55 mins total**)*

☐ Participate in ALL components including Anthropometric Precision, as well as the Anthropometric – Dual Energy X-Ray Absorptiometry (DXA) Precision assessment

- *(50 mins for the anthropometric precision & 55mins for the Anthropometric-Dual Energy X-Ray Absorptiometry(DXA) Precision Assessment = **105 mins total or 1hour 45 mins**.)*

Responsibilities of Volunteer

Note: All anthropometric measurements will be taken by the primary investigator Daryl Hurrie, and will require the athlete to wear a sports bra and sport shorts to allow access to the right hand side of the body where triplicate measures can be made at landmarked locations. As well, subjects will be required to take a pregnancy test (as even small amounts of x-rays may be harmful to child development) and provide a urine sample (to screen for hydration). All sampling can be done at home the night prior to testing.

Volunteers will be asked to follow the "Pre- Assessment Procedures"

Pre - Assessment Procedures

Subjects should arrive 10 minutes prior to their scheduled assessment time. Prior to testing the subject should follow the pre-assessment instructions in order to maximize the accuracy of the testing procedures.

- have filled out all necessary forms (Informed Consent / Medical Background Questionnaire)
- have fasted for a minimum of 8 hours, (as foodstuff will affect DXA reading)

- have NOT trained for a minimum of 8 hours prior to assessment
- have NOT undergone barium tests/exams, or had a nuclear medicine scan or injections with an x-ray dye within the last two weeks
- wear clothing that permits access to the measurement sites located in the "standardized anthropometric assessment" protocol (i.e., sports bra and athletic shorts)
- remove all metallic objects including zippers (as metal may affect bone density values, thus, affecting body composition values)
- be reasonably hydrated
- have brought urine sample and pregnancy test results from 1st urine sample of the day.*

* If menstruating, please indicate on medical background questionnaire and urine and pregnancy test will not be necessary prior to assessment.*

If this is something you might be interested in, or if you have any questions, do not hesitate to contact me for further information!

Daryl Hurrie
(204) xxx

Appendix D

Presentation - Recruitment Script

This script describes what will be explained during the session between the researcher and athletes.

Script:

My name is Daryl Hurrie, I'm a graduate student in the Faculty of Kinesiology and Recreation Management, at the University of Manitoba. Currently I'm doing research to fulfill my thesis requirements for my masters' degree.

I'm here today to recruit subjects for my project as well as provide you with an opportunity for an assessment that not many people have access to.

The purpose of my study is to investigate existing body composition techniques focusing on predictive methods using anthropometrics (skinfolds, girths and breadths) and comparing them to dual energy x-ray absorptiometry (DXA). My hope is that by assessing the precision, reliability, and validity of these methods, I will be able to create, validate, and disseminate the best practical methods and recommendations to practitioners trying to track body composition changes in athletes.

Potential benefits for participating in this study are the opportunity to get a free body composition assessment that not many people have access to and in the event they do, it would cost anywhere from $80-200 dollars depending on where you get assessed. As

well, the participation in this study has the potential to contribute to our understanding of how to assess changes in female athletes composition, with the possibility of changing the way in which athletes are assessed in the future, and how assessment results are interpreted.

My study will require 105 female athletes competing at an elite level. I am currently recruiting athletes from CIS sports teams as well as the Canadian Sport Centre Manitoba. If you choose to participate in the study you will be assigned to one of 3 groups. All groups will receive 2 body composition assessments. One will be an anthropometric approach that takes approximately 20 minutes and the other will be a DEXA scan that takes approximately 20 minutes as well. Both procedures will be completed on the same day and due to logistical reasons will require 90 – 120 minutes of your time.

DXA scans will be taken at Richardson Centre for Functional Foods and Neutraceuticals. Briefly, a DXA scan uses a very low dose of two distinct x-ray signals. These signals are passed through your body while you lay still on an open bed. The scanning arm quantifies the amount of x-rays absorbed by different tissues allowing us to quantify the composition of your body. The dose of x-ray is extremely small and safe. Put into context, you would need 95 DXA scans to reach the energy exposure of a single x-ray (courtesy of Richardson Neutraceuticals). You may wear loose fitting comfortable clothing for this assessment.

The anthropometric assessment will be conducted in the performance lab Rm 128 Frank Kennedy Centre. The assessment consists landmarking measurement sites with an eyeliner pencil, and then the measurement of 8 girths, 9 skinfolds, and 2 bone breadths taken two to three times. The assessment requires that you wear sports bras, and exercise shorts. You will be assessed by an experienced anthropometrist (me) who has conducted over 800 anthropometric assessments on elite athletes.

The procedures for the DXA precision group will be identical to the one I just mentioned, but will require you to get scanned by the DXA machine twice instead of once during the same assessment. In reality, this amounts to an approximate addition of 25 minutes of your time. For this component of the study I will need 30 athletes.

The procedure for the Anthropometric precision group will be the same as the first groups. The only difference is that instead of getting 2 DXA scans you will have 2 anthropometric assessments. However, these assessments will NOT be conducted consecutively, but will require you to return to the lab either later in the evening (4 hours separation), or within a 3 day period, but the sooner the better. This will require an approximate addition of 20 minutes of your time.

There are minimal risks involved in participating in this study. Both tests are of minimal invasiveness and only require your passive compliance. The exposure to extremely small doses of X-rays during the DXA scan is the only risk involved. If you choose not to participate in the study there will be no ramifications for doing so.

Furthermore, should you choose to participate, you may withdraw from the research at any time without penalty. All testing results will be kept secure and confidential, and will not be disseminated to anyone except the subject themselves. In the event that the subject would like their results they may indicate this at the time of assessment and the researched will send them a copy of the results via email. To ensure confidentiality, subjects will be assigned experiment ID numbers and all documents will refer to only the subject experiment number. Only the researcher will have the master copy linking the subject identity to their assigned experiment number. Subject information will be kept for 5 years upon completion of the study and stored in a locked filing cabinet in the researchers office which has restricted access. All data used in the analysis will be void of any personal information. If you wish to be notified of the final research findings, please inform the researcher at any time within the study. You may contact me at any time through email at xxx

At this time I will hand out consent forms which you may read in your own time. For additional questions you may contact me through email at all times. If you wish to participate in this study, please complete the consent form and contact me to schedule your assessment. If you would like to speak with me in private, please email me to set up a meeting.

I appreciate your time and hope to hear from you soon.

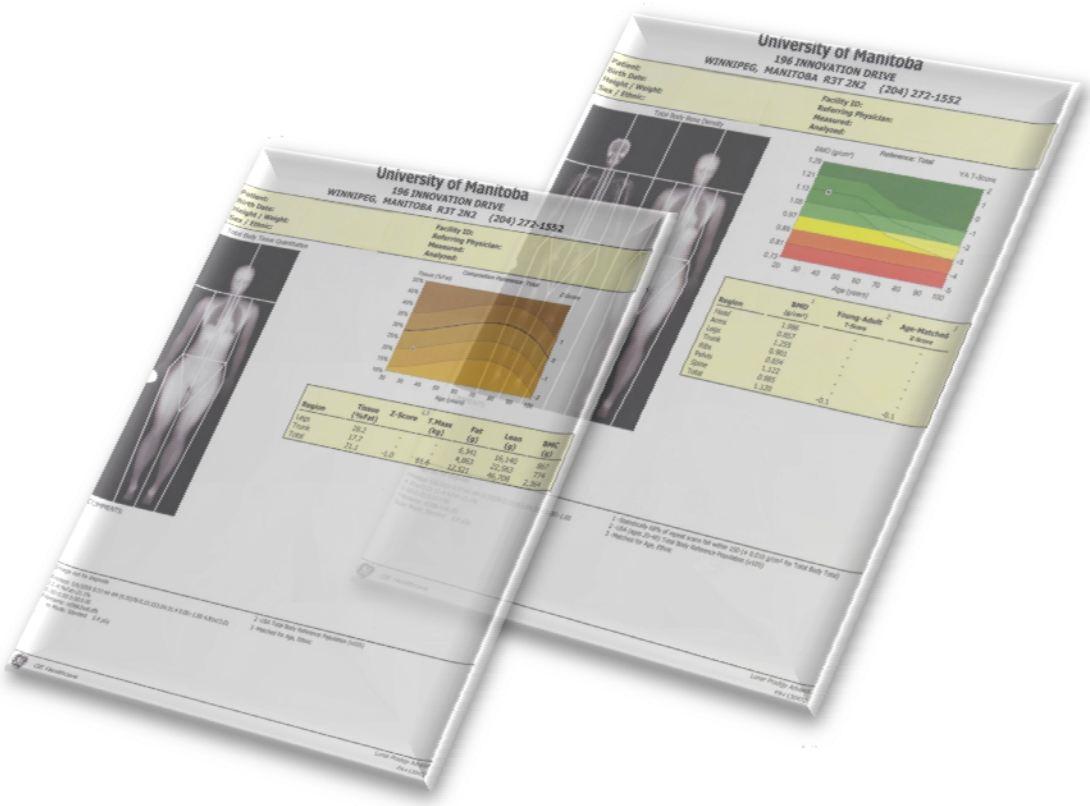
Appendix E

RESEARCH STUDY ON:

Anthropometry and DXA Methods for defining athletic body composition

Have your body composition assessed at NO CHARGE

($80-$200 value) using the latest technologies in as little as 35 minutes!



Female Athletes competing at CIS level or higher, 18-30 years, are eligible!

Contact Daryl Hurrie

University of Manitoba

XXX

xxx@cc.umanitoba.ca

Appendix F

Phone Script

Contact person: Daryl Hurrie

Script:

Introduction:

a) For potential subjects who are responding to recruitment posters or word of mouth.

*If you have approximately 5 minutes, I'd like to tell you a little bit about the upcoming research project. After, if you're still interested, I can ask you some questions to see if you are eligible to participate in the study and provide you with information on how to proceed. Feel free to ask questions at any time. Keep in mind that you are not obliged to participate in the study, and that all questions and answers are confidential. (**Read section "Information About the Study")***

b) For potential subjects responding to the recruitment email.

If you have approximately 10 minutes, I can ask you some questions to see if you are eligible to participate in the study. If you don't have time right now, perhaps you could provide me with a more suitable time to contact you and we can go through some questions which will allow me to assess your eligibility. Would you like to answer the questions right now?

If no, ask for suitable time to call back, and acquire contact number.

If yes, proceed;

*Feel free to ask questions at any time. Keep in mind that you are not obligated to participate in the study. The information being collected today will be used for screening purposes. You have the right to not answer any questions you don't want to and you may stop the interview at any time. As with all aspects of the study, all responses to questions and all personal data are confidential. **(Proceed to read & collect answers to "General Information Questionnaire")***

Information about the study:

Purpose

The purpose of my study is to investigate existing body composition techniques focusing on how well methods using skinfolds, girths, breadths and lengths work compared to dual energy x-ray absorptiometry (DXA), the gold standard. By assessing existing methods, I will be able to create, validate, and disseminate the best practical

methods and recommendations to practitioners trying to track body composition changes in athletes.

Population

The study requires 105 female athletes competing at an elite level. I am currently recruiting athletes from University CIS sports teams as well as the Canadian Sport Centre Manitoba.

Benefits

Potential benefits for participating in this study are the opportunity to get free body composition assessments that few people have access to and in the event they do, it would cost anywhere from $80-$200 depending on where you get assessed. As well, the participation in this study has the potential to contribute to our understanding of how to assess changes in female athletes composition, with the possibility of changing the way in which athletes are assessed in the future, and how assessment results are interpreted.

Risks

There are minimal risks involved in participating in this study. The dose of x-ray is extremely small and safe ranging from .02 to 1.5mrem. Put into context, you would need 95 DXA scans to reach the energy exposure of a single x-ray. Similarly, you would be exposed to more x-rays by flying from Vancouver to St. Johns.

If you choose not to participate in the study there will be no ramifications for doing so. Furthermore, should you choose to participate, you may withdraw from the research at any time without penalty. All testing results will be kept secure and confidential, and will not be disseminated to anyone except you.

If this is something you might be interested in, or have any questions please feel free to ask. If you are interested we can proceed and I will ask a few questions to see if you are eligible. Would you like to proceed now?

Appendix G

Pre - Assessment Procedures

Subjects should arrive 10 minutes prior to their scheduled assessment time. Prior to testing the subject should follow the pre-assessment instructions in order to maximize the accuracy of the testing procedures.

- have filled out all necessary forms (Informed Consent / Medical Background Questionnaire)
- have fasted for a minimum of 8 hours, (as foodstuff will affect DXA reading)
- have NOT trained for a minimum of 8 hours prior to assessment
- have NOT undergone barium tests/exams, or had a nuclear medicine scan or injections with an x-ray dye within the last two weeks
- wear clothing that permits access to the measurement sites located in the "standardized anthropometric assessment" protocol (i.e., sports bra and athletic shorts)
- remove all metallic objects including zippers (as metal may affect bone density values, thus, affecting body composition values)
- be reasonably hydrated
- have brought sample and pregnancy test results from 1st urine sample of the day as per "Urinalysis" section of methodology*

* if menstruating, please indicate on medical background questionnaire and urine and pregnancy test will not be necessary prior to assessment.*

Appendix H

Training History Questionnaire

Competition Status (please check one)

□Have competed this week

□Have competed in last 4 weeks

□Have competed 5-8 weeks ago

□Have competed 9-12 weeks ago

□Have NOT Competed in over 12 Weeks

Training Recall – *please recall your training history over the last 4 weeks (1 month) and fill in the spaces below to the best of your recollection.*

Avg. Practices Per Week_____ Avg. Duration of Practices_____ (mins)

Average intensity of practices (please circle one)

Low Low – Moderate Moderate Moderate-High High

Avg. Aerobic Sessions Per Week _____Avg. Duration of Aerobic session____ (mins)

Average intensity of aerobic sessions (please circle one)

Low Low – Moderate Moderate Moderate-High High

Average Mode of Aerobic Sessions_____ (i.e. running, ellipse, etc.)

Avg. Resistance Sessions Per Week_____ Avg. Duration of Resistance_____ (mins)

Please describe your weekly weightlifting cycle

Examples:

	4 day split	2 day split
Monday	Chest & Tri's	Upper body circuit
Tuesday	Back & Bi's	
Wednesday	Legs	
Thursday	Off	Lower Body Circuit
Friday	Shoulders	
Saturday	Off	
Sunday	Off	

Monday_____

Tuesday_____

Wednesday_____

Thursday_____

Friday _____

Saturday_____

Sunday_____

Please describe your primary training goals for the last 4 weeks (1 month) (please check one)

□Hypertrophy or muscle tissue growth (usually moderate loads, 8-12 rep range)

□Maximal Strength (usually high loads, 1-6 rep range)

□Maximal Power (usually low or high loads with explosive movements, 1-6 rep range)

□Muscular Endurance (usually low to moderate loads, >12 rep range)

□other if so, please describe_____

Appendix I

An Applied Approach to Defining Within Subject Body Composition Change

Pre-Assessment Medical Background Questionnaire

Have you followed the pre-assessment guidelines? Yes No

Are you currently taking any medications? (Includes birth control) Yes No
If "Yes" please indicate which medications you are taking

_____

_____

Have you undergone barium tests/exams in the last 14 days? Yes No

Have you had a nuclear medicine scan in the last 14 days? Yes No

Have you had injections with an x-ray dye in the last 14 days? Yes No

Have you taken the provided pregnancy test (pre-DXA)? Yes No

Results ___Positive ___ Negative

Researcher
Signature_____

TO BE FILLED OUT BY RESEARCHER

Urine Specific Gravity: Time of Collection_____ T1_____ T2_____

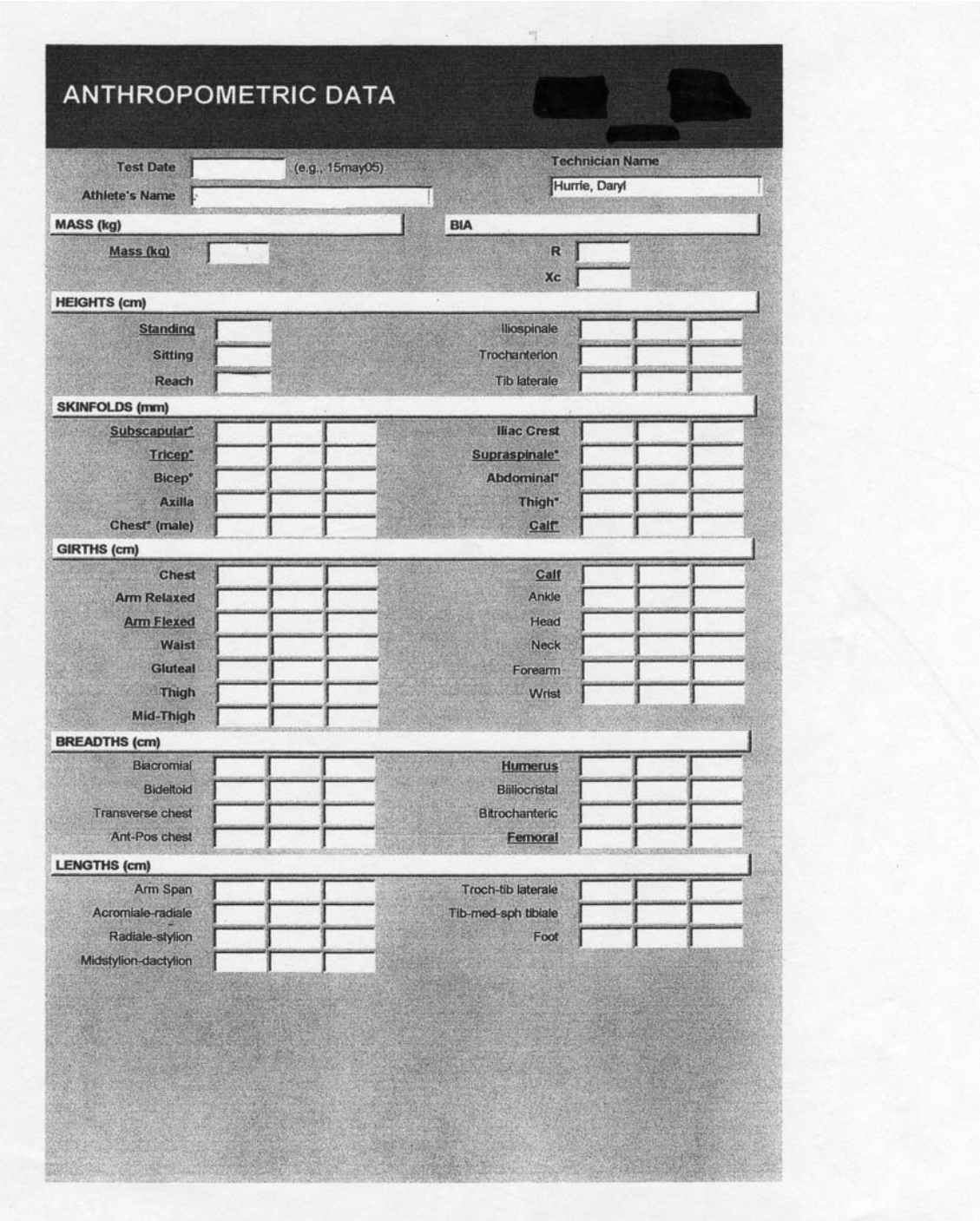
Is this person qualified to proceed in the assessment? Yes No

If not, why?

_____

Appendix J

Anthropometric Performa



ANTHROPOMETRIC DATA

Test Date (e.g., 15may05)

Technician Name: Hurrie, Daryl

Athlete's Name

MASS (kg)

Mass (kg)

BIA

R

Xc

HEIGHTS (cm)

Standing
Sitting
Reach
Iliospinale
Trochanterion
Tib laterale

SKINFOLDS (mm)

Subscapular*
Tricep*
Bicep*
Axilla
Chest* (male)
Iliac Crest
Supraspinale*
Abdominal*
Thigh*
Calf*

GIRTHS (cm)

Chest
Arm Relaxed
Arm Flexed
Waist
Gluteal
Thigh
Mid-Thigh
Calf
Ankle
Head
Neck
Forearm
Wrist

BREADTHS (cm)

Biacromial
Bideltoid
Transverse chest
Ant-Pos chest
Humerus
Biiliocristal
Bitrochanteric
Femoral

LENGTHS (cm)

Arm Span
Acromiale-radiale
Radiale-stylion
Midstylion-dactylion
Troch-tib laterale
Tib-med-sph tibiale
Foot

Appendix K

Regression equations created from various methods, and combinations of anthropometric measurements.

Derivation Method	Equation
dh#1 fwd & bck step.var.	% BF DXA = -23.683 + (0.761 * SSpMed3) + (0.279 * ThiMed3) + (0.345 * GluMed3)
dh#2 best subsets model #5	% BF DXA = -17.680-(.219*AxMed3)+(.841*SspMed3)+(.257*ThiMed3)+(.352*GluMed3)-(.250*Age)
dh#3 sum 5 Ind. Entry	% BF DXA = 6.901+(.0970*SSc)+(.368*tri)+(.231*Bi)+(.314*IC)+(.292*Calf)
dh#4 sum 5	% BF DXA = 7.145 + (0.274 * sum5)
dh#5 sum 5 & age	% BF DXA = 13.889 + (0.255 * sum5) - (0.259 * Age)
dh#6 log (10) sum 5	% BF DXA = -48.909 + (41.091 * Log10(sum5)
dh#7 log (10) sum 5 & age	% BF DXA = -39.170 + (38.549 * log10(col(88))) - (0.241 * Age)
dh#8 quadratic sum 5 *	% BF DXA = 16.386 + (0.00187 * sum5sqr)
dh#8 quadratic sum 5 & age*	% BF DXA = 23.538 + (0.00173 * sum5sqr) - (0.303 * Age)
dh#9 polynomial sum 5 *	% BF DXA= -2.889 + (0.587 * sum5) - (0.00227 * sum5sqr)
dh#10 poly. sum 5 & age	% BF DXA= 3.958 - (0.00216 * sum5sqr) + (0.553 * sum5) - (0.243 * Age)

Appendix L

Analysis of Assumption Checking & Outlier Identification for the various new regression equations

	Assumption Tests				Outlier Diagnostics	
Derivation Method	**MCL**	**Durbin-Watson**	**Normality**	**Constant Variance**	**Std Res**	**Stud Del Res**
dh#1 fwd & bck step.var.	yes	pass	pass	fail	3	3,30,31,32,90
dh#2 best subsets model #5	yes	pass	pass	pass	3,67	3,52,67,72
dh#3 sum 5 Ind. Entry	yes	pass	fail	pass	3	3,67,73
dh#4 sum 5	no	pass	fail	pass	3	3,52,67,72
dh#5 sum 5 & age	no	pass	fail	pass	3,67	3,19,52,67,72, 73,
dh#6 log (10) sum 5	no	pass	pass	fail	3	3
dh#7 log (10) sum 5 & age	no	pass	fail	pass	3	3,52,67,73
dh#8 quadratic sum 5 *	no	pass	pass	pass	3,72	3,52,60,67,72
dh#8 quadratic sum 5 & age*	no	pass	pass	pass	3,67	3,52,67,72
dh#9 polynomial sum 5 *	no	pass	pass	pass	3	NA
dh#10 poly. sum 5 & age	no	pass	fail	pass	3	3,67,73

Appendix M

APPLIED COMPOSITION ANALYSIS - WORKSHEET

Practitioners should follow the steps below and utilize the tools in the appendices to provide interpretation to body composition in female athletes.

OBJECTIVES

1. Establish TEM values using desired anthropometric performa as per ISAK guidelines
2. Calculate or use attached charts to identify magnitude of change in summed skinfolds necessary to denote “true change” as designated by SRD (95% confidence level). In addition calculate or use attached charts to identify magnitude of change in summed skinfolds necessary to denote “true change” has occurred with 90%, or 80% confidence, as these are still important indices of change in applied settings.
3. Use the reliability and likelihood of change scores to interpret the % BF results.
4. Only test over time intervals when it is likely that these magnitudes of change have been met.

APPLIED WORKSHEET

1) For this step you will need two values. Record your %TEM value for your sum5 here_____%. In addition you will need the mean of your sum5 skinfolds for the group of subjects assessed to establish your TEM here_____ (mm).

2) Using these two values, and the chart below, highlight your %TEM value or the closest value on the left axis denoted by numbers 1-10, to identify the corresponding co-ordinate for your reliability.

		Mean sum5 skinfolds												
		25	30	35	40	45	50	55	60	65	70	75	80	85
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
0.5%	**1**	0.13	0.15	0.18	0.20	0.23	0.25	0.28	0.30	0.33	0.35	0.38	0.40	0.43
1.0%	**2**	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85
1.5%	**3**	0.38	0.45	0.53	0.60	0.68	0.75	0.83	0.90	0.98	1.05	1.13	1.20	1.28
2.0%	**4**	0.50	0.60	0.70	0.80	0.90	1.00	1.10	1.20	1.30	1.40	1.50	1.60	1.70
2.5%	**5**	0.63	0.75	0.88	1.00	1.13	1.25	1.38	1.50	1.63	1.75	1.88	2.00	2.13
3.0%	**6**	0.75	0.90	1.05	1.20	1.35	1.50	1.65	1.80	1.95	2.10	2.25	2.40	2.55
3.5%	**7**	0.88	1.05	1.23	1.40	1.58	1.75	1.93	2.10	2.28	2.45	2.63	2.80	2.98
4.0%	**8**	1.00	1.20	1.40	1.60	1.80	2.00	2.20	2.40	2.60	2.80	3.00	3.20	3.40
4.5%	**9**	1.13	1.35	1.58	1.80	2.03	2.25	2.48	2.70	2.93	3.15	3.38	3.60	3.83
5.0%	**10**	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50	3.75	4.00	4.25

%TEM

Ex. A TEM of 1% is denoted by row 2.

On the horizontal access locate and highlight the mean sum5 or the closest value denoted by letters A-M, to identify the corresponding co-ordinate for your reliability.

Ex. Sample mean for sum5 = 55mm is denoted by letter G.

Where the two co-ordinates intersect the value represents your TEM in units of measurement. Check to ensure that this value is similar to your calculated TEM value. Combine the two co-ordinates (vertical and horizontal axis) to get a single alphabetical-numerical (alpha-numeric) co-ordinate.

Ex. TEM sum5 co-ordinate = G and a 1% TEM co-ordinate = 2, therefore, combined reliability co-ordinate is G2.

Highlight the numbers associated with your alpha-numeric co-ordinate for each category of alternative confidence intervals (Alternative confidence intervals table indicating sum5 change values in mm). These highlighted values represent the smallest real difference or

change in (mm) that must be exceeded for each level of certainty when interpreting two serial measurements.

Alternative confidence intervals table indicating sum5 change values in mm

		A	B	C	D	E	F	G	H	I	J	K	L	M
POSSIBLE 68% chance, 2:1 odds	**1**	0.18	0.21	0.25	0.28	0.32	0.35	0.39	0.42	0.46	0.49	0.53	0.56	0.60
	2	0.35	0.42	0.49	0.56	0.63	0.71	0.78	0.85	0.92	0.99	1.06	1.13	1.20
	3	0.53	0.63	0.74	0.85	0.95	1.06	1.16	1.27	1.37	1.48	1.59	1.69	1.80
	4	0.71	0.85	0.99	1.13	1.27	1.41	1.55	1.69	1.83	1.97	2.12	2.26	2.40
	5	0.88	1.06	1.23	1.41	1.59	1.76	1.94	2.12	2.29	2.47	2.64	2.82	3.00
	6	1.06	1.27	1.48	1.69	1.90	2.12	2.33	2.54	2.75	2.96	3.17	3.38	3.60
	7	1.23	1.48	1.73	1.97	2.22	2.47	2.71	2.96	3.21	3.45	3.70	3.95	4.19
	8	1.41	1.69	1.97	2.26	2.54	2.82	3.10	3.38	3.67	3.95	4.23	4.51	4.79
	9	1.59	1.90	2.22	2.54	2.86	3.17	3.49	3.81	4.12	4.44	4.76	5.08	5.39
	10	1.76	2.12	2.47	2.82	3.17	3.53	3.88	4.23	4.58	4.94	5.29	5.64	5.99
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
LIKELY 80% chance, 4:1 odds	**1**	0.23	0.27	0.32	0.36	0.41	0.45	0.50	0.54	0.59	0.63	0.68	0.72	0.77
	2	0.45	0.54	0.63	0.72	0.81	0.91	1.00	1.09	1.18	1.27	1.36	1.45	1.54
	3	0.68	0.81	0.95	1.09	1.22	1.36	1.49	1.63	1.76	1.90	2.04	2.17	2.31
	4	0.91	1.09	1.27	1.45	1.63	1.81	1.99	2.17	2.35	2.53	2.72	2.90	3.08
	5	1.13	1.36	1.58	1.81	2.04	2.26	2.49	2.72	2.94	3.17	3.39	3.62	3.85
	6	1.36	1.63	1.90	2.17	2.44	2.72	2.99	3.26	3.53	3.80	4.07	4.34	4.62
	7	1.58	1.90	2.22	2.53	2.85	3.17	3.48	3.80	4.12	4.43	4.75	5.07	5.38
	8	1.81	2.17	2.53	2.90	3.26	3.62	3.98	4.34	4.71	5.07	5.43	5.79	6.15
	9	2.04	2.44	2.85	3.26	3.67	4.07	4.48	4.89	5.29	5.70	6.11	6.52	6.92
	10	2.26	2.72	3.17	3.62	4.07	4.53	4.98	5.43	5.88	6.34	6.79	7.24	7.69
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
LIKELY 90% chance, 9:1 odds	**1**	0.29	0.35	0.41	0.47	0.52	0.58	0.64	0.70	0.76	0.82	0.87	0.93	0.99
	2	0.58	0.70	0.82	0.93	1.05	1.17	1.28	1.40	1.51	1.63	1.75	1.86	1.98
	3	0.87	1.05	1.22	1.40	1.57	1.75	1.92	2.10	2.27	2.45	2.62	2.80	2.97
	4	1.17	1.40	1.63	1.86	2.10	2.33	2.56	2.80	3.03	3.26	3.50	3.73	3.96
	5	1.46	1.75	2.04	2.33	2.62	2.91	3.20	3.50	3.79	4.08	4.37	4.66	4.95
	6	1.75	2.10	2.45	2.80	3.15	3.50	3.84	4.19	4.54	4.89	5.24	5.59	5.94
	7	2.04	2.45	2.85	3.26	3.67	4.08	4.49	4.89	5.30	5.71	6.12	6.52	6.93
	8	2.33	2.80	3.26	3.73	4.19	4.66	5.13	5.59	6.06	6.52	6.99	7.46	7.92
	9	2.62	3.15	3.67	4.19	4.72	5.24	5.77	6.29	6.82	7.34	7.86	8.39	8.91
	10	2.91	3.50	4.08	4.66	5.24	5.83	6.41	6.99	7.57	8.16	8.74	9.32	9.90
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
VERY LIKELY (SRD) 95% chance, 19:1 odds	**1**	0.35	0.42	0.48	0.55	0.62	0.69	0.76	0.83	0.90	0.97	1.04	1.11	1.18
	2	0.69	0.83	0.97	1.11	1.25	1.39	1.52	1.66	1.80	1.94	2.08	2.22	2.35
	3	1.04	1.25	1.45	1.66	1.87	2.08	2.29	2.49	2.70	2.91	3.12	3.32	3.53
	4	1.39	1.66	1.94	2.22	2.49	2.77	3.05	3.32	3.60	3.88	4.16	4.43	4.71
	5	1.73	2.08	2.42	2.77	3.12	3.46	3.81	4.16	4.50	4.85	5.19	5.54	5.89
	6	2.08	2.49	2.91	3.32	3.74	4.16	4.57	4.99	5.40	5.82	6.23	6.65	7.06
	7	2.42	2.91	3.39	3.88	4.36	4.85	5.33	5.82	6.30	6.79	7.27	7.76	8.24
	8	2.77	3.32	3.88	4.43	4.99	5.54	6.09	6.65	7.20	7.76	8.31	8.86	9.42
	9	3.12	3.74	4.36	4.99	5.61	6.23	6.86	7.48	8.10	8.73	9.35	9.97	10.60
	10	3.46	4.16	4.85	5.54	6.23	6.93	7.62	8.31	9.00	9.70	10.39	11.08	11.77

3) Final interpretation is made by reporting to client the qualitative likelihood that the observed change is true, as indicated by the far left axis.

> Ex. Assume two successive measurements were taken and the sum5 was calculated for both. The first sum5 was 55 mm; the second sum5 was 56.2 mm. The change between the two successive measures is 1.2 mm (56.2-55 = 1.2). Looking at your alternative confidence interval the client would be told that it is likely there has been an increase in fatness. If near absolute certainty is needed, then the client would have had to change 1.52 mm, and would have needed to be assessed 56.6 mm upon the second test.

Similarly, to assess the % BF change necessary when using the dh#9 equation for alternative confidence intervals use the your alpha-numeric co-ordinate and the "Alternative Confidence Intervals Table & % BF detectable using dh#9 equation" below.

> Ex. If we want to assess body composition changes that are likely with 90% certainty in our subjects (mean sum5 =55, 1% TEM) then the magnitude of change necessary in body fat will have to be approximately 0.43 or just under ½ a percent.

Alternative Confidence Intervals Table & % BF detectable using dh#9 equation

		A	B	C	D	E	F	G	H	I	J	K	L	M
POSSIBLE 68% chance, 2:1 odds	**1**	0.08	0.10	0.11	0.11	0.12	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.12
	2	0.17	0.19	0.21	0.23	0.24	0.25	0.26	0.26	0.27	0.26	0.26	0.25	0.24
	3	0.25	0.29	0.32	0.34	0.36	0.38	0.39	0.40	0.40	0.39	0.39	0.37	0.35
	4	0.33	0.38	0.42	0.45	0.48	0.50	0.52	0.53	0.53	0.52	0.51	0.49	0.47
	5	0.42	0.47	0.52	0.57	0.60	0.63	0.65	0.66	0.66	0.65	0.64	0.61	0.58
	6	0.50	0.57	0.63	0.68	0.72	0.75	0.77	0.78	0.79	0.78	0.76	0.73	0.69
	7	0.58	0.66	0.73	0.79	0.84	0.87	0.90	0.91	0.91	0.90	0.88	0.85	0.80
	8	0.66	0.76	0.84	0.90	0.96	1.00	1.02	1.04	1.04	1.03	1.00	0.96	0.91
	9	0.75	0.85	0.94	1.01	1.07	1.12	1.15	1.16	1.17	1.15	1.12	1.08	1.02
	10	0.83	0.94	1.04	1.13	1.19	1.24	1.27	1.29	1.29	1.27	1.24	1.19	1.12
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
LIKELY 80% chance, 4:1 odds	**1**	0.11	0.12	0.14	0.15	0.16	0.16	0.17	0.17	0.17	0.17	0.17	0.16	0.15
	2	0.21	0.24	0.27	0.29	0.31	0.32	0.33	0.34	0.34	0.34	0.33	0.32	0.30
	3	0.32	0.37	0.40	0.44	0.46	0.48	0.50	0.51	0.51	0.50	0.49	0.48	0.45
	4	0.43	0.49	0.54	0.58	0.62	0.64	0.66	0.67	0.67	0.67	0.65	0.63	0.60
	5	0.53	0.61	0.67	0.73	0.77	0.80	0.83	0.84	0.84	0.83	0.81	0.78	0.74
	6	0.64	0.73	0.81	0.87	0.92	0.96	0.99	1.00	1.00	0.99	0.97	0.93	0.88
	7	0.74	0.85	0.94	1.01	1.07	1.12	1.15	1.16	1.16	1.15	1.12	1.08	1.02
	8	0.85	0.97	1.07	1.16	1.22	1.27	1.31	1.32	1.32	1.31	1.27	1.22	1.15
	9	0.95	1.09	1.20	1.30	1.37	1.43	1.47	1.48	1.48	1.46	1.42	1.36	1.28
	10	1.06	1.21	1.33	1.44	1.52	1.58	1.62	1.64	1.64	1.61	1.57	1.50	1.41
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
LIKELY 90% chance, 9:1 odds	**1**	0.14	0.16	0.17	0.19	0.20	0.21	0.22	0.22	0.22	0.22	0.21	0.21	0.20
	2	0.28	0.31	0.35	0.38	0.40	0.42	0.43	0.44	0.44	0.43	0.42	0.41	0.39
	3	0.41	0.47	0.52	0.56	0.60	0.62	0.64	0.65	0.65	0.65	0.63	0.61	0.58
	4	0.55	0.63	0.69	0.75	0.79	0.83	0.85	0.86	0.86	0.85	0.83	0.80	0.76
	5	0.68	0.78	0.86	0.93	0.99	1.03	1.06	1.07	1.07	1.06	1.03	0.99	0.94
	6	0.82	0.94	1.03	1.12	1.18	1.23	1.26	1.28	1.28	1.26	1.23	1.18	1.11
	7	0.96	1.09	1.20	1.30	1.37	1.43	1.47	1.48	1.48	1.46	1.42	1.36	1.28
	8	1.09	1.24	1.37	1.48	1.57	1.63	1.67	1.69	1.69	1.66	1.61	1.54	1.45
	9	1.23	1.40	1.54	1.66	1.76	1.82	1.87	1.89	1.88	1.85	1.80	1.72	1.61
	10	1.36	1.55	1.71	1.84	1.94	2.02	2.07	2.09	2.08	2.04	1.98	1.89	1.77
	1	0.16	0.19	0.21	0.22	0.24	0.25	0.26	0.26	0.26	0.26	0.25	0.25	0.23
	2	0.33	0.37	0.41	0.45	0.47	0.49	0.51	0.52	0.52	0.51	0.50	0.48	0.46
	3	0.49	0.56	0.62	0.67	0.71	0.74	0.76	0.77	0.77	0.76	0.75	0.72	0.68
	4	0.65	0.74	0.82	0.89	0.94	0.98	1.01	1.02	1.02	1.01	0.99	0.95	0.90
	5	0.81	0.93	1.02	1.11	1.17	1.22	1.25	1.27	1.27	1.25	1.22	1.17	1.11
	6	0.97	1.11	1.23	1.32	1.40	1.46	1.49	1.51	1.51	1.49	1.45	1.39	1.31
	7	1.13	1.29	1.43	1.54	1.63	1.69	1.73	1.75	1.75	1.72	1.67	1.60	1.50
	8	1.29	1.47	1.63	1.75	1.85	1.92	1.97	1.99	1.98	1.95	1.89	1.81	1.69
	9	1.45	1.65	1.82	1.96	2.08	2.16	2.21	2.23	2.22	2.18	2.11	2.01	1.88
	10	1.61	1.83	2.02	2.18	2.30	2.38	2.44	2.46	2.44	2.40	2.32	2.20	2.05