Development of a Viable Bedside Ultrasound Protocol to Accurately Predict Appendicular Lean Tissue Mass

by

Michael Paris

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Masters of Science
in
Kinesiology

Waterloo, Ontario, Canada, 2016

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Background: Quantifying lean tissue or muscle mass in aging and clinical populations is of increasing importance due to emerging associations between low muscle mass and poor physical function, as well as increased rates of morbidity and mortality. Lean tissue or muscle mass can be quantified using accurate and precise modalities, such as dual-energy x-ray absorptiometry (DXA), computerized tomography (CT), and magnetic resonance imaging (MRI) scans; but these modalities have a number of practical limitations, including limited accessibility in clinical settings for body composition analysis, high cost and in some cases, radiation exposure. Ultrasound is emerging as a modality that can accurately predict muscle mass from measures of muscle thickness and may circumvent many of these limitations associated with DXA, CT and MRI. Numerous ultrasound protocols for acquiring muscle thickness measures, such as a previously developed 9-site protocol, utilize several anatomical landmarks to enhance the accuracy in prediction of muscle mass. However, these protocols: 1) may be a time-burden for clinical staff and patients, and 2) are performed in a standing posture and include posterior muscle thickness measures, which may not be feasible in many hospitalized patients who may be less mobile. Viable bedside ultrasound protocols, such as the 4-site protocol (measures the quadriceps muscle thickness), have been developed and utilized in the intensive care unit, but have yet to comprehensively assessed for accuracy in predicting muscle mass.

Objectives: The primary objectives of this thesis were to: 1) compare the agreement between the 4-site ultrasound protocol and appendicular lean tissue mass measured by DXA, 2) develop an optimized bedside-friendly protocol to predict appendicular lean tissue, using the 4-site protocol and additional accessible muscle thicknesses and easily obtained covariates, and 3) assess the ability of the optimized ultrasound protocol to identify individuals with lower than normal lean tissue mass. The secondary objectives were to: 1) compare the accuracy of predicting lean tissue mass using

minimal and maximal compression of the 4-site protocol, 2) apply the 9-site protocol in a supine posture to obtain additional accessible muscle thicknesses and to compare the accuracy of lean tissue predictions to the 4-site and optimized ultrasound protocols, and 3) assess the reliability of the 4-site protocol.

Methods: Healthy adults (≥18 years) were recruited for whole body DXA scans and ultrasound assessments on a single day. Whole body DXA scans were used to quantify appendicular lean tissue mass, the lean soft tissue in the upper and lower limbs, for each participant. Participants were identified as having lower than normal lean tissue if their appendicular lean tissue mass (kg) divided by their height (m) squared, was below previously established cut-points of 7.26 kg/m^2 and 5.45 kg/m^2 for males and females respectively. The 4-site and 9-site ultrasound protocols were performed on participants in a supine or prone position, depending on the muscle thickness measured. The 4-site protocol quantifies the muscle thickness of the rectus femoris and vastus intermedialis, at the mid-point and lower third, between the anterior superior iliac spine and the upper pole of the patella. The 9-site protocol quantifies anterior and posterior muscle thicknesses of the upper arm, trunk, upper leg and lower leg and the anterior surface of the forearm. Inter-rater and intra-rater reliability of the 4-site protocol was performed in a subset of participants.

Results: We recruited 96 participants (57% females), with a median (interquartile range) age of 36.5 (24.0-72.0) years, BMI of 24.3 (22.3-27.3) kg/m^2 and body fat of 30.2 (24.3-36.8) %. Significant differences for appendicular lean tissue mass and the 4-site muscle thicknesses were observed between males and females ($p<0.001$) and young and older adults ($p<0.001$). Regression analysis revealed a strong association between the 4-site muscle thickness and appendicular lean tissue mass, $r^2=0.72$ ($p<0.001$), but accounting for age, sex and the additional muscle thickness of the anterior upper arm, improved the association to $r^2=0.91$ ($p<0.001$). Using DXA based low lean tissue mass, 18% of participants were identified as below their sex specific cut-points. The optimized ultrasound

protocol demonstrated a strong ability (area under the curve=0.89) to identify individuals with lower than normal lean tissue mass.

Conclusions: This thesis demonstrated that a previously developed 4-site protocol strongly predicts appendicular lean tissue mass, but accounting for additional muscle thicknesses of the anterior upper arm and age and sex, greatly improves the predictive accuracy. Furthermore the optimized protocol strongly identifies individuals with lower than normal lean tissue mass. These results demonstrate that this viable bedside protocol may be useful for assessing lean tissue mass in clinical settings, but external validation in clinical populations is necessary to ensure the robustness of these findings.

Acknowledgements

I would like to begin by first thanking my advisor, Dr. Marina Mourtzakis, not only for her guidance and support throughout this thesis, but also for the numerous opportunities you have provided me to expand my knowledge and network. Secondly, I would like to thank my committee members, Dr. Heather Keller, and Dr. Jack Callaghan for your advice and guidance throughout this thesis.

I cannot be more thankful for our fantastic DXA technicians, Janice Skafel and Stephanie Auer, without your flexible scheduling, this thesis would not have been possible. I also wish to thank Benoit Lafleur, without your help in recruitment and collections, this thesis would be half of what it is. I would also like to thank Dr. Joel Dubin, for all your time and statistical expertise in enhancing the results of this thesis.

I also wish to thank my lab mates, for those weeks of long statistically discussions in our office. Lastly, I wish to express my gratitude to all participants for their time, thank you.

Table of Contents

AUTHOR'S DECLARATION ii
Abstract iii
Acknowledgements vi
Table of Contents vii
List of Figures x
List of Tables xii
List of Abbreviations xiii
List of Body Composition Terminologies xv
Chapter 1 Thesis Overview 1
Chapter 2 Literature Review 3
2.1 Clinical importance of quantifying lean tissue 3
2.2 Body composition modalities to assess body composition 3
2.3 Established body composition modalities to identify individuals with lower than normal muscularity 9
2.4 Limitations and challenges with CT, MRI and DXA 10
2.5 Assessing lean tissue quantity: the role of ultrasound 11
2.5.1 An overview of ultrasound 11
2.5.2 Using ultrasound to assess muscle mass 11
2.5.3 Existing protocols for measuring muscle thickness with ultrasound and estimating fat-free, lean tissue or muscle mass 12
2.5.4 Identifying lower than normal lean tissue or muscle mass with ultrasound 18
2.5.5 Limitations of ultrasound measures of muscle thickness 21
Chapter 3 Rationale, Objectives and Hypotheses 23
3.1 Rationale 23
3.2 Objectives 25
3.2.1 Primary objectives 25
3.2.2 Secondary objectives 26
3.3 Hypotheses 26
3.3.1 Primary hypotheses 26
3.3.2 Secondary hypotheses 27
Chapter 4 Methodology 28

4.1 Study design 28
4.2 Participants 28
4.3 Anthropometric data 29
4.4 DXA scan 29
4.5 Ultrasound protocols 31
4.6 Reliability of the 4-site protocol 34
4.7 Sample size 35
4.8 Statistical analysis 35
4.8.1 General analysis 35
4.8.2 Cross-validation analysis 36
4.8.3 Regression and ROC analysis 37
4.8.4 Bland-Altman analysis 37
Chapter 5 Results 39
5.1 Physical and demographic description of participants 39
5.2 Agreement between the 4-site protocol and DXA measures of lean tissue 42
5.3 Agreement between the 9-site protocol and DXA measures of lean tissue 44
5.4 Agreement between optimized protocol and DXA measures of lean tissue 45
5.5 Reliability of the 4-site protocol 48
Chapter 6 Discussion, Future Directions and Conclusions 50
6.1 Discussion 50
6.1.1 The 4-site protocol may have limited accuracy in predicting appendicular lean tissue mass 50
6.1.2 Distribution differences of lean tissue mass in a heterogeneous participant cohort increase the variability of 4-site protocol predictions 54
6.1.3 The 9-site protocol agrees with previous literature demonstrating high accuracy in predicting lean tissue mass 55
6.1.4 The optimized 5-site protocol can accurately predict appendicular lean tissue mass 56
6.1.5 Optimized 5-site protocol can accurately identify individuals with lower than normal lean tissue mass 57
6.1.6 The 4-site protocol demonstrates good reliability within and between raters 58
6.1.7 Limitations 60
6.2 Future directions 62

6.3 Conclusions 64
References 66
Appendix 77

List of Figures

Figure 1. General overview of whole body, molecular and tissue-organ levels of body composition, multiple compartments within a given level and modalities frequently used for assessment.

Figure 2. Schematic overview of x-ray attenuation, fundamental principle of DXA scanner**s**.

Figure 3. Pre and post CSA analysis of axial CT scans of the 3rd lumbar vertebra for skeletal muscle and adipose tissue depots. Red – muscle CSA (Hounsfield units of -29 to +150), yellow – visceral adipose CSA (Hounsfield units of -150 to -50), green – intramuscular adipose CSA (Hounsfield units of -190 to -30), subcutaneous adipose CSA (Hounsfield units of -190 to -30).

Figure 4. Transverse cross-sectional ultrasound image of the quadriceps muscle layer thickness. Rectus femoris CSA – 6.55 cm^2; muscle thickness – 4.59 cm.

Figure 5. Multiple stages of evaluation for a clinically accessible body composition modality

Figure 6. Regional assessment using DXA software for site to site comparison of muscle thickness and lean tissue mass.

Figure 7. Comparison between minimal and maximal compression protocols. 1 – Region highlighting ample acoustic coupling gel to ensure no contact occurs between skin and ultrasound probe. 2 – Highlighting convex nature of the skin and muscle belly, required for minimal compression definition. 3 – Skin and muscle belly loose convex shape during maximal compression.

Figure 8. Schematic of 3-fold cross-validation for model development and testing

Figure 9. Regression analysis between site specific lean tissue mass derived from DXA and the 4-site ultrasound protocol using A) minimal compression, r^2=0.82 and B) maximal compression, r^2= 0.66.

Figure 10. Bland-Altman plot comparing DXA derived and 4-site predicted appendicular lean tissue mass, utilizing participants from all folds

Figure 11. Bland-Altman plot comparing DXA derived and the optimized 5-site predicted appendicular lean tissue mass, utilizing participants from all folds.

Figure 12. ROC curves for folds 3, 2 and 1, utilizing the optimized 5-site protocol to identify lower than normal lean tissue mass identified by DXA cut-points

Figure 13. Bland-Altman plots for inter-rater reliability using the 4-site protocol for A) minimal and B) maximal compression.

Appendix A3. Bland-Altman plot comparing DXA derived and the 9-site protocol predicted appendicular lean tissue mass, utilizing all participants from all folds.

Appendix A4. Demonstration of a non-deal muscle thickness measured using the 4-site protocol

Appendix A5. Bland-Altman plots for intra-rater reliability using the 4-site protocol for minimal compression.

Appendix A6. Bland-Altman plots for intra-rater reliability using the 4-site protocol for maximal compression.

List of Tables

Table 1. Comprehensive overview of literature comparing ultrasound protocols to accurate and established whole body measures of fat-free, lean tissue or muscle mass.

Table 2. Physical and demographic description of participant cohort

Table 3. Descriptive summary of DXA measures of lean tissue

Table 4. Descriptive summary of ultrasound measured muscle thickness using the 9-site protocol

Table 5. Descriptive summary of ultrasound measured muscle thickness using the 4-site protocol

Table 6. Linear regression analysis to predict appendicular lean tissue using the 4-site protocol

Table 7. Multi-linear regression analysis to predict appendicular lean tissue using the optimized 5-site protocol

Table 8. Logistic regression analysis to predict presence of low appendicular lean tissue using the optimized 5-site protocol

Table 9. Intra-rater and inter-rater reliability of the 4-site protocol

Appendix A1. Comparison of multiple body composition modalities

Appendix A2. Linear regression analysis to predict appendicular lean tissue using the 9-site protocol

List of Abbreviations

ADP	Air displacement plethysmography
ALT	Appendicular lean tissue
ANOVA	Analysis of variance
AUC	Area under the curve
BIA	Bioelectrical impedance analysis
BMI	Body mass index
c-index	Concordance index
CI	Confidence interval
COPD	Chronic obstructive pulmonary disease
CSA	Cross-sectional area
CT	Computerized tomography
CV	Coefficient of variation
DXA	Dual-energy x-ray absorptiometry
FFM	Fat-free mass
ICC	Intra-class correlation coefficient
ICU	Intensive care unit
IQR	Interquartile range
MRI	Magnetic resonance imaging
ROC	Receiver operating characteristics

SEE	Standard error of the estimate
TE	Total error

List of Body Composition Terminologies

Whole body lean tissue – This tissue depot includes skeletal muscles, organs, connective tissues, skin, smooth muscles, bowel contents of the same density as lean tissue, glycogen and soft tissue minerals.

Appendicular lean tissue – Lean tissues of the upper and lower limbs. This depot consists primarily of skeletal muscle mass, with minor contributions from skin and connective tissues.

Appendicular lean tissue index – Lean tissues of the upper and lower limbs normalized by height squared (kg/m^2). Used for the identification of individuals with lower than normal lean tissue mass.

Muscle mass – Refers specifically to skeletal muscle mass, which excludes smooth and cardiac muscles.

Fat-free mass – Combination of whole body lean tissue mass and bone mass.

Chapter 1
Thesis Overview

Ultrasound has been used for a number of decades to image various structures in a clinical setting for diagnostic purposes, but it has been undervalued in its ability to assess muscle mass. Ultrasound is a noninvasive, portable and safe tool to measure muscle thickness of specific muscles or groups of muscles (1). Numerous protocols have been developed to quantify muscle thickness from multiple anatomical landmarks and have been compared against magnetic resonance imaging (MRI), dual-energy x-ray absorptiometry (DXA) and body densitometry for accuracy in predicting muscle, lean tissue or fat-free mass. However many of these ultrasound protocols are time consuming to conduct and are performed in standing posture, utilizing posterior muscle thicknesses, which may not be feasible for all clinical populations who are partially or completely immobile. Developing the ideal protocol to maintain accuracy, yet still be practical, is crucial in establishing ultrasound as a bedside tool for body composition analysis, but literature assessing the accuracy of viable bedside protocols in predicting muscle mass is lacking. In a clinical setting, accurately quantifying muscle mass and reliably tracking longitudinal changes is important in assessing risk of malnutrition or muscle atrophy, and determining success or failure of nutrition and/or exercise interventions. Thus, it is essential to assess the ability for ultrasound to predict muscle mass and identify patients with lower than normal muscle mass using a viable bedside protocol. Here, I compared a practical 4-site protocol (previously used in the intensive care unit (ICU) (2; 3), which is based on 2 bilateral sites of the quadriceps muscles) with regional and appendicular measures of lean tissue mass derived from DXA. I also evaluated other bedside accessible muscle thicknesses and easily obtainable parameters (age, sex, body mass index) that may optimize the current 4-site protocol in its ability to predict appendicular lean tissue mass. Finally, I evaluated the ability of the optimized protocol to identify

individuals with lower than normal lean tissue mass; these were based on previously established cut-points for DXA scans (4). This work will build a foundation upon which future studies can aim to validate the use of ultrasound and develop predictive equations for lean tissue mass in specific clinical populations.

Chapter 2
Literature Review

2.1 Clinical importance of quantifying lean tissue

Emerging literature has demonstrated that low lean tissue or muscle mass is associated with increased rates of mortality in a multitude of clinical populations (5–9). Although the association between mortality and lean tissue or muscle mass is an important aspect to investigate, lower than normal muscle mass has many other functional and clinical implications such as physical impairment, frailty, increased risk of falls and fractures and increased hospital length of stay and rates of readmission (5; 9). These negative implications of low lean tissue or muscle mass are observed in many populations, including diabetes (10), cancer (8; 11), chronic obstructive pulmonary disease (COPD) (12), liver cirrhosis (13), ICU (6; 14) and aged individuals (4; 15). Lower than normal lean tissue mass does not only have important implications in terms of outcomes for the patient, it also results in an increased cost and burden on the health care system (16; 17). Taken together, these factors are fundamental in the growing interest for accurately measuring or predicting lean tissue mass, and more specifically skeletal muscle in a clinical setting.

2.2 Body composition modalities to assess body composition

Lean tissue or muscle mass can be quantified from a variety of different body composition modalities (modalities compared in Appendix A1), such as bioelectrical impedance analysis (BIA), MRI, computerized tomography (CT), DXA, ultrasound, stable isotope infusions and body densitometry techniques. These modalities assess an individual's body composition by utilizing multiple compartments at a variety of different levels (Figure 1); these compartments are based on the fundamental principle of the technique and mode of analysis (i.e. hydrostatic weighing to measure body density, which is used to estimate fat mass and fat-free mass based on previously defined

molecular densities). The current benchmark techniques are considered to be DXA, CT and MRI, due to their high accuracy and reliability in assessing lean tissue or muscle mass (5).

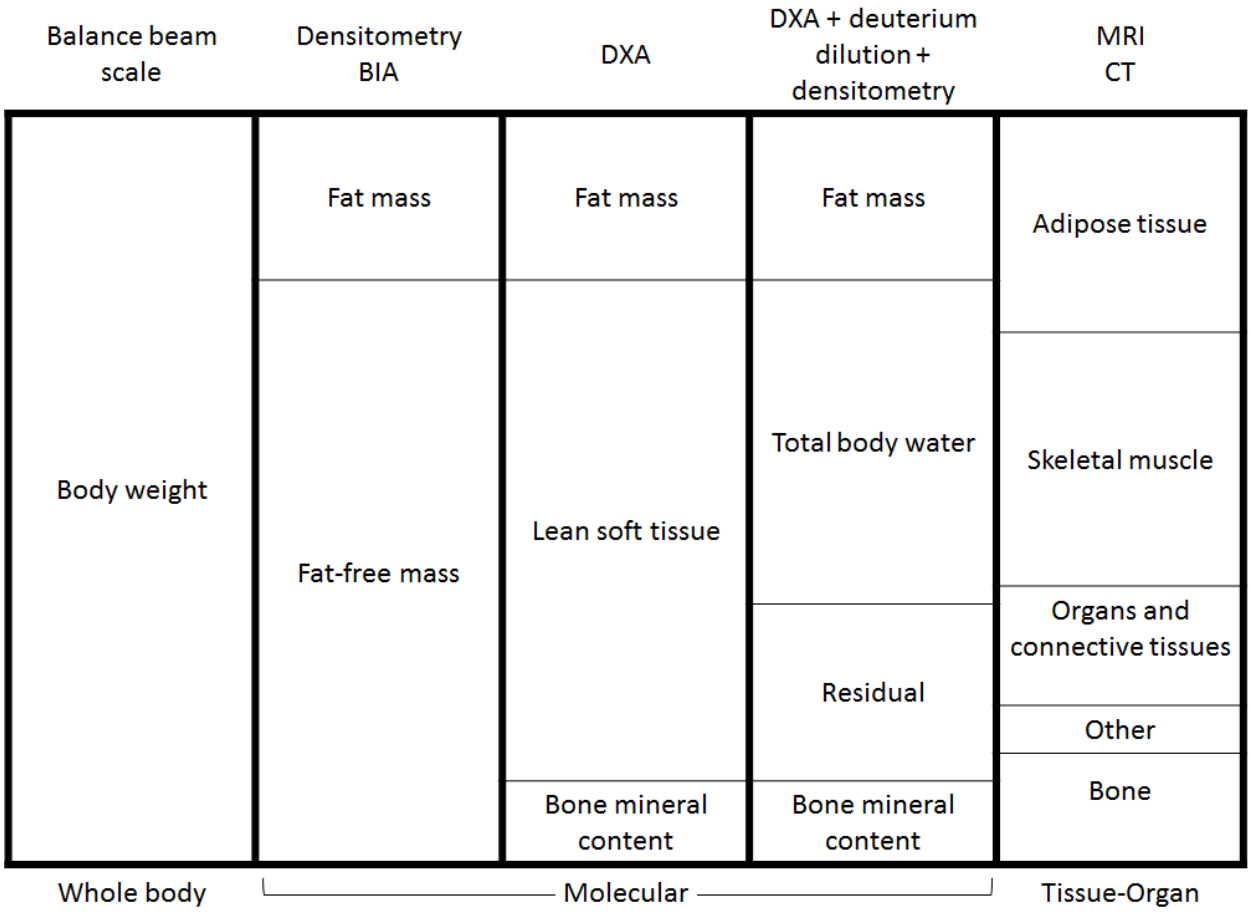


Figure 1. General overview of whole body, molecular and tissue-organ levels of body composition, the compartments within a given level and modalities frequently used for assessment.

Of these benchmark techniques, DXA is the more commonly used modality for body composition purposes in a research setting because it can provide a cost-effective whole body measurement of body composition with minimal radiation exposure to the participant. Originally developed for measuring bone mineral density in the assessment and study of osteoporosis, DXA scanners and software have been refined and are now widely accepted as accurate measures of body composition (5). DXA assesses body composition at the molecular level (Figure 1), based on a 3-

compartment model of fat mass, lean soft tissue mass and bone mineral content. Whole body lean soft tissue measured using DXA includes skeletal muscles, organs, connective tissues, skin, smooth muscles, some bowel contents, glycogen and soft tissue minerals, and therefore cannot strictly assess

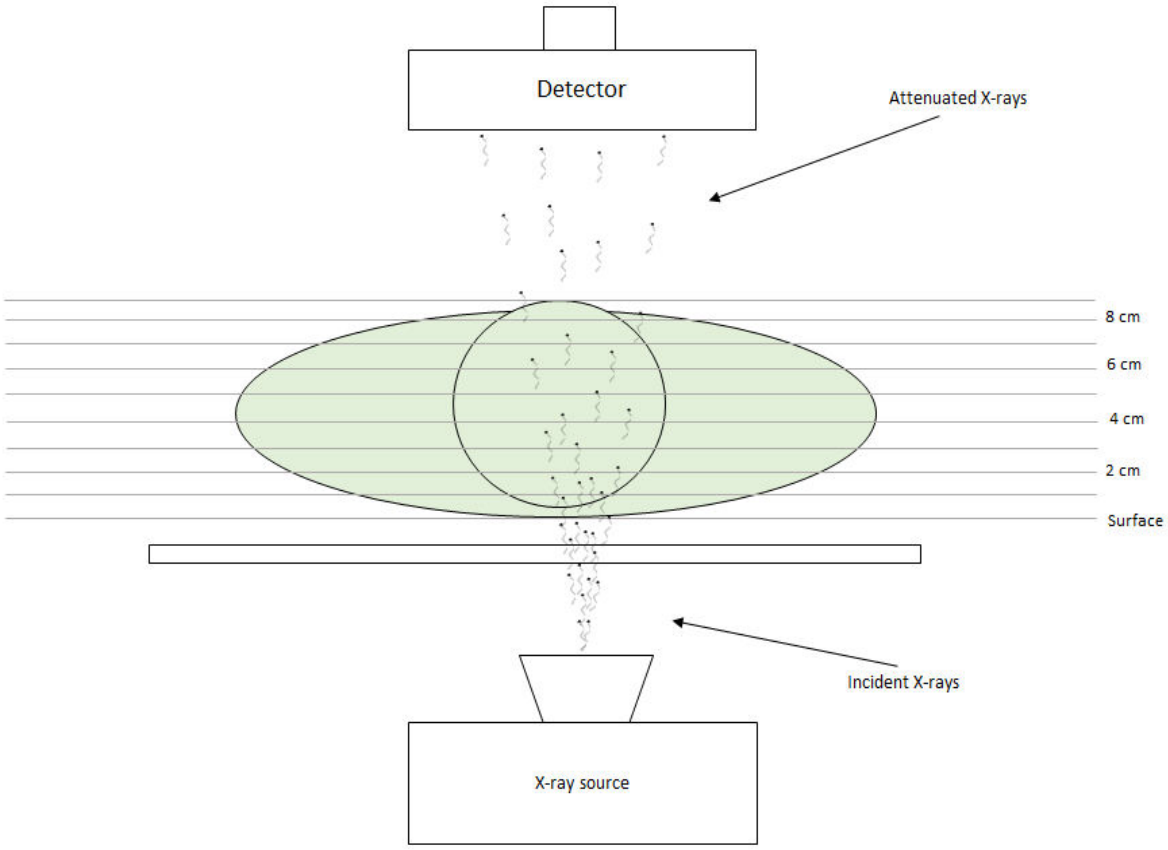


Figure 2. Schematic overview of x-ray attenuation, fundamental principle of DXA scanners.

whole body muscle mass. DXA scanners use two different energy x-rays for the analysis of body composition and is based on the fundamental principle of attenuation (absorption or scattering) of the x-ray beam as it passes through anatomical structures of different densities and thicknesses (Figure 2) (18). For example, low density tissues (muscle, liver, kidneys, fat) allow more photons to pass through (less attenuation) compared to more dense tissues (bone), where more attenuation of the x-ray beam occurs. The amount of attenuation between the two energy x-rays from a whole body scan, in conjunction with previously established attenuation coefficients for specific tissue compositions,

allows DXA to quantify fat mass, lean soft tissue and bone mineral content (19). While DXA measures body composition at the molecular level, many other accurate modalities, such as CT and MRI, evaluate body composition at the tissue-organ level, making direct comparisons difficult. Despite comparing different body composition levels (molecular vs. tissue-organ), DXA fat and lean soft tissue mass strongly associate with whole body adipose and muscle mass using CT and MRI scans in males, females, younger and older adults (20–23). However, DXA scans for individuals who are obese or have a large body thickness (>25cm), a process known as beam hardening occurs (increased attenuation of lower energy x-rays), which may lead to underestimation of fat mass in these individuals, but it is unclear if this process alters lean tissue estimations (5; 24). Therefore, although accurate for lean tissue estimates, body fat estimates may have to be interpreted with caution in obese individuals.

A major advantage of DXA is the ability to perform regional analysis, in which the body is segmented into multiple sections; the most common being the left and right upper and lower limbs, the torso and the head. Using those body segments, lean soft tissue in the upper and lower limbs, also known as appendicular lean tissue mass (25), can be quantified. Appendicular lean tissue is primarily muscle mass (with exception of skin and connective tissues) and is strongly associated with whole body muscle mass measured using MRI (23). Appendicular lean tissue has therefore been an advantageous and a commonly used measure to assess low muscularity and its associations with poor physical function in older adults (4; 26; 27). These regional assessments are also crucial when investigating lean tissue mass in certain clinical populations, such as advanced cancer patients, in which visceral lean tissue compartments (liver, spleen) can become enlarged and mask low muscle mass if only whole body lean tissue was assessed (28). DXA is also a very precise modality for assessing body composition. If care is taken to ensure optimal preparation before the scan (fasting, refrain from intense exercise) and standardization of patient positioning on the table, the coefficient of

variation (CV) for lean soft tissue, fat mass and bone mineral content is generally accepted as <2%, <1% and 2% respectively (18).

Despite the high accuracy and precision, DXA does have a number of practical limitations with its use. DXA exposes the individual to radiation, which is contra-indicated for a few populations (pregnancy), although the dose received is minor (3.5 μSv – less radiation than living a day in Toronto, ON) and is generally accepted as a safe level of exposure for both children and adults (29; 30). Taller or more obese individuals may not fit properly onto the table, and may require two half body scans for assessment. One of the bigger issues is agreement of measures between DXA manufacturers, software versions and different x-ray production techniques (fan beam, pencil beam, narrow-fan beam), as significant differences have been observed while scanning the same individuals on different scanners (31). Even with newer generation DXA scanners, there is need for cross-calibration equations and use of standard calibration phantoms to ensure good agreement across scanners and software versions (32). Although DXA scanners are becoming more widely available in both research and clinical settings (for bone mineral assessments), these scanners are rarely used for body composition purposes outside of a research setting.

Because of the poor accessibility of DXA scanners in clinical settings for body composition analysis, MRI and CT scans have been commonly used in clinical literature for body composition analysis; particularly CT scans, as they are routinely performed for clinical diagnosis in many populations such as cancer, liver cirrhosis and ICU (33–35). CT and MRI scans use high dose radiation or strong magnetic fields, to produce high resolution axial cross-sectional images and analyze body composition at the tissue-organ level, quantifying skeletal muscle or adipose tissue directly (Figure 3).

Full body CT and MRI scans can be analyzed using specialized software to quantify whole body muscle mass and adipose tissues. The cross-sectional area (CSA) of muscle or adipose tissue is

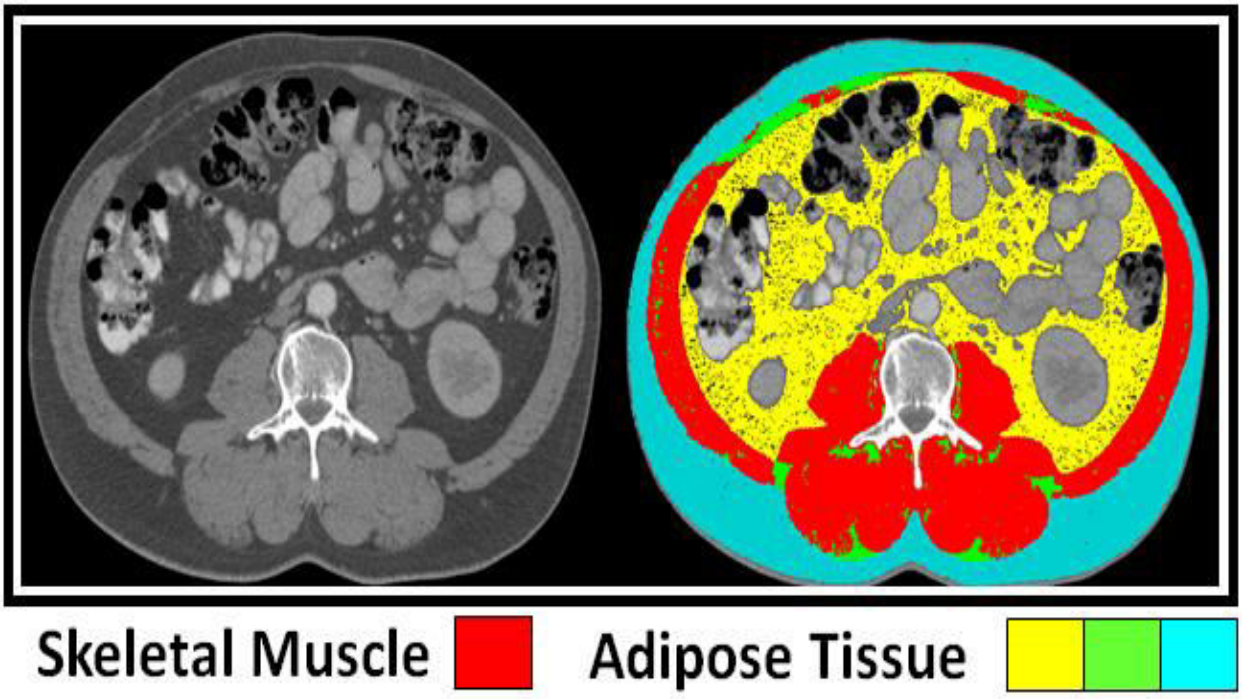


Figure 3. Pre and post CSA analysis of axial CT scans of the 3rd lumbar vertebra for skeletal muscle and adipose tissue depots. Red – muscle CSA (Hounsfield units of -29 to +150), yellow – visceral adipose CSA (Hounsfield units of -150 to -50), green – intramuscular adipose CSA (Hounsfield units of -190 to -30), subcutaneous adipose CSA (Hounsfield units of -190 to -30).

analyzed for each scan, combined with the thickness of the scan, a volume can be estimated; which has very strong associations with kilogram measures of muscle and adipose tissue from cadaver dissections (36). However, full body CT or MRI scans for diagnostic purposes are not performed as frequently as regional scans of the body, which often include the abdominal region. Taking advantage of this fact, Shen et al. (2004) examined the agreement between skeletal muscle and adipose tissue CSA from a single MRI scan at various anatomical landmarks in the abdominal region and whole body muscle and adipose mass from full body MRI scans, observing the strongest associations using scans from around the 3rd lumbar vertebra (37). A single scan of the 3rd lumbar vertebra is now a commonly used landmark for skeletal muscle and adipose tissue analysis in many clinical

populations, including cancer (38), liver cirrhosis (34) and ICU (3; 6; 39) populations. Using these technologies and expedient modes of analysis, literature investigating the associations between low muscularity and clinical outcomes has become more widely accessible and investigated.

Single slice CT or MRI scans have very precise analysis using specialized software, and have CV of 2% for muscle CSA in many clinical populations (3; 40). Although more clinically accessible than DXA in many populations, CT and MRI scan analyses still have many limitations (further discussed below in section 2.4). Analyzing the scans is time-consuming and requires specialized software, which may not be accessible in clinical settings. Moreover, CT and MRI scans are not routinely performed for body composition purposes, limiting investigations to retrospective analysis (33).

2.3 Established body composition modalities to identify individuals with lower than normal muscularity

DXA, CT and MRI scans are often used to identify individuals with lower than muscle mass. These modalities have cut-points specific to their analysis and have been applied in many clinical populations (4; 6; 34; 39; 41; 42). Baumgartner el al. (1998) developed sex-specific low lean tissue cut-points and applied those cut-points in a large (n=883) cohort of elderly Hispanic and non-Hispanic white males and females using DXA appendicular lean tissue mass normalized to their height squared (4). In that study, lower than normal lean tissue was statistically defined as being less than two standard deviations below a healthy young reference group and these cut-points were significantly associated with self-reported physical disability, independent of age, obesity, ethnicity and health behaviors in the elderly adult cohort (4). These same cut-points have also been applied in lung and colorectal cancer populations using DXA scans, which were then transferred to CT analysis using the 3rd lumbar vertebra landmark in those same patients (38). Subsequently, many studies have utilized those CT cut-points in many cancer, liver cirrhotic and ICU populations, finding a number of

associations between low muscle mass and poor clinical outcomes (11; 13; 28; 35). ICU specific cut-points using CT analysis have also been developed by performing a receiver operator characteristics (ROC) curve and observing higher rates of mortality in those with low muscle CSA (<110 cm^2 for females and <170 cm^2 for males) (39). The growing evidence supporting the significance of muscle mass or lean tissue on clinical outcomes highlights the need to further investigate the implications of interventions aimed at improving lean tissue and to determine if meaningful clinical or functional outcomes are improved.

2.4 Limitations and challenges with CT, MRI and DXA

Although DXA, CT and MRI are highly sought after due to their accuracy and precision in quantifying lean tissue or muscle mass, there are a number of practical limitations associated with their use, such as cost, availability of these scanners in clinical settings for body composition analysis, body size limits to fit within the scanner and in some cases, radiation exposure (5). These issues taken together, constrain the wide spread application of these comprehensive body composition techniques in many clinical settings. This often limits clinical investigations to anthropometric measures, such as weight and height; which are unable to quantify specific tissues, leading to variable and inaccurate measures of muscle. Even investigations utilizing CT or MRI analysis for body composition measures are generally limited to retrospective analysis, given that prospective analysis would be expensive, potentially burdensome in terms of time-commitment and, in the case of CT, would expose the patient to additional radiation. It would therefore be advantageous to use a modality that can circumvent many of these limitations and prospectively assess lean tissue or muscle mass in clinical investigations. Ultrasound may elude many of these limitations and has the ability to prospectively estimate lean tissue or muscle mass (1), while being portable, noninvasive, easy to use and readily available in most clinical settings (43).

2.5 Assessing lean tissue quantity: the role of ultrasound

2.5.1 An overview of ultrasound

Ultrasound devices generate high frequency sound waves (1-14 MHz) through an electrically stimulated piezoelectric crystal in the head of the transducer. When an alternating current is applied to the piezoelectric crystal, high frequency vibrations occur, producing ultrasonic waves. In conjunction with acoustic coupling gel, these ultrasound waves propagate through the skin and are then partly reflected and partly transmitted through the underlying subcutaneous tissues (44). The amount of reflection and transmittance that occurs is dependent upon changes in acoustic impedance and is determined by the characteristics of the underlying tissues; these reflections and transmissions occur at transitional interfaces between two different tissues (i.e. adipose-muscle interface) (44). The waves that are reflected back to the transducer are received by the piezoelectric crystals, processed based on timing, frequency and amplitude of the reflected waves, finally being displayed as a 2-dimensional image on the ultrasound screen. However, since the ultrasound waves are travelling through multiple different tissues (skin, adipose tissue, muscle), and the tissue properties change the velocity of the ultrasound waves, there is an assumption with ultrasound imaging that the wave propagates at approximately 1540 m/s across soft tissues (45). In reality, the ultrasound wave propagates at approximately 1450 m/s in adipose tissue and 1580 m/s in muscle, and may therefore overestimate adipose tissue thickness and underestimate muscle thickness (45). However, cadaver dissection has demonstrated strong agreement with ultrasound measured thickness, suggesting the velocity assumption is most likely a negligible issue (46).

2.5.2 Using ultrasound to assess muscle mass

Ultrasound can be used for muscle mass estimations by obtaining transverse cross-sectional images of predefined landmarks and then analyzing those images for the thickness or CSA of the underlying muscle groups (Figure 4). These thickness or CSA measures, in conjunction with

prediction equations, can give estimates of whole body or regional measures of lean tissue or muscle mass (47–51). While muscle CSA may provide a more comprehensive analysis of muscle architecture

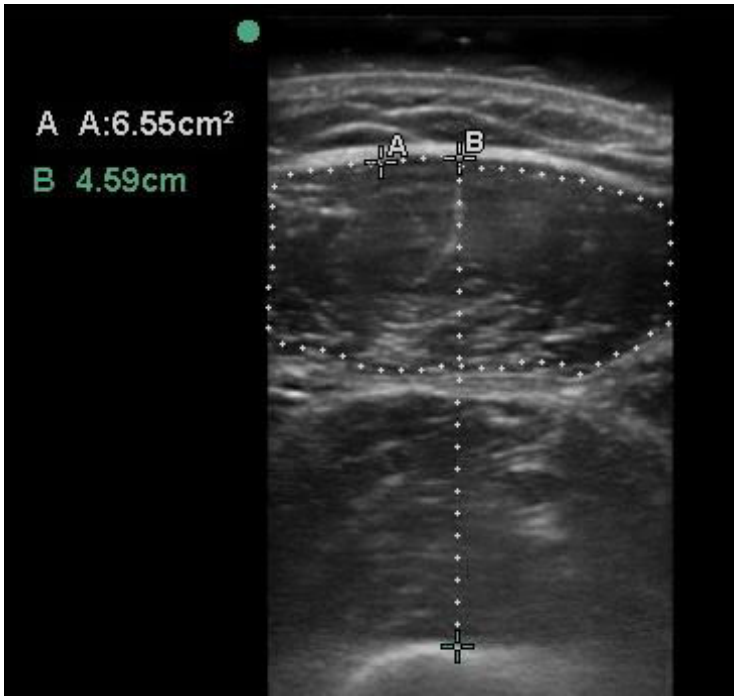


Figure 4. Transverse cross-sectional ultrasound image of the quadriceps muscle layer thickness. Rectus femoris CSA – 6.55 cm^2; muscle thickness – 4.59 cm.

compared to thickness, thickness is more commonly utilized for lean tissue mass estimations because analysis is less time-consuming, it is easier to delineate the muscle-bone interface compared to the entire fascia of the muscle and many muscle groups, such as the vastus lateralis, have CSA that may be too large to visualize in a single image using linear ultrasound probes.

2.5.3 Existing protocols for measuring muscle thickness with ultrasound and estimating fat-free, lean tissue or muscle mass

To date, many ultrasound protocols have been developed to measure either muscle thickness or CSA of a wide range of muscles using a variety of different landmarks. Some of these protocols include measuring muscle thicknesses for: 9 anterior and posterior sites (47–49; 52–54), 7 anterior and posterior sites (50; 55), 2 site bilateral of the quadriceps (2; 56–58), 4 anterior and posterior sites

on the lower limbs (59), 3 anterior sites on upper and lower limbs (60) and for measuring CSA of the: rectus femoris (12; 61; 62) and vastus lateralis (63; 64). However, only the 9 site protocol has been more extensively validated for lean tissue or muscle mass estimations (Table 1) (65).

The 9-site protocol has been compared to hydrostatic weighing (48), MRI (49; 52; 66) and DXA (47; 59; 65) measures of fat-free, lean tissue or muscle mass and has regression coefficients that range from 0.75 to 0.99 (Table 1). These regression coefficients have been observed in both males and females, younger (<60 years of age) and older (>60 years of age) adults and between different ethnicities (Japanese and Caucasian). These studies comparing the accuracy of the 9-site protocol have also performed Bland-Altman plot analysis (67) and the largest observed limits of agreement were ±4 kg when compared with either DXA lean tissue mass or MRI muscle mass. Importantly, there were no observable biases in the cross-validation groups within any given study. Taken together, there is strong agreement between the 9-site ultrasound protocol and DXA or MRI measures of lean tissue or muscle mass. However, while studies developing regression equations did not observe any bias, Abe et al. (2014) externally assessed the accuracy of four previously published regression equations using the 9-site protocol. These regression equations, which were originally developed in Japanese populations, were applied in 77 middle aged and older Caucasian adults (aged 50-78) and significant systematic or proportional bias was observed in three of the four regression equations tested, with only one equation having no bias and acceptable limits of agreement (65). Overall, ultrasound may have a high degree of accuracy for estimating lean tissue or muscle mass using the 9-site protocol, but regression equations may be population specific.

While the 9-site protocol may be accurate, it is not be feasible in many clinical settings because it requires the inclusion of posterior measures in patients who are typically in supine position, sedated/unconscious and/or difficult to move. Not to mention, that a standing protocol may result in

Table 1. Comprehensive overview of literature comparing ultrasound protocols to accurate and established whole body measures of fat-free, lean tissue or muscle mass.

Authors	Population and Sample Size	Ultrasound Protocol	Reference Modality	Agreement with Reference Method	Limitations/Merits
Abe etl al. (1994)	n=117 Males n=62 Females n=55 Age 18-51 Healthy volunteers Japanese population	**9-site muscle thickness protocol:** anterior and poster upper arm, anterior forearm, abdomen, subscapular, anterior and posterior upper and lower leg Posture: standing	Hydrostatic weighing: fat-free mass	Regression analysis Male r=0.81 SEE=4.39 kg Female r=0.75 SEE=2.37 kg	Did not consider body proportions
Campbell et al. (1995)	n=36 Males n=14 Females n=22 Age 20-74 Healthy volunteers undergoing bone mineral density test North American population	**3-site muscle thickness protocol:** anterior upper arm, anterior forearm and anterior upper leg Posture: supine	DXA scan: whole body lean soft tissue mass	Regression analysis Combined r=0.87 SEE=5.19 kg	Did not consider body proportions Performed on non-dominant side
Sanada et al. (2006)	n=72 Males n=38 Females n=34 Age 18-61 Healthy volunteers Japanese population	**9-site muscle thickness protocol:** anterior and poster upper arm, anterior forearm, abdomen, subscapular, anterior and posterior upper and lower leg Posture: standing	Whole body MRI scan: whole body muscle mass	Regression analysis Model development r= 0.98 Males SEE=2.24 kg Females SEE=2.75 kg Cross-validation r= 0.97 T-test p-value > 0.05 Non-significant difference between muscle mass estimations	Accounted for body proportions by multiplying muscle thickness by height
Utter et al. (2008)	n=70 Age 14-18 High school male wrestlers North American population	**3-site muscle thickness protocol:** posterior upper arm, subscapular and abdomen Posture: not stated	Hydrostatic weighing: fat-free mass	Regression analysis r=0.97 SEE= 2.40 kg T-test p-value > 0.05 Non-significant difference between fat-free mass estimations	Did not consider body proportions

Abe et al. (2013)	n=36 Males n=18 Females n=18 Age 19-65 Healthy volunteers North American population	**9-site muscle thickness protocol:** anterior and poster upper arm, anterior forearm, abdomen, subscapular, anterior and posterior upper and lower leg Posture: standing	DXA scan: appendicular lean soft tissue mass	Regression analysis Males r=0.93 Females r=0.93 Combined r=0.98 SEE=1.46 kg	Accounted for body proportions by multiplying muscle thickness by height
Takai et al. (2014)	n=77 Males n=33 Females n=44 Age 53-77 Healthy volunteers Japanese population	**9-site muscle thickness protocol:** anterior and poster upper arm, anterior forearm, abdomen, subscapular, anterior and posterior upper and lower leg Posture: standing	DXA scan: whole body fat-free mass	Regression analysis 4-site + sex r=0.96 SEE=2.50 kg 4-site x limb length + sex r=0.98 SEE=2.00 kg	Accounted for body proportions by multiplying muscle thickness by limb length Accounted for sex
Abe et al. (2014)	n=79 Males n=40 Females n=39 Age 50-78 Healthy volunteers Caucasian population	**9-site muscle thickness protocol:** anterior and poster upper arm, anterior forearm, abdomen, subscapular, anterior and posterior upper and lower leg Posture: standing	DXA scan: whole body and appendicular lean soft tissue mass	Regression analysis using previously published ultrasound equations [1] Males r=0.85, SEE=5.00 kg; Females r=0.84, SEE=4.20 kg [2] Males r=0.92, SEE=1.6 kg; Females r=0.90, SEE=1.4 kg [3] Males r=0.94, SEE=1.00 kg; Females r=0.94, SEE=0.80 kg [4] Males r=0.87, SEE=4.60 kg; Females r=0.81, SEE=4.50 kg	Applied Japanese developed equations in Caucasian population Accounted for body proportions by multiplying muscle thickness by limb length or height
Smith-Ryan et al. (2014)	n=47 Males n=20 Females n=27 Age 18-55 Overweight and obese (BMI $\geq$ 25 kg/m^2) North American population	**7-site muscle thickness protocol:** posterior upper arm, subscapular, abdomen, suprailiac, midaxillary, chest and anterior upper leg. Posture: standing	ADP and BIA: Total body water, body weight and body density to estimate fat-free mass	T-test p-value 0.001 Significant difference between fat-free mass estimations	Comparison methods not performed within single day

Abe et al. (2015)	n=102 Males n=59 Females n=43 Age 50-76 Healthy Caucasian population	**10-site muscle thickness protocol:** anterior and poster upper arm, anterior and ulnar forearm, abdomen, subscapular, anterior and posterior upper and lower leg Posture: standing	DXA scan: appendicular lean soft tissue mass	Regression analysis developing multiple models Age + Sex + 1 to 7-site thickness r=0.95 – 0.99 (range) SEE=1.95 – 1.13 kg (range)	Accounted for body proportions by multiplying muscle thickness by height
Midorikawa et al. (2015)	n=145 Males n=89 Females n=56 Age 6-12 Healthy volunteers Japanese population	**9-site muscle thickness protocol:** anterior and poster upper arm, anterior forearm, abdomen, subscapular, anterior and posterior upper and lower leg Posture: standing	Whole body MRI scan: whole body muscle mass	Regression analysis Model development Males r= 0.93, SEE=0.72 kg Females r=0.89, SEE=0.80 kg T-test p-value > 0.05 Non-significant difference between muscle mass estimations	Accounted for body proportions by multiplying muscle thickness by height
Ismail et al. (2015)	n=20 Age 18-75 Females n=10 low muscle n=10 normal muscle North American population	**5 site muscle thickness protocol**: trapezius, anterior upper arms, deltoid, pectoralis, rectus femoris Posture: seated	DXA scan: appendicular lean soft tissue mass index	Regression Analysis 5 site thickness r=0.64 BMI + 5-site thickness + age r=0.93 SEE=0.482 kg/m^2	Did not consider body proportions
Berger et al. (2015)	n=105 Males n=52 Females n=53 Age 20-55 (n=54) Age >60 (n=51) Healthy South American population	**2 site muscle thickness protocol:** Left and right rectus femoris Posture: seated	DXA scan: whole body and appendicular fat-free mass	Correlation analysis Whole body fat-free mass r=0.71 Appendicular fat-free mass r=0.74	Did not consider body proportions Muscle strength tests prior to muscle thickness measures
Roelofs et al. (2015)	n=36 Males n=21 Females n=15 Age Division I cross-country athletes North American population	**1 site muscle cross-sectional area:** vastus lateralis Posture: supine	DXA scan: whole body lean soft tissue mass	Correlation analysis Males r=-0.01 Females r=0.46	Did not consider body proportions

Johnson et al. (2016)	n=74 Males n=33 Females n=41 College-aged Healthy volunteers Caucasian population	**7-site muscle thickness protocol:** posterior upper arm, subscapular, abdomen, suprailiac, midaxillary, chest and anterior upper leg. Posture: not stated	DXA scan and ADP: fat-free mass	ANOVA Ultrasound-ADP FFM p-value 1.00 Non-significant difference between fat-free mass estimations Ultrasound-DXA FFM p-value <0.001 Significant difference between fat-free mass estimations	Did not consider body proportions
Kendall et al. (2016)	n=23 Age 22-30 Elite male rowers Caucasian population	**7-site muscle thickness protocol**: anterior and poster upper arm, chest, abdomen, anterior and poster upper leg and posterior lower leg Posture: not stated	DXA scan, ADP and BIA: bone mineral content, total body water and body density to estimate fat-free mass	Regression analysis r=0.93 SEE=3.80 kg	Did not consider body proportions
Schoenfeld et al. (2016)	n=20 Age 18-35 Females Healthy volunteers North American population	**4-site muscle thickness protocol:** posterior upper arm, anterior upper leg, suprailiac and abdomen Posture: standing	ADP: fat-free mass	Regression analysis Pre diet intervention r=0.87 SEE=2.68 Post diet intervention r=0.90 SEE=2.31	Did not consider body proportions
Abe et al. (1994)	n=158 Male n=72 Females n=86 Age 50-79 Healthy volunteers Japanese population	**1-site muscle thickness protocol:** ulnar forearm Posture: standing	DXA scan: appendicular lean soft tissue mass	Regression analysis using previously published equation 1-site thickness r=0.88 TE=2.60 kg 1-site thickness x limb length r=0.94 TE=1.38 kg	Accounted for body proportions by multiplying muscle thickness by limb length
Toda et al. (2016)	n=61 Young male athletes: Lacrosse, rugby, football Japanese population	**9-site muscle thickness protocol:** anterior and poster upper arm, anterior forearm, abdomen, subscapular, anterior and posterior upper and lower leg Posture: standing	Whole body MRI scan: whole body muscle mass	T-test p-value <0.001 Significant difference between muscle mass estimations using previously published Japanese based equation Regression analysis r=0.96	Accounted for body proportions by multiplying muscle thickness by height

some variability depending on the amount of time the participant is standing, which may alter fluid distribution (68). This is especially important since validation studies often use DXA, where participants are supine, as the reference method. Therefore, viable bedside protocols have been developed to assess and track changes in muscle thickness and their relation to functional and clinical outcomes (2; 12; 62; 69), but they have yet to be extensively tested for accuracy in predicting lean tissue mass. To date, two studies have compared viable bedside protocols against measures of lean tissue mass. Campbell et al. (1995) developed an ultrasound protocol that measures the muscle thicknesses of the anterior upper arm, anterior forearm and anterior upper leg in a supine position and observed in 36 healthy volunteers that these summed muscle thicknesses were strongly associated (r=0.87) with DXA measured lean tissue mass. More recently, Berger et al. (2015) compared rectus femoris muscle thickness, albeit in a seated position with the knee bent, to DXA measured appendicular and whole body lean tissue in a group of 105 younger and older adults, observing moderate associations (r=0.74). Although the latter of the two studies is not directly applicable to bedside implementation, these two studies demonstrate that readily accessible supine landmarks can produce fairly strong associations with lean tissue mass and that they may be useful surrogates of muscle mass at the bedside.

2.5.4 Identifying lower than normal lean tissue or muscle mass with ultrasound

Given the potential for ultrasound to prospectively assess muscle mass in clinical settings, it is valuable to examine whether ultrasound has the capacity to accurately identify individuals with lower than normal muscle mass. Recently, a number of studies have assessed the ability of ultrasound to discriminate between normal and low levels of lean tissue or muscle mass, using a variety of different protocols and previously established cut-points from DXA, BIA, MRI and CT measures of lean tissue or muscle mass (3; 41; 54; 70–72). Ismail et al. (2015) investigated the ability of a 5-site ultrasound protocol (in seated position) to distinguish between 10 normal and 10 lower than normal

lean tissue mass women, identified using previously published DXA cut-points. Significant differences in the summed muscle thicknesses were observed between the normal and low lean tissue groups, indicating these anterior and posterior sites (trapezius, brachioradialis, deltoid, pectorals and rectus femoris) may be useful markers of low lean tissue. However, in addition to the small sample size, no further specificity or sensitivity analysis was performed, making it difficult to draw conclusions on whether this ultrasound protocol can accurately discriminate low from normal lean tissue on an individual basis. Minetto et al. (2015) established muscle specific cut-points for the rectus femoris and vastus lateralis thickness using ultrasound in a group of 60 young participants (20-36 years of age) by taking 2 standard deviations below the mean thickness, after which they applied those cut-points and previously established BIA cut-points in a group of 44 older adults (67-93 years of age). Using rectus femoris and vastus lateralis cut-points, 86% of the older adults were identified as having lower than normal muscle thickness, whereas BIA identified between 2% and 75% of individuals as having lower than normal fat-free mass, depending on which BIA specific cut-points were used. This large discrepancy in the classification of low fat-free mass using BIA makes interpretation of these ultrasound cut-points challenging. Since there is no definitive identification of individuals with lower than normal fat-free mass, we cannot adequately assess utility of the ultrasound cut-points in identifying these individuals. However, these results are interesting, as they demonstrate that site specific (i.e. quadriceps) low muscle mass may be an important aspect to investigate. This is further supported by literature indicating a greater loss of lower limb musculature relative to upper limb musculature in aged individuals (73). Furthermore, there is evidence of preferential atrophy of the quadriceps musculature compared to the hamstrings musculature in older adults (74–76), suggesting that it may be advantageous to use a modality that can assess site specific muscle atrophy (ultrasound), compared to modalities that measure whole body or even lower limb lean tissue mass (DXA). Abe et al. (2015), used the previously developed 9-site ultrasound protocol

to predict DXA appendicular lean tissue mass and compared the agreement in identifying individuals with low lean tissue categorized by DXA cut-points. Although a small proportion of their participants had low lean tissue mass, they described that 6 of 7 participants were correctly identified as low lean tissue mass using ultrasound. Kuyumcu et al. (2016) used ultrasound assessment of the gastrocnemius muscle groups to identify sarcopenic individuals (low fat-free mass index from BIA and low handgrip strength) in a group of 100 older adults. Of the 100 individuals, 8 males and 6 females were identified as being sarcopenic, and a ROC analysis, using gastrocnemius thickness or fascicle length, resulted in an area under the curve (AUC) between 0.78 and 0.83. However, sex was not accounted for during ROC analysis, making these results difficult to interpret, since a lower than normal muscle thickness for females, should be lower than males. Conservatively, considering these data, in community dwelling older adults, the ability of ultrasound to identify low muscle is moderate (based on the ROC analysis from Kuyumcu et al. (2016)). However, with the wide range of ultrasound protocols applied, reference techniques used and statistical analyses performed, there is still much work required to ensure an optimal and consistent approach for identifying low muscle with the use of ultrasonography.

Recently, in clinical populations, ultrasound has also been used to identify lower than normal muscle mass in ICU and liver cirrhotic patients. Within the ICU population, the 4-site protocol, using maximal compression of the ultrasound probe against the skin, has been applied and compared to abdominal CT analysis for identification of low muscle in 145 mixed medical and surgical patients near ICU admission (3). A ROC analysis revealed an AUC of 0.67 for ultrasound thickness alone, which improved to 0.77 with the addition of covariates age, sex and body mass index (BMI). In liver cirrhotic patients, the same 4-site protocol was applied, but using both minimal and maximal compression of the ultrasound probe against the skin, and was compared to CT and MRI analysis in 159 patients. Minimal compression resulted in better associations with low muscle mass and in

conjunction with BMI, developed an AUC of 0.78 for males and 0.89 for females using a ROC analysis. Although few studies to date have used ultrasound to identify low muscle in clinical populations, the results thus far demonstrate that ultrasound has moderate to strong capabilities to identify lower than normal muscle mass, similar to community dwelling older adults.

2.5.5 Limitations of ultrasound measures of muscle thickness

While ultrasound has potential to assess muscle mass, there are a number of limitations associated with its use such as error in identification of bony landmarks, consistency in the placement of the probe (centered on landmark, force applied against the skin and tilt of the probe), accurate and consistent caliper placement for thickness assessment and the need for prediction equations for lean tissue mass. Examiners using different levels of force on the ultrasound probe can induce varying levels of tissue depression, creating potential deviations in muscle thickness measures (77), making it difficult to directly compare studies using different approaches. These deviations may also result in high variability, as muscle thickness measures are relatively small in magnitude (generally less than 5 cm for larger muscle groups) and small differences between examiners may result in relatively large CV, especially in individuals who may already have lower than normal muscle thicknesses. In addition, when analyzing the image, placement of the electronic calipers may vary between examiners; specifically, determining the muscle border can be confounding due to the multiple fascial layers surrounding the muscle (epimysium vs perimysium) and infiltration of adipose tissue into the muscle results in increased attenuation of the ultrasound wave, producing a poorer quality image.

However if care is taken to ensure correct identification of landmarks, neutral probe tilt, minimal depression of the subcutaneous tissues and correct identification of the muscle fascia, ultrasound literature has demonstrated good to excellent reproducibility both within and between examiners (56). For measures of peripheral muscle thickness, intra-rater reliability has produced CV values between 2.3 to 3.0 % (2; 78–80) and intra class correlation coefficients (ICC) between 0.71 to

0.99 (56; 57; 68). Fewer studies have investigated inter-rater reliability for muscle thickness, but, those that have, obtained ICC values between 0.73 to 0.99 (56; 81). While it may seem that ultrasound is rather reliable in assessing muscle thickness (82), both within and between raters, a major limitation associated with most literature that examines the reliability of ultrasound for thickness measures is, that more often than not, these investigations focus on using previously defined landmarks, for both intra-rater and inter-rater reliability. That is to say, the entire ultrasound protocol, from landmarking to image acquisition to caliper placement is not being investigated; most studies, assess the latter two aspects. To date, only a single investigation has employed methodology to assess the reliability of the entire ultrasound protocol (83). Using the 9-site protocol in acute stroke patients, investigators observed strong reliability (ICC >0.80) in many, but not all landmarks. Of these landmarks, one of the more reliable sites was the anterior thigh (ICC >0.9), but the bony landmarks used in the 9-site protocol for anterior thigh thickness are different than those used in the 4-site protocol (83). Although promising, the robustness of these measures need to be further explored, in healthy, obese and clinical populations, to ensure specific challenges associated with those populations (i.e. identification of bony landmarks in obese individuals) do not confound the reliability of muscle thickness measures.

Chapter 3
Rationale, Objectives and Hypotheses

3.1 Rationale

Despite the potential utility of ultrasound in not only quantifying muscle mass, but also identifying those individuals with lower than normal muscle mass in a clinical setting, research investigating the accuracy of viable bedside protocols is lacking. The majority of research has focused on protocols (such as the 9-site protocol) that lack the feasibility of being conducted in most clinical settings. Viable bedside protocols (such as the 4-site protocol) have been employed and have be shown to be feasible in both healthy (56), liver cirrhotic (41) and ICU populations (2; 3), but they have yet to be comprehensively evaluated for accuracy against whole body measures of lean tissue or muscle mass. Although two studies have compared the agreement of viable bedside ultrasound protocols to accurate measures of lean tissue in healthy participants (58; 60), recent literature has highlighted potential improvements to ultrasound methodology that was not previously utilized in those studies, such as accounting for limb length or height to improve accuracy (84). Furthermore, a recent multicentre study in the ICU investigated the associations between the 4-site ultrasound protocol and 3rd lumbar vertebra muscle CSA from CT scans, which albeit is a non-whole body measure of muscle mass, but observed significant improvements when physical covariates such as age, sex and BMI were accounted for. Viable bedside protocols need to be evaluated for accuracy against an well-established measures of lean tissue or muscle mass in a heterogeneous participant cohort, accounting for physical characteristics, to assess the utility of these measures in a population that is not confounded by clinical factors, such as edema, before applying these protocols in future clinical investigations.

Multiple stages of evaluation (Figure 5) are needed before these protocols are applicable and meaningful in a clinical setting. First, the protocol needs to be compared with a previously established, precise and accurate body composition modality (such as DXA) in healthy individuals to ensure that it is a valid approach in estimating lean tissue or muscle mass. Second, the accuracy of the

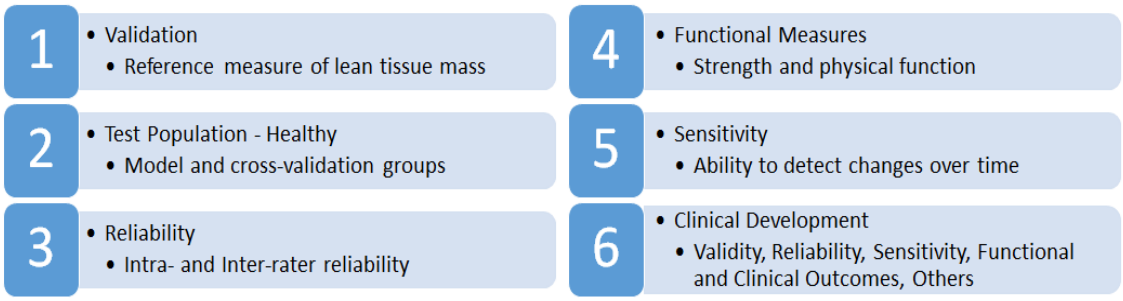


Figure 5. Multiple stages of evaluation for a clinically accessible body composition modality

previously developed models need to be internally or externally validated, to ensure the generalizability of the model. Third, the intra-rater and inter-rater reliability of the entire ultrasound protocol needs to be assessed, in multiple different populations (young, old, obese), to ensure that these measures can be confidently obtained across multiple clinical centres. Fourth, the associations between ultrasound measures of muscle mass and measures of strength and physical function need to be elucidated to ensure these measures are valid at identifying individuals who may be sarcopenic, as current guidelines on the definition of sarcopenia include both low muscle mass and poor physical function or strength (26). Fifth, the sensitivity of ultrasound based muscle thickness measures need to be assessed to determine how well they track changes in muscle mass over time. Lastly, all of these steps will need to be re-evaluated in specific clinical populations, to ensure the validity within those populations. This thesis primarily focused on steps one and two, with secondary objectives assessing step three. This work creates a foundation upon which viable bedside protocols can be optimized for accuracy and reliability, prior to application in clinical populations.

A potentially confounding factor for measures of muscle thickness in clinical populations is the presence of edema. Edema is the abnormal accumulation of interstitial fluid and is thought to

artificially increase measures of muscle thickness (56). Edema should be considered before developing ultrasound guidelines for use in a clinical setting. Although maximal compression of the ultrasound probe against the skin has been suggested as a potential strategy to account for edema (56), there has only been a single investigation comparing maximal and minimal compression, determining that minimal compression was superior, but this was performed in liver cirrhotic patients and compared to a non-whole body measure of muscle mass (41). Before being further applied in a clinical setting, maximal compression and minimal compression should be compared in a healthy population to determine which of these approaches better estimates lean tissue mass.

The 4-site ultrasound protocol has been used in clinical populations that are less mobile, and may be a versatile method for quantifying muscle thickness. Thus, I assessed the agreement between 4-site protocol (using minimal and maximal compression), performed in a supine position, and DXA appendicular lean tissue mass. Additional anatomical landmarks from the 9-site protocol, obtained in a supine position, and easily obtained covariates (sex, age, and BMI) were used to optimize the existing 4-site protocol, to develop a protocol that more accurately predicts lean tissue, but is still applicable at the bedside. Lastly, previously published low lean tissue mass DXA cut-points were used to categorize participants and the ability of the optimized ultrasound protocol to identify these individuals was assessed.

3.2 Objectives

3.2.1 Primary objectives

In a cohort of participants that have a wide range of low to high lean tissue mass (approximately 30 – 80 kg of lean tissue mass), we proposed:

1. To evaluate the agreement between muscle thickness measured by ultrasound using the 4-site protocol and regional and appendicular measures of lean tissue mass derived from DXA

2. To improve the accuracy of the 4-site protocol for predicting appendicular lean tissue mass by incorporating easily obtained covariates and additional bedside accessible muscle thicknesses from the 9-site protocol
3. To assess the ability of the optimized ultrasound protocol (objective #2) to discriminate between individuals with normal and low lean tissue mass using established cut-points for appendicular lean tissue mass index derived by DXA.

3.2.2 Secondary objectives

1. To compare the accuracy of using minimal and maximal compression of the ultrasound probe in measuring muscle thickness to assess lean tissue mass
2. To obtain additional accessible muscle thicknesses from the 9-site protocol and to compare the accuracy of the 9-site protocol to the 4-site and optimized ultrasound protocols for appendicular lean tissue mass predictions.
3. To evaluate the intra-rater reliability of the image acquisition and caliper placement of the 4-site protocol and inter-rater reliability of the entire 4-site protocol

3.3 Hypotheses

3.3.1 Primary hypotheses

1. There will be a strong ($r > 0.8$) linear association between the 4-site ultrasound muscle thicknesses and DXA-derived regional and appendicular lean tissue mass.
2. Anterior muscle thicknesses of the upper limbs (anterior upper arm and anterior forearm) and covariates sex, age and BMI will significantly improve predictive accuracy of DXA derived appendicular lean tissue mass.
3. There will be a strong (AUC ≥ 0.85) ability to identify individuals with low lean tissue mass using the optimized ultrasound protocol.

3.3.2 Secondary hypotheses

1. Minimal compression will produce a stronger association between ultrasound muscle thickness and DXA lean tissue than maximal compression
2. There will be a strong ($r > 0.9$) linear association between 9-site muscle thickness protocol and DXA-derived appendicular lean tissue mass. Further comparison with the 4-site and optimized protocols will demonstrate that the 9-site protocol is more accurate in predicting appendicular lean tissue mass.
3. There will be a strong (ICC>0.80, CV<5%) degree of intra-rater and inter-rater reliability for the 4-site protocol.

Chapter 4
Methodology

4.1 Study design

This observational thesis recruited 96 participants to attend a single data collection session. Participants underwent anthropometric measures, a whole body DXA scan and ultrasound assessments using the 9-site and 4-site protocols in supine and prone positions. A small subset (n=16) of participants had inter-rater reliability performed using the 4-site protocol for a single leg (alternating between dominant and non-dominant legs). This study was reviewed and cleared by a University of Waterloo Clinical Research Ethics Committee. Written informed consent was obtained from all participants in accordance with established protocols for human research.

4.2 Participants

Participants (≥18 years of age) were recruited from the University of Waterloo student population, from the Kitchener-Waterloo community and from the Waterloo Research Aging Participant Pool. Participants were screened using a health questionnaire and excluded if they: 1) had a previous history of neuromuscular disorders, 2) were currently or suspect they may be pregnant, 3) had undergone a barium swallow or nuclear medicine scan within the past three weeks, 4) had a stroke within the past five years, and 5) had a prosthetic joint replacement. These exclusion criteria were in place to avoid factors that would have confounded DXA measured lean tissue, ultrasound measured muscle thickness or the comparison between them. To obtain a heterogeneous cohort of participants with a wide range of lean soft tissue mass and to increase the generalizability of these results, we attempted to recruit a minimum of 20 participants in each of the following BMI (weight (kg)/height^2 (m^2)) categories: <25 kg/m^2, 25 - 30 kg/m^2 and >30 kg/m^2, but were only successful for the <25 kg/m^2 and 25-30 kg/m^2 groups (only 14 participants were recruited into the >30 kg/m^2

group). Participants were instructed to refrain from alcohol consumption for 24 hours and moderate to vigorous physical activity for 48 hours prior to their scheduled data collection session.

4.3 Anthropometric data

Height and weight was obtained with participants in lightweight clothing (loose shorts and t-shirt) or a cloth hospital gown in their socks or bare feet. Weight was obtained using a balance beam scale (Mechanical Beam Scale, Health o meter, McCook, IL) while the participant stood still, with both feet on the weighing platform, measured to the nearest 0.1 kg. Height was obtained using a standing stadiometer with participant's heels against the wall and feet as close together as the participant was able to maintain, height was measured to the nearest 0.01 m. Circumferences of the upper thigh were taken using a flexible tape measure, to the nearest 0.5 cm, at the midpoint and lower third between the anterior superior iliac spine and the upper pole of the patella for both left and right legs. All anthropometric measures were taken once.

4.4 DXA scan

Certified Medical Radiation Therapists conducted one to two whole body DXA scans (Hologic Discovery QDR 4500, Hologic, Toronto, ON) on each participant. The second scan was required if the participant did not fit within the limits of the scanning table. If two scans were required, the scan containing the upper limb within the table limits was used for lean tissue, fat mass and bone mineral content for that limb, whereas all other body segments (head, trunk and lower limbs) were averaged across both scans. Using Hologic software (version 13.2), the whole body scan was segmented into the head, trunk, left and right upper limbs and left and right lower limbs by a single investigator according to a standardized protocol (25). The lean tissue of the upper and lower limbs (appendicular lean tissue mass) was summed and used as the criterion outcome for all linear regression analysis. Appendicular was chosen over whole body lean mass as the reference criteria

since DXA cannot distinguish skeletal muscle from other muscle and organ compartments within whole body lean tissue measures. This is important to distinguish since whole body lean tissue measures may confound the association with ultrasound measured skeletal muscle thickness. Lastly and most importantly, the appendicular lean tissue depot is the criteria used to determine if an individual has lower than normal lean tissue mass (26); therefore utilizing appendicular lean tissue mass as our criterion measure, we were able to determine which muscle thicknesses and covariates are most useful for identification of individuals with low lean tissue mass. Individuals were identified as have lower than normal lean tissue mass by dividing their appendicular lean tissue mass by their height squared (kg/m^2) and using previously published cut-points of $\leq$ 7.26 kg/m^2 for males and $\leq$ 5.45 kg/m^2 for females (4).

Regional measures were performed for direct comparison of the 4-site protocol muscle thicknesses to the corresponding site on the DXA scan for lean tissue measures (Figure 6). Lean

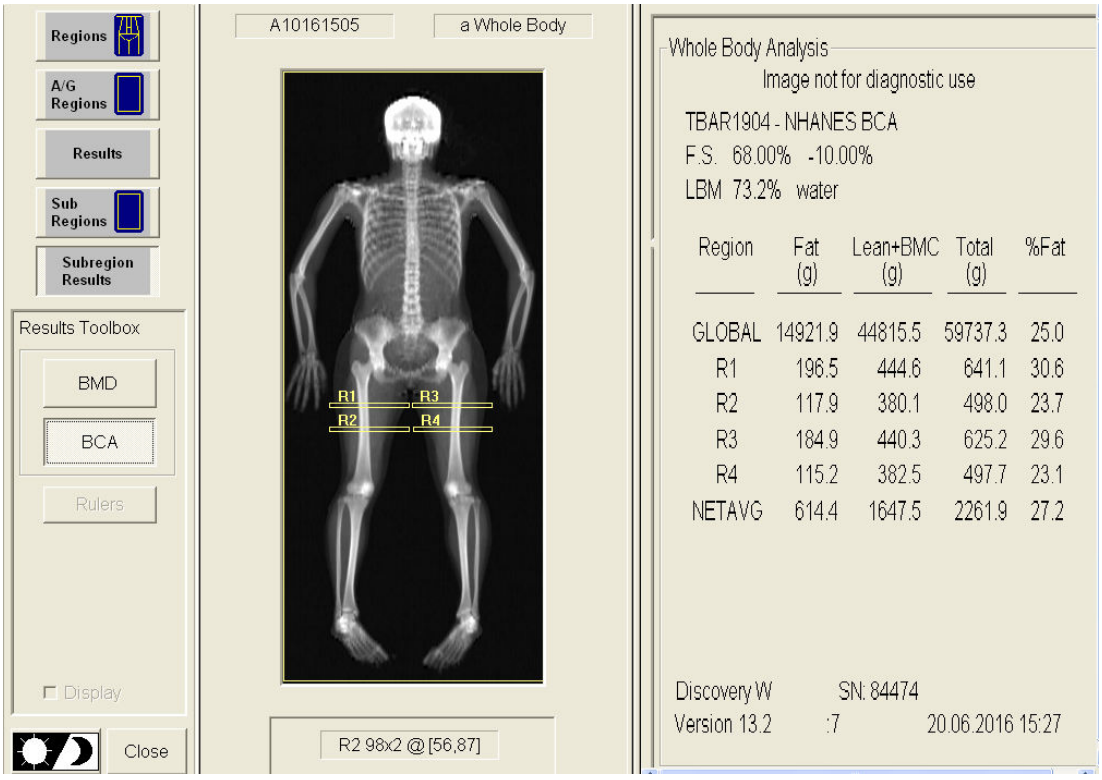


Figure 6. Regional assessment using DXA software for site to site comparison of muscle thickness and lean

.

tissue at the midpoint and lower third between the anterior superior iliac spine and upper patella was measured using the DXA software. The pixels in the y-axis pertaining to the anterior superior iliac spine and the upper patella was determined visually and used to calculate the relative distances and pixels associated with the midpoint and lower third sites. Subsequently, rectangular boxes, 2 pixels in height and wide enough to encompass the skin on both sides of the upper thigh, were placed to quantify the lean soft tissue. Lean tissue was averaged across all 4 sites and compared to average muscle thickness from the 4-site protocol using both minimal and maximal compression. A height of 2 pixels was chosen as this is the smallest height a region of interest box can be made, and is equivalent to approximately 2.6 cm, in comparison to a 1 cm wide ultrasound probe imaging size.

4.5 Ultrasound protocols

A real-time B-mode ultrasound imaging device (M-Turbo, SonoSite, Markham, ON) equipped with a multi-frequency linear array transducer (5-10 MHz) was used to obtain transverse images of muscle groups at predefined sites. Adjustable parameters: gain, time gain compensation and compression (dynamic range - neutral) were held constant throughout the imaging process, whereas depth was adjusted as required in order to obtain a complete image of the muscle thickness. Muscle thickness measures were obtained from frozen images using onscreen calipers, measuring the distance between the upper margin of the underlying bone and the lower boundary of the ventral fascia of the muscle group of interest (85). For example, the quadriceps muscle thickness measurement was the distance between the upper margin of the femoral bone and the lower boundary of the ventral fascia of the rectus femoris (8), incorporating both the rectus femoris and the vastus intermedialis (Figure 4, page 12). For muscle thickness measures that do not use a bony surface, the lower and upper boundaries of the muscle fascia was used for analysis. All measurements were made with the participant in the prone or supine position with their ankles wrapped with an adjustable strap to ensure neutral rotation of the lower limbs. All bony landmarks were identified by palpation and

specific sites to be imaged identified with the use of a flexible tape measure and marked using easily removable ink. During image analysis and caliper placement, all raters were blinded to thickness measures by the use of a removable sticker on the ultrasound screen to ensure subsequent thickness measures were not influenced by prior image analysis.

Ultrasound images were obtained using two protocols, the 4-site and the 9-site protocol. The 4-site protocol imaged the anterior surface of the left and right rectus femoris and vastus intermedialis muscles from i) the mid-point between the anterior superior iliac spine and the upper pole of the patella and ii) the border of the lower third and upper two-thirds between the anterior superior iliac spine and the upper pole of the patella (56). The 4-site protocol had two phases, minimal and maximal

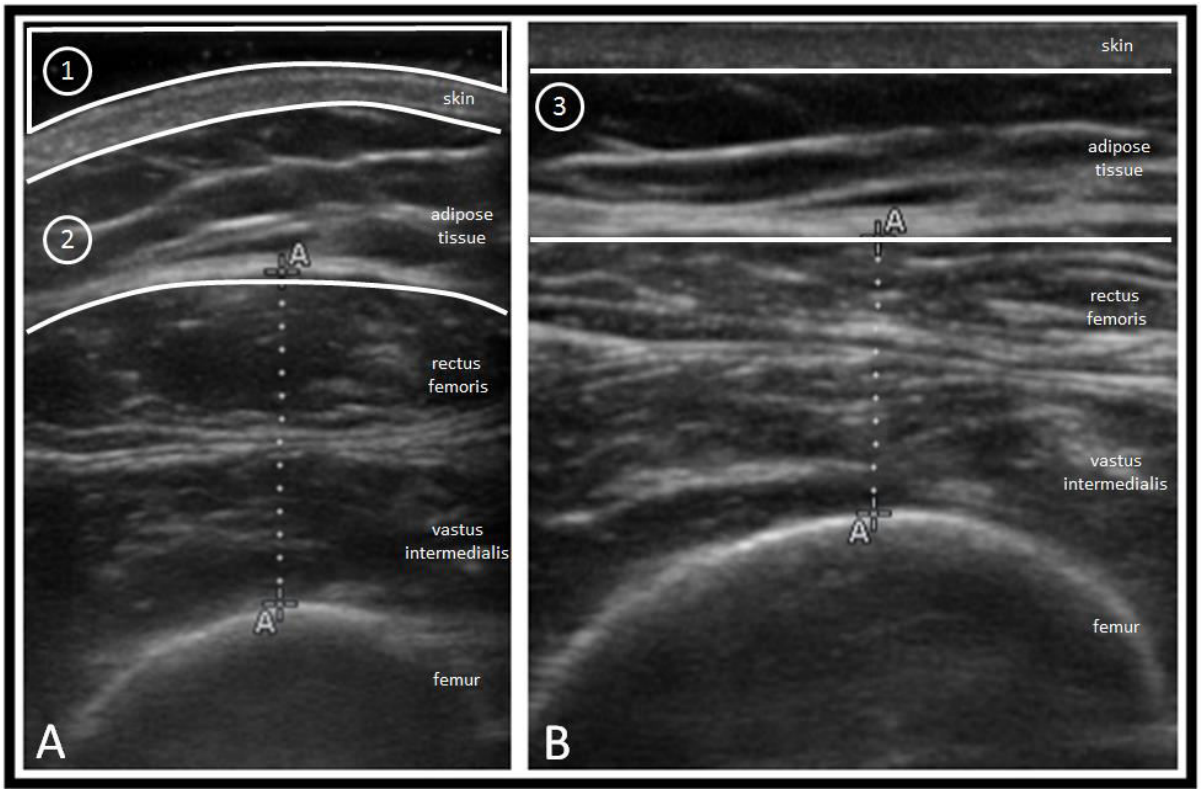


Figure 7. Comparison between minimal (A) and maximal compression (B) protocols within a single participant. 1 – Region highlighting ample acoustic coupling gel to ensure no contact occurs between skin and ultrasound probe. 2 – Highlighting convex nature of the skin and muscle belly, required for minimal compression. 3 – Skin and muscle belly loose convex shape during maximal compression.

compression, and each landmark was imaged twice, using each level of compression. Minimal compression was achieved by coating the transducer with ample water-soluble transmission gel and positioned to obtain an image containing the highest density of cortical bone of the femur (neutral probe tilt). A thick layer of ultrasound gel was maintained between the probe-skin interface to ensure there is no tissue depression and the operator ensured the muscle belly and skin maintained its convex shape prior to freezing the image (Figure 7A). In contrast, maximal compression was considered to be maximal attainable compression of the underlying tissue with the transducer (Figure 7B).

The 9-site protocol was performed only on the right side of the body and each landmark was imaged a single time, as previously described by Takai et al. (2014). The upper arm muscle thickness measures were taken on the anterior and posterior surface, 60% distal from the acromial process of the scapula to the lateral epicondyle of the humerus. Forearm muscle thickness was taken from the anterior surface 30% distal from the radial head to the styloid process of the radius. Abdominal measures of muscle thickness were taken 3 cm from the right of the umbilicus. Subscapular muscle thickness measures were taken 5 cm directly below the inferior angle of the scapula. Measures of thigh muscle thickness were taken on the anterior and posterior surfaces midway between the lateral condyle of the femur and the greater trochanter. Finally, the lower leg muscle thickness measures were taken on the anterior and posterior surface, 30% distal from the head of the fibula to the lateral malleolus. Only minimal compression as described above was used for the 9-site protocol.

Previous literature has shown improved accuracy in predicting lean tissue when accounting for body size differences by multiplying the muscle thicknesses by the corresponding limb length or height. Miyatani et al. (2004) compared muscle thicknesses of the upper and lower limbs with their corresponding muscle volumes from MRI and demonstrated that multiplying the muscle thickness by limb length significantly improved the accuracy in predicting muscle volume. This is further enforced by So et al. (2004), who showed that MRI volume of the quadriceps demonstrates a strong association

with whole body muscle mass. To obtain a surrogate measure of muscle volume, which was used to predict appendicular lean tissue mass, all muscle thicknesses were either multiplied by their corresponding limb length or height.

4.6 Reliability of the 4-site protocol

Inter-rater reliability was performed for the 4-site protocol by a second rater, experienced with the application of musculoskeletal ultrasound, but who had minimal training applying the 4-site protocol. The order of inter-rater reliability measures by the two raters was random, but not randomized. Reliability was assessed on a single leg for each participant and alternated between dominant and non-dominant legs, in a sequential order. Inter-rater reliability assessment occurred according to the following steps: 1) landmarking with ink by rater one, 2) acquisition of image and placement of calipers for muscle thickness measures for both minimal and maximal compression protocols, with removal and repositioning of the probe between images, by rater one, 3) removal of ink landmarks using 70% ethanol wipes by rater 1, and 4) rater 2 performs steps 1-3, blinded to all measures from rater 1. Therefore, the second rater was conducting the entire 4-site protocol, on a single leg, alternating between dominant and non-dominant legs for each participant. Inter-rater reliability was performed on the last 16 individuals participating (31% female), with median (interquartile range (IQR)) age, BMI and body fat of 70.5 (67.3-78), 26.4 (22.9-29.9) and 31.3 (26.3-36.7), respectively. Assessing reliability on the last 16 individuals may have displayed the ideal situation for intra-rater reliability, as the primary rater will have gained considerable experience in applying the 4-site protocol; but, this would not be the case for inter-rater reliability, as the second rater was inexperienced with the 4-site protocol at the start of these 16 individuals. Intra-rater reliability was assessed by comparing the 1st measure to the 2nd measure for all 4-site landmarks, according to steps 1 and 2, as described above.

4.7 Sample size

Utilizing online software developed by Dr. David Schoenfeld (Professor in Department of Biostatistics at Harvard T.H. Chan School of Public Health and Professor of Medicine, Harvard Medical School) for sample size calculations, a two tailed test (α=0.05) was performed using standard deviations of 374 cm^3 (84) and 0.9 cm (47) for knee extensor muscle volume measured using MRI and muscle thickness measured using ultrasound. A sample size of 97 was calculated, based on setting the power at 80%, with a minimal detectable difference of 120 cm^3. The minimal detectable difference was calculated using a regression equation to predict knee extensor volume (84), using a minimal change in muscle thickness of 0.3 cm.

4.8 Statistical analysis

4.8.1 General analysis

Normality was assessed for continuous variables using the Shapiro-Wilks test. Normality was violated for numerous variables and therefore, descriptive statistics are reported as median and IQR (Q1-Q3) and differences between males and females or young and older adults was analyzed using the Mann-Whitney U test. Differences in proportion of low lean tissue was assessed using a Chi-square test. A Fisher z transformation was utilized to assess if there was a significant difference between regression coefficients comparing minimal or maximal compression to site specific measures of lean tissue mass. Intra-rater and inter-rater reliability was assessed using CV, ICC and Bland-Altman plots (described below in section 4.8.4). ICC equation (1,1) was used for both intra-rater and inter-rater reliability, as it is considered the most conservative approach (87). All analysis was performed using Sigma Plot 13 (Systat Software, San Jose, CA) and the level of significance was set at $p \leq 0.05$.

4.8.2 Cross-validation analysis

In an attempt to account for overfitting of our models, a 3-fold cross-validation analysis was used for all linear and logistic regression and the ROC curve analysis. A 3-fold cross-validation splits

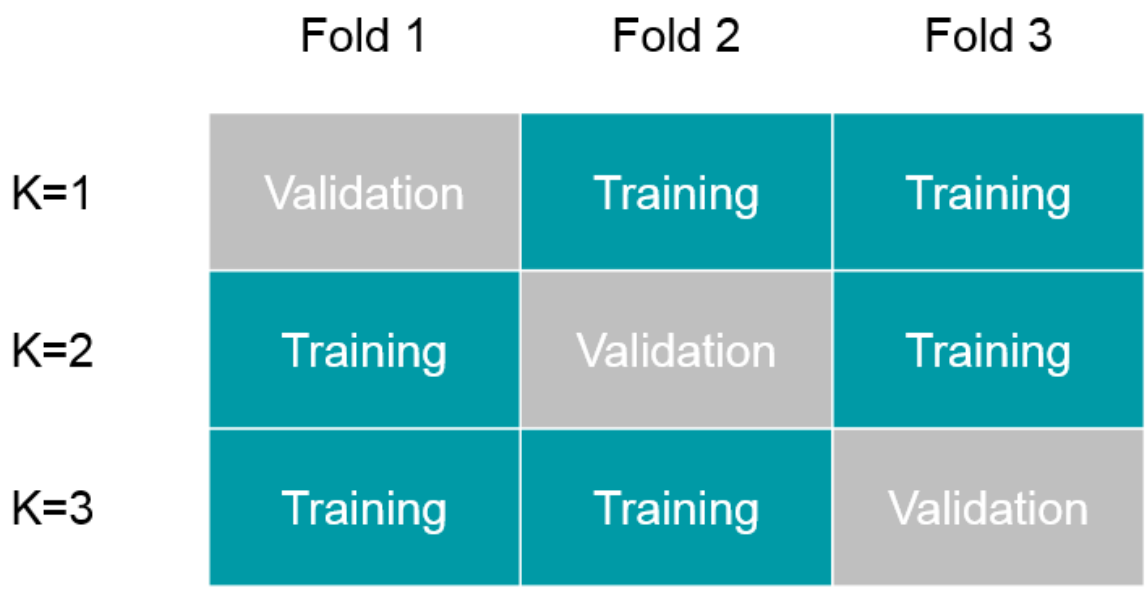


Figure 8. Schematic of 3-fold cross-validation for model development and testing

the participant cohort into 3 equally distributed folds (or groups), where model training (development) occurs with 2 of the 3 folds, with subsequent validation in the left out fold; this is performed 3 times, utilizing all permutations (Figure 8) (88). All participants were stratified by their appendicular lean tissue mass index (kg/m^2) and then randomly allocated to folds 1, 2 or 3, to ensure equal proportions of lower than normal lean tissue mass individuals in each fold. For linear regression, model parameters coefficient of determination (r^2) and standard error of the estimate (SEE) were averaged across each cross-validation fold to determine average model performance. For logistic regression and ROC analysis, model parameters odds ratio and c-index (area under ROC curve) were averaged across model development and cross-validation folds as an average of model performance.

4.8.3 Regression and ROC analysis

The 4-site protocol muscle thicknesses were averaged across all sites, multiplied by limb length and was used predict appendicular lean tissue mass using linear regression analysis. The 9-site protocol muscle thicknesses were summed, multiplied by height and used to predict appendicular lean tissue mass using linear regression analysis. To determine the variables for the optimized protocol, backwards stepwise regression, incorporating all 96 participants, was performed using the 4-site muscle thicknesses, a-priori defined accessible muscle thicknesses of the anterior upper arm, anterior forearm and anterior lower leg from the 9-site protocol and easily obtained covariates age, sex and BMI. Multiple linear regression using variables identified in the backwards stepwise regression was performed to predict appendicular lean derived from DXA. To assess the ability of the optimized protocol to identify low lean tissue mass individuals, we performed a ROC analysis within each fold. In order to increase the sample of low lean tissue mass individuals, we combined males and females within each fold by developing a multiple logistic regression model to identify low lean tissue mass, using variables identified in the backwards stepwise regression. The log odds from logistic regression for each cross-validation fold was used as input for the ROC analysis.

4.8.4 Bland-Altman analysis

Bland-Altman plots were used to assess the absolute agreement at the individual level for the 4-site, 9-site and the optimized ultrasound protocol by comparing predicted and DXA measured appendicular lean tissue mass and for assessing reliability of the 4-site protocol, as previously suggested (87). Bland-Altman analysis examines the spread of the differences between two measures, compared against their averages, to determine if the absolute differences between the two measures is consider an acceptable level of error (limits of agreement: mean difference ± 1.96 SD) for the new technique to be considered an accurate surrogate of the reference technique or to accept the technique as reliable (89). While the limits of agreement will contain 95% of the differences for the current

participant cohort, if the technique that is being validated is intended to be applied to other populations as a surrogate of the reference technique, it has been suggested to state the tolerance limits (upper and lower 95% CI for the limits of agreement), as these would contain 95% of future predictions for a new participant cohort (90). To ensure the validity of the limits of agreement, a regression analysis was performed for the differences against averages to assess for proportional bias and homoscedasticity of the differences was assessed visually by examining a plot of the residuals of the regression analysis against averages (89; 90). Lastly, a major, but often neglected step of Bland-Altman analysis (91), is to state whether the limits of agreement are clinically acceptable.

Chapter 5
Results

5.1 Physical and demographic description of participants

Ninety-six participants were recruited (86% Caucasian), with a median (IQR) age of 36.5 (24.0-72.0) years, BMI of 24.3 (22.3-27.3) kg/m^2 and body fat of 30.2 (24.3-36.8) % (Table 2). Of the 96 participants, 57% were female, and compared with males, significant differences were observed for BMI ($p=0.021$) and body fat ($p<0.001$), but not age ($p=0.279$). Comparing young and older adults, significant differences were observed for age ($p<0.001$), BMI ($p<0.001$) and body fat ($p<0.001$) (Table 2).

Table 2. Physical and demographic description of participant cohort.

Variable Median (IQR)	All (n=96)	Males (n=41)	Females (n=55)	p-value	Young (<60, n=55)	Older (≥60, n=41)	p-value
Age, years	36.5 (24.0-72.0)	39.0 (25.0-73.0)	35.0 (23.0-69.0)	0.279	24.0 (23.0-30.0)	73.0 (69.0-78.0)	<0.001
Sex, % female	57%	-	-	-	60%	54%	-
Height, m	1.69 (1.62-1.77)	1.77 (1.71-1.81)	1.64 (1.58-1.69)	<0.001	1.70 (1.64-1.76)	1.66 (1.59-1.78)	0.173
Weight, kg	70.5 (62.8-82.3)	82.0 (71.9-88.1)	65.2 (58.5-70.9)	<0.001	68.0 (60.3-76.8)	75.8 (64.7-84.2)	0.063
BMI, kg/m^2	24.3 (22.3-27.3)	25.7 (23.9-27.5)	23.7 (21.8-26.5)	0.021	23.7 (21.6-25.7)	26.6 (23.7-29.2)	<0.001
Underweight ≤18.5 kg/m^2	0	0	0	-	0	0	-
Normal 18.5–24.9 kg/m^2	52	16	36	-	38	14	-
Overweight 25.0–29.9 kg/m^2	30	19	11	-	12	18	-
Obese ≥30.0 kg/m^2	14	6	8	-	5	9	-
Body fat, %	30.2 (24.3-36.8)	25.0 (20.4-29.8)	34.9 (29.3-41.0)	<0.001	27.3 (21.7-33.8)	35.1 (29.0-41.5)	<0.001
Ethnicity	-	-	-	-	-	-	-
Caucasian, n	83	35	48	-	42	41	-
Asian, n	12	5	7	-	12	0	-
Other, n	1	1	0	-	1	0	-

IQR, interquartile range.

For all participants, median whole body lean tissue mass, appendicular lean tissue mass and appendicular lean tissue mass index was 44.7 (37.6-54.7) kg, 19.2 (15.6-24.2) kg and 6.92 (5.82-7.73) kg/m^2, respectively (Table 3). In total, 17 of 96 (18%) participants were identified as having lower than normal lean tissue (males ≤7.26 kg/m^2; females ≤5.45 kg/m^2), despite that no participants were classified as underweight according BMI. Median appendicular lean tissue was significantly different between males and females (p<0.001) and between young and older adults (p=0.041), with no differences seen in the proportion of low muscle mass between males and females (p=0.341) or between young and older groups (p=0.864) (Table 3).

Table 3. Descriptive summary of DXA measures of lean tissue

Variable Median (IQR)	All (n=96)	Males (n=41)	Females (n=55)	p-value	Young (<60, n=55)	Older (≥60, n=41)	p-value
Whole body lean soft tissue, kg	44.7 (37.6-54.7)	55.4 (51.3-59.5)	37.9 (35.9-42.2)	<0.001	44.8 (37.9-55.5)	42.6 (36.1-53.3)	0.310
Appendicular lean tissue, kg	19.2 (15.6-24.2)	24.7 (21.7-27.2)	15.8 (14.7-17.8)	<0.001	19.8 (16.3-25.3)	17.5 (14.7-23.4)	0.041
Appendicular lean tissue index, kg/m^2	6.92 (5.82-7.73)	7.75 (7.38-8.56)	6.03 (5.50-6.72)	<0.001	7.26 (6.03-8.07)	6.60 (5.52-7.43)	0.033
Proportion low lean tissue, n/N and %	17/96 18%	5/41 12%	12/55 22%	0.341	6/55 11%	11/41 27%	0.864
Site specific lean soft tissue, kg	0.40 (0.34-0.47)	0.45 (0.41-0.54)	0.36 (0.31-0.40)	<0.001	0.45 (0.37-0.51)	0.38 (0.31-0.41)	<0.001

IQR, interquartile range.

Muscle thicknesses were significantly different between males and females for both the 9-site (p<0.05) (Table 4) and 4-site ultrasound protocols (p<0.001) (Table 5). Comparing younger and older adults, significant differences were observed for all muscle thicknesses using the 4-site protocol (p<0.001) (Table 5), but using the 9-site protocol, significant differences were only seen for the

muscle thicknesses of the posterior upper arm, abdominal, anterior forearm, anterior upper leg, posterior upper leg and combined total ($p<0.05$), but not for the subscapular, anterior upper arm, anterior lower leg and posterior lower leg ($p>0.05$) (Table 4).

Table 4. Descriptive summary of ultrasound measured muscle thickness using the 9-site protocol

Variable Median (IQR)	All (n=96)	Males (n=41)	Females (n=55)	p-value	Young (<60, n=55)	Older (≥60, n=41)	p-value
Anterior upper arm, cm	2.76 (2.31-3.50)	3.55 (3.20-3.86)	2.37 (2.20-2.70)	<0.001	2.99 (2.34-3.45)	2.67 (2.29-3.53)	0.364
Posterior upper arm, cm	2.32 (1.75-3.03)	3.06 (2.56-3.61)	1.93 (1.56-2.31)	<0.001	2.51 (1.98-3.38)	1.95 (1.50-2.57)	<0.001
Anterior forearm, cm	1.58 (1.29-1.84)	1.80 (1.55-2.06)	1.36 (1.16-1.61)	<0.001	1.69 (1.31-1.94)	1.47 (1.27-1.70)	0.046
Abdominal, cm	1.02 (0.82-1.34)	1.16 (0.84-1.59)	0.99 (0.77-1.19)	0.039	1.23 (1.01-1.49)	0.80 (0.65-0.95)	<0.001
Subscapular, cm	0.81 (0.63-1.02)	0.96 (0.76-1.24)	0.70 (0.55-0.92)	<0.001	0.83 (0.64-1.01)	0.75 (0.61-1.05)	0.612
Anterior upper leg, cm	3.34 (2.57-3.89)	3.77 (3.12-4.39)	2.96 (2.46-3.52)	<0.001	3.77 (3.12-4.38)	2.65 (2.25-3.23)	<0.001
Posterior upper leg, cm	5.15 (4.53-5.69)	5.46 (4.80-5.93)	4.83 (4.43-5.37)	0.006	5.39 (5.04-5.98)	4.55 (4.19-5.17)	<0.001
Anterior lower leg, cm	2.57 (2.33-2.81)	2.81 (2.61-3.11)	2.40 (2.27-2.58)	<0.001	2.55 (2.33-2.79)	2.60 (2.40-2.89)	0.201
Posterior lower leg, cm	5.68 (5.09-6.22)	6.19 (5.60-6.89)	5.14 (4.84-5.55)	<0.001	5.81 (5.37-6.26)	5.48 (4.78-6.19)	0.060
9-site total, cm	24.69 (22.12-28.61)	28.64 (27.08-31.54)	22.83 (21.28-24.53)	<0.001	27.46 (23.53-30.11)	22.83 (21.06-26.68)	<0.001

IQR, interquartile range.

Table 5. Descriptive summary of ultrasound measured muscle thickness using the 4-site protocol

Variable Median (IQR)	All (n=96)	Males (n=41)	Females (n=55)	p-value	Young (<60, n=55)	Older (≥60, n=41)	p-value
4-site protocol – minimal compression							
Right mid-point, cm	4.07 (3.36-4.80)	4.59 (3.98-5.32)	3.51 (2.95-4.26)	<0.001	4.43 (4.06-5.12)	3.37 (2.75-3.89)	<0.001
Right lower third, cm	3.04 (2.43-3.72)	3.53 (3.00-4.29)	2.71 (2.12-3.25)	<0.001	3.48 (3.01-4.18)	2.46 (2.01-2.97)	<0.001
Left mid-point, cm	3.97 (3.29-4.70)	4.67 (3.95-4.99)	3.67 (2.96-4.05)	<0.001	4.30 (3.92-4.97)	3.26 (2.83-3.95)	<0.001
Left lower third, cm	3.08 (2.41-3.71)	3.65 (2.91-4.24)	2.72 (2.16-3.21)	<0.001	3.45 (3.00-4.17)	2.39 (2.06-2.91)	<0.001
Average, cm	3.53 (2.82-4.24)	4.17 (3.44-4.58)	3.12 (2.59-3.72)	<0.001	3.92 (3.45-4.52)	2.80 (2.42-3.44)	<0.001
4-site protocol – maximal compression							
Right mid-point, cm	1.68 (1.32-2.14)	1.97 (1.56-2.35)	1.51 (1.22-1.96)	<0.001	2.04 (1.59-2.21)	1.40 (1.19-1.59)	<0.001
Right lower third, cm	1.53 (1.20-1.92)	1.79 (1.37-2.03)	1.37 (1.07-1.69)	<0.001	1.77 (1.32-2.03)	1.28 (0.97-1.56)	<0.001
Left mid-point, cm	1.68 (1.41-2.14)	1.92 (1.62-2.31)	1.48 (1.26-1.93)	<0.001	1.99 (1.61-2.30)	1.47 (1.23-1.66)	<0.001
Left lower third, cm	1.45 (1.19-1.85)	1.69 (1.42-2.10)	1.24 (1.11-1.72)	<0.001	1.72 (1.37-2.02)	1.22 (0.98-1.47)	<0.001
Average, cm	1.56 (1.30-1.98)	1.93 (1.48-2.18)	1.37 (1.17-1.85)	<0.001	1.86 (1.51-2.18)	1.32 (1.15-1.53)	<0.001

IQR, interquartile range.

5.2 Agreement between the 4-site protocol and DXA measures of lean tissue

To assess how well ultrasound measured muscle thicknesses using the 4-site protocol represents DXA derived lean tissue, and to evaluate the level of compression (minimal or maximal) that was most accurate, site specific comparisons were performed as described above (section 4.4). Both minimal and maximal compression of the 4-site protocol are strongly associated ($r>0.80$) (92) with site specific DXA measures of lean tissue mass, but the r^2 for minimal compression, $r^2=0.82$,

was significantly different compared to maximal compression, r^2=0.66 (p<0.001) (Figure 9). Therefore, all further analysis was performed using minimal compression for the 4-site protocol.

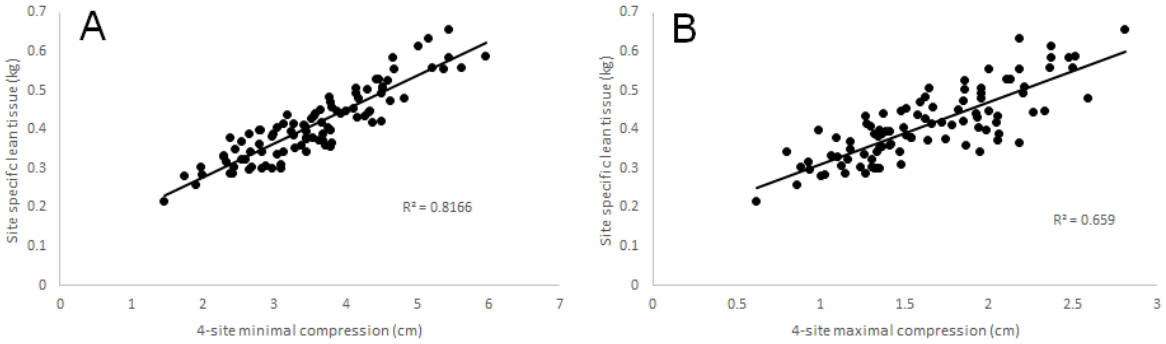


Figure 9. Regression analysis between site specific lean tissue mass derived from DXA and the 4-site ultrasound protocol using A) minimal compression, r^2=0.82 and B) maximal compression, r^2= 0.66.

To determine the agreement between the 4-site protocol and DXA measured appendicular lean tissue, linear regression and Bland-Altman analysis was performed. Linear regression analysis, using the 4-site protocol, averaged across 3-fold cross-validation to predict appendicular lean tissue mass resulted in an adjusted r^2 of 0.72 and SEE of 2.88 kg (Table 6). Bland-Altman analysis revealed normally distributed and homoscedastic differences (DXA – 4-site protocol) for appendicular lean tissue. A significant (r^2=0.08, p=0.005) proportional bias was present, with the 4-site protocol overestimating at the lower end and underestimating at the higher end of appendicular lean tissue; therefore, hyperbolic limits of agreement were constructed (Figure 10). The average range for the

Table 6. Linear regression analysis to predict appendicular lean tissue using the 4-site protocol

Model development	Appendicular lean tissue prediction (kg)	Cross-validation fold	Unadjusted r^2	Adjusted r^2	SEE (kg)	p-value model
Folds 1+2	4.061+(0.100X_1)	3	0.72	0.71	2.93	<0.001
Folds 1+3	4.037+(0.099X_1)	2	0.73	0.72	2.84	<0.001
Folds 2+3	3.587+(0.102X_1)	1	0.75	0.74	2.86	<0.001
Average	3.895+(0.100X_1)	-	0.73	0.72	2.88	-

ALT, appendicular lean tissue; SEE, standard error of the estimate; X_1 = 4-site muscle thickness x limb length (cm x cm).

hyperbolic limits of agreement, on the extreme ends (widest limits) of 10 and 35 kg of appendicular lean tissue, was 11.33 kg (11.35 for 10 kg and 11.3 for 35 kg).

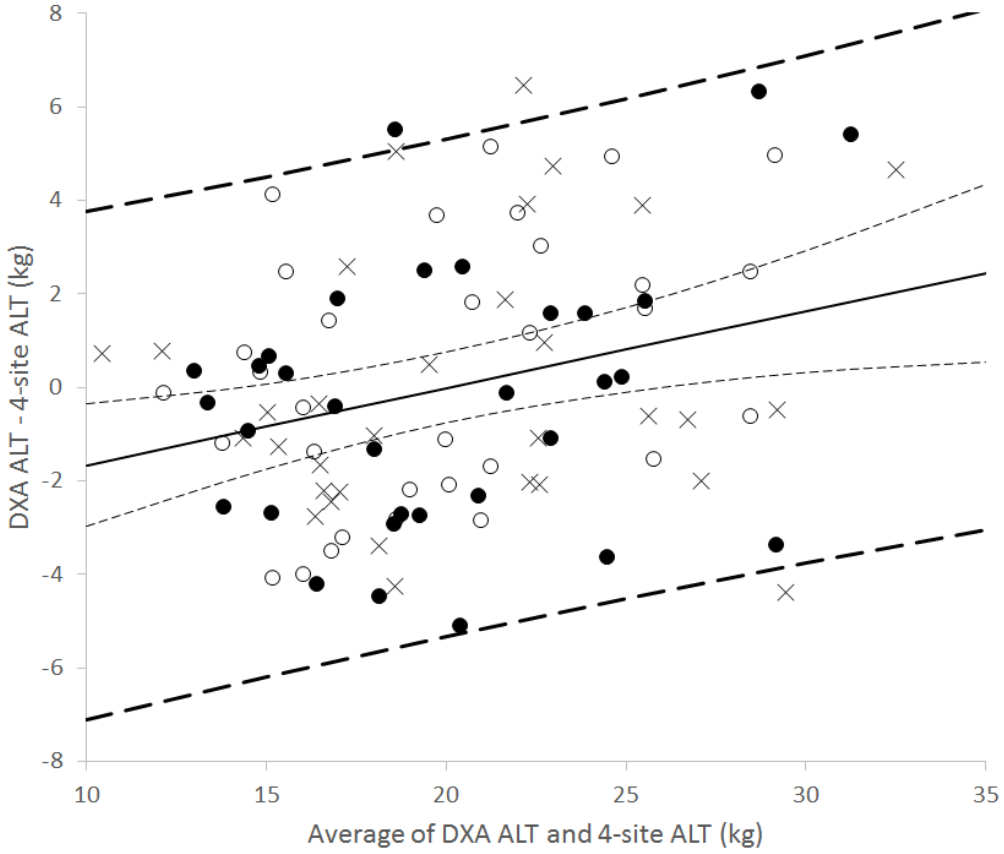


Figure 10. Bland-Altman plot comparing DXA derived and 4-site predicted appendicular lean tissue mass, utilizing participants from all folds. A significant ($p<0.05$) proportional bias was present (solid black line, 95% CI - inner curved dashed lines), with 95% prediction intervals (outer curved dashed lines) with an average range on the highest (35 kg) and lowest (10 kg) ends of appendicular lean tissue of 11.33 kg. Crosses (×) represent fold 1, closed circles (●) represent fold 2, and open circles (○) represent fold 3.

5.3 Agreement between the 9-site protocol and DXA measures of lean tissue

Linear regression analysis, using the 9-site protocol, across the 3 cross-validation folds to predict appendicular lean tissue resulted in an average adjusted r^2 of 0.90 and SEE of 1.73 kg (Appendix A2); this was the standard to achieve when optimizing the 4-site protocol. Bland-Altman

analysis revealed normally distributed and homoscedastic differences (DXA – 9-site protocol) for appendicular lean tissue, with no proportional bias present ($p>0.05$). A non-significant fixed bias of 0.04 kg [95% CI: -0.34, 0.35 kg] was present with limits of agreement of -3.32 and 3.32 kg, and tolerance limits of -3.91 and 3.92 kg (Appendix A3).

5.4 Agreement between optimized protocol and DXA measures of lean tissue

To determine the variables to be included in the optimized ultrasound protocol, backwards stepwise regression was performed. Backwards stepwise regression identified the anterior upper arm x limb length, 4-site x limb length, age and sex as the covariates that were significantly associated with appendicular lean tissue. The anterior upper arm and 4-site muscle thicknesses were combined as a single variable (summed and multiplied by height) to simplify use; summed or separate muscle thickness inputs did not alter the results. Hereafter, this protocol (anterior upper arm + 4-site muscle thicknesses x height, age and sex) is referred to as the optimized 5-site protocol.

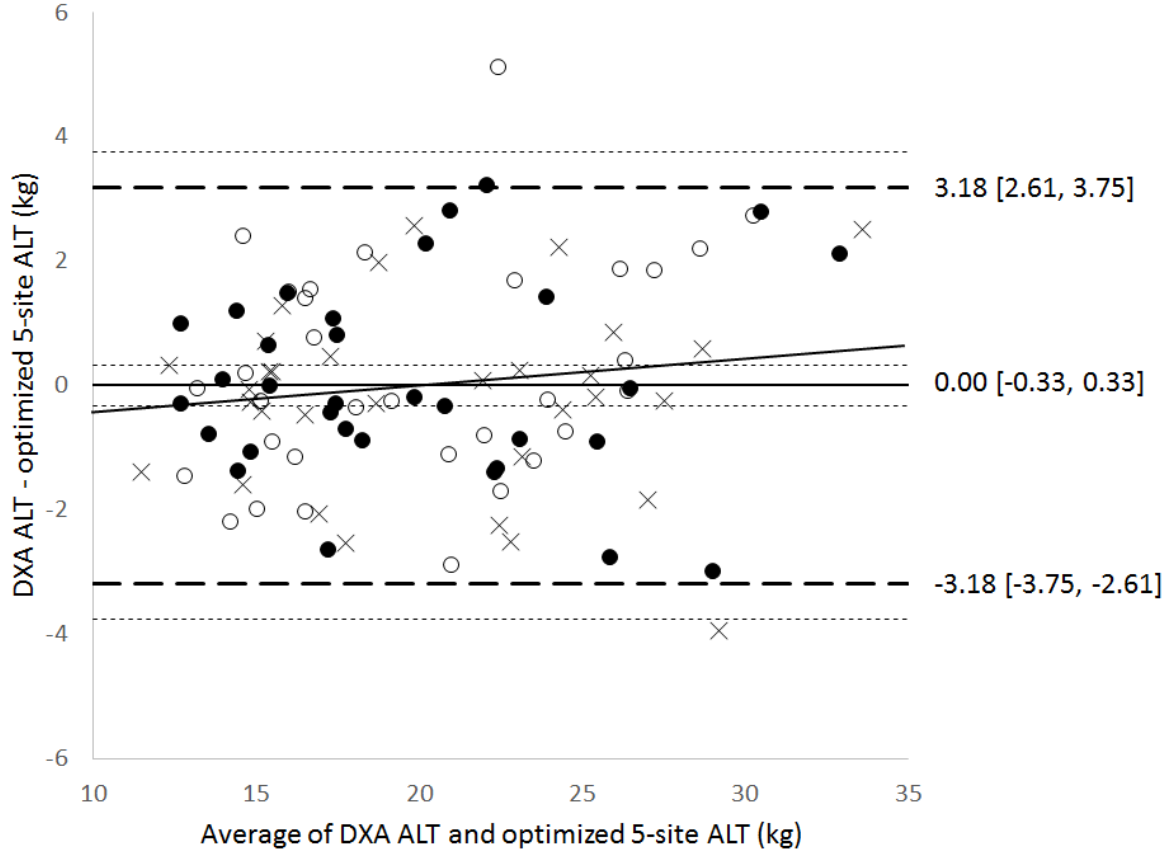
Multi-linear regression analysis, using the variables for the optimized 5-site protocol, averaged across 3-fold cross-validation to predict appendicular lean tissue mass resulted in an r^2 of 0.91 and SEE of 1.62 kg (Table 7); identical to results obtained using the 9-site protocol. Bland-Altman analysis revealed normally distributed and homoscedastic differences (DXA – optimized 5-

Table 7. Multi-linear regression analysis to predict appendicular lean tissue using the optimized 5-site protocol

Model development	Appendicular lean tissue prediction (kg)	Cross-validation fold	Unadjusted r^2	Adjusted r^2	SEE (kg)	p-value model
Folds 1+2	$2.801+(1.564*X_2)-(2.011*X_3)+(0.0253*X_4)$	3	0.92	0.91	1.62	<0.001
Folds 1+3	$3.509+(1.507*X_2)-(2.105*X_3)+(0.0238*X_4)$	2	0.90	0.90	1.72	<0.001
Folds 2+3	$2.477+(1.594*X_2)-(1.838*X_3)+(0.0250*X_4)$	1	0.93	0.93	1.54	<0.001
Average	$2.929+(1.555*X_2)-(1.985*X_3)+(0.0247*X_4)$	-	0.91	0.91	1.62	-

SEE, standard error of the estimate. X_2 = 5 site muscle thickness x height (cm x m), X_3 = sex (male=0, female=1), X_4 = age (years)

site protocol) for appendicular lean tissue, with no proportion bias present ($p<0.05$). A non-significant fixed bias of 0.00 kg [95% CI: -0.33, 0.33 kg] was present with limits of agreement of -3.18 and 3.18



kg, and tolerance limits of -3.75 and 3.75 kg (Figure 11).

Figure 11. Bland-Altman plot comparing DXA derived and the optimized 5-site predicted appendicular lean tissue mass, utilizing participants from all folds. No fixed (0.00 [-0.33, 0.33]) or proportional bias was present (solid black line, 95% CI – inner short dashed line), with limits of agreement (1.96 SD) of -3.18 and 3.18 (middle long dashed lines) and tolerance limits of -3.75 and 3.75 (outer short dashed lines). Crosses (×) represent fold 1, closed circles (●) represent fold 2, and open circles (○) represent fold 3.

Multi-logistic regression to identify individuals with lower than normal lean tissue mass, using the optimized 5-site protocol, across the model development folds resulted in an average odds ratio [95% CI] for the variables 5-site*height, sex and age of 0.255 [0.107, 0.613], 0.02 [0.001, 0.669]

and 0.972 [0.929, 1.017], respectively (Table 8). Across the cross-validation folds, an average concordance index of 0.89 [0.78, 1.00] was observed (Figure 12).

Table 8. Logistic regression analysis to predict presence of low appendicular lean tissue using the optimized 5-site protocol

Model Development	Log Odds – low lean tissue	Odds Ratio [95% CI]			Cross-validation fold	c-index [95% CI]
		5-site	Sex (females=1)	Age		
Folds 1+2	13.209-(1.157*X_2)-(3.415*X_3)-(0.0217*X_4)	0.314 [0.147, 0.674]	0.033 [0.001, 0.857]	0.979 [0.939, 1.020]	3	0.96 [0.89, 1.00]
Folds 1+3	17.914-(1.554*X_2)-(4.380*X_3)-(0.0355*X_4)	0.211 [0.079, 0.569]	0.013 [0.000, 0.501]	0.965 [0.917, 1.016]	2	0.85 [0.71, 0.99]
Folds 2+3	16.431-(1.431*X_2)-(4.198*X_3)-(0.0275*X_4)	0.239 [0.096, 0.596]	0.015 [0.000, 0.650]	0.973 [0.932, 1.016]	1	0.87 [0.74, 1.00]
Average	15.851-(1.381*X_2)-(3.998*X_3)-(0.0282*X_4)	0.255 [0.107, 0.613]	0.020 [0.001, 0.669]	0.972 [0.929, 1.017]	-	0.89 [0.78, 1.00]

c-index, concordance index; CI, confidence interval. X_2 = 5 site muscle thickness x height (cm x m), X_3 = sex (male=0, female=1), X_4 = age (years)

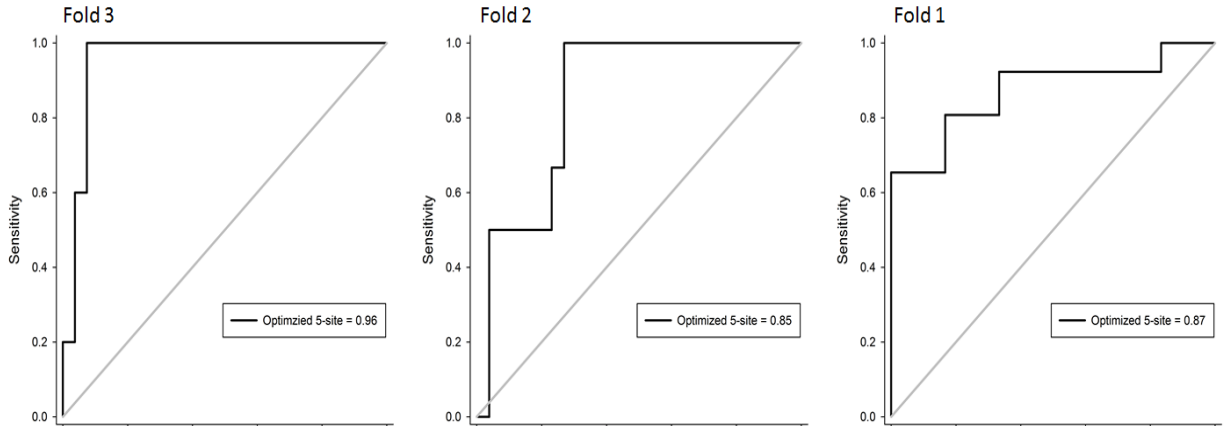


Figure 12. ROC curves for folds 3, 2 and 1, utilizing the optimized 5-site protocol to identify lower than normal lean tissue mass identified by DXA cut-points. AUC for folds 3, 2 and 1 were 0.96, 0.85 and 0.87, respectively, average of 0.89.

5.5 Reliability of the 4-site protocol

Intra-rater reliability for the 4-site protocol, assessing image acquisition and caliper placement, revealed CV and ICC [95% CI] using minimal compression of 1.1% and 0.998 [0.996, 0.998] respectively, and for maximal compression, 2.5% and 0.989 [0.983, 0.993] (Table 9). Inter-rater reliability of the 4-site protocol, assessing the entire ultrasound protocol, revealed CV and ICC using minimal compression of 3.7% and 0.988 [0.966, 0.996], respectively, and for maximal compression, 9.0% and 0.945 [0.843, 0.981] (Table 9).

Table 9. Intra-rater and inter-rater reliability of the 4-site protocol

	Intra-rater	
Compression	Minimal	Maximal
ICC (1,1) [95% CI]	0.998 [0.996, 0.998]	0.989 [0.983, 0.993]
CV (%)	1.1	2.5
	Inter-rater	
Compression	Minimal	Maximal
ICC (1,1) [95% CI]	0.988 [0.966, 0.996]	0.945 [0.843, 0.981]
CV (%)	3.7	9.0

CV, coefficient of variation; CI, confidence interval; ICC, intra-class correlation coefficient.

Reliability was further assessed using Bland-Altman plots. Intra-rater plots, for both minimal and maximal compression revealed normal and homoscedastic differences (1^{st} measure – 2^{nd} measure) for the 4-site protocol, with proportional bias in a single plot (minimal compression, $p<0.05$). Minimal compression average bias [95% CI] for all intra-rater plots (except plot with proportional bias) was -0.04 [-0.03, 0.01 cm] and for maximal compression, a significant fixed bias of -0.02 [-0.04, -0.01 cm] (Appendix A5). Minimal compression average limits of agreement and tolerance limits were -0.14 and 0.12 cm and -0.16 and 0.14 cm and for maximal compression, -0.16 and 0.11 cm and -0.18 and 0.14 cm (Appendix A6).

Inter-rater Bland-Altman plot for minimal compression and maximal compression demonstrated normally distributed and homoscedastic differences (Rater BL – Rater MP) for the 4-site protocol, with no proportional bias ($p>0.05$). For minimal compression, non-significant fixed bias

[95% CI] of -0.05 cm [-0.15, 0.05 cm], with limits of agreement and tolerance limits of -0.41 and 0.31 cm and -0.58 and 0.49 cm, respectively (Figure 13). For maximal compression, non-significant fixed

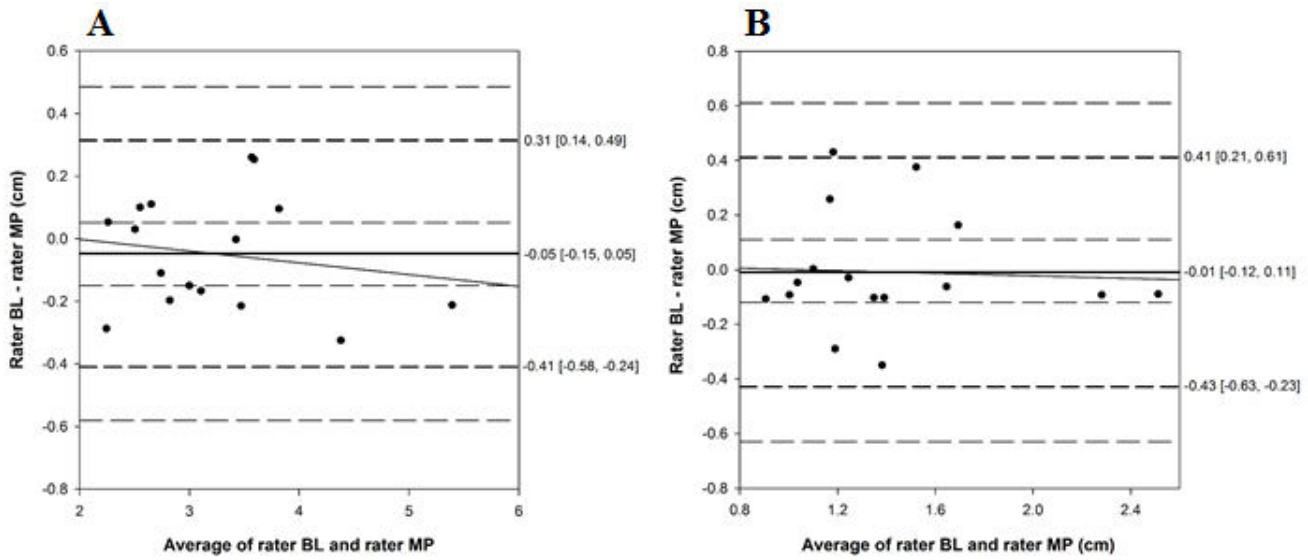


Figure 13. Bland-Altman plots for inter-rater reliability using the 4-site protocol for A) minimal and B) maximal compression. A) Minimal compression: no fixed (-0.05 [-0.15, 0.05]) or proportional bias was present (solid black line, 95% CI – inner long dashed lines), with limits of agreement (1.96 SD) of -0.41 and 0.31 (middle short dashed lines) and tolerance limits of -0.58 and 0.49 (outer long dashed lines). B) Maximal compression: no fixed (-0.01 [-0.12, 0.11]) or proportional bias was present (solid black line, 95% CI – inner long dashed lines), with limits of agreement (1.96 SD) of -0.43 and 0.31 (middle short dashed lines) and tolerance limits of -0.63 and 0.61 (outer long dashed lines).

bias of -0.01 cm [-0.12, 0.11 cm], with limits of agreement and tolerance limits of -0.43 and 0.41 cm and -0.63 and 0.61 cm, respectively (Figure 13).

Chapter 6
Discussion, Future Directions and Conclusions

6.1 Discussion

This thesis sought to develop and internally validate a viable bedside ultrasound protocol to predict appendicular lean tissue mass and identify individuals with lower than normal lean tissue mass in a heterogeneous cohort of healthy participants. We demonstrated that a previously developed 4-site protocol (2) is strongly (r^2=0.72, SEE=2.88 kg) associated with DXA derived appendicular lean tissue mass and has excellent intra-rater and inter-rater reliability. However, by accounting for the anterior upper arm muscle thickness, sex and age in the regression analysis for the 4-site protocol, we observed improved model performance (r^2=0.91, SEE=1.62 kg) for predicting appendicular lean tissue mass, while still developing a protocol that is feasible at the bedside. Interestingly, this optimized 5-site protocol matched the accuracy of the 9-site protocol (r^2=0.90, SEE=1.73 kg), within this thesis, and demonstrated excellent ability to discriminate between low and normal lean tissue mass individuals (AUC=0.89).

6.1.1 The 4-site protocol may have limited accuracy in predicting appendicular lean tissue mass

As a first step, it was critical to demonstrate that ultrasound measured muscle thickness using the 4-site protocol was a valid representation of site specific lean tissue assessed using DXA. We observed strong associations when comparing average muscle thickness from the 4-site protocol (rectus femoris and vastus intermedialis) with site specific measures of lean tissue from DXA (average upper leg lean tissue mass for regions 1-4, Figure 6, page 30). This is in agreement with previous work which demonstrated that the muscle thickness of specific muscle groups (anterior upper arm, posterior upper arm, anterior upper leg and posterior lower leg) are strongly associated with their corresponding muscle volume measured using MRI (84). While both minimal and maximal

compression were strongly associated with site specific lean tissue, minimal compression displayed a significantly stronger association compared to maximal compression, which has also recently been demonstrated in liver cirrhotic patients using the 4-site protocol (41). These results suggest that minimal compression measures of muscle thickness are valid estimates of lean tissue and may be more useful for predicting appendicular lean tissue mass compared with maximal compression.

While a number of bedside applicable ultrasound protocols have been developed and utilized in various clinical populations, few have been compared to accurate whole body measures of lean tissue or muscle mass. Berger et al. (2015) observed that the muscle thickness at the midpoint of the rectus femoris, averaged across both legs, strongly associated (r^2=0.55) with appendicular lean tissue mass measured by DXA. While similar landmarks were used in this thesis, the 4-site protocol is more comprehensive, imaging both the midpoint and lower third of the anterior thigh, representing rectus femoris and vastus intermedialis muscle thicknesses. These additional landmarks and depth of measurement (inclusive of the vastus intermedialis), in addition to accounting for limb length, may partly explain why we observed a stronger association (r^2=0.72); especially since the participant cohorts recruited were quite similar, a combination of younger and older adults, of both sexes, with wide ranges of BMI (58). A more direct comparison of ultrasound protocols applied can be made to a previous collaborative investigation we published, in which the 4-site protocol was compared with CT based abdominal measures of muscle mass in 145 mixed medical and surgical critically ill patients (3). A moderate (r^2 =0.20) correlation was observed, which may seem in direct opposition to the results from this thesis. However, this deviation may be explained by: 1) comparing a peripheral limb muscle thickness to an abdominal measure of muscle mass, which albeit was the most feasible approach for muscle mass validations in critically ill patients, and 2) utilizing maximal compression of the ultrasound probe against the skin, which this thesis and others (41) have demonstrated, may be less accurate than minimal compression (3).

While regression analysis is a necessary step since we are comparing two distinct tissue compartments, there are limitations with relying upon regression analysis as the sole indicator of model performance. Regression analysis examines how two measures relate to each other, and does not assess the absolute agreement at the individual level. It is therefore suggested to assess model performance using both regression and Bland-Altman analysis (93). Bland-Altman analysis involves clinically interpreting if the limits of agreement for the differences between methods represent an acceptable level of error. This is a rather difficult task for measures of lean tissue mass, as there is no currently defined clinically acceptable limits of agreement, and therefore, depending on what the authors consider an acceptable level of error, there may be large discrepancies in the interpretation of a Bland-Altman plots.

One potential avenue to determine acceptable limits of agreement for ultrasound assessments of lean tissue is to examine literature utilizing Bland-Altman analysis of the 9-site protocol; as this can be considered the most accurate and comprehensive ultrasound protocol to date. Except for the initial study developing and cross-validating the 9-site protocol (48) (reference technique - hydrostatic weighing), every publication that has developed and tested the 9-site, or a subset of muscle thicknesses from the 9-site protocol, revealed limits of agreement no larger than ±4 kg (47; 49; 52; 94). Therefore, limits of agreement that fall within ±4 kg may be a potential level of error to be considered acceptable for newly developed ultrasound protocols that predict lean tissue or muscle mass. However, the approach we are utilizing in this thesis is to accept limits of agreement based on the average difference between normal and low appendicular lean tissue mass individuals in a large cohort of community dwelling older adults. Scott et al. (2016), used previously published appendicular lean tissue cut-points (96) and observed that the average difference between normal and low lean tissue in 1100 Caucasian older adults (50-79 years old) was 3.3 kg, which was associated with an increased risk of falls. Therefore, when comparing DXA measured and ultrasound predicted

appendicular lean tissue mass (from the 4-site, 9-site and optimized 5-site protocols), Bland-Altman limits of agreement below ±3.3 kg will be considered an acceptable level of error. Within this participant cohort, 3.3 kg of appendicular lean tissue mass, when normalized average height squared, equates to 1.23 kg/m^2 and 1.05 kg/m^2 for females and males respectively. Comparing these values with previously established low lean tissue cut-points, 5.45 kg/m^2 for females and 7.26 kg/m^2 for males, demonstrates that a change corresponding to 3.3 kg of appendicular lean tissue may result in large discrepancies in the classification of individuals as low or normal lean tissue mass. Therefore, even if Bland-Altman analysis reveals limits of agreement below 3.3 kg, the ability to identify individuals with low lean tissue mass should be assessed, and ultimately, compared with measures of strength or physical function, to determine the usefulness of these measures for assessing functional status.

Bland-Altman analysis for the 4-site protocol resulted in a proportional bias, and therefore hyperbolic limits of agreement were constructed as suggested previously (90; 93), making interpretation challenging. Here, we took a conservative approach and observed the range of differences for the 95% prediction interval (hyperbolic limits of agreement) at the widest points (highest and lowest values of appendicular lean tissue mass), observing ranges of 11.33 kg, or if we view this range similar to traditional limits of agreement, ±5.67 kg. Limits of agreement, of that magnitude, are quite similar to many single frequency BIA estimates of fat-free mass, often considered accurate for population averages (no fixed bias), but far too inaccurate for individual measures of fat-free mass (wide limits of agreement) (97). These limits of agreement are wider than our clinically acceptable limits of ±3.3 kg and, therefore, despite the strong associations (r^2=0.72), the 4-site protocol estimates of appendicular lean tissue would be considered an unacceptable level of error.

6.1.2 Distribution differences of lean tissue mass in a heterogeneous participant cohort increase the variability of 4-site protocol predictions

Although the 4-site protocol, in this thesis, demonstrated strong associations with appendicular lean tissue (r^2=0.72), there were wide limits of agreement and much unexplained variance. This level of error and unexplained variance for lean tissue predictions may partly be attributed to the participant cohort we recruited; as they were heterogeneous in terms of physical and demographic characteristics, such as age, sex and BMI. These physical and demographic characteristics may result in altered distribution of lean tissue mass between participants, which quadriceps muscle thickness is unable to account for. For example, there are sex-based differences in the distribution of skeletal muscle mass, such that females have approximately 70% of the muscle mass in the lower limbs in comparison to males, but only 50-60% for the upper limb musculature (73; 98). Since the main outcome being predicted is appendicular lean tissue mass (upper and lower limb lean tissue), a lower limb muscle thickness will be unable to account for sex-based differences in the musculature of the upper limbs. On the other hand, aged individuals typically display increased atrophy of the lower limb musculature, relative to the upper limbs (73; 76; 99), and relying solely upon a lower limb muscle thickness, we would be unable to account for upper limb differences. BMI may also present a problem, particularly if DXA is used as the reference technique. Since DXA cannot distinguish muscle from lean tissue mass, and approximately 15-20% of adipose tissue is considered lean tissue (connective tissues) (5; 100), individuals with excess adiposity will have lean tissue that is unaccounted for from a measure of muscle thickness. Our participant cohort was intended to recruit males and females, with wide ranges of age, BMI and lean tissue mass for greater generalizability and this may have increased the variability seen with lean tissue mass predictions.

6.1.3 The 9-site protocol agrees with previous literature demonstrating high accuracy in predicting lean tissue mass

Since the 9-site protocol is a comprehensive assessment of many muscle groups, it is not surprising that this protocol may be able to account for differences in lean tissue mass distributions based on sex and age. Here, we observed strong associations (r^2=0.90, SEE=1.73) between the 9-site protocol and appendicular lean tissue mass; which is slightly lower, but similar (r^2=0.94-0.97) to previous literature comparing the 9-site protocol to accurate measures of lean tissue or muscle mass (47; 49; 52). This slightly lower association may be attributed to our heterogeneous participant cohort, as previous studies have been performed in more homogeneous cohorts, focusing mainly on older or younger adults. Another potential issue that may have contributed to our lower association is with the subscapular landmark. Originally the 9-site protocol is performed in a standing posture, whereas all of our landmarking occurred in a supine or prone position, which may have imposed more variability on the position of the inferior angle of the scapula (more lateral translocation compared to a standing posture). This change in location of the landmark often led to lower than expected muscle thicknesses, which cannot be entirely accounted for due to a change in position from standing to supine (68). While the subscapular muscle thickness may have been altered due to landmarking issues, changes in posture, from standing to sitting, also alters measured muscle thickness for a number of landmarks for the 9-site protocol (68). English et al. (2009) determined that seven of the nine muscle thicknesses from the 9-site protocol are significantly reduced when measured in supine compared with standing position. Taken together, these issues may partly explain why a slightly lower association was observed in this thesis compared with previous literature. Although the regression analysis was slightly lower than previous publications, the limits of agreement (-3.32 and 3.32 kg) were similar, as they fall below the ±4 kg generally seen with the 9-site protocol and below

our clinically defined limits of ±3.3 kg limits. Overall, the 9-site protocol, in a supine posture, can accurately predict appendicular lean tissue mass.

6.1.4 The optimized 5-site protocol can accurately predict appendicular lean tissue mass

While the 9-site protocol may be accurate, it cannot be easily applied at the bedside due to the inclusion of posterior landmarks, which would require a patient to be mobile or be moved to acquire those muscle thicknesses. We sought to optimize the accuracy of the 4-site protocol using bedside accessible muscle thicknesses from the 9-site protocol and easily obtained covariates that may influence appendicular lean tissue mass. The final optimized model included accounting for sex, age and the muscle thicknesses of the anterior upper arm and the 4-site protocol; interestingly, this model performed (r^2=0.91) similar to the 9-site protocol (r^2=0.90) for appendicular lean tissue predictions. A similar approach has been undertaken previously, Campbell et al. (1995), determined that the anterior thigh, anterior upper arm and anterior forearm, in a supine posture, strongly (r^2=0.76) associated with whole body lean tissue mass measured by DXA. While similar muscle groups were imaged, we likely observed stronger associations because we accounted for age, sex and body proportions (height) and our comparison was against appendicular lean tissue mass, whereas Campbell and colleagues (1995), compared against whole body lean tissue mass.

Optimization of the 4-site protocol with easily obtainable covariates has also been attempted previously. Two studies, comparing the 4-site protocol to abdominal muscle CSA, in ICU (3) and liver cirrhotic (41) patients, performed multi-variate regression using easily obtainable variables to improve model performance; within both models, BMI was included. Although potentially useful in a model to predict lean tissue, as weight is moderately associated with lean tissue mass (101), weight may be a rather difficult parameter to accurately obtain, depending on patient mobility/consciousness or the equipment available within a specific clinic. Therefore, a major advantage of our model is that

it is weight-independent, and can be entirely performed with the patient in a supine position, without requiring an accurate assessment of weight. However, a caveat for our model is that all participants' heights were measured in a standing posture; whereas for a patient unable to stand, heights would be measured supine or estimated from anthropometric equations, which may result in additional variability of model performance.

When examining Bland-Altman analysis for the optimized 5-site protocol, no significant fixed or proportional bias was present, and the limits of agreement (-3.18, 3.18 kg) are within the ±3.3 kg clinically acceptable limits as described above. While the tolerance limits (upper and lower 95% confidence interval for the limits of agreement; -3.75, 3.75 kg) do not fall within the ±3.3 kg limits of agreement, they are still within the ±4 kg generally seen with the 9-site protocol. While the limits of agreement are generally used as the final decision of being acceptable, these wider tolerance limits demonstrate the need for external validation of the optimized 5-site protocol. We can therefore accept that the optimized 5-site protocol can accurately predict appendicular lean tissue mass on par with the 9-site protocol. While the worst case scenario (for this participant cohort), a difference of 5.11 kg, would represent an unacceptable level of error and alter how this individual is identified as low or normal lean tissue, a difference of 5.11 kg represents over 3 standard deviation (less than 0.3% of individuals) difference from the mean, indicating a rather rare occurrence and would not alter the application of this protocol.

6.1.5 Optimized 5-site protocol can accurately identify individuals with lower than normal lean tissue mass

Using the optimized 5-site protocol, we observed a strong (AUC=0.89) ability to discriminate low and normal lean tissue individuals. Low lean tissue mass was identified using appendicular skeletal muscle index cut-points previously published by Baumgartner et al. (1998). While we recruited an expected proportion of low lean tissue mass individuals in the older adults (27%) (4;

101), a larger than expected proportion of individuals with low lean tissue mass was observed in our younger cohort (~11%), as we would expect approximately 2.5% of young individuals to fall below 2 standard deviations of the average appendicular lean tissue of a young healthy reference group. Although surprising, this discrepancy may be partly explained by the fact that the original cut-points were developed in a cohort of young adults from 1986-1992, who may have led more active lifestyles in comparison to current physical activity levels (102). It is therefore difficult to exactly determine how these cut-points relate to poor physical function.

Two previously discussed studies, using the 4-site protocol in ICU and liver cirrhotic patients, have also performed ROC curves to identify low muscle mass using CT based cut-points (3; 41). A wider variation in performance was observed, AUC=0.77 – 0.89, depending on patient cohort and sex, but still demonstrated moderate to strong ability to identify low muscle. While promising, these studies exhibit similar challenges to those presented in this thesis, the ultrasound models used to identify those individuals were developed and applied within the same cohort of individuals. Although we attempted to account for overfitting with cross-validation, external validation of these models is needed to ensure the usefulness of these measures in accurately identifying individuals with low lean tissue or muscle mass.

6.1.6 The 4-site protocol demonstrates good reliability within and between raters

All of the measures used for model training and cross-validations were performed by a single ultrasound operator, and it is therefore necessary to assess the reliability of these measures both within and between raters, as any given rater has influence on the landmarking, image acquisition and caliper placement, all of which may alter the final result. Reliability testing was only performed for the 4-site protocol because the development of the optimized 5-site protocol occurred following completion of data collection. Our results, ICC's and CV's for both minimal and maximal compression, agree with other studies demonstrating the strong reliability of the 4-site protocol (3;

56) and were consistent with a recent systematic review suggesting that ultrasound is a reliable modality to measure muscle thickness for numerous landmarks (82). For this thesis, intra-rater reliability was performed on previously defined landmarks, and therefore these results do not apply to the entire protocol To determine whether the intra-rater Bland-Altman limits of agreement (-0.14, 0.12 cm) are acceptable, we examined the average change in the rectus femoris and vastus intermedialis muscle thicknesses after 3 days in the ICU from a recent ultrasound study, and determined that limits of agreement less than the average change over 3 days would be considered acceptable. Parry et al. (2016), observed that after 3 days in an ICU, a 10% reduction occurred in the rectus femoris and vastus intermedialis muscle thicknesses using similar landmarks to this thesis; which corresponded to a 0.435 cm loss compared to baseline values. Since our limits of agreement for minimal compression are smaller than the loss observed after 3 days in the ICU, we conclude that these are acceptable. Furthermore, utilizing the 4-site regression equation, changes in muscle thickness of 0.14 and 0.12, correspond to 0.63 kg and 0.54 kg of appendicular lean tissue mass, or 3% of the average appendicular lean tissue mass; further demonstrating the acceptable nature of these limits of agreement. This approach cannot be directly applied to the limits of agreement using maximal compression, as no studies have tracked changes in muscle thickness using maximal compression. But if we assume a theoretical 10% reduction over 3 days, it would result in a reduction of 0.13 cm (baseline maximal compression muscle thickness using the 4-site protocol in ICU patients (3)), which is smaller than the observed limits of agreement for maximal compression of the 4-site protocol (-0.16, 0.11 cm) and would be considered unacceptable.

Our inter-rater results should be interpreted with caution due to the small sample size. But interestingly, we observed strong ICC's and CV's for both minimal and maximal compression, despite performing inter-rater reliability for the entire 4-site protocol. Similar results have been seen in acute stroke patients, where the entire protocol for the anterior upper leg using the 9-site protocol,

was shown to be reliably measured within the same rater (83). Although not assessed in this thesis, English et al. (2012) also demonstrated that the intra-rater reliability for the entire anterior upper arm protocol was also reliable. While the inter-rater reliability of the anterior upper arm needs to be assessed, these results are still promising for the reliability of the optimized 5-site protocol. Inter-rater Bland-Altman analysis revealed limits of agreement of -0.41 and 0.31 cm for minimal compression. Using the same approach as described above, our limits of agreement are again smaller than the 3 day reduction in muscle thickness for ICU patients and therefore are considered acceptable. Applying these limits of agreement to the 4-site regression equation results in changes of 1.8 kg and 1.4 kg of appendicular lean tissue mass, or 9% of the average appendicular lean tissue mass; which is below a commonly suggested cutoff of 10% (93). However, for maximal compression, the limits of agreement, -0.43 and 0.41 cm, are much wider than the theoretical muscle thickness loss of 0.13 cm and would be considered unacceptable. Overall, minimal compression demonstrates stronger reliability than maximal compression, in addition to being more accurate, and should therefore be utilized moving forward.

The participant cohort that had inter-rater reliability measures performed was a group of predominately older adults, alternating between dominant and non-dominant legs to ensure the potentially higher quality muscle was not over represented. While these results may only apply to older adults, we suspect the results may be further improved if performed on a younger group of individuals, since they tend to have lower body fat, allowing easier identification of bony landmarks.

6.1.7 Limitations

While these results are promising, this thesis has a number of limitations, in addition to those discussed above. When landmarking for the 4-site protocol, a straight line connecting the anterior superior iliac spine and the mid-upper patella is used as central vertical line for the ultrasound transducer to image. In some cases, this landmarking approach may not line up directly on the rectus

femoris and vasutus intermedialis (Appendix A4), providing a less than ideal muscle thickness (i.e. not measuring the greatest thickness of the muscle belly). In future investigations, allowing lateral movement of the ultrasound transducer to image the ideal landmark may further reduce the variability in predicting appendicular lean tissue. Second, no participants in this thesis were categorized as being underweight (BMI<18.5 kg/m^2), which may limit the applicability of these results to those individuals who may be on the lowest end of the appendicular lean tissue spectrum; but, even in the majority of clinical populations, typically less than 10% of patients fall below 18.5 kg/m^2 (103). Furthermore, the older adult cohort that was recruited was a high functioning group of individuals and therefore these results may not be generalizable to those older adults that are potentially at higher risk of poor physical function. Fourth, the reliability results should be interpreted with caution, due to both a small sample size and because all analysis was performed using two fixed raters. Due to having two fixed raters, with averaged muscle thicknesses for each landmark, it has been suggested to apply the ICC (3,2) equation, which give the highest values, but is the least generalizable (87), whereas others suggest applying the more conservative ICC (1,1) equation (104). Regardless of which equation is applied, the design of using two fixed raters would only allow generalization of these results to these specific raters, and not others, limiting the applicability of these reliability results. Lastly, we assume that any unexplained variation or error is associated with the ultrasound measures, and that there is no inherent error or variability with DXA measures of appendicular lean tissue. While DXA is rather precise (<2% CV for lean tissue) technique, there are still a number of limitations associated with its use. Specific to this thesis, we had a number of individuals with high body fat, which may have added lean tissue mass that muscle thickness measures will not account for, as discussed above (section 6.1.2).

6.2 Future directions

While this thesis has focused on developing an accurate, yet viable, bedside ultrasound protocol to assess appendicular lean tissue and identify individuals who may be at risk of low lean tissue mass, ultimately these measures need to be validated in specific clinical populations, to ensure they accurately and reliability identify those individuals who require nutritional or rehabilitative therapies targeted to improve muscle mass or function. Prior to that end goal, to ensure that the approach taken in this thesis is an accurate model, external validation of the optimized 5-site protocol would be ideal, both for prediction of appendicular lean tissue and identification of low lean tissue mass. Alongside that external validation study, a more robust reliability study design should be undertaken, involving much larger sample sizes and incorporating a wider array of randomly selected raters to ensure the generalizability of those findings. Following those external validation and reliability studies, it would be necessary to assess if the optimized 5-site protocol can: 1) determine if the participants identified as having lower than normal lean tissue mass display poorer physical function and 2) assess the ability to detect changes in muscle or lean tissue mass overtime.

External validation of the optimized 5-site protocol should be ideally performed in a similar heterogeneous participant cohort; but active recruitment of individuals in the underweight BMI ($\leq$ 18.5 kg/m^2) category may increase the proportion of individuals identified as being lower than normal lean tissue or muscle mass, which would improve the validity of the ROC analysis and increase the generalizability of the model. One aspect that should be accounted for in future designs, is ethnicity, as equations developed in Japanese populations perform poorly in Caucasian adults (49). While it is difficult to speculate on the potential causes of this discrepancy, it is not an uncommon issue, as Asian specific thresholds for BMI categories and CT based low muscle mass have been suggested (105; 106). Future external validation studies should attempt to stratify their participants by ethnicity or develop population specific equations. This external validation study could also further improve

upon the protocol, by allowing lateral movement of the 4-site landmarks to obtain the ideal muscle thickness for all individuals. Alongside that external validation study, a robust reliability study could be performed as well, in which numerous raters are utilized to assess the entire optimized 5-site protocol. The protocol should be broken down into three steps, landmarking, image acquisition and caliper placement. By using a digitizing probe (i.e. Optotrak) for landmarking, agreement for both palpation of bony surfaces and final landmark location can be assessed between raters. For image acquisition, a predefined landmark can be used for raters to acquire ultrasound scans of the underlying muscle groups. A single trained investigator can then quantify each scan for muscle thickness measures, to assess the agreement of raters in obtaining similar scans. For caliper placement, a single image, acquired from a single trained rater can be analyzed using offline image analysis software by different raters, to compare agreement in muscle thickness values. This comprehensive reliability setup would allow generalizability of these results, and determine which steps require further standardization to improve the applicability of these measures across multiple centres.

While accurate and reliable identification of individuals with low lean tissue mass using the optimized 5-site protocol is important, it is also crucial to determine if these individuals also display poor physical function. Comparing measures of strength and physical function between those individuals identified as low or normal lean tissue mass would determine whether the optimized 5-site protocol can identify not only low lean tissue, but also sarcopenic individuals (low muscle and poor strength or function). Ideally this cohort of individuals would consist of males and females, with a wide range of age and BMI, to ensure the generalizability of these results. The age factor is important to account for since older adults will have poorer strength and function, and these results may only be applicable to these individuals. BMI on the other hand may be important to account for due to an interesting theme that is gaining more attention, accounting for fat mass, alongside muscle mass, for

identifying individuals with poor function (107; 108). Previous studies, incorporating appendicular fat mass, into the assessment of low lean tissue using DXA resulted in better identification of individuals with poor physical function (108). This may be due to requiring more muscle mass for a given function if there is excess adiposity, due to poorer quality muscle in those individuals who are overweight or obese, or perhaps even associated with the limitation of DXA assessing the lean tissue components of adipose tissue. Regardless, moving forward, a potentially useful variable to include when attempting to identify individuals with poor physical function may be adipose tissue thickness; as it has been demonstrated that using adipose thickness from a variety of landmarks, can accurately predict whole body fat mass (44).

Lastly, it is necessary to accurately assess changes in muscle or lean tissue mass overtime. The sensitivity of the optimized 5-site protocol to detect meaningful changes should be assessed; this would allow researchers and clinicians to determine if nutritional or rehabilitative interventions result in improvements in muscle mass and quality or to determine if there is further deterioration of musculature. Assessing the agreement in changes of muscle quantity using the optimized 5-site ultrasound protocol, during both disuse induced atrophy (bed rest) and exercise related hypertrophy (resistance training), with changes quantified using accurate modalities (i.e. DXA, MRI), would reveal whether changes in the muscle thicknesses of the anterior upper arm and 4-site protocol (landmarks utilized in the optimized 5-site protocol) are representative of whole body changes in muscle or lean tissue mass.

6.3 Conclusions

This thesis developed and internally validated an accurate viable, bedside ultrasound protocol to predict appendicular lean tissue mass in a heterogeneous cohort of young and older adults. We have shown that minimal compression is more accurate than maximal compression of the 4-site

protocol for prediction of lean tissue mass; moving forward, for lean tissue estimates, minimal compression should be used. The previously developed 4-site protocol is strongly associated with appendicular lean tissue mass, but Bland-Altman analysis revealed wide limits of agreement and proportional bias; demonstrating that the 4-site protocol alone may not accurately predict appendicular lean tissue mass. We determined that the muscle thicknesses of the anterior upper arm and 4-site quadriceps protocols, in addition to sex and age, are important variables in predicting lean tissue and demonstrated that this optimized 5-site protocol can accurately predict appendicular lean tissue mass and identify individuals with lower than normal lean tissue mass. These results demonstrate that this viable bedside protocol may be useful for assessing lean tissue mass in clinical settings, but external validation in clinical populations is necessary to ensure the robustness of these findings.

References

1. Mourtzakis M, Wischmeyer P (2014): Bedside ultrasound measurement of skeletal muscle. *Curr Opin Clin Nutr Metab Care*. 17: 389–95.

2. Gruther W, Benesch T, Zorn C, Paternostro-Sluga T, Quittan M, Fialka-Moser V, *et al.* (2008): Muscle wasting in intensive care patients: ultrasound observation of the M. quadriceps femoris muscle layer. *J Rehabil Med*. 40: 185–9.

3. Paris M, Mourtzakis M, Day A, Leung R, Watharkar S, Kozar R, *et al.* (2016): VALIDation of bedside Ultrasound of Muscle layer thickness of the quadriceps in the critically ill patient (VALIDUM study): A prospective multicenter study. *J Parenter Enter Nutr*. . doi: 10.1177/0148607116637852.

4. Baumgartner RN, Koehler KM, Gallagher D, Romero L, Heymsfield SB, Ross RR, *et al.* (1998): Epidemiology of sarcopenia among the elderly in New Mexico. *Am J Epidemiol*. 147: 755–63.

5. Prado C, Heymsfield SB (2014): Lean Tissue Imaging: A New Era for Nutritional Assessment and Intervention. *JPEN J Parenter Enteral Nutr*. 38: 940–53.

6. Moisey LL, Mourtzakis M, Cotton B a, Premji T, Heyland DK, Wade CE, *et al.* (2013): Skeletal muscle predicts ventilator-free days, ICU-free days, and mortality in elderly ICU patients. *Crit Care*. 17: R206.

7. Filippin LI, Teixeira VNDO, da Silva MPM, Miraglia F, da Silva FS (2015): Sarcopenia: a predictor of mortality and the need for early diagnosis and intervention. *Aging Clin Exp Res*. 27: 249–54.

8. Psutka SP, Carrasco A, Schmit GD, Moynagh MR, Boorjian S a, Frank I, *et al.* (2014): Sarcopenia in patients with bladder cancer undergoing radical cystectomy: impact on cancer-specific and all-cause mortality. *Cancer*. . doi: 10.1002/cncr.28798.

9. Greening N, Harvey-Dunstan T, Chaplin E, Vincent E, Morgan M, Singh S, Steiner M (2015): Bedside assessment of quadriceps muscle using ultrasound following admission for acute exacerbations of chronic respiratory disease. *Am J Respir Crit Care Med*. 192: 810–816.

10. Leenders M, Verdijk LB, van der Hoeven L, Adam JJ, van Kranenburg J, Nilwik R, Van Loon LJC (2013): Patients with type 2 diabetes show a greater decline in muscle mass, muscle

strength, and functional capacity with aging. *J Am Med Dir Assoc*. 14: 585–592.

11. Di Sebastiano KM, Yang L, Zbuk K, Wong RK, Chow T, Koff D, *et al.* (2013): Accelerated muscle and adipose tissue loss may predict survival in pancreatic cancer patients: the relationship with diabetes and anaemia. *Br J Nutr*. 109: 302–12.

12. Seymour JM, Ward K, Sidhu PS, Puthucheary Z, Steier J, Jolley CJ, *et al.* (2009): Ultrasound measurement of rectus femoris cross-sectional area and the relationship with quadriceps strength in COPD. *Thorax*. 64: 418–423.

13. Hanai T, Shiraki M, Nishimura K, Ohnishi S, Imai K, Suetsugu A, *et al.* (2014): Sarcopenia impairs prognosis of patients with liver cirrhosis. *Nutrition*. 31: 193–199.

14. Reid C, Campbell I, Little R (2004): Muscle wasting and energy balance in critical illness. *Clin Nutr*. 23: 273–80.

15. Morley JE (2012): Sarcopenia in the elderly. *Fam Pract*. 29: 44–48.

16. Kirk PS, Friedman JF, Cron DC, Terjimanian MN, Wang SC, Campbell D a., *et al.* (2015): One-year postoperative resource utilization in sarcopenic patients. *J Surg Res*. . doi: 10.1016/j.jss.2015.04.074.

17. Janssen I, Shepard DS, Katzmarzyk PT, Roubenoff R (2004): The Healthcare Costs of Sarcopenia in the United States. *J Am Geriatr Soc*. 52: 80–85.

18. Heymsfield SB, Wang Z, Baumgartner RN, Ross R (1997): Human body composition: advances in models and methods. *Annu Rev Nutr*. 17: 527–558.

19. Lee R, Wang Z, Heymsfield SB (2001): Skeletal Muscle Mass and Aging: Regional and Whole-Body Measurement Methods. *Can J Appl Physiol*. 26: 102–122.

20. Chen Z, Wang Z, Lohman T, Heymsfield SB, Outwater E, Nicholas JS, *et al.* (2007): Dual-energy X-ray absorptiometry is a valid tool for assessing skeletal muscle mass in older women. *J Nutr*. 137: 2775–2780.

21. Geisler C, Pourhassan M, Braun W, Schweitzer L, Müller MJ (2015): The prediction of total skeletal muscle mass in a Caucasian population - comparison of Magnetic resonance imaging (MRI) and Dual-energy X-ray absorptiometry (DXA). *Clin Physiol Funct Imaging*. 2. doi: 10.1111/cpf.12282.

22. 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.

23. Kim J, Wang Z, Heymsfield SB, Baumgartner RN, Gallagher D (2002): Total-body skeletal muscle mass: estimation by a new dual-energy X-ray absorptiometry method. *Am J ClinNutr*. 76: 378–383.

24. LaForgia J, Dollman J, Dale MJ, Withers RT, Hill AM (2009): Validation of DXA body composition estimates in obese men and women. *Obesity (Silver Spring)*. 17: 821–826.

25. Heymsfield SB, Smith R, Aulet M, Bensen B, Lichtman S, Wang J, Pierson Jr. RN (1990): Appendicular skeletal muscle mass: Measurement by dual- photon absorptiometry. *Am J Clin Nutr*. 52: 214–218.

26. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, *et al.* (2010): Sarcopenia: European consensus on definition and diagnosis: Report of the European Working Group on Sarcopenia in Older People. *Age Ageing*. 39: 412–23.

27. Visser M, Kritchevsky S, Goodpaster B, Newman AB, Nevitt MC, Stamm E, Harris TB (2002): Leg Muscle Mass and Composition in Relation to Lower Extremity Performance in Men and Women Aged 70 to 79: The Health, Aging and Body Composition Study. *J Am Geriatr Soc*. 50: 897–904.

28. Lieffers JR, Mourtzakis M, Hall KD, McCargar LJ, Prado CMM, Baracos VE (2009): A viscerally driven cachexia syndrome in patients with advanced colorectal cancer: Contributions of organ and tumor mass to whole-body energy demands. *Am J Clin Nutr*. 89: 1173–1179.

29. Lifshitz F, Hecht JP, Bermúdez EF, Gamba CA, Reinoso JM, Casavalle PL, *et al.* (2016): Body composition analysis by dual-energy X-ray absorptiometry in young preschool children. *Eur J Clin Nutr*. 1–7.

30. Nana, Alisa; Slater, Gary; Stewart, Arthur & Burke L (2014): Methology review: Uisng dual-energy x-ray absorptiometry (DXA) for the assessment of body composition in athletes and active people. *Int J Sport Nutr Exerc*. 1–44.

31. Oldroyd B, Smith a H, 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.

32. Toombs RJ, Ducher G, Shepherd JA, De Souza MJ (2012): The impact of recent technological advances on the trueness and precision of DXA to assess body composition. *Obesity (Silver Spring)*. 20: 30–9.

33. Paris M, Mourtzakis M (2016): Assessment of skeletal muscle mass in critically ill patients. *Curr Opin Clin Nutr Metab Care*. 19: 125–130.

34. Tandon P, Ney M, Irwin I, Ma M, Gramlich L, Bain V, *et al.* (2012): Severe Muscle Depletion in Patients on the Liver Transplant Wait List: Its Prevalence and Independent Prognostic Value. *Liver Transplant*. 18: 1209–1216.

35. Tamandl D, Paireder M, Asari R, Baltzer PA, Schoppmann SF, Ba-Ssalamah A (2015): Markers of sarcopenia quantified by computed tomography predict adverse long-term outcome in patients with resected oesophageal or gastro-oesophageal junction cancer. *Eur Radiol*. . doi: 10.1007/s00330-015-3963-1.

36. Mitsiopoulos N, Baumgartner RN, Heymsfield SB, Lyons W, Gallagher D, Ross R (1998): Cadaver validation of skeletal muscle measurement by magnetic resonance imaging and computerized tomography. *J Appl Physiol*. 85: 115–122.

37. Shen W, Punyanitya M, Wang Z, Gallagher D, St-Onge M-P, Albu J, *et al.* (2004): Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J Appl Physiol*. 97: 2333–8.

38. Mourtzakis M, Prado CMM, Lieffers JR, Reiman T, McCargar LJ, Baracos VE (2008): A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl Physiol Nutr Metab*. 33: 997–1006.

39. Weijs PJM, Looijaard WGPM, Dekker IM, Stapel SN, Girbes AR, Oudemans-van Straaten HM, Beishuizen A (2014): Low skeletal muscle area is a risk factor for mortality in mechanically ventilated critically ill patients. *Crit Care*. 18: 1–7.

40. Baracos V, Reiman T, Mourtzakis M, Gioulbasanis I, Antoun S (2010): Body composition in patients with non–small cell lung cancer: a contemporary view of cancer cachexia with the use of computed tomography image analysis. *Am J Clin Nutr*. 91: 1133–1137.

41. Tandon P, Low G, Mourtzakis M, Zenith L, Myers R, Abraldes J, *et al.* (2016): A Model to Identify Sarcopenia in Patients with Cirrhosis. *Clin Gastroenterol Hepatol*. . doi: 10.1016/j.toxicon.2008.05.012.

42. Prado CM, Lieffers JR, McCargar LJ, Reiman T, Sawyer MB, Martin L, Baracos VE (2008): Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol*. 9: 629–635.

43. Zaidman CM, Wu JS, Wilder S, Darras BT, Rutkove SB (2014): Minimal training is required to reliably perform quantitative ultrasound of muscle. *Muscle Nerve*. 50: 124–8.

44. Wagner DR (2013): Ultrasound as a tool to assess body fat. *J Obes*. 2013. doi: 10.1155/2013/280713.

45. Topchyan A, Tatarinov A, Sarvazyan N, Sarvazyan A (2006): Ultrasound velocity in human muscle in vivo: Perspective for edema studies. *Ultrasonics*. 44: 259–264.

46. Cartwright MS, Demar S, Griffin LP, Balakrishnan N, Harris JM, Walker FO (2013): Validity and reliability of nerve and muscle ultrasound. *Muscle Nerve*. 47: 515–21.

47. Takai Y, Ohta M, Akagi R, Kato E, Wakahara T, Kawakami Y, *et al.* (2014): Applicability of ultrasound muscle thickness measurements for predicting fat-free mass in elderly population. *J Nutr Health Aging*. 18: 579–85.

48. Abe T, Kondom M, Kawakami Y, Fukunaga T (1994): Prediction Equations for Body Composition of Japanese Adults by B-mode Ultrasound. *Am J Hum Biol*. 170: 161–170.

49. Sanada K, Kearns CF, Midorikawa T, Abe T (2006): Prediction and validation of total and regional skeletal muscle mass by ultrasound in Japanese adults. *Eur J Appl Physiol*. 96: 24–31.

50. Johnson KE, Miller B, Juvancic-Heltzel J a, Agnor SE, Kiger DL, Kappler RM, Otterstetter R (2014): Agreement between ultrasound and dual-energy X-ray absorptiometry in assessing percentage body fat in college-aged adults. *Clin Physiol Funct Imaging*. 34: 493–6.

51. Schoenfeld BJ, Aragon AA, Moon J, Krieger JW, Tiryaki-Sonmez G (2016): Comparison of amplitude-mode ultrasound versus air displacement plethysmography for assessing body composition changes following participation in a structured weight-loss programme in women. *Clin Physiol Funct Imaging*. . doi: 10.1111/cpf.12355.

52. Toda Y, Kimura T, Taki C, Kurihara T, Homma T, Hamaoka T, Sanada K (2016): New

ultrasonography-based method for predicting total skeletal muscle mass in male athletes. *J Phys Ther Sci*. 28: 1556–1559.

53. Abe T, Dabbs NC, Nahar VK, Ford MA, Bass M a., Loftin M (2013): Relationship between Dual-Energy X-Ray Absorptiometry-Derived Appendicular Lean Tissue Mass and Total Body Skeletal Muscle Mass Estimated by Ultrasound. *Int J Clin Med*. 04: 283–286.

54. Abe T, Thiebaud RS, Loenneke JP, Young KC (2015): Prediction and validation of DXA-derived appendicular lean soft tissue mass by ultrasound in older adults. *Age (Omaha)*. 37: 1–10.

55. Kendall KL, Fukuda DH, Hyde PN, Smith-Ryan AE, Moon JR, Stout JR (2016): Estimating fat-free mass in elite-level male rowers: a four-compartment model validation of laboratory and field methods. *J Sports Sci*. 0414: 1–10.

56. Tillquist M, Kutsogiannis DJ, Wischmeyer PE, Kummerlen C, Leung R, Stollery D, *et al.* (2014): Bedside ultrasound is a practical and reliable measurement tool for assessing quadriceps muscle layer thickness. *JPEN J Parenter Enteral Nutr*. 38: 886–90.

57. Agyapong-Badu S, Warner M, Samuel D, Narici M, Cooper C, Stokes M (2014): Anterior thigh composition measured using ultrasound imaging to quantify relative thickness of muscle and non-contractile tissue: a potential biomarker for musculoskeletal health. *Physiol Meas*. 35: 2165–76.

58. Berger J, Bunout D, Barrera G, de la Maza MP, Henriquez S, Leiva L, Hirsch S (2015): Rectus femoris (RF) ultrasound for the assessment of muscle mass in older people. *Arch Gerontol Geriatr*. . doi: 10.1016/j.archger.2015.03.006.

59. Takai Y, Ohta M, Akagi R, Kato E, Wakahara T, Kawakami Y, *et al.* (2013): Validity of ultrasound muscle thickness measurements for predicting leg skeletal muscle mass in healthy Japanese middle-aged and older individuals. *J Physiol Anthropol*. 32: 12.

60. Campbell IT, Watt T, Withers D, England R, Sukumar S, Keegan M, *et al.* (1995): Muscle thickness, measured with ultrasound, may be an indicator of lean tissue wasting in multiple organ failure in the presence of edema. *Am J Clin Nutr*. 62: 533–539.

61. Rosenberg JG, Ryan ED, Sobolewski EJ, Scharville MJ, Thompson BJ, King GE (2014): Reliability of panoramic ultrasound imaging to simultaneously examine muscle size and quality of the medial gastrocnemius. *Muscle Nerve*. 49: 736–40.

62. Puthucheary Z a, Rawal J, McPhail M, Connolly B, Ratnayake G, Chan P, *et al.* (2013): Acute skeletal muscle wasting in critical illness. *JAMA*. 310: 1591–600.

63. Reeves ND, Maganaris CN, Narici M V (2004): Ultrasonographic assessment of human skeletal muscle size. *Eur J Appl Physiol*. 91: 116–8.

64. Franchi M V, Atherton PJ, Reeves ND, Flück M, Williams J, Mitchell WK, *et al.* (2014): Architectural, functional and molecular responses to concentric and eccentric loading in human skeletal muscle. *Acta Physiol (Oxf)*. 210: 642–54.

65. Abe T, Loenneke JP, Young KC, Thiebaud RS, Nahar VK, Hollaway KM, *et al.* (2014): Validity of Ultrasound Prediction Equations for Total and Regional Muscularity in Middle-aged and Older Men and Women. *Ultrasound Med Biol*. 1–8.

66. Midorikawa T, Ohta M, Hikihara Y, Torii S, Sakamoto S (2015): Prediction and validation of total and regional skeletal muscle volume using B-mode ultrasonography in Japanese prepubertal children. *Br J Nutr*. 114: 1209–1217.

67. Altman DG, Bland JM (1983): Measurement in Medicine : the Analysis of Method Comparison Studies. *Statistician*. 32: 307–317.

68. Thoirs K, English C (2009): Ultrasound measures of muscle thickness: intra-examiner reliability and influence of body position. *Clin Physiol Funct Imaging*. 29: 440–6.

69. Parry SM, El-Ansary D, Cartwright MS, Sarwal A, Berney S, Koopman R, *et al.* (2015): Ultrasonography in the intensive care setting can be used to detect changes in the quality and quantity of muscle and is related to muscle strength and function. *J Crit Care*. 30: 1151.e9–1151.e14.

70. Ismail C, Zabal J, Hernandez HJ, Woletz P, Manning H, Teixeira C, *et al.* (2015): Diagnostic ultrasound estimates of muscle mass and muscle quality discriminate between women with and without sarcopenia. *Front Physiol*. 6: 1–10.

71. Minetto MA, Caresio C, Menapace T, Hajdarevic A, Marchini A, Molinari F, Maffiuletti NA (2015): Ultrasound-Based Detection of Low Muscle Mass for Diagnosis of Sarcopenia in Older Adults. *PM R*. 1–10.

72. Kuyumcu ME, Halil M, Kara Ö, Çuni B, Çağlayan G, Güven S, *et al.* (2016): Ultrasonographic Evaluation of the Calf Muscle Mass and Architecture in Elderly Patients With and Without

Sarcopenia. *Arch Gerontol Geriatr*. 65: 218–224.

73. Janssen I, Heymsfield SB, Wang Z, Ross R (2000): Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr. *J Appl Physiol*. 89: 81–88.

74. Abe T, Thiebaud RS, Loenneke JP, Loftin M, Fukunaga T (2014): Prevalence of site-specific thigh sarcopenia in Japanese men and women. *Age (Dordr)*. 36: 417–26.

75. Abe T, Patterson KM, Stover CD, Geddam D a R, Tribby AC, Lajza DG, Young KC (2014): Site-specific thigh muscle loss as an independent phenomenon for age-related muscle loss in middle-aged and older men and women. *Age (Dordr)*. 36: 9634.

76. Abe T, Loenneke JP, Thiebaud RS, Fukunaga T (2014): Age-related site-specific muscle wasting of upper and lower extremities and trunk in Japanese men and women. *Age (Dordr)*. 36: 813–21.

77. Harris-Love MO, Monfaredi R, Ismail C, Blackman MR, Cleary K (2014): Quantitative ultrasound: measurement considerations for the assessment of muscular dystrophy and sarcopenia. *Front Aging Neurosci*. 6: 172.

78. Koskelo E-K, Kivisaari LM, Saarinen UM, Siimes M a. (1991): Quantitation of Muscles and Fat by Ultrasonography: A Useful Method in the Assessment of Malnutrition in Children. *Acta Paediatr*. 80: 682–687.

79. Ema R, Wakahara T, Mogi Y, Miyamoto N, Komatsu T, Kanehisa H, Kawakami Y (2013): In vivo measurement of human rectus femoris architecture by ultrasonography: Validity and applicability. *Clin Physiol Funct Imaging*. 33: 267–273.

80. Jenkins NDM, Miller JM, Buckner SL, Cochrane KC, Bergstrom HC, Hill EC, *et al.* (2015): Test–Retest Reliability of Single Transverse versus Panoramic Ultrasound Imaging for Muscle Size and Echo Intensity of the Biceps Brachii. *Ultrasound Med Biol*. 41: 1584–1591.

81. Schneebeli A, Egloff M, Giampietro A, Clijsen R, Barbero M (2014): Rehabilitative ultrasound imaging of the supraspinatus muscle: Intra- and interrater reliability of thickness and cross-sectional area. *J Bodyw Mov Ther*. 18: 266–272.

82. English C, Fisher L, Thoirs K (2012): Reliability of real-time ultrasound for measuring skeletal muscle size in human limbs in vivo: a systematic review. *Clin Rehabil*. 26: 934–944.

83. English CK, Thoirs KA, Fisher L, McLennan H, Bernhardt J (2012): Ultrasound is a reliable

measure of muscle thickness in acute stroke patients, for some, but not all anatomical sites: a study of the intra-rater reliability of muscle thickness measures in acute stroke patients. *Ultrasound Med Biol*. 38: 368–76.

84. Miyatani M, Kanehisa H, Ito M, Kawakami Y, Fukunaga T (2004): The accuracy of volume estimates using ultrasound muscle thickness measurements in different muscle groups. *Eur J Appl Physiol*. 91: 264–72.

85. Fukumoto Y, Ikezoe T, Yamada Y, Tsukagoshi R, Nakamura M, Mori N, *et al.* (2012): Skeletal muscle quality assessed from echo intensity is associated with muscle strength of middle-aged and elderly persons. *Eur J Appl Physiol*. 112: 1519–25.

86. So JL, Janssen I, Heymsfield SB, Ross R (2004): Relation between whole-body and regional measures of human skeletal muscle. *Am J Clin Nutr*. 80: 1215–1221.

87. Rankin G, Stokes M (1998): Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil*. 12: 187–199.

88. Kohavi R (1995): A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *Int Jt Conf Artif Intell*. 14: 1137–1143.

89. Bland JM, Altman DG (1999): Measuring agreement in method comparison studies. *Stat Methods Med Res*. 8: 135–160.

90. Ludbrook J (2010): Confidence in Altman-Bland plots: A critical review of the method of differences. *Clin Exp Pharmacol Physiol*. 37: 143–149.

91. Chhapola V, Kanwal SK, Brar R (2015): Reporting standards for Bland-Altman agreement analysis in laboratory research: a cross-sectional survey of current practice. *Ann Clin Biochem*. 52: 382–6.

92. Steyerberg EW (2008): *Clinical prediction models*. *Stat Biol Heal*. . doi: 10.1007/978-0-387-77244-8.

93. Earthman CP (2015): *Body Composition Tools for Assessment of Adult Malnutrition at the Bedside: A Tutorial on Research Considerations and Clinical Applications*. *J Parenter Enter Nutr*. (Vol. 39). doi: 10.1177/0148607115595227.

94. Abe T, Fujita E, Thiebaud RS, Loenneke JP, Akamine T (2016): Ultrasound-Derived Forearm Muscle Thickness Is a Powerful Predictor for Estimating DXA-Derived Appendicular Lean

Mass in Japanese Older Adults. *Ultrasound Med Biol*. 1–4.

95. Scott D, Chandrasekara SD, Laslett LL, Cicuttini F, Ebeling PR, Jones G (2016): Associations of Sarcopenic Obesity and Dynapenic Obesity with Bone Mineral Density and Incident Fractures Over 5–10 Years in Community-Dwelling Older Adults. *Calcif Tissue Int*. 99: 1–13.

96. Scott D, Sanders KM, Aitken D, Hayes A, Ebeling PR, Jones G (2014): Sarcopenic obesity and dynapenic obesity: 5-year associations with falls risk in middle-aged and older adults. *Obesity*. 22: 1568–1574.

97. Janssen I, Heymsfield SB, Baumgartner RN, Ross R (2000): Estimation of skeletal muscle mass by bioelectrical impedance analysis. *J Appl Physiol*. 89: 465–471.

98. Abe T, Kearns CF, Fukunaga T (2003): Sex differences in whole body skeletal muscle mass measured by magnetic resonance imaging and its distribution in young Japanese adults. *Br J Sport Med*. 37: 436–440.

99. Abe T, Kawakami Y, Suzuki Y, Gunji A, Fukunaga T (1997): Effects of 20 days bed rest on muscle morphology. *J Gravit Physiol*. 4: S10–4.

100. Abe T, Patterson KM, Stover CD, Young KC (2015): Influence of adipose tissue mass on DXA-derived lean soft tissue mass in middle-aged and older women. *Age (Omaha)*. 37. doi: 10.1007/s11357-014-9741-1.

101. Cawthon PM, Peters KW, Shardell MD, McLean RR, Dam TTL, Kenny AM, *et al.* (2014): Cutpoints for low appendicular lean mass that identify older adults with clinically significant weakness. *Journals Gerontol - Ser A Biol Sci Med Sci*. 69 A: 567–575.

102. Church TS, Thomas DM, Tudor-Locke C, Katzmarzyk PT, Earnest CP, Rodarte RQ, *et al.* (2011): Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. *PLoS One*. 6: 1–7.

103. Paris M, Mourtzakis M (2016): Assessment of skeletal muscle mass in critically ill patients: considerations for the utility of computed tomography imaging and ultrasonography. *Curr Opin Clin Nutr Metab Care*. In Press.

104. Moseley A, Adams R (1991): Measurement of passive ankle dorsiflexion: Procedure and reliability. *Aust J Physiother*. 37: 175–81.

105. James WPT, Chunming C, Inoue S (2002): Appropriate Asian body mass indices? *Obes Rev*. 3:

139.

106. Kim EY, Kim YS, Park I, Ahn HK, Cho EK, Jeong YM, Kim JH (2016): Evaluation of sarcopenia in small-cell lung cancer patients by routine chest CT. *Support Care Cancer*. 0–5.

107. Newman AB, Kupelian V, Visser M, Simonsick E, Goodpaster B, Nevitt M, *et al.* (2003): Sarcopenia: Alternative Definitions and Associations with Lower Extremity Function. *J Am Geriatr Soc*. 51: 1602–1609.

108. Delmonico MJ, Harris TB, Lee JS, Visser M, Nevitt M, Kritchevsky SB, *et al.* (2007): Alternative definitions of sarcopenia, lower extremity performance, and functional impairment with aging in older men and women. *J Am Geriatr Soc*. 55: 769–774.

109. Walker DK, Thaden JJ, Deutz NEP (2015): Application of gas chromatography-tandem mass spectrometry (GC/MS/MS) for the analysis of deuterium enrichment of water. *J Mass Spectrom*. 50: 838–843.

110. Utter AC, Hager ME (2008): Evaluation of ultrasound in assessing body composition of high school wrestlers. *Med Sci Sports Exerc*. 40: 943–949.

111. Smith-Ryan AE, Fultz SN, Melvin MN, Wingfield HL, Woessner MN (2014): Reproducibility and validity of A-mode ultrasound for body composition measurement and classification in overweight and obese men and women. *PLoS One*. 9. doi: 10.1371/journal.pone.0091750.

112. Roelofs EJ, Smith-Ryan AE, Melvin MN, Wingfield HL, Trexler ET, Walker N (2015): Muscle size, quality, and body composition: characteristics of division I cross-country runners. *J strength Cond Res / Natl Strength Cond Assoc*. 29: 290–296.

113. Johnson KE, Miller B, Gibson AL, McLain TA, Juvancic-Heltzel JA, Kappler RM, Otterstetter R (2016): A comparison of dual-energy X-ray absorptiometry, air displacement plethysmography and A-mode ultrasound to assess body composition in college-age adults. *Clin Physiol Funct Imaging*. . doi: 10.1111/cpf.12351.

Appendix

Appendix A1. Comparison of multiple body composition modalities

Modality	Fundamental Principle	Advantages	Disadvantages	References
DXA	Uses 2 different energy x-ray beams to differentiate soft tissue and bone. Adipose tissue and lean soft tissue are further differentiated based on attenuation factors of the tissues.	Differentiates body into 3 compartments (adipose, lean and bone tissue). Can obtain whole body, regional and segment specific measures. Low dose radiation x-rays allows for safe repeated measures High accuracy and repeatability	Cannot further differentiate adipose (subcutaneous, visceral, and intramuscular) and lean tissue (muscle, organs, connective tissue). Not safe for pregnant women due to radiation exposure	(5; 32)
CT	Uses attenuation coefficients of x-rays that pass through tissues. Attenuation is based on different density of tissues and cross-sectional images are produced.	Highly accurate and reliable quantitative measures of body composition (muscle and adipose tissue) Can differentiate between intramuscular, subcutaneous and visceral adipose tissue.	Large radiation exposure – generally limited to quantifying images take as part of routine care High cost and technical skills required to operate	(5; 18)
MRI	Strong magnetic field is generated and hydrogen atoms in tissues align with magnetic field. Aligned hydrogen atoms are activated via a radio frequency signal and the subsequent decay signal is used to generate cross-sectional images.	Safe for all ages – no radiation exposure Very high image resolution Accurate measures of muscle mass, intramuscular, visceral and subcutaneous adipose tissue	High cost and technical skills required to operate Limited availability	(5; 18)
BIA	Alternating electrical current is passed through proximal and distal electrodes,	Widely available, safe for all ages, no technical expertise required	Requires population specific equations Limited applicability in obese patients	(5)

	the voltage drop is measured. Fat-free mass has higher water and electrolyte content and allows easier transmission of current.		High variability in fat-free mass estimations	
Hydrostatic weight and air displacement plethysmography	Body volume is measured using relationships between pressure and volume. Body volume and weight are used to calculate body density.	Very precise and accurate for assessment of body fat. Easily accommodates most individuals.	Questionable validity for fat-free mass estimates Less widely available	(5)
Deuterium oxide dilution	Volume of compartment is measured by assessing dilution of isotopically labelled water. Total body water is used to estimate fat-free mass	Safe and easily applied during field research	Several assumptions for tracer metabolism Assumed hydration status of fat-free mass Expensive	(109)
B-mode ultrasound	High frequency sound waves are transmitted and partly reflected within tissues. Different tissues have different reflection coefficients. Reflected waves are detected and used to generate on screen images.	Portable, no radiation exposure, lower cost relative to other modalities Accurate and reliable measures of cross sectional area or thickness of subcutaneous tissues	Measures depend greatly on the operator (tissue depression, landmarking, tissue border identification) Hydration status and/or tissue swelling can confound measures No standardized protocol for assessment of tissues	(33)

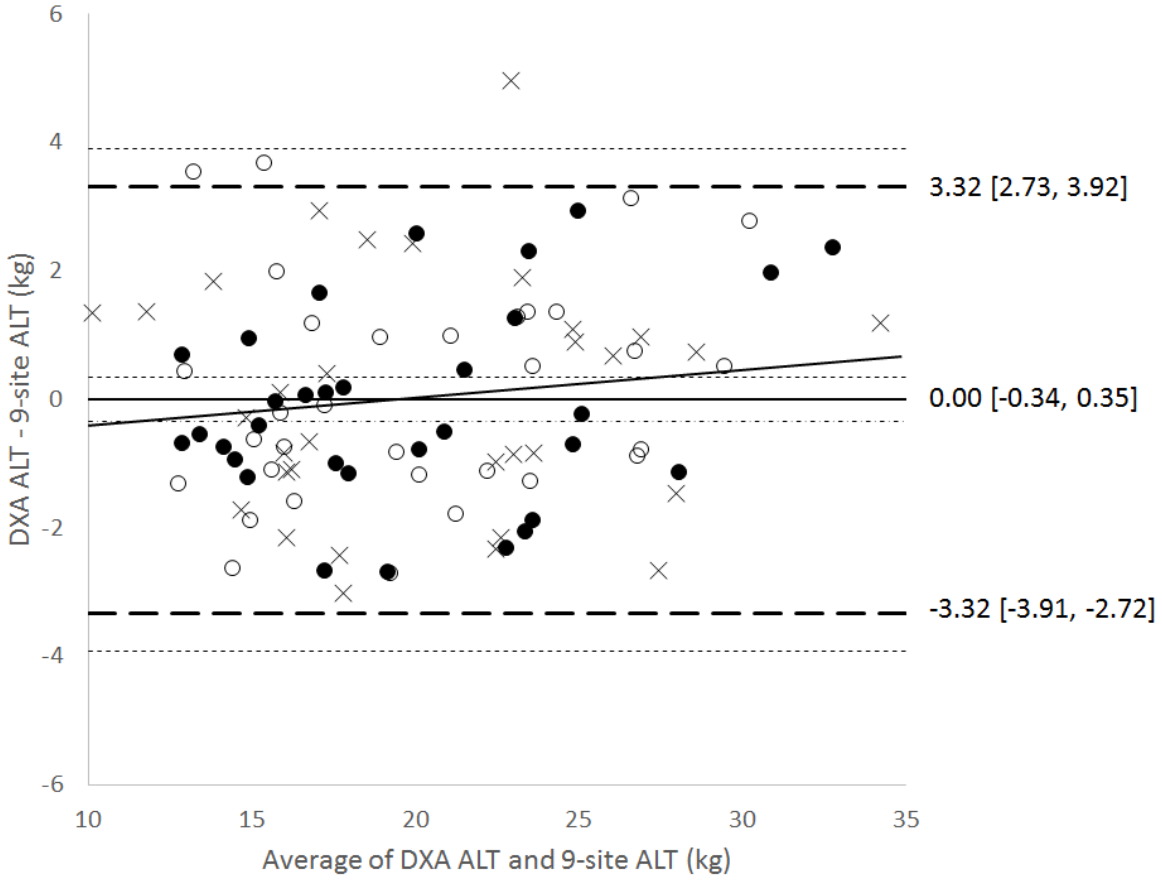
BIA, bioelectrical impedance analysis; CT, computerized tomography; DXA, dual-energy x-ray absorptiometry; MRI, magnetic resonance imaging, MRI.

Appendix A2. Linear regression analysis to predict appendicular lean tissue using the 9-site protocol

Model development	Appendicular lean tissue prediction (kg)	Cross-validation fold	Unadjusted r^2	Adjusted r^2	SEE (kg)	p-value model
Folds 1+2	$-4.320+(0.563X_5)$	3	0.92	0.92	1.53	<0.001
Folds 1+3	$-4.671+(0.569X_5)$	2	0.90	0.90	1.76	<0.001
Folds 2+3	$-5.155+(0.581X_5)$	1	0.89	0.89	1.89	<0.001
Average	$-4.715+(0.571X_5)$	-	0.90	0.90	1.73	-

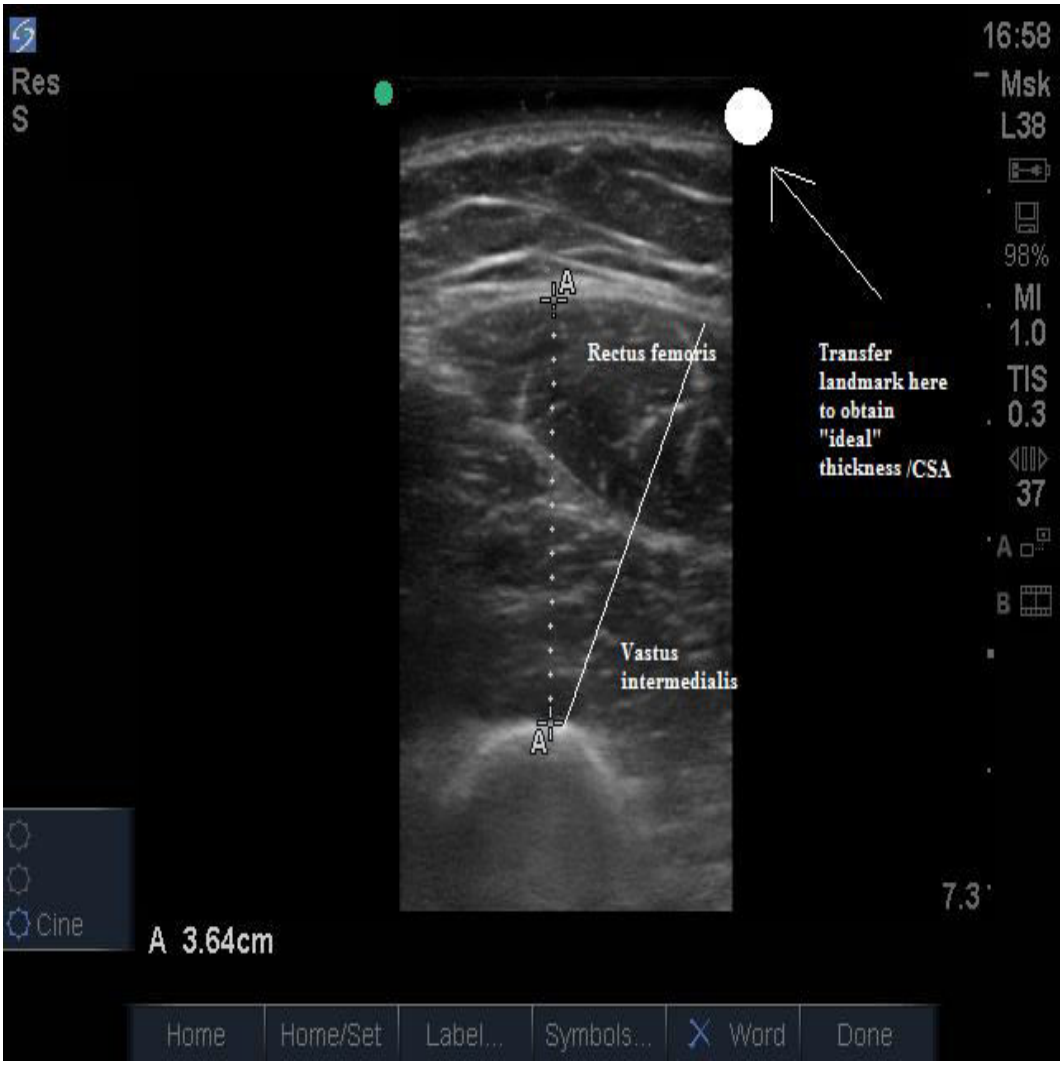
SEE, standard error of the estimate. X_5 = 9 site muscle thickness x height (cm x m).

Appendix A3.



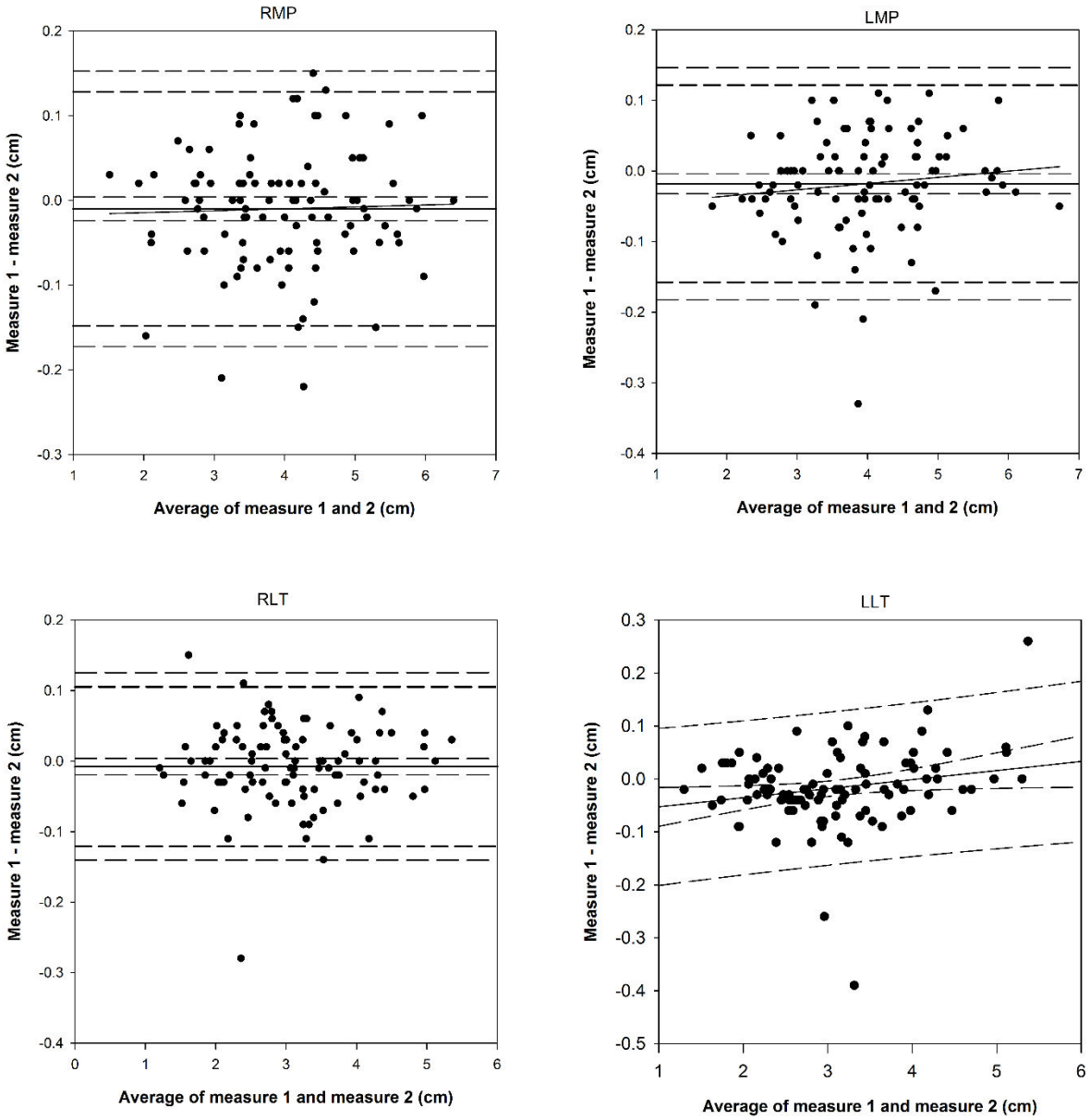
Appendix A3. Bland-Altman plot comparing DXA derived and the 9-site protocol predicted appendicular lean tissue mass, utilizing all participants from all folds. No fixed (0.00 [-0.34, 0.35]) or proportional bias was present (solid black line, 95% CI – inner short dashed line), with limits of agreement (1.96 SD) of -3.32 and 3.32 (middle long dashed lines) and tolerance limits of -3.91 and 3.92 (outer short dashed lines). Crosses (×) represent fold 1, closed circles (●) represent fold 2, and open circles (○) represent fold 3.

Appendix A4.



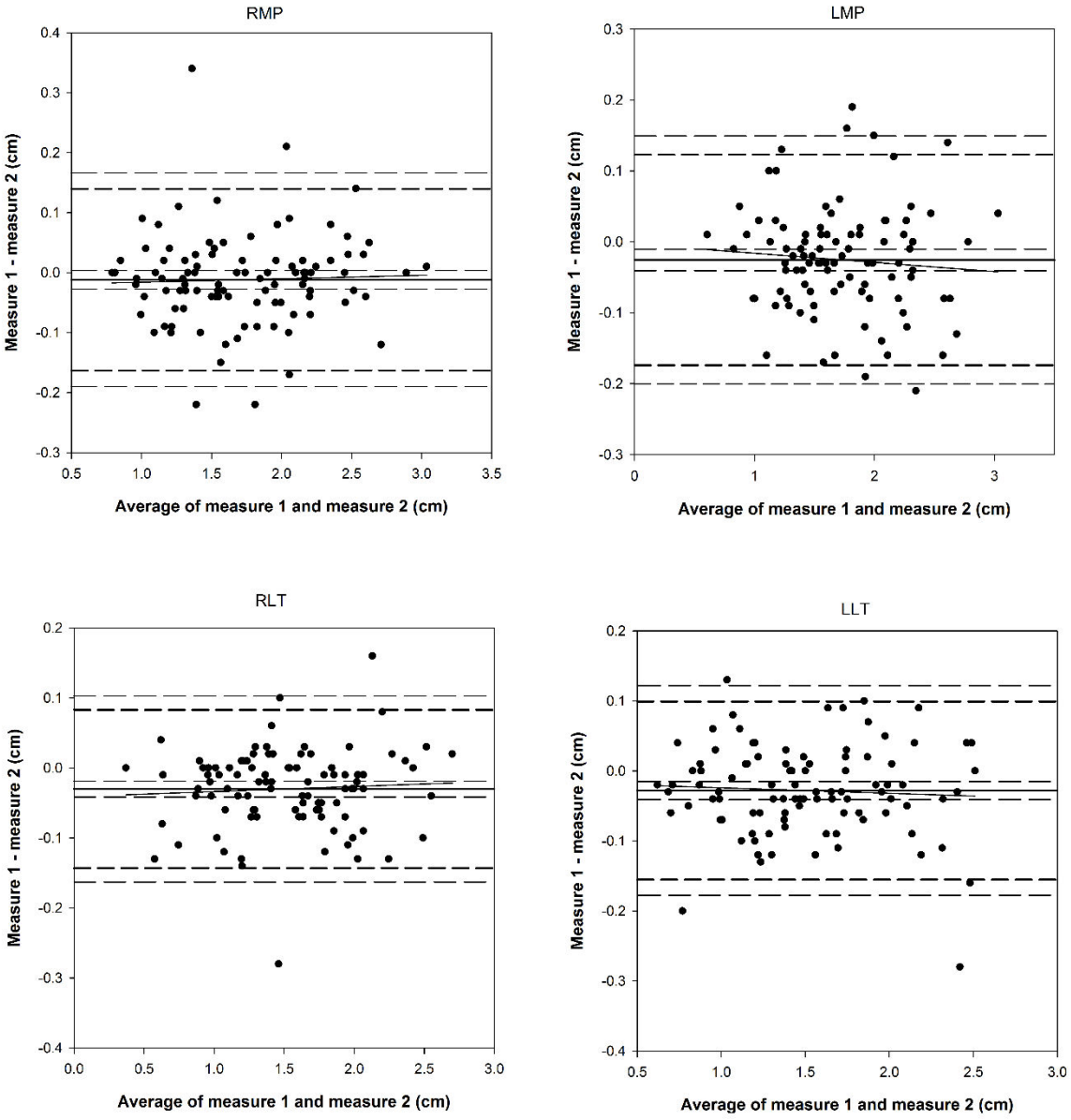
Appendix A4. Demonstration of a non-ideal muscle thickness measured using the 4-site protocol.

Appendix A5.



Appendix A5. Bland-Altman plots for intra-rater reliability using the 4-site protocol for minimal compression. Minimal compression average bias [95% CI] for all intra-rater plots (except plot with proportional bias) was -0.04 [-0.03, 0.01] with average limits of agreement and tolerance limits were -0.14 and 0.12 and -0.16 and 0.14.

Appendix A6.



Appendix A6. Bland-Altman plots for intra-rater reliability using the 4-site protocol for maximal compression. Maximal compression presented a significant fixed bias of -0.02 [-0.04, -0.01] with average limits of agreement and tolerance limits were -0.14 and 0.12 and -0.16 and 0.14 and for maximal compression, -0.16 and 0.11 and -0.18 and 0.14.