Factors Influencing Measures of Trochanteric Soft Tissue Thickness

by

Benoit Lafleur

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

Waterloo, Ontario, Canada, 2016

AUTHOR'S DECLARATION

Abstract

Introduction: Trochanteric soft tissue thickness (TSTT), which attenuates the force applied to the hip upon impact, has emerged as a potential factor that may improve the assessment of the risk of a hip fracture. However, a gold standard technique has not been established for the measurement of TSTT. Different measurement techniques using ultrasound (US) and dual energy x-ray absorptiometry (DXA) have been used, but the accuracy and reliability of these techniques have not been extensively assessed.

Objectives and Hypotheses: The first objective was to determine the concordance validity between US and DXA measurements of TSTT. The second objective was to determine if there were significant differences in TSTT between standing, supine, and side-lying measurements, as well as between internal hip rotation of 25°, 0° rotation, external hip rotation of 25°, and possible interactions. The third objective was to determine the intra and inter-rater reliability across DXA and US measurements made in a standing, supine, and side-lying position. The corresponding hypotheses for the first and third objectives are that the intraclass correlation will be strong ($ICC > 0.8$), and that the coefficient of repeatability (i.e. the 95% confidence interval of the differences) will be below a clinical threshold of 0.96 cm ($CR < 0.96$ cm). The corresponding hypotheses for the second objective were that there will be a significant difference between the TSTT collected across: a) postures including standing, supine, and side-lying positions; b) hip rotation angles of 25° internal, 0°, and 25° external, and c) that there would be no significant interaction between body position and hip rotation ($p > 0.05$).

Methods: Forty-five community-dwelling older adults (20 males, 25 females) were recruited for this study. Mean (SD) age was 70.2 (10.8) years, and BMI was 27.5 (4.3) kg/m^2. TSTT were obtained using a QDR Discovery™ DXA Scanner (Hologic Inc., MD, USA) and by using a curvilinear ultrasound transducer (C60x, 2-5MHz) in combination with an M-Turbo Ultrasound Unit 1.0.6™ (Sonosite Inc., WA, USA). Ultrasound measurements were made with each participant in a standing, supine, and side-lying position. Within each position, the hip was rotated internally at 25°, at 0°, or externally at 25° by using standardized foam triangles. Repeat measurements, as well as measurements by another investigator, were taken for reliability analyses. The investigators were blinded to the TSTT value during the collection, and the protocol of landmarking the greater trochanter was repeated for every measurement. The concordance validity between US and DXA, and the different reliabilities were tested by calculating two-way random ICCs and CRs. Also, Bland-Altman plots were used to visualize the results. A two-factor repeated measures ANOVA ($\alpha=0.05$) was used to determine the main effects of body position and hip rotation on TSTT and a potential interaction.

Results: The ICC (2,1) of TSTT measurements between US and DXA was 0.898, but the CR was 2.15 cm and the mean bias of the differences was 0.46 cm. There was a main effect of body position ($p < 0.001$) and hip rotation ($p < 0.001$) on TSTT. Specifically, mean (SD) standing TSTT was 4.33 (2.1) cm, the supine TSTT was 5.57 (2.8) cm, the side-lying TSTT was 3.29 (1.7) cm. Mean (SD) TSTT for the 25° internal hip rotation was 4.17 (2.4) cm, no rotation was 4.33 (2.4) cm, and the 25° external rotation was 4.69 (2.5) cm. There was also a significant ordinal interaction between body position and hip rotation ($p = 0.018$). The US intra-rater ICC (2,1) was 0.980, 0.972, and 0.977 for the standing, supine, and side-lying measurements, and their respective CR was 0.87 cm, 1.32 cm, and 0.69 cm. The US inter-rater ICC (2,1) was 0.970, 0.939, and 0.977 for the standing, supine, and side-lying measurements, and their respective CR was 1.17 cm, 2.08 cm, and 1.10 cm. The DXA intra-rater and inter-rater reliability analyses produced ICCs (2,1) of 0.995 and 0.995, and CRs of 0.45 cm and 0.38 cm. For all of the reliability analyses, the mean bias of the differences was under 0.2 cm.

Discussion/Conclusion: When compared to US, DXA underestimates TSTT, which goes against common belief in the literature. As main effects of body position and hip rotation were found, future researchers should strongly consider tightly controlling / standardizing these factors when measuring TSTT. From a reliability perspective, the side-lying US and DXA analyses were superior to US measurement in standing or supine postures. This thesis provides important information about TSTT characteristics across measurement modalities and body postures. Future research should be conducted to determine what approach for measuring TSTT is most effective as part of models that predict hip fracture risk.

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Andrew Laing. I could not have asked for a better mentor. His guidance throughout my entire degree was best to none, and he supported all of my ideas when developing this thesis study. His hard-working yet enjoy-life attitude is very contagious, and I know it will serve me well in the future. I would also like to thank my committee members Dr. Marina Mourtzakis and Dr. Lora Giangregorio for the constructive feedback and being great collaborators in this exciting new field of research.

A huge acknowledgement goes out to Alyssa Tondat and Mike Paris, as they helped me with the ultrasound training, participant recruitment, data collection, and data analysis of this study. It is no exaggeration when I say that the study would not be completed without them, so they deserve a large amount of credit for all of the work presented in this thesis. I would also like to acknowledge Janice and Stephanie for conducting all of the DXA scans in this study.

I am extremely proud to say that I am an IBAL graduate. I would like to thank Tyler, Taylor, Fred, Iris, Dan, Steve and Emily for hearing me out when I needed it, and for creating an environment that was an absolute pleasure to work in. We had great scientific discussions that have fueled our research, but I will never forget all the great times sports–ing.

I would like to thank all of the great friends that I made in the UW Department of Kinesiology for the fun times during orientation week, conferences, intramural sports, board game breaks, Aftab Patla Memorial Game, Oktoberfest, and various nights out. A specific shout out goes to Ben, Jordan, Jeff, Mike Glinka, Erin, Maureen, and Kristen for including me in their circle of friends, and for making me feel at home in Waterloo.

Next, I would like to thank my family. Mom, Dad, Steph, and Al, thank you for making sure that I was doing well while I was away, and for being there when I needed it. Lastly, I would like to thank Gennie Eaton for supporting my decision to chase my dreams. You inspire me every day to be the best that I can be, and I am excited to be back home with you.

Table of Contents

List of Figures

List of Tables

List of Acronyms/Abbreviations

Alyssa Tondat -	AT
Benoit Lafleur -	BL
Body Mass Index -	BMI
Bone Mineral Content -	BMC
Bone Mineral Density -	BMD
Coefficient of Repeatability -	CR
Coefficient of Variation	CV
Computerized Tomography -	CT
Dual Energy X-Ray Absoptiometry -	DXA
Fracture (cases) -	Fx
Intraclass correlation -	ICC
Michael Paris -	MP
Trochanteric Soft Tissue Thickness -	TSTT
Ultrasound -	US
World Health Organisation -	WHO

Chapter 1

Thesis Overview

Approximately 30 000 Canadians suffer a hip fracture per year, and these rates are expected to rise with our aging population (Leslie et al., 2010). The total annual cost to care for hip fractures was estimated to be 622 million dollars (Tarride et al., 2012), and this cost is projected to rise to 2.4 billion dollars in the year 2040 (Wiktorowicz et al., 2001). From an individual perspective, hip fractures are a serious injury as they are highly linked to functional impairment, disability, loss of independence, reduction of quality of life, and death (Korhonen et al., 2013). This is unfortunate as hip fractures are an injury, and injuries are preventable with the proper tools and education.

Per the 2010 Canadian guidelines, clinicians are recommended to use the FRAX tool to assess the risk of hip fracture (Papaioannou et al., 2010). FRAX has been validated in several countries and has revolutionized the way clinicians assess the risk of hip fracture in patients. However, one of its limitations is that it only uses BMI to estimate the force applied to the hip from a fall, which is equally important in predicting the risk of hip fracture. One variable that can improve the estimate of applied force to the hip in models that assess hip fracture risk is trochanteric soft tissue thickness (TSTT).

Several different approaches have been used to measure TSTT in the literature, so it is difficult to make a recommendation as to how to measure it. The measurement devices used were US, DXA, CT, and a BMI regression equation, and the participants in these studies were in different positions: standing, supine, and lying on their side. Although it is assumed that there would be differences between the three different positions, no study has quantified these TSTT differences. The effects of hip rotation have also never been assessed, even though it has been shown to affect BMD measurements at the hip (Goh et al., 1995; Lekamwasam and Lenora, 2003). Moreover, the agreement between the two most popular TSTT measurement devices – US and a whole body DXA scan – has never been evaluated. Accordingly, the first objective of this study is to determine the concordance validity of TSTT between the supine US and DXA measurements. The associated hypotheses are:

i. The intraclass correlation between the supine US and DXA TSTT measurement will be strong (ICC > 0.8)

ii. The CR between the supine US and DXA TSTT measurements will be below 0.96 cm

The second purpose is to determine if there are any significant differences in TSTT between standing, supine, and side-lying measurements, as well as between an internal hip rotation of 25°, 0° rotation, and an external hip rotation of 25°. Additionally, a possible interaction between position and hip rotation will be tested. The associated hypotheses are:

iii. There will be no significant interaction between body position and hip rotation ($p > 0.05$).

iv. There will be a significant difference between the TSTT collected in a standing, supine, and side-lying position ($TSTT_{Stand} \neq TSTT_{Side} \neq TSTT_{Sup}$, $p<0.05$)

v. There will be significant differences between the TSTT collected in a 25° internal hip rotation, a 0° rotation, and a 25° external hip rotation ($TSTT_{25int} \neq TSTT_{0rot} \neq TSTT_{25ext}$, $p<0.05$).

The reliability of these different approaches has not been completely assessed. Determining the reliability of each approach is important so future researchers or clinicians can be confident in their measurement of TSTT. Hence, the third objective is to determine the intra and inter-rater reliability across DXA and US measurements made in a standing, supine, and side-lying position. Also, supine US image inter-day reliability will be determined in this study. The associated hypotheses are:

vi. The intraclass correlation within raters, between raters, and within images will be strong for US and DXA analyses ($ICC > 0.8$)

vii. The CR will be below 0.96 cm within rater, between raters, and within images for US and DXA analyses.

In Canada, there are more osteoporotic fractures in the moderate risk group than the high-risk group because there are more individuals classified in the moderate-risk group (Papaioannou et al., 2010). TSTT, combined with BMD and other clinical risk factors, may help differentiate the moderate risk group into a high-risk or low-risk group. Although this theory cannot be tested in this thesis, the first step is to provide guidance towards improving current approaches for measuring TSTT.

Chapter 2

General Introduction and Literature Review

2.1 Hip Fractures: A National and International Problem

2.1.1 Prevalence of Hip Fractures

Fall-related hip fractures in the aging population is a major burden for all economic and health care systems (Cooper et al., 1992; Johnell and Kanis, 2004). Projections into 2050 estimate that there will be 6.26 million hip fractures worldwide, almost a fourfold increase when compared with 1.66 million hip fractures that occurred in 1990 (Cooper et al., 1992).

Approximately 30 000 Canadians that suffer a hip fracture per year (Leslie et al., 2009). The total number of annual hip fractures is expected to rise in the future due to the aging population, both nationally and internationally. Osteoporotic hip fracture rates are similar to the annual incidence of heart attack, stroke, and breast cancer combined (Osteoporosis Canada, 2015)

2.1.2 Economic Cost of Hip Fractures

From a monetary perspective, hip fractures are very costly for patients and for healthcare systems around the world. In the USA and Europe, each hip fracture is approximately an acute cost of $20 000 (US) (Cotter et al., 2005; Roudsari et al., 2005). For Canada, each hip fracture averaged a direct healthcare system cost of $27 000 (CD), ranging from $21 000 to about $47 000, depending if the patient was discharged home or to a long-term care facility (Wiktorowicz et al., 2001). In 2007/2008, the total annual cost to care for hip fractures was estimated to be $650 million in Canada, and $25 billion in the USA (Braithwaite et al., 2003; Tarride et al., 2012). The 2040 and 2041 cost is expected to be $47 billion in the USA, and $2.4 billion in Canada, respectively (Braithwaite et al., 2003; Wiktorowicz et al., 2001). The cost to treat hip fractures seems to be unsustainable with the rising rates, so it would be beneficial for all parties to focus our attention on the prevention of hip fractures.

2.1.3 Severity of Hip Fractures

There are several repercussions associated with fall-related hip fractures that really increase the severity of the injury. For example, death, functional impairment, disability, loss of independence, and quality of life reduction are all interrelated with hip fractures in older adults (Korhonen et al., 2013). An international study estimated 740 000 deaths, 4.48 million disabilities, and 1.75 million disability adjusted life-years lost associated with hip fractures in 1990 (Johnell and Kanis, 2004).

There are several complications that come with hip fractures. Osteoporotic hip fractures require more hospital bed-days than stroke, diabetes, or heart attack (Osteoporosis Canada, 2015). Papaioannou et al. (2009) measured the health-related quality of life of hip fracture patients and found an overall decrease in both men and women. They also found that mobility, ambulation, and self-care attributes were affected by hip fractures. After a year following a hip fracture, 15.5% of community residents were transferred to a long-term facility, and 21.6% had passed away (Wiktorowicz et al., 2001). Overall, there is an increased risk of death (hazard ratio of 3.2) during the first year after a hip fracture (Ioannidis et al., 2009). Of those who returned to the community, a large number of these patients was still dependent on home care or informal from their family or friends (Wiktorowicz et al., 2001). Additionally, nearly 14% of hip fracture patients were hospitalized within a year due to a complication related to the injury. Lastly, approximately 11% and 6% of hip fracture patients suffered another fracture or a second hip fracture, respectively. All in all, hip fractures are a serious injury that has long lasting effects with regards to quality of life and life expectancy. They strongly affect patients from a physical and emotional standpoint. Avoiding this injury can be a huge relief for patients, family, and friends.

2.1.4 Section Summary

Preventing hip fractures is of high priority as they are highly prevalent, very costly to the economy, and detrimental to the lives that are affected by this injury. The next sections of this literature review present guidelines for hip fracture risk assessment, and a variable (trochanteric soft tissue thickness) that influences hip fracture risk and will be examined in depth in this thesis.

2.2 Current Assessment of Hip Fracture Risk

2.2.1 Assessment of Osteoporosis

In 2010, the Canadian clinical practice guidelines for the diagnosis and management of osteoporosis were updated. Compared to the 2002 guidelines, the new focus is preventing fragility fractures, rather than treating low BMD (Papaioannou et al., 2010). The guidelines were developed by surveying specialists from various disciplines, and by conducting systematic reviews of literature that evaluate assessments of risk of fracture and therapies for osteoporosis. After that, an expert panel reviewed the recommendations in the guidelines. Overall, these guidelines aim to improve the assessment and management of women and men that are at a high risk of fracture. Although these guidelines are for the diagnosis and management of osteoporosis, clinicians would follow the same steps to assess the risk of hip fracture.

The first step is to assess for osteoporotic fracture risk and to determine who should undergo BMD testing. According to the 2010 Canadian guidelines, all men and women over the age of 65 should undergo DXA BMD testing and be assessed for fracture risk. For women and men aged 50-64, a specific clinical assessment is recommended. This assessment consists of identifying clinical risk factors associated with low BMD, future fractures, and falls, and conducting a physical examination to screen for vertebral fractures. Additionally, the Get-Up and Go Test is recommended to assess the individual's fall risk (the person is timed getting out of a chair without using their arms, walk three meters and return to the chair). Anyone who has a one or more risk factors is recommended for DXA BMD testing. For adults under the age of 50, DXA BMD testing is recommended for those with a previous fragility fracture, prolonged use of glucocorticoids, the use of other high-risk medications (e.g. aromatase inhibitors or androgen deprivation therapy), hypogonadism or premature menopause (before age 45), malabsorption syndrome, primary hyperparathyroidism, or other disorders strongly associated with rapid bone. Basic bone health (regular weight-bearing exercise, high calcium and vitamin D intake, fall-prevention strategies) are still recommended for those who are not recommended to undergo BMD testing.

2.2.2 FRAX and CAROC

After determining a person's femoral neck BMD, the second step is to calculate their 10-year major osteoporotic fracture risk (clinical spine, forearm, hip, or shoulder) by using the CAROC or FRAX tool. The CAROC tool uses age, femoral neck BMD, and sex to stratify individuals in a low-risk zone (10-year major osteoporotic fracture risk <10%), a moderate risk zone (10% - 20%), or a high-risk zone (>20%) (Figure 2-1). Moreover, if the individual prolonged use of systemic glucocorticoids, then they are automatically increased to the next risk zone (i.e from low to moderate or moderate to high). If the individual has a fragility fracture, they are automatically categorized in the high risk zone. It is important to note that the CAROC tool only calculates 10-year major osteoporotic fracture risk, and not hip fracture risk.

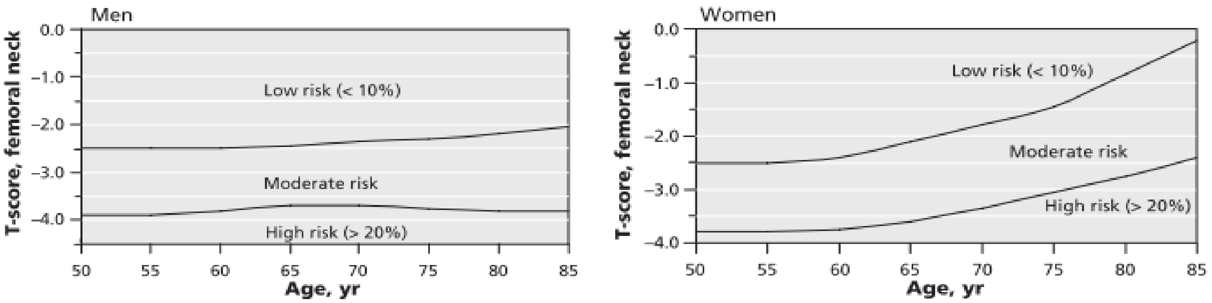
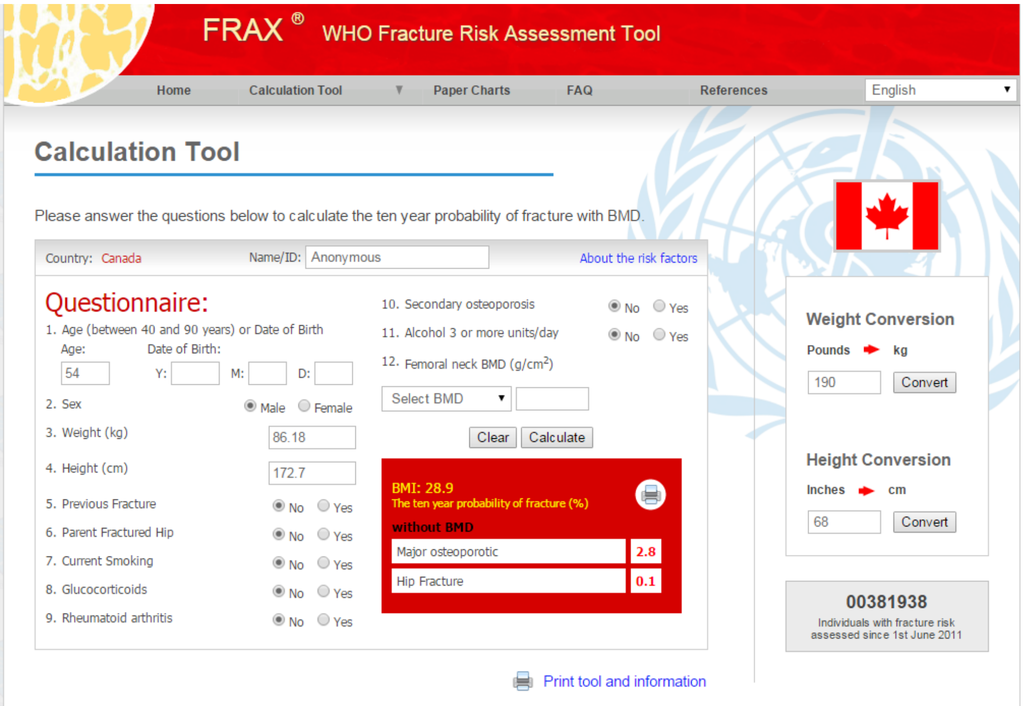


Figure 2-1. The CAROC tool used to assess osteoporotic fracture risk. The left tool is for men, the right tool is for women

The FRAX tool developed by the World Health Organization calculates both a 10-year major osteoporotic fracture risk and a 10-year hip fracture risk. To calculate these fracture risks, the following inputs are needed: age (years), sex (male or female), weight (kg), height (cm), previous fracture (yes or no), parental hip fracture history (yes or no), current smoking (yes or no), glucocorticoids (yes or no), rheumatoid arthritis (yes or no), secondary osteoporosis (yes or no), alcohol intake (>3 units/day, yes or no), and femoral neck BMD (Figure 2-2). These risk factors were identified by nine prospective studies from various countries and shown to be relevant risk factors via meta-analyses (Kanis et al., 2008). At present, there are 50 different FRAX models that have been calibrated for 45 different countries (McCloskey et al., 2012). Each model represents a different race/ethnicity/nationality. An increase of age increases the risk of hip fracture, and women experience higher hip fracture risks than men. The presence of other clinical risk factors will increase the risk of hip fracture, and the magnitude will depend on the age and sex. Weight and height are used to calculate BMI, and an increase in BMI is associated with an exponential decrease of lower hip fracture risk. Lastly, a decrease of BMD will increase the risk of hip fracture. What makes FRAX unique is that BMD is an optional input, and it has recently been shown that the use of clinical risk factors alone is comparable to the use of BMD alone (Kanis et al., 2012). Prior fracture, glucocorticoid use, family history, and BMI added additional information to the 10-year hip fracture risk (Cosman et al., 2014; Kanis et al., 2008).



Calculation Tool

Please answer the questions below to calculate the ten year probability of fracture with BMD.

Country: Canada   Name/ID: Anonymous   About the risk factors

Questionnaire:

1. Age (between 40 and 90 years) or Date of Birth
Age: 54   Date of Birth: Y: M: D:
2. Sex ● Male ○ Female
3. Weight (kg) 86.18
4. Height (cm) 172.7
5. Previous Fracture ● No ○ Yes
6. Parent Fractured Hip ● No ○ Yes
7. Current Smoking ● No ○ Yes
8. Glucocorticoids ● No ○ Yes
9. Rheumatoid arthritis ● No ○ Yes
10. Secondary osteoporosis ● No ○ Yes
11. Alcohol 3 or more units/day ● No ○ Yes
12. Femoral neck BMD (g/cm^2)
Select BMD

Clear   Calculate

BMI: 28.9
The ten year probability of fracture (%)
without BMD

Major osteoporotic	2.8
Hip Fracture	0.1

Figure 2-2. A screenshot of the Canadian FRAX tool. From https://www.shef.ac.uk/FRAX/tool.aspx?country=19

In Canada, the intervention threshold will depend on the 10 year major osteoporotic fracture risk of FRAX, and not the 10-year hip fracture risk (Papaioannou et al., 2010). When the 10-year fracture risk is low (<10%), the clinician is recommended to re-assess fracture risk in 5 years. Patients in the moderate risk zone (10% - 20%) should be considered for pharmacologic therapy if an additional risk factor is present. These risk factors include a previous wrist fracture in adults over the age of 65 or with a T-score ≤ -2.5, a lumbar spine T-score much lower than femoral neck T-score, rapid bone loss, androgen deprivation therapy, aromatase inhibitor therapy, long-term or repeated systemic glucocorticoid use, more than 2 falls in the past 12 months, or other disorders strongly associated with osteoporosis, rapid bone loss or fractures. When the 10-year fracture risk is high (>20%), pharmacotherapy is highly recommended, but the patient preference needs to be considered first. In addition, the National Osteoporosis Foundation recommends that anyone with a 10-year risk of hip fracture above 3% should be prescribed an intervention (Cosman et al., 2014).

2.2.3 Section Summary

A direct quote from the 2010 Canadian clinical practice guidelines for the diagnosis and management of osteoporosis: “More osteoporotic fractures occur in the moderate-risk group than the high-risk group (because there are more individuals in the moderate-risk group), even though the individual risk of fracture is higher in the high-risk group. Therefore, patients who are at moderate risk should undergo a careful clinical evaluation to identify additional risk factors that are not considered in the risk assessment system, and certain of these individuals should be offered pharmacologic therapy". Accordingly, the risk of hip fracture can be better managed by aiding clinicians to identify who to treat in the moderate risk category. With my understanding of the etiology of hip fractures, I believe that the additional risk factors that can aid clinicians are the ones associated with the fall-induced impact force. The current clinical risk factors have a strong link with bone strength and the risk of falling, but they have a weak link with the fall-induced impact force. This is well explained in a review article called 'A biomechanical sorting of clinical risk factors affecting osteoporotic hip fracture' by Luo (2015). For example, BMI is the only clinical risk factor in FRAX that is associated with the applied load. BMI is a moderate surrogate for estimating the applied force at the hip (a higher BMI means a larger impact force, but it also means more attenuation because the individual is likely to have more soft tissue on their hip). A better estimation of the impact force would be to use height and weight individually, and a better estimation of the attenuation would be to use trochanteric soft tissue thickness (TSTT). The next sections are a review of the literature surrounding TSTT. TSTT may be an additional risk factor that helps clinicians identify those in the moderate risk category who might benefit from an intervention to decrease hip fracture risk.

2.3 TSTT: A Promising Variable for Hip Fracture Assessment

2.3.1 Introduction and Summary of TSTT Studies

Trochanteric soft tissue thickness (TSTT) is the total amount of soft tissue overlying the greater trochanter, and it can be used to assist in estimating the impact force on a hip from a fall. From exterior to interior, the soft tissues that overlie the greater trochanter are skin, fat, fascia lata, muscles (gluteus medius, gluteus minimus, vastus lateralis, possibly piriformis), and the trochanteric bursa (Robinovitch et al., 1995). It was first measured by Robinovitch et al. (1991), who demonstrated that increased TSTT decreases the effective stiffness of the body, thus decreasing the peak force applied to the hip during an impact. Since then, it has been measured or estimated in 13 different studies whose purposes can be categorized into 3 groups: i) understanding the mechanical role of TSTT during impact, ii) using TSTT to assist in predicting the risk of hip fracture, and iii) comparing different measurement techniques. A general description of these studies and their findings is presented in Table 2-1.

Table 2-1. General description of studies that measured or used TSTT

Authors/Year	Category	Sample Size	Age (years)	TSTT (mm)	Method	Major Contribution (for TSTT)
Robinovitch et al. (1991)	‘Mechanical role’	7 males and 7 females	26.9 ± 5.5, 20 to 35	26.1 ± 12.7, 9 to 50	US (standing)	Pioneer study that first demonstrated that TSTT influences pelvic stiffness
Robinovitch et al. (1995)		3 male and 6 female cadavers	72 ± 4, 60 to 102	24 ± 13, 8 to 45	Direct measurement	First to use surrogate pelvis test system, made attenuation equation of 71N/mm
Etheridge et al. (2005)		10 female pelvises	75.9 ± 8.6, 53 to 82	41.3 ± 18.8, 13.4 to 79	CT	Showed that TSTT energy dissipation changes at different velocities
Majumder et al. (2008)		Male FE model of pelvis-femur-TSTT	N/A	5, 14, 17, 23, and 26	N/A	Made a FE model that can be used to simulate different fall scenarios
Majumder et al. (2013)		7 FE males	N/A	5, 14, 17, 23, and 26	N/A	Demonstrated that TSTT was a more dominant parameter than weight and height (when it comes to normalized peak force, time to peak force, and strain ratio)
Choi et al. (2015)		17 young females 17 older females	21.2 ± 2.7, N/A 69.9 ± 4.7, N/A	32.1 ± 7.2, N/A 30.4 ± 14.9,N/A	US (side-lying)	Demonstrated that TST stiffness and damping is decreased in older women versus young women
Bouxsein et al. (2007)	‘TSTT and Fx risk’	42 controls 21 Fx cases (all females)	73.9 ± 8, N/A 73.9 ± 8.3, N/A	49.8, ± 16.8, N/A 40.4, ± 16.7, N/A	Whole-body DXA	Showed significant difference in TSTT between fracture cases and controls, and was trending to being a predictor of hip fracture (independent of BMD)

Nielson et al. (2009)		222 controls 70 Fx cases (all males)	74.2 ± 6.1 79.7 ± 6.	31 ± 11.5 29.1± 11.9 (13.3 to 78)	Whole-body DXA and subset of QCT	TSTT was not significantly different between the male groups, but the Factor of Risk was demonstrated to be
Roberts et al. (2010)		48 female and 25 male cadavers	74.38 ± 8.91, 55 to 98	41.86 ± 30.84, N/A	BMI regression equations	Showed that the Factor of Risk has better predictive capabilities than using solely BMD T-scores
Dufour et al. (2012)		425 males (26 Fx) 675 women (110 Fx)	76 ± 5.1, 67 to 95	30.1 ± 9.3 29.5 ± 9.9 55.3 ± 16.8 49.5 ± 16.8	BMI regression equations	Factor of Risk was significantly associated to hip fracture risk in a population-based cohort study (in men and women) // Showed that fall force and TSTT was predictive of hip fracture in women, independent of BMD
Maitland et al. (1993)	'Measurement technique'	50 females	72 ± 4, (all over 65)	N/A, N/A, ~15 to 85	US (standing) and "DXA"	Found significant correlations between US TSTT, DXA, BMI, BIA and hip circumference
Minns et al. (2007)		12 controls 20 Fx cases (all females)	82, N/A, 69-88 79, N/A, 76-93	18.1, N/A 27.9, N/A	US (standing)	Provided insights as to how hip protectors should be made// GT is 12cm postero-lateral from ASIS
Schacter and Leslie (2014)		2 cohorts, 188 each 83% female	56.6 ± 20.8 54.8 ± 20.1 (all over 20)	49 ± 23 48 ± 21 ~3 to 140 (all)	Whole-body DXA	Made regression model of TSTT that uses regional DXA scan info of the spine and hip
Levine et al. (2015)		10 females 10 males ALL	22.3 ± 1.1 22.2 ± 1.9	33.3 ± 6.6 22.8 ± 9.7 28.1 ± 9.7	US (standing)	Demonstrated that changes of posture (flex, ext, flex+add) influence TSTT

2.3.2 Mechanical Role of TSTT during an Impact

As mentioned, Robinovitch et al. (1991) first found that TSTT reduced the effective stiffness of the body when conducting pelvis release experiments (Figure 2-3). At the moment of impact, there is minimal contact between the ground and the hip, explaining the initial low stiffness. As the contact area increases, more soft tissue compresses. The force increases non-linearly (or in other words, the stiffness rises) due to the viscous properties of the soft tissue. The physical properties of a viscoelastic material (i.e. soft tissue) will depend on the rate and duration of the applied stress (Nigg and Herzog, 2007). Once TSTT "bottoms out", the pelvis begins to deform, which largely contributes to the effective stiffness of the body (Figure 2-3). At this point, the stiffness remains constant, and the force seems to increase in a linear fashion. This can be attributed to the elastic nature of bones (it is technically viscoelastic, but it has much more elastic properties than viscous). These authors found that 82% of their final predicted stiffness occurred after 230N, meaning that TSTT was (probably) fully compressed at that value, and the pelvic system began to significantly contribute to the effective stiffness (Robinovitch et al., 1997). Laing and Robinovitch (2010) also found that the non-linearity of pelvic stiffness occurred below 300N when performing pelvis release experiments. However, these experiments used young lean participants, so it was not known if these notions apply to bigger people. Levine et al. (2013) used pelvis release experiments in a low BMI (<22.5 kg/m^2) and high BMI group (>28 kg/m^2), and found no significant differences in pelvis stiffness. This contrasts the results of Robinovitch et al. (1991) as greater TSTT decreased effective stiffness, and BMI is highly correlated to TSTT (Maitland et al., 1993). The authors hypothesize that using linear estimates of stiffness may be an inappropriate method for people with a high BMI or lots of TSTT. This is because the greater amount of viscoelastic soft tissue trumps the 'elastic' properties of bones, making the response to a lateral impact more non-linear (Figure 2-3).

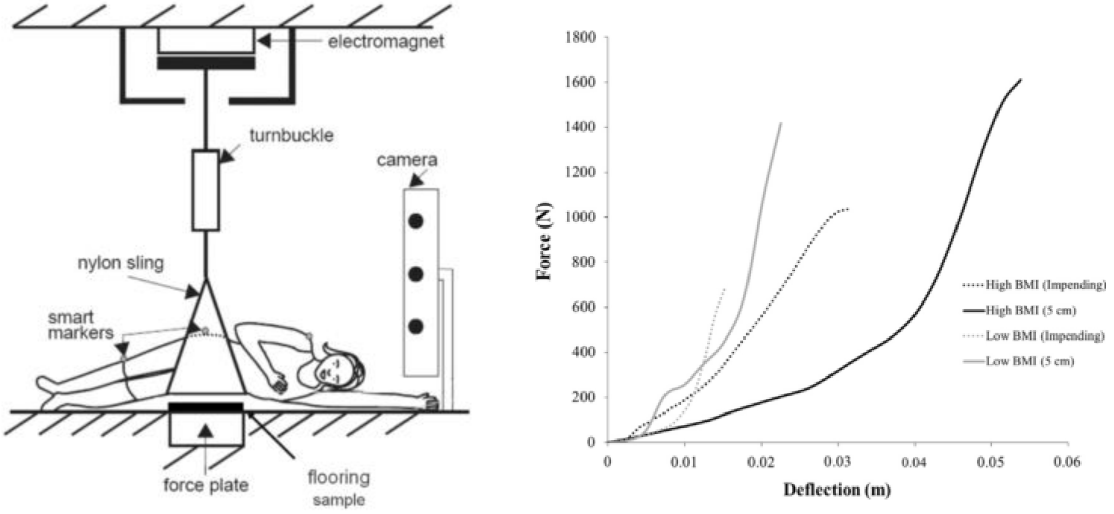


Figure 2-3. Left: Schematic of a pelvis release experiment. Right: Force-deflection profiles of a 0cm and 5cm drop with low and high BMI participants. Note that the high-BMI 5cm drop demonstrated the most non-linear response: it seems as though higher velocity impacts has a greater viscoelastic response with participants who have more TSTT. From *Bhan et al. (2014), Journal of Biomechanics, 47(10): 2391-2397*

Due to its viscoelastic properties, it is likely that there is an interaction between TSTT, effective stiffness and impact velocity. For example, a cadaveric study simulating lateral motor-vehicle accidents showed that TSTT dissipated greater amounts of energy at higher impact velocities (Etheridge et al., 2005). Although the impact velocities were greater than a lateral fall (they used impacts of 3.35 m/s to 8.34 m/s), their low impact velocities are similar to the velocity of a high severity fall (Feldman and Robinovitch, 2007). In brief, people with a large amount of TSTT will demonstrate greater effective viscoelasticity, meaning that their applied force will be more dependent on velocity than people who have a small amount of TSTT (Figure 2-3). Additionally, there is a likely interaction between TSTT, effective stiffness, and age. Choi et al. (2015) demonstrated that trochanteric soft tissue stiffness and damping was significantly smaller in older females than young females. These age-related decreases in soft tissue properties will reduce their capacity to dissipate impact energy from a fall. These differences may be attributed to a decrease in elastin and collagen content in the skin, but it may also be attributed to composition differences of skin, fat, and muscle layers.

The energy absorbing capabilities of TSTT were first quantified by Robinovitch et al. in 1995. They excised the TSTT from nine elderly cadavers (3 males and 6 females) and used it in a surrogate pelvis/impact pendulum test system that simulated a 44kg fall at 2.5m/s. Each TSTT sample measured 10 x 10 cm in surface area, and it ranged from 8 to 45 mm, with a mean (SD) of 24 (13) mm. The peak impact force correlated negatively and very strongly with TSTT ($r^2 = 0.91$), with a resulting equation of -71N * TSTT (in mm) (Figure 2-4). It is very important to note that this equation is only 'validated' for a range of 8 to 45 mm with 9 samples and that it is technically impossible to have a linear relationship because you cannot have a negative impact force. This equation crosses the x-axis at 101.4 mm, meaning that anyone with more than 100 mm of TSTT will experience a 'negative' impact force. Even with its flaws, this equation is highly used in epidemiological studies that want to estimate soft tissue attenuation. In a more recent finite element simulation, Majumder et al. (2008) found a greater negative and non-linear correlation ($\text{Fmax/body weight} = 11.363 * \text{TSTT}^{-0.2011}$, $r^2 = 0.972$) between normalized peak impact force and TSTT. However, this simulation was only done with 5 samples (TSTT of 5, 14, 17, 23, and 26 mm). Although these studies were very well done, there is still the question of bio-fidelity when it comes to in-vitro testing and finite element simulations.

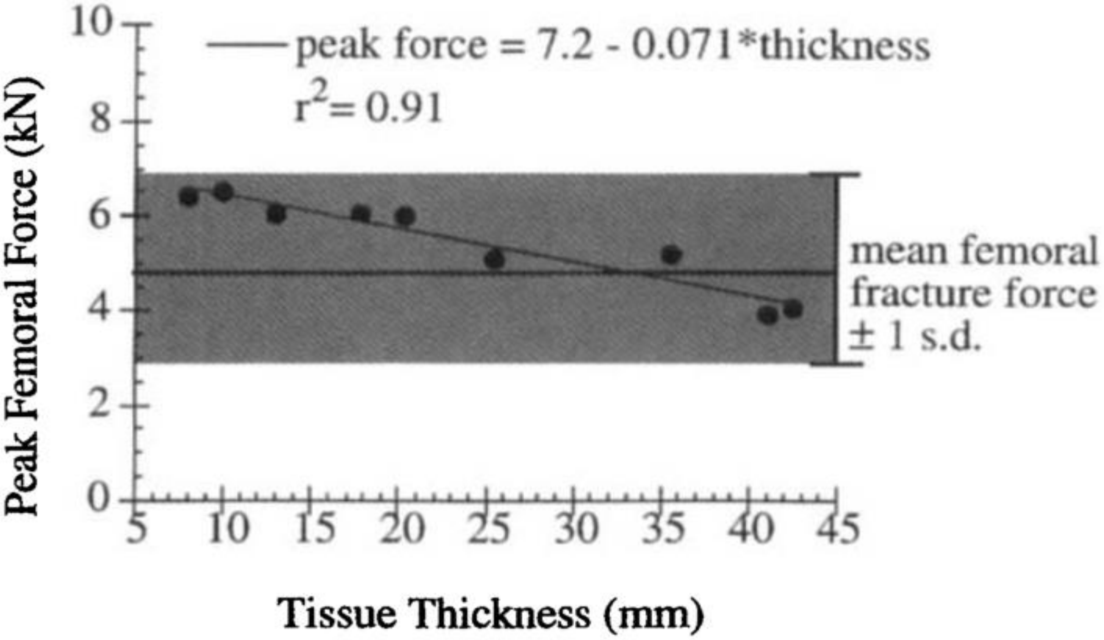


Figure 2-4. The relationship between peak force applied to the hip and TSTT. Note that this equation is only 'validated' for a range of 8 to 45 mm with 9 samples, and that it is technically impossible to have a linear relationship because you cannot have a negative impact force (which in this case would occur for anyone with a soft tissue thickness greater than 100 mm).

2.3.3 The Factor of Risk and TSTT - A Predictor of Hip Fractures

The Factor of Risk (φ) is an engineering principle that is used to ensure the safety of a structure while providing a reasonable economic solution to its design (Hibbeler, 2011). Generally speaking, it is the ratio between the applied load and the failure load of the structure. Specifically for hip fractures, it is the ratio between the impact force applied at the hip and femoral strength. Impact force is typically calculated by using height, weight, stiffness and TSTT estimates, whereas femoral strength is typically estimated by using BMD. If φ is >1, failure is more likely to occur at the proximal femur, and vice-versa if φ is ≤1. The larger the ratio, the greater the probability of failure, and vice-versa for a smaller ratio.

The Factor of Risk principle was first applied to the hip by Hayes et al. (1991), where they used QCT data to estimate femoral strength (they used QCT because it can acquire volumetric properties). At the time, the force applied to the hip was not known, but they made estimations and determined that the Factor of Risk at the hip would be close to 1 in the elderly (Hayes et al., 1991). Later on, the force applied to the hip during impact was discovered, and the authors were able to demonstrate a strong and significant interaction between Factor of Risk and hip fractures (Hayes et al., 1996). In a study of 231 subjects (98 hip fracture cases, 133 non-fractured fallers), the Factor of Risk was more associated with hip fractures than BMD alone, which demonstrates the importance of including impact forces in hip fracture predictions (Hayes et al., 1996). This notion was further proven recently, as the Factor of Risk predicted fractures in 86% of cadaveric femora, whereas osteoporosis (a T-score below ≤ -2.5) predicted fractures in only 74% of cadaveric femora (Roberts et al., 2010). Also, among the cadaveric femora that were not osteoporotic, 52% of the femora would be predicted to fracture by the Factor of Risk (Roberts et al., 2010). The results of these studies suggest that including fall force predictions may improve hip fracture assessment than using BMD measurements alone.

Bouxsein et al. (2007) also tested the Factor of Risk in a subset sample from a female cohort study (21 hip fracture cases and 42 controls). They calculated two types of Factor of Risk: a "peak φ" version that uses the person's mass and height to estimate the peak force applied to the hip and an "attenuated φ" version that included the soft tissue attenuation by using TSTT. These authors showed that the force applied to hip was reduced by an average of 50% and 61% in hip fracture cases and controls when including TSTT. More importantly, the peak φ was greater than the theoretical threshold of 1 in both groups, but when including TSTT, the attenuated φ was near 1 in the fracture cases, and well below 1 for the controls. Although both peak and attenuated φ was significantly associated with the risk of a hip fracture (a 1 SD increase lead to ORs (95% CI) of 1.8 (0.92 – 3.5) and 1.85 (0.96 – 3.6), respectively), the results of this study suggest that including TSTT will likely improve the hip fracture prediction probabilities when using the Factor of Risk. Lastly, this study also demonstrated that a one SD decrease in TSTT was associated with 1.8-fold (1.01 – 3.31) increase in the risk of a hip fracture, but this association was not significant when adjusting for BMD (p=0.25).

A similar study was conducted in a male subset cohort study (70 fracture cases, 222 non-cases), but there were no significant differences in TSTT between fracture cases and controls (29 mm versus 31 mm, p = 0.2), and TSTT was not associated with the risk of hip fracture (Nielson et al., 2009). Additionally, attenuation was estimated to only be up to 26% and 27% in male fracture cases and controls. However, the Factor of Risk was still significantly associated with the risk of hip fracture in men and attenuated φ made more theoretical sense than peak φ (mean attenuated φ in non-cases was still above 1, but was not as high as peak).

More recently, Dufour et al. (2012) showed that an estimated version of the Factor of Risk was a significant predictor of hip fracture in men and women in the large Framingham cohort study. The peak fall force was estimated using a mass-spring model:

$$Peak\ fall\ force\ (N) = \sqrt{2gh_{com}mk}$$

where g is the gravitational constant (9.81m/s^2), h_{com} is the centre of mass height (height*0.51, in m), m is the effective mass (kg), and k is the stiffness constant derived from the Robinovitch et al. (1991) (71 060N/m for females, 90.440N/m in males).

The attenuated fall force was estimated by subtracting the soft tissue attenuation equation from Robinovitch et al. (1995) from the peak fall force:

$$Attenuated\ fall\ force\ (N) = Peak\ fall\ force - (71 * TSTT)$$

where TSTT is the trochanteric soft tissue thickness (mm).

As the Framingham study did not measure TSTT, they estimated it using the following regression equations from Bouxsein et al. (2007) and Nielson et al. (2009):

$$TSTT_{males}\ (mm) = 3.4795 * BMI - 38.015$$
$$TSTT_{females}\ (mm) = 2.3415 * BMI - 33.444$$

where the subscript denotes the male or female formula, and BMI is the body mass index (kg/m^2)

Lastly, the femoral strength was estimated by using a regression formula from Roberts et al. (2010):

$$Femoral\ strength\ (N) = 8207 * BMD - 568.62$$

where BMD is the femoral neck BMD (g/cm^2).

Altogether, the peak and attenuated factor of risk can be calculated as the ratio between fall force and femoral strength:

$$\varphi_{peak} = \frac{Peak\ Fall\ Force}{Femoral\ Strength}$$

$$\varphi_{atten} = \frac{Attenuated\ Fall\ Force}{Femoral\ Strength}$$

Even with the crude estimations, the Factor of Risk was a significant predictor of hip fractures in males and females. An SD increase of peak φ was associated with a 1.88-fold (1.38 – 2.55) and 1.23-fold (1.10 – 1.37) increase in hip fracture risk in men and women, whereas an SD increase of attenuated φ was associated with a 1.78-fold (1.30 – 2.44) and 1.41-fold (1.26 – 1.58) increase. This study also showed that fall force and TSTT was predictive of hip fracture in women, independent of femoral strength (i.e. BMD). Similar to the Nielson study, TSTT seems to be more important in women than men. However, this study comes with a significant limitation. For 13 women who had a BMI greater than 35 kg/m^2, their estimated fall force was negative. The authors believe it was because of their BMI regression equations being validated only in women with a BMI smaller than 35 kg/m^2. It could also be a limitation of the 71N/mm attenuation equation that they used (Robinovitch et al., 1995). Nevertheless, this study demonstrates the importance of obtaining an accurate good measurement of TSTT when predicting hip fracture risk using a Factor of Risk approach.

2.3.4 Measurement Techniques of TSTT

2.3.4.1 Ultrasound

A diagnostic ultrasound device will create inaudible soundwaves between 1 to 20 MHz by converting electrical energy into mechanical energy. More specifically, an alternating current is applied to the ultrasound transducer that contains piezo-electric crystals (usually zirconate titanate, PZT). Most transducers contain arrays of thin, rectangular PZT slabs, which are also called elements (Fairhead and Wittingham, 2012). When these PZT slabs undergo voltage changes, it compresses and expands, which creates ultrasound waves (Thayalan, 2014). The PZT slabs deform in synchronization, making it comparable to pistons pumping ultrasound waves at the same frequency as the applied voltage (Fairhead and Wittingham, 2012). The returning ultrasound waves will have the opposite effect, or in other words, the returning pressure variations will cause the PZT slab to deform, which will ultimately generate voltage variations (Fairhead and Wittingham, 2012). These voltage variations caused by the returning echo is amplified and processed (in several ways) to produce the image on the screen (Fairhead and Wittingham, 2012). There are different types of ultrasound transducers – linear array, phased array, annular array, and endo-probes – but only the linear array probes are relevant to this thesis, so the other transducers will not be presented. There are linear array probes that cover a rectangular field of view (Figure 2-5a), and there are also curvilinear probes that cover a cone-like field of view (Figure 2-5b). Curvilinear probes produce a wider field of view and greater depths of penetration, but pushing the convex front face into full contact causes distortion of superficial structures (Fairhead and Wittingham, 2012). Accordingly, using a linear probe is probably best suited to measure TSTT, as it can better image the skin-air interface, and have a better echo off the greater trochanter due to having more perpendicular ultrasound waves. However, if there is too much soft tissue, a curvilinear probe will be required to obtain a greater penetration depth.

When the ultrasound waves cross a new medium, it can reflect, refract, scatter or be absorbed by matter (Thayalan, 2014). Reflection is when the ultrasound wave “bounces” back towards the transducer (i.e. 180° phase shift, Figure 2-5c), whereas refraction occurs when the ultrasound waves are not perpendicular to the surface, so not all of the waves “bounce” back towards the transducer (Figure 2-5d). Reflection produces an ideal good echo, whereas refraction produces artifacts in the ultrasound signal (Thayalan, 2014). For example, there is 30% reflection at the bone-tissue interface, explaining why there is a shadow underneath it (Thayalan, 2014). Ultrasound waves can also be scattered, resulting in more energy spread out in different directions (Figure 2-5e). Scattering is a common occurrence when imaging bones as they typically have irregular surfaces, Interference from non-targeted material (e.g. blood corpuscle, tissue parenchyma) scatters echoes, which creates an artificial speckle pattern in the image (Thayalan, 2014). Lastly, the ultrasound wave can be absorbed, meaning that the mechanical energy is converted to heat due to frictional and viscous forces (Thayalan, 2014). There are many techniques that can be used to avoid these artifacts – but the most important technique is to ensure that the probe is perpendicular to the desired surface. Other optimization techniques on the ultrasound device, such as adjusting the gain, frequency selection, and tissue harmonics imaging, can be used to acquire a better image. These techniques are well summarized at: http://www.providianmedical.com/ultrasound-imaging-guide.

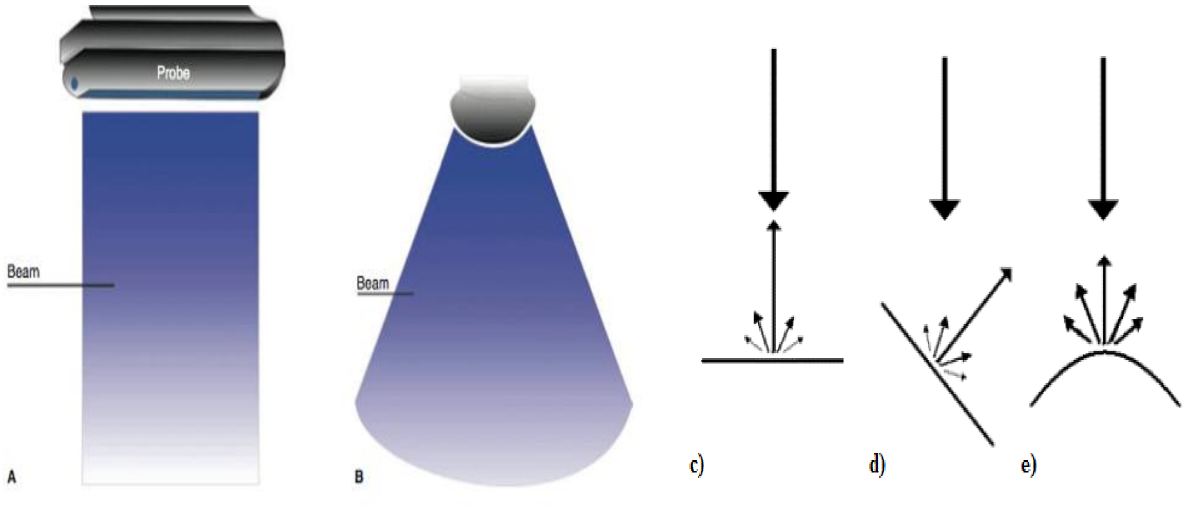


Figure 2-5a) Ultrasound beam produced from a linear probe. b) Ultrasound beam produced by a curvilinear probe. c) An example of reflection. d) An example of refraction. e) an example of scattering. From http://www.nysora.com/mobile/regional-anesthesia/foundations-of-us-guided-nerve-blocks-techniques/3084-ultrasound-physics.html and http://folk.ntnu.no/stoylen/strainrate/Basic_ultrasound

US can be considered the commonly used technique to acquire soft tissue thicknesses in a non-invasive manner. It has been shown to be very strongly correlated with ruler measurements of abdominal thickness in patients undergoing surgery (r = 0.99) (Balta et al., 1981), and it has been shown to have a high intra-trial and inter-day reliability when measuring various body sites (Katch, 1983). When it comes to measuring soft tissue thickness over the greater trochanter, there are different protocols that can be used. In musculoskeletal ultrasound textbooks, it is recommended that the lateral hip is measured when the patient is lying on their side, with a pillow between their knees for comfort (Hill and Leiszler, 2013; Martinoli and Bianchi, 2007; O’Neill, 2008). Depending on what needs to be measured, the transducer can be placed longitudinally or transversely. A 9 to 15 MHz linear array transducer can be used for most patients, but a 5 MHz transducer may have to be used for obese or muscular patients (Hill and Leiszler, 2013; O’Neill, 2008). Ironically, only one research study has collected TSTT with the participants lying on their side (Choi et al., 2015). The remainder of the research studies measured TSTT while standing upright (Levine et al., 2015; Maitland et al., 1993; Minns et al., 2007). Although there is no soft tissue movement in a standing position, it could be difficult for older adults to stand still for a prolonged amount of time (Maitland et al., 1993). Interestingly, no study has measured trochanteric soft tissue thickness using an US in a supine position, even though TSTT is measured this way when using radiography. Moreover, the intra-rater reliability of TSTT measurements has only been documented in a standing position (Levine et al., 2015; Maitland et al., 1993). It is not known if ultrasonic TSTT measurements can be taken more reliably in another position, or if a certain position has better reliability measures between different evaluators.

2.3.4.2 DXA/CT Measurements

Dual Energy X-ray Absorptiometry (DXA or DEXA) and computerized tomography (CT) are other imaging modalities that can be used to acquire soft tissue thicknesses. DXA is the most popular assessment of bony tissue in both research and clinical practice. It is a machine that produces two X-ray beams of a distinct energy to separate bone from soft tissue (Bonnick, 2004). By assuming that the transmission of the X-ray beams can be accurately described by a mono-exponential attenuation process and that the body is a two component system of bone and soft tissue, the density of bone is essentially calculated by determining the ratio of X-ray intensity before and after it passes through the body (Njeh and Shepard, 2004). DXAs can also calculate body fat percentage by using the same core principles. DXA assumes a three component model of bone mass, fat mass, and bone mineral-free lean mass (Wilmore et al., 2009). These are the same principles used by the predecessor Dual Photon Absorptiometry (DPA), but DPA used radiographic isotopes (usually Gadolinium-153) instead of X-ray tubes (Bonnick, 2004). Depending on the manufacturer, the X-rays of two distinct energies are either produced using filters or using pulsed power sources (Bonnick, 2004). Currently, there are three types of scanners in DXA devices: pencil-beam, fan-beam, and cone-beam. Pencil-beam scanners project a narrow X-ray beam that moves rectilinearly with the detector (Bonnick, 2004). Fan-beam scanners project a fan-shaped beam onto multiple projectors, allowing a whole scan line to be instantly quantified and improving resolution (Bonnick, 2004). Cone-beam scanners project a cone beam to the region of interest, thus reducing scanning time and correcting for the effects of scattered radiation at the detector (Behari, 2009). Major manufacturers of DXA scanners are Hologic Inc., GE Medical Systems (Lunar), and Norland of the CooperSurgical Company (Bonnick, 2004).

DXA was first used to measure TSTT, and it correlated very strongly with ultrasound measurements ($r^2 = 0.815$, (Maitland et al., 1993)). This study used a Hologic QDR – 1000 bone densitometer and performed a hip scan that went from the surface of the skin to the medial border of the acetabulum. A plastic block was placed 1cm away from the patient to help distinguish the skin-air interface, and TSTT was measured by counting the 1.006mm square pixels between the greater trochanter and the skin. A coefficient of variation (CV) was calculated by repeating the measurement 9 times in a subject, and the CV was 3.7% (Maitland et al., 1993).

A slightly different technique has also been used in the literature. Bouxsein et al. (2007) used a whole body analysis algorithm (“Enhanced array whole body”, V5.54A; QDR2000, Hologic) to determine TSTT. The skin-air interface was improved by manually adjusting the luminosity and contrast. The CV of this method was 3.9% (used a repeated analysis of 20 whole body scans on three separate occasions). Another similar study used the same technique, but they also compared this DXA method to a QCT method and found a high correlation between the two (r = 0.8, (Nielson et al., 2009)). This study calculated an intra-analyzer CV of 2.6% and an inter-scan CV of 6.4%. Although it is more commonly used for diagnoses in other regions of the body, CT is another imaging tool that can be used to assess bone structure and health. Lastly, Schacter and Leslie (2014) used a similar protocol as the other studies, but they used a GE Prodigy DXA scanner and it’s respective whole body analysis software. This study tested inter-analyzer reliability and got a CV of 3.5%, and intra-analyzer CV of 0.93%

These studies have noted interesting limitations. An issue with using DXA to assess TSTT is that the patients are supine, thus the soft tissue from the buttocks may laterally spread and alter measurements (Bouxsein et al., 2007). This issue is assumed to be magnified with obese patients because they have greater amounts of soft tissue (Maitland et al., 1993). Moreover, current methods to assess TSTT using DXA/CT must be done manually, which are labor intensive (Nielson et al., 2009). Nevertheless, DXA/CT measurements of TSTT are easy to conduct (if the equipment is available), explaining why DXA is the most common way to measure TSTT in the literature.

2.3.4.3 Anthropometric Measurements

Certain anthropometric measurements can also be used as estimators of TSTT (Table 2-2). There are several simple measurements (e.g. BMI, body fat percentage, hip circumference) that can be used as a surrogate for TSTT when US or DXA is unavailable. For example, Dufour et al. (2012) used BMI regression equations to estimate TSTT in the Framingham cohort study. Using a backward stepwise multivariable linear regression model, Schacter and Leslie (2014) found that significant contributors to a TSTT prediction equation were sex, BMI, spine average thickness, and hip average thickness. This study found a modest relationship ($r^2 = 0.6$) between this equation and TSTT measured from a whole body DXA scan (Schacter and Leslie, 2014). The authors of this study conclude that TSTT can be well estimated from a BMD hip scan, avoiding the need to do a whole-body scan. Maitland et al. (1993) found that a zoomed in DXA scan measure of TSTT and hip circumference explained 89% of the variance of US TSTT (while standing). To the best of my knowledge, a regression-based model of TSTT that doesn't need a DXA scanner, and one that includes more than BMI, has never been made (or used) in the literature.

Table 2-2. Linear correlations between anthropometric measurements and TSTT

Measurement	Study	Correlation (r) with US TSTT (standing)	Correlation (r) with DXA TSTT
DXA	Maitland et al. (1993)	0.903	(1)
Hip Circumference	Maitland et al. (1993)	0.9	0.837
BMI	Maitland et al. (1993) Nielson et al. (2009) Schacter and Leslie (2014)	0.849 -	0.78 0.75 0.67
BF%	Maitland et al. (1993)[1] Nielson et al. (2009)[2]	0.862 -	0.781 0.65
Waist/hip circumference	Maitland et al. (1993)	0.434	0.343
Age	Nielson et al. (2009) Schacter and Leslie (2014)	- -	-0.16 -0.14
Sex (male)	Schacter and Leslie (2014)	-	-0.25
Height (cm)	Nielson et al. (2009) Schacter and Leslie (2014)	-	0.06 -0.11
Weight (kg)	Nielson et al. (2009) Schacter and Leslie (2014)	-	0.66 0.55
Total body fat mass (kg)	Nielson et al. (2009)	-	0.75
Total body lean mass (kg)	Nielson et al. (2009)	-	0.43
Leg fat mass (kg)	Nielson et al. (2009)	-	0.81
Leg lean mass (kg)	Nielson et al. (2009)	-	0.43
Leg fat %	Nielson et al. (2009)	-	0.69
Total hip BMD (g/cm^2)	Nielson et al. (2009)	-	0.23
Trochanteric BMD (g/cm^2)	Nielson et al. (2009)	-	0.14
Femoral neck BMD (g/cm^2)	Nielson et al. (2009)	-	0.22
Hip fat	Schacter and Leslie (2014)	-	0.44

Spine fat	Schacter and Leslie (2014)	-	0.5
Hip AP tissue thickness (cm)	Schacter and Leslie (2014)	-	0.64
Spine AP tissue thickness (cm)	Schacter and Leslie (2014)	-	0.45

[1] Collected BF% using a body impedance analyzer

[2] Collected BF% from the DXA scan

2.3.5 Section Summary

TSTT influences the effective stiffness of the body when subjected to a lateral fall, which ultimately modulates the amount of force applied to the hip if the deformation remains constant. A recent finite element study demonstrated that TSTT was a more dominant parameter than weight and height when it came to normalized peak force, time to peak force, and strain ratio at the femoral neck (Majumder et al., 2013). Studies have also shown that velocity and age will affect energy dissipating capabilities of TSTT, or in other words, it's capability to influence the effective stiffness of the body (Choi et al., 2015; Etheridge et al., 2005). Consequently, it is extremely difficult to predict someone's effective stiffness (which would predict the applied fall force to the hip) in a non-laboratory setting, so researchers have developed TSTT attenuation equations to help estimate the applied force to the hip (Majumder et al., 2008; Robinovitch et al., 1995). Although the 71 N/mm attenuation equation has only been validated in cadavers with less than 45 mm, it has been used in epidemiological studies to estimate the impact force from a fall (Bouxsein et al., 2007; Dufour et al., 2012; Nielson et al., 2009). As demonstrated in Table 2-1, different measurement techniques of TSTT have been used in the literature. The measurement technique may influence the TSTT obtained, which will then influence its attenuation estimation, which may ultimately influence its predictive capabilities for hip fracture risk. Consequently, these techniques should be assessed and compared to help develop a standard protocol method for future studies and practice.

2.4 Key Knowledge Gaps

2.4.1 Concordance validity between TSTT measurement techniques

Several different approaches have been used to measure TSTT in the literature. The measurement devices used were US, DXA, CT, and a BMI regression equation, and the participants in these studies were in different positions: standing, supine, and lying on their side. To assess pathologies of the lateral hip (e.g. trochanteric bursitis, gluteus medius/minimus tendinitis), it is recommended to use US while the patient is lying on their side (Hill and Leiszler, 2013; Martinoli and Bianchi, 2007; O’Neill, 2008). However, only one US study has measured TSTT with the participant in a side-lying position; the rest employed a standing position. US measurements while standing is feasible, but frail older adults may not tolerate it because of the prolonged standing while being undressed and having gel applied (Maitland et al., 1993). Supine measurements would be more comfortable , but compressed gluteal soft tissue may overestimate the true amount of TSTT. Although it is assumed that there would be differences, no study has quantified TSTT differences between the three different positions. The effects of hip rotation have also never been assessed, even though it has been shown to affect BMD measurements at the hip (Goh et al., 1995; Lekamwasam and Lenora, 2003). Moreover, the agreement between the two most popular TSTT measurement devices – US and a whole body DXA scan – has never been evaluated. Accordingly, it is not known whether the results from the different studies can be compared to each other. In the event that differences exist, equations to convert TSTT across devices/postures could be a valuable tool for future researchers or clinicians.

2.4.2 Accuracy, reliability, acceptability and availability

When trying to determine the ideal measurement method, it is important to consider the technique’s accuracy, reliability, acceptability, and availability. Accuracy is very difficult to assess in-vivo with TSTT, but studies have shown that US is excellent at measuring thickness in cadaveric specimens,. The accuracy of the BMI regression equation used in the literature is not certain as it has never been validated in a separate cohort. DXA and US reliability have been well assessed, with the exception of US inter-rater reliability, and the reliability between positions. Patient acceptability is likely to be high for each measurement tool, but it may differ between positions. The availability of US and DXA should be high in clinics or research facilities that assess hip fracture risk, but the BMI regression equation is universally available. Filling in these knowledge gaps regarding accuracy, reliability, acceptability and availability may help future clinicians and researchers choose the measurement tool and person position that is most appropriate for their situation. Additionally, it would help make a “gold standard” recommendation for measuring TSTT.

2.5 Corresponding Thesis Objectives

The first objective of this study was to determine the concordance validity of TSTT between the supine US and DXA measurements. The second purpose was to determine if there are any significant differences in TSTT between standing, supine, and sideline measurements, as well as between an internal hip rotation of 25°, 0° rotation, and an external hip rotation of 25°. The third objective was to determine the intra and inter-rater reliability across DXA and US measurements made in standing, supine, and side-lying position, in addition to supine US image inter-day reliability.

Chapter 3

Thesis Research Study

3.1 Introduction

Approximately 30 000 Canadians suffer a hip fracture per year, and these rates are expected to rise with our aging population (McGlasson et al., 2011). The total annual cost to care for hip fractures was estimated to be 650 million dollars, and this cost is projected to rise to 2.4 billion dollars in the year 2040 (Wiktorowicz et al., 2001). From a personal perspective, hip fractures are a serious injury as they are highly linked to functional impairment, disability, loss of independence, reduction of quality of life, and death (Korhonen et al., 2013).

There are three major factors that influence the risk of a hip fracture: the risk of falling, the fall-induced impact force, and the proximal femur strength. When reviewing the 2010 Canadian guidelines, I found they recommended a good assessment of the risk of falling and the proximal femur strength, but did not do an adequate assessment to determine the fall-induced impact force. I suspect that including TSTT could improve the risk of hip fracture assessment, especially for those who are classified in the medium risk category. However, there is no established technique when it comes to measuring TSTT, as studies have used different methodologies. Accordingly, it is not known which measurement technique of TSTT would have the most predictive capabilities of hip fracture.

3.1.1 Rationale for Thesis

The measurement technique of TSTT will affect the estimate of force attenuation and ultimately influence its ability to predict hip fractures. Accordingly, it is important to compare current measurement techniques to determine if there are any significant differences. If there are differences, researchers and clinicians should be wary when choosing their measurement technique of TSTT, as it can affect the risk assessment of a hip fracture. Also, the accuracy, reliability, acceptability and availability of these different techniques should be determined, so that these researchers and clinicians can make an informed decision when choosing how to measure TSTT.

3.1.2 Purpose and Hypotheses

Overall, the general purpose of this thesis is to provide guidance towards improving current approaches for measuring TSTT. More specifically, the first objective of this study is to determine the concordance validity of TSTT between the supine US and DXA measurements. The associated hypotheses were:

i. The intraclass correlation (ICC) between the supine US and DXA TSTT measurement will be strong (ICC > 0.8)

ii. The coefficient of repeatability(CR) between the supine US and DXA TSTT measurements will be below 0.96 cm

The second purpose was to determine if there are any significant differences in TSTT between standing, supine, and side-lying measurements, as well as between an internal hip rotation of 25°, 0° rotation, and an external hip rotation of 25°. The associated hypotheses were:

iii. There will be no significant interaction between body position and hip rotation ($p > 0.05$).

iv. There will be a significant difference between the TSTT collected in a standing, supine, and side-lying position ($TSTT_{Stand} \neq TSTT_{Side} \neq TSTT_{Sup}$, $p<0.05$)

v. There will be significant differences between the TSTT collected in a 25° internal hip rotation, a 0° rotation, and a 25° external hip rotation ($TSTT_{25int} \neq TSTT_{0rot} \neq TSTT_{25ext}$, $p<0.05$).

The third objective is to determine the intra and inter-rater reliability across DXA and US measurements made in a standing, supine, and side-lying position. Also, supine US image inter-day reliability will be determined in this study. The associated hypotheses were:

vi. The intraclass correlation within raters, between raters, and within images will be strong for US and DXA analyses (ICC > 0.8)

vii. The CR will be below 0.96 cm within rater, between raters, and within images for US and DXA analyses.

3.2 Methodology

3.2.1 Participants

Forty-five participants (25 females and 20 males) between the ages of 40 to 90 were recruited from the local Waterloo community. Participants were excluded from the study if they:

a) Suspected or were known to be pregnant
b) Ingested a contrast solution, or had recent injections for a radiologic investigation, within the last month (e.g. CT scan or nuclear medicine test)
c) Had undergone bariatric surgery, lost or gained more than 25 lbs in the last year

d) Is unable to stand for a continuous 15 minutes with a supporting aid (i.e. table)

As this thesis' primary question was to determine the agreement between US and DXA, a large range of TSTT was needed. The BMI (a surrogate of TSTT) of the first 20 participants were categorized, and then missing or underrepresented groups were targeted. As shown in Figure 3-1, the distribution of the participants was normal with the exception of the 21 to 23 kg/m^2 group. However, individuals with a lower BMI are at a higher risk of a hip fracture, so it was important to capture extra information about this sub-group. Descriptive statistics of the participants are presented in Table 3-1.

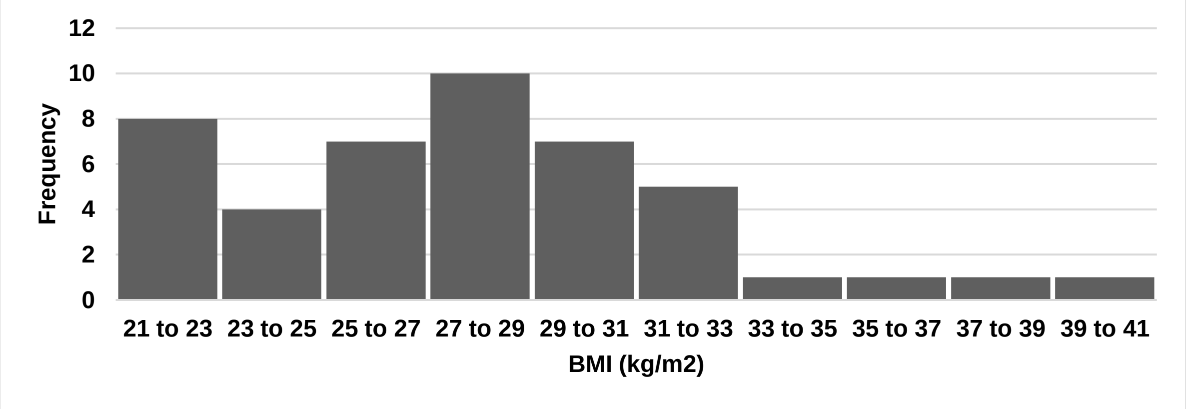


Figure 3-1. Frequency plot of the BMI of the participants in this study.

Table 3-1. Descriptive statistics for all participants, and for male and female participants separately. Values presented are mean (SD, Range)

	Both (n = 45)	**Men (n =20)**	**Women (n= 25)**
Age (years)	70.2 (10.8, 40 to 85)	72.5 (9.1, 43 to 85)	68.5 (11.9, 40 to 82)
Height (m)	1.68 (0.1, 1.5 to 1.86)	1.76 (0.07, 1.66 to 1.86)	1.61 (0.1, 1.5 to 1.72)
Weight (kg)	77.4 (14.2, 51 to 116.6)	84.5 (13.4, 59 to 116.6)	71.8 (12.3, 51 to 98.9)
BMI (kg/m^2)	27.5 (4.3, 20.8 to 40.2)	27.1 (3.6, 20.8 to 37.1)	27.9 (4.8, 21.6 to 40.2)
Waist Circumference (cm)	95.0 (13.3, 64.4 to 130.9)	101 (11.8, 84 to 130.9)	90.1 (12.7, 64.4 to 112.3)
Hip Circumference (cm)	105.9 (8.6, 92.4 to 133.1)	104.1 (5.9, 95.5 to 120.8)	107.4 (10.1, 92.4 to 133.1)
Fem Neck BMD (T-score)	-0.69 (1.1, -3.4 to 1.5)	-0.5 (1.4, -3.4 to 1.5)	-0.8 (0.9, -2.5 to 0.8)

3.2.2 Recruitment Protocol

Participants were recruited in a variety of ways. First, I conducted a presentation about my research work at a local condominium complex, and the attendees were provided with my contact information. Second, I contacted potential participants by using the Waterloo Research and Aging Pool (WRAP), which is a pre-established list of interested older-adults from the community. Lastly, recruitment occurred by using “word of mouth” and posters that were placed throughout the region and the campus of the University of Waterloo. Participants that contacted me underwent an initial screening to ensure that they met the inclusion criteria listed above.

3.2.3 General Experimental Protocol

After obtaining informed oral and written consent, all participants completed two questionnaires pertaining to their general health, and previous occupational activity. This process took approximately 15 minutes.

Participants were instructed to wear loose shorts, and they were rolled up to access the lateral right hip. If they did not bring loose shorts, participants were asked to change into medical scrubs that had a closable patch on the right hip before taking any measurement. After (possibly) changing into the scrubs, the following anthropometric measurements were made: mass, height, waist circumference, hip circumference, thigh circumference, thigh length, standing width (i.e. heel to heel), and standing hip angle. Each measurement was taken twice, and occasionally a third measurement was taken if the first two were not similar. The anthropometric measurements took approximately 15 minutes to conduct.

Afterward, participants underwent two DXA scans that were conducted by a certified Medical Radiation Technologist. The first DXA scan was a hip scan needed to acquire BMD of the hip, whereas the second DXA scan was a whole body scan needed to measure TSTT. The two scans took approximately 30 minutes.

After the DXA scans, US measurements of the participant’s right hip were taken by the principal and secondary investigators. In brief, the greater trochanter was palpated, the probe (with ultrasound gel on it) was applied gently to the skin over the greater trochanter, and an image was taken. There were a total of 18 measurements of TSTT, which took approximately 60 minutes to acquire. Each measurement was unique in the sense that it was made by a different investigator, in a different position, in a different hip rotation, or with a different transducer.

The whole experimental protocol took approximately 2 hours for each participant. It is important to note that experimental protocol was not always performed in the order mentioned above. Specific details of each mini-protocol will be presented in the next sections.

3.2.4 DXA Scan

The DXA scanner used in this study was a Hologic QDR4500 Discovery. All whole body DXA scans were conducted by a certified Medical Radiation Technologist. Outputs of this scan are lean soft tissue mass and body fat percentages, but the whole body scan was needed in this study to measure TSTT. A protractor was used to ensure that the participant's hip was rotated 25° internally. The participant was also told to avoid any type of movement during the DXA scans. Each image was checked immediately after the scan to make sure there were no abnormalities.

3.2.5 Ultrasound

For each US measurement, the following procedure was used. First, the lateral hip was palpated by the investigator, and the greater trochanter was located. When the greater trochanter was difficult to palpate, the following techniques were used:

- The investigator asked the participant to "squish a bug with their foot" (i.e. make internal/external rotations of the hip). During this motion, the investigator palpated the hip in order to locate the moving greater trochanter. A finger was kept on the location until the transducer was placed. This technique was useful for the standing posture condition.
- The investigator palpated the Anterior Superior Iliac Spine (ASIS) with the thumb (with permission of the participant), and then placed the index finger 12 cm postero-laterally. Minns et al. (2007) demonstrated that the greater trochanter was approximately 12 cm postero-lateral from the ASIS, and this technique was useful for supine measurements.
- The investigator placed the transducer longitudinally on the leg and located the femur. Once the femoral shaft was located, the participant moved superiorly (relative to the participant) until they found the greater trochanter. This technique was useful for standing, supine, and side-lying measurements, in addition to for participants with greater amounts of TSTT (Figure 3-2).
- When palpating, the participant was asked if they felt like the investigator was palpating their greater trochanter. Many participants had a good awareness of when their greater trochanter was being palpated. This was a useful technique for standing, supine, and side-lying measurements.

- When examining the ultrasound screen, there were certain landmarks that helped identify the greater trochanter. Examples of these landmarks are non-circular bone, gluteus minimus or medius tendons right above the bone, or even the iliotibial band halfway between the greater trochanter and the skin (Figure 3-2).

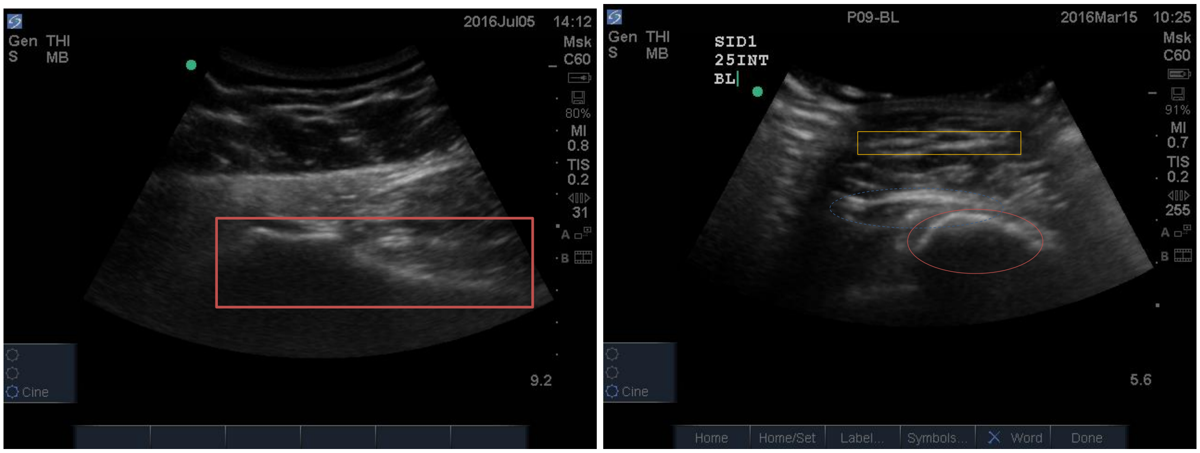


Figure 3-2. Left: An ultrasound image in the longitudinal (or sagittal) view of the greater trochanter and femoral shaft (in the red box). Once the centre was located, the transducer was turned to the transverse axis. Right: The non-circular greater trochanter (red circle), the gluteus minimus and medius tendons (blue dashed circle), and the iliotibial band (orange box) helped the investigators identify the greater trochanter.

Once the greater trochanter was located, US images were taken using a curvilinear probe (C60x, 2-5MHz) in combination with the M-Turbo Ultrasound Unit 1.0.6, a Lite II Mobile Docking System, with 2D visualization and processing software (SonoSite, Bothell, WA). The default frequency selected was "General", but for some participants with substantial soft tissue thickness the "Penetrate" frequency was selected to get a better echo off the greater trochanter. Sufficient ultrasound gel was placed on the probe to help distinguish the skin-air interface, and to avoid compressing the tissue. The probe was always placed transversally in relation with the participant, and best attempts were made to be perpendicular with the greater trochanter. Moreover, image depth and gain was modified to improve the image quality. After the image was taken, a text code was typed on the image to associate the image with the specific investigator, position, and hip rotation angle. Next, a built-in caliper function of the ultrasound system was used to determine the TSTT, which is defined as the distance between the proximal layer of skin and the proximal layer of the greater trochanter (all relative to the probe). One image was saved with the calipers on them, and another image was saved without the calipers on them. The image with the calipers will be used to determine the 'full' inter and intrarater reliability, whereas the image without the calipers will be used later to determine 'image' interday reliability. It is important to note that a sticky note covered the bottom left aspect of the ultrasound screen so that the investigators could not see the TSTT value at any point during the collection. Thorough training (>15 hours) occurred prior to the thesis data collection.

3.2.5.1 Participant Positioning

The three body positions tested in this thesis were standing, supine, and side-lying and the three rotations were 25° internal rotation, 0° rotation, and 25° external rotation, totaling to nine different configurations (Figure 3-3). To ensure that only internal/external hip rotation was changing at the hip, and no hip adduction/abduction/flexion/extension occurred, foam triangles were used to standardize positions across all participants (Figure 3-3). The researchers ensured that the heels were always at the back of the triangle so that abduction remained constant. A goniometer was also used to ensure that the participants' trunk angle and knee angles were maintained at zero degrees.

Table 3-2 outlines how many measures were collected by the investigators with the participant in each specific position. More measures were taken when the patient was supine or 25° internally rotated as it is a more clinically relevant position. This set-up allows a full-factorial analysis to determine if positioning or hip rotation affects TSTT.

Table 3-2. Number of US images taken by BL in the three different positions and hip rotation conditions. Values in parentheses indicate the number of trials conducted by a co-investigator for a subset of 30 participants.

	25° internal rotation	0 rotation	25° external rotation
Standing	2 (1)	1	1
Supine	3 (1)	2	2
Side-lying	2 (1)	1	1

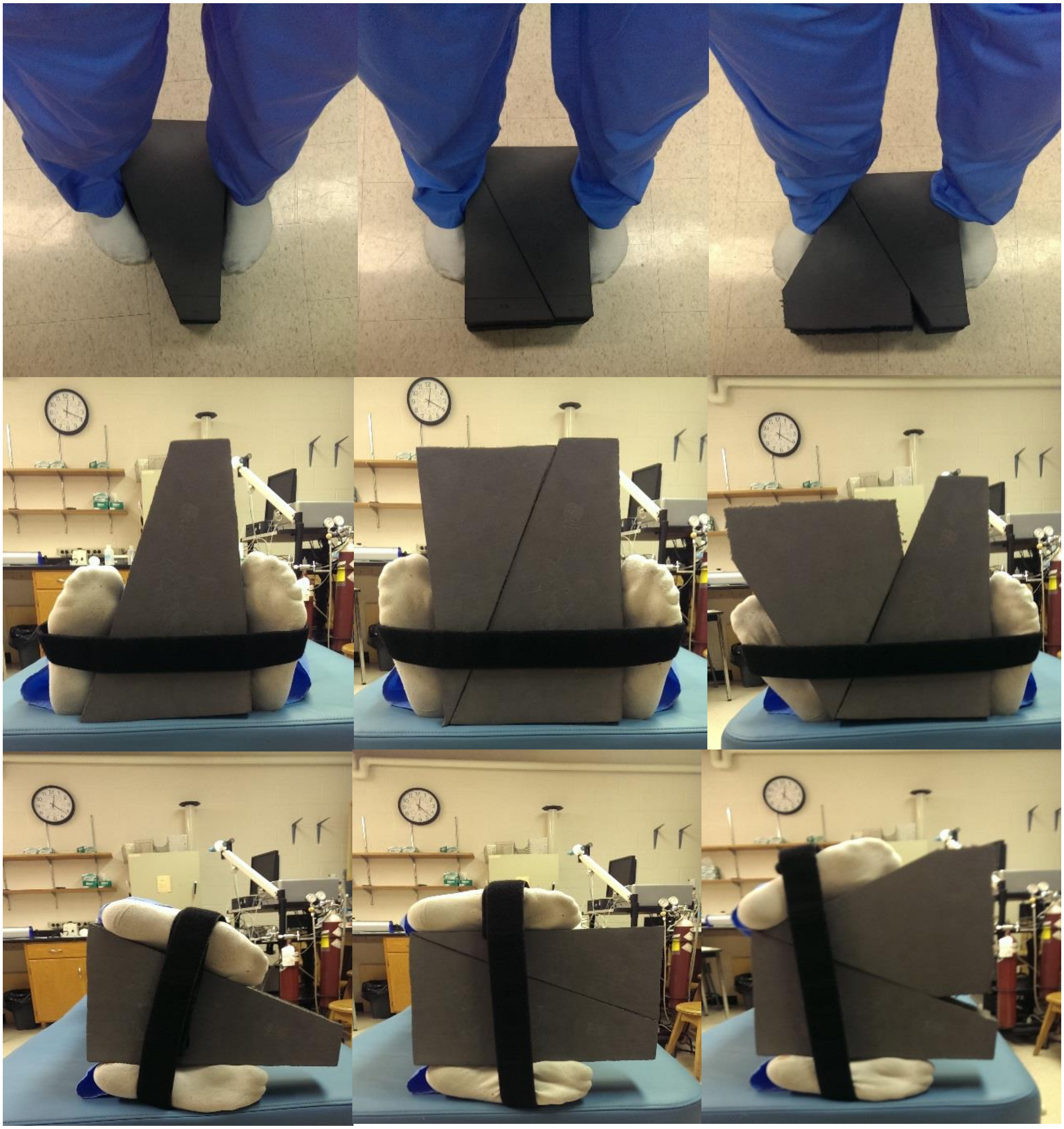


Figure 3-3. The nine different positions used in this thesis. The 1st row demonstrates the standing position, the 2nd row demonstrates the supine position, and the 3rd row demonstrates the side-lying position. The 1st column demonstrates 25° internal rotation, the 2nd column demonstrates 0 rotation, and the 3rd column demonstrates 25° external rotation. Care was taken to make sure the heels were always at the back of the triangle.

3.2.5.2 Investigators' Roles/Reliability Protocol

A specific order was used when taking the US measurements (see Appendix D - Additional Details of Methodology). This study employed one primary investigator, and a secondary investigator to support inter-rater reliability analyses. Measurements were alternated between the primary and secondary investigator to induce a mini washout period. When one investigator was taking the US image, the other investigator was assisting the participant with tasks including: changing their hip rotation angle, communicating with the participant, keeping track of the measurement order, etc. The investigator not taking the image avoided looking at the screen at any time. After each trial, the image-taker returned the image depth to its maximum of 9 cm, and to the "General" penetration setting to avoid biasing the reliability results. They also wiped off the ultrasound gel so that the next investigator would have to re-landmark. This protocol was chosen so that every source of error can be introduced into the reliability analysis: the error associated with landmarking the greater trochanter on the person, the error associated with changing the image depth or gain, the error associated with the transducer orientate or pressure placed on the person, the error associated with landmarking on the image, etc.

3.2.6 Data Analysis

TSTT from the ultrasound was measured during the collection by using the built-in caliper function (SonoSite, Bothell, WA). Although the value was saved, TSTT was never identified during the collections as the corner of the computer screen (where TSTT was displayed) was covered by a sticky note. After the collection session, US TSTT details were input into a Microsoft Excel spreadsheet.

After the DXA data input of all of the participants, the first supine 25° internal rotation image (without calipers) was analyzed in ImageJ. The supine position was chosen as it was expected to be the most difficult position to obtain TSTT, thus providing greater insight into reliability. More specifically, TSTT was reanalyzed in Image J, without prior knowledge of the thickness because of a two-week washout period. Unlike the caliper function with the ultrasound device, ImageJ allows the user to zoom in on the specific landmarks (i.e. the greater trochanter and the epidermis layer). This process was needed to calculate the image inter-day reliability. Preliminary training was completed to avoid a learning effect. The TSTT from the image reanalysis was inputted (and date-stamped) in the same Microsoft Excel spreadsheet as above.

After the data from all participants had been collected, TSTT was measured from the whole body DXA scans by using the Hologic Physician Viewer program (Version 6.2). The contrast and zoom of the scan were manipulated to help the investigators better identify the shortest distance between the greater trochanter and the overlying skin. The step-by-step is explained in Figure 3-4. All 45 DXA scans were analyzed within the same day (i.e. session). The principal investigator (BL) took three measures of DXA TSTT. The first measure was taken after two weeks of the last data collection session, and the second and third measure of DXA TSTT were taken 4 and 7 days after the first measure. The DXA scans were randomized for each session in order to avoid measurement bias. The secondary investigator (AT) performed the same DXA session at another time in order to conduct inter-rater reliability analyses.

Figure 3-4. Illustration of the DXA TSTT analysis. First, TSTT is measured in the whole body scan (left image). Second, the right hip is magnified and adjustments are made (middle image). Third, the contrast is manipulated to properly identify the bone pixels and skin pixels (right image). Fourth, the image is zoomed out and TSTT is visually checked before the measurement value is recorded.

Bland-Altman plots were created by using a customized MATLAB program (version 8.2, Mathworks Inc., MA, USA) that plotted the difference of two TSTT measurements (y-axis) against the mean of these two TSTT differences (x-axis). These differences were between US or DXA, between positions, or between raters. The mean of all the differences was plotted as a horizontal line, along with the 95% confidence intervals of these differences.

3.2.7 Statistical Analyses

To test the reliability within and between raters, the coefficient of repeatability (CR) was utilized. The CR is the standard deviation of the differences multiplied by 1.96 to give the 95% confidence interval (Bland and Altman, 1986). Previous literature has shown that the TSTT difference between a fracture and non-fracture case is approximately 0.96 cm (0.94 cm for Bouxsein et al. (2007); 0.98 cm for Minns et al., 2007). Therefore, I am considering 0.96 cm to be a clinically significant TSTT difference, and the CR of the differences should fall within this threshold to be considered good agreement (for inter-device comparison) or reliable (for intra and inter-rater reliability analyses). In other words, a CR > 0.96 cm is considered wide limits of agreement, whereas a CR < 0.96 cm is considered narrow limits of agreement. The mean TSTT of hip fracture patients are between 3 cm to 4 cm (Table 2-1), so 0.96 cm represents 25 to 33% of that value.

For the Bland-Altman plots, the DXA TSTT was considered the reference method for the device agreement analysis. The first TSTT measure by BL will be considered the reference method for the reliability analyses as he was the primary investigator. Proportional bias will be tested by determining if the regression between the TSTT means and differences are significant ($p < 0.05$) (Earthman, 2015). Lastly, the percentage of differences that exceed this threshold will be determined. The 0.96 cm boundaries are displayed in all of the Bland-Altman plots as dotted lines.

Lastly, coefficients of variation (CV) of US and DXA measurements of TSTT were calculated for each participant. The coefficient of variation is calculated by dividing the standard deviation by the mean of the two or three TSTT measures, and multiplying that by 100 to get a percentage. To maintain consistency, all CVs were calculated by using the first two measures. The CV will be determined for each participant, and the mean and range of all the participant CVs will be reported for each approach. CV will not be used in any statistical tests, but it will be compared to CVs calculated in other studies.

For each of the hypotheses, a separate analysis was performed using statistical software (SPSS Version 22, SPSS Inc., Chicago, IL, USA) with an alpha value of 0.05.

i. The correlation between the supine US and DXA TSTT measurement will be strong (ICC > 0.8).

   - A two-way random absolute agreement intraclass correlation (ICC (2,1)) will be calculated to test this hypothesis. The two-way random ICC was chosen over the two-way mixed ICC as this technique is meant to generalize the results for several raters, and not just the raters in this study (Rankin and Stokes, 1998). Only the first measure was compared as it is more clinically relevant.

ii. The CR between the supine US and DXA TSTT measurements will be below 0.96 cm
   - The CR will be calculated and a Bland-Altman plot will be used to visualize the results.

iii. There will be no significant interaction between body position and hip rotation.
   - To test this hypothesis, a two-factor repeated measure ANOVA will be conducted. If Mauchly's test of sphericity is violated ($p > 0.05$), the Greenhouse-Geisser correction will be used.

iv. There will be a significant difference between the TSTT collected in a standing, supine, and side-lying position.
   - To test this hypothesis, a two-factor repeated measure ANOVA will be conducted. Furthermore, pairwise comparisons will be conducted to determine specific differences. If Mauchly's test of sphericity is violated ($p > 0.05$), the Greenhouse-Geisser correction will be used.

v. There will be significant differences between the TSTT collected in a 25° internal hip rotation, a 0° rotation, and a 25° external hip rotation.
   - To test this hypothesis, a two-factor repeated measure ANOVA will be conducted. Furthermore, pairwise comparisons will be conducted to determine specific differences. If Mauchly's test of sphericity is violated ($p > 0.05$), the Greenhouse-Geisser correction will be used.

vi. The correlation within raters, between raters, and within images will be strong for US and DXA analyses ($ICC > 0.8$)
   - A two-way random absolute agreement intraclass correlation (ICC (2,1)) will be calculated to test this hypothesis.

vii. The CR will be below 0.96 cm within rater, between raters, and within images for US and DXA analyses.

- The CR will be calculated and a Bland-Altman plot will be used to visualize the results.

3.3 Results

All of the US and DXA measurements presented below can be found in Appendix B.

3.3.1 Concordance Validity between US and DXA

When using the first measure of the three supine US and DXA measurements, the ICC between the two imaging devices was 0.898. The same analyses were conducted with the average measures and produced similar results. Accordingly, only the first supine measure results were presented, as it is consistent with the other analyses. The mean bias of the differences (with DXA as the reference method) was 0.46 cm, the CR was 2.15 cm, and 62.2% of the differences were within the 0.96cm bounds (Figure 3-5). There was an increasing proportional bias between the two imaging devices as the regression between the differences and means was significant ($y = 0.212x - 0.6005$, $r^2 = 0.242$, $p < 0.001$). The smallest difference between the two imaging devices was 0.05 cm, whereas the largest difference was 2.77 cm. The mean (range) CV was 13.4 (1.9 – 33.3) %.

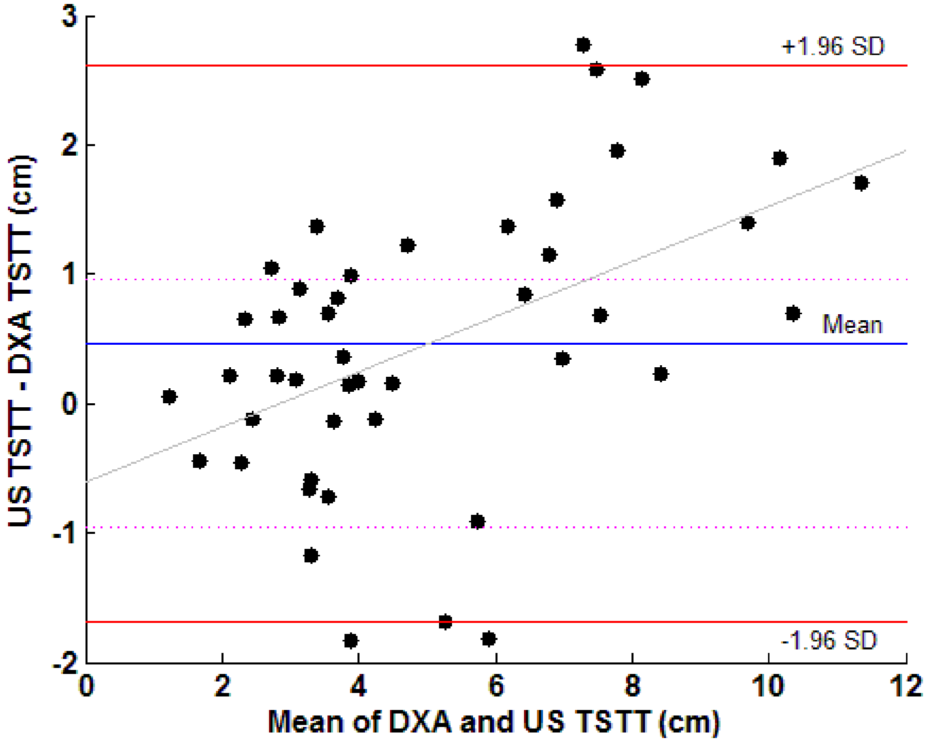


Figure 3-5. Bland-Altman plot for the supine US and DXA measure. The mean bias is 0.46 cm (blue line) and the CR is 2.15 cm (red lines). 62.2% of the US-DXA differences were within the 0.96 cm bounds (pink lines). A proportional bias was found (grey line), and the regression equation is $y = 0.212x – 0.6005$ ($r^2 = 0.242$, $p < 0.001$).

3.3.2 TSTT Interaction between Position and Hip Rotation

The two-factor repeated measure ANOVA results demonstrate a significant interaction between body position and hip rotation ($p = 0.018$). As demonstrated by Figure 3-6, the interaction is ordinal, where the slopes of the lines do not cross and are not parallel. Independent of hip rotation, the supine position produced the largest TSTT, followed by the standing position, and finishing with the side-lying position. Independent of body position, the 25 external rotation produced the largest TSTT, followed by a zero-degree rotation, and finishing with a 25 internal rotation. The mean (SD) of TSTT in each body position and hip rotation are: 4.16 (2.2) cm for standing/25° internal rotation, 4.22 (2.1) cm for standing/no rotation, 4.61 (2.2) cm for standing/25° external rotation; 5.24 (2.8) for supine/25° internal rotation, 5.46 (2.8) for supine/no rotation, 5.99 for supine/25° external rotation; 3.1 (1.6) for side-lying/25° internal rotation, 3.3 (1.8) for side-lying/no rotation, and 3.47 (1.7) for side-lying/25° external rotation.

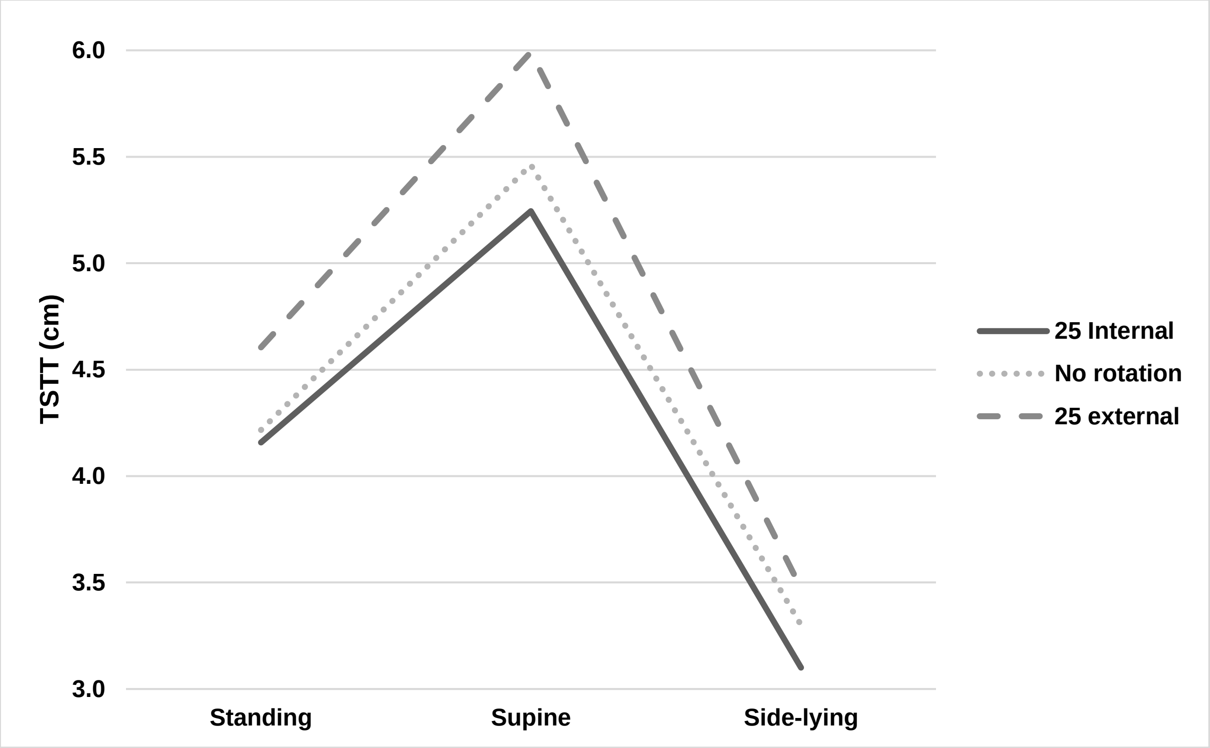


Figure 3-6. Mean TSTT in each body position and hip rotation

3.3.3 Main Effect of Body Position on TSTT

The two-factor repeated measure ANOVA results demonstrate a main effect of body position on TSTT ($p < 0.001$). With a significant interaction between body position and hip rotation, it could be argued that main effects should not be interpreted. However, this interaction can be open to interpretation as hip rotation always had the same effect on each body position (TSTT increased with external rotation of the hip, Figure 3-6). Additionally, the differences between each position are too large to not be considered a main effect. The mean (SD) TSTT was largest for the supine position (5.57 (2.8) cm), followed by the standing position (4.33 (2.1) cm), and smallest for the sideline position (3.29 (1.7) cm) (Figure 3-7). Pairwise comparisons with Bonferroni corrections demonstrate significant differences between standing and supine TSTT (-1.24 cm, $p < 0.001$), between standing and side-lying TSTT (1.04 cm, $p < 0.001$), and between supine and side-lying TSTT (2.28 cm, $p < 0.001$) .

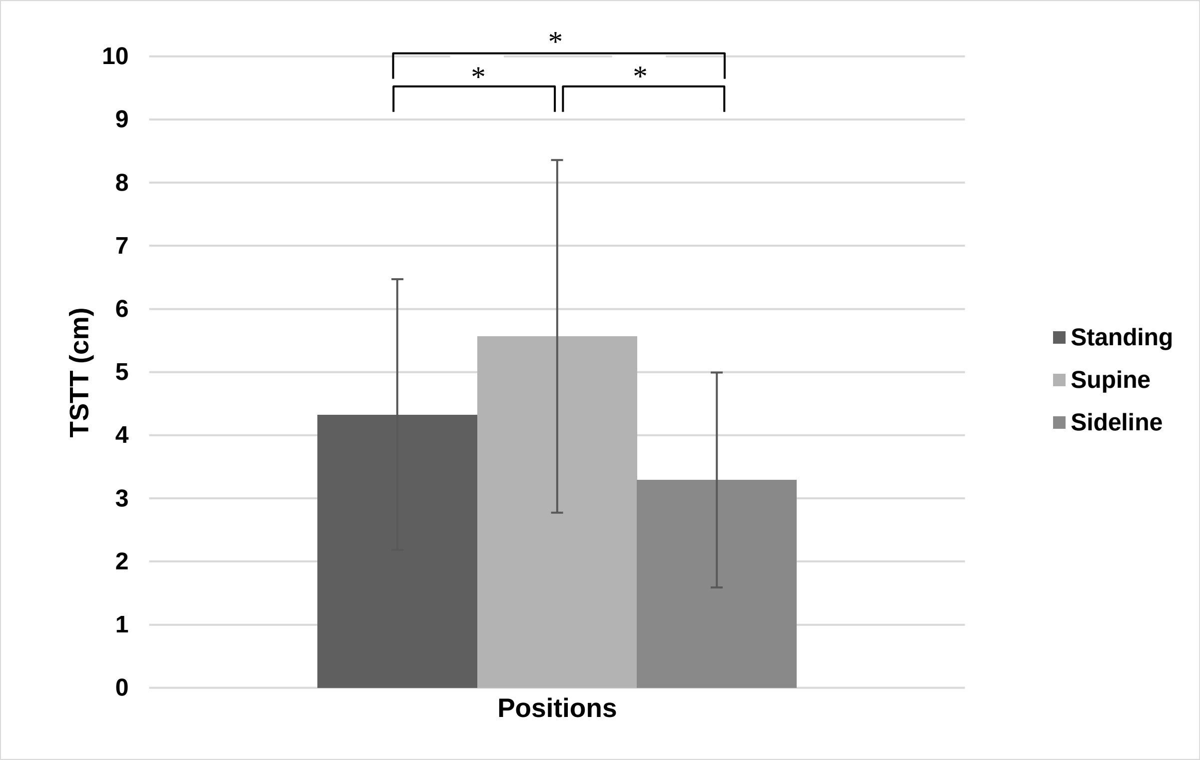


Figure 3-7. Mean (SD) TSTT in each body position (* denotes p < 0.001)

3.3.4 Main Effect of Hip Rotation on TSTT

The two-factor repeated measure ANOVA results demonstrate a main effect of body position on TSTT ($p < 0.001$). Unlike the main effect of body position, it is not clear whether the main effect of hip rotation can be interpreted because of the interaction. Consequently, three one-factor ANOVAs (the factor being hip rotation) were conducted at each body position. Each one-factor ANOVA demonstrated a main effect of hip rotation on TSTT. Furthermore, all pairwise comparisons with Bonferroni corrections showed significant differences ($p < 0.05$), with the exception of 25 internal rotations and no rotation in the standing and supine position. Accordingly, these subsequent analyses of one-factor ANOVAs support the results of the two-factor ANOVA. The mean (SD) of the first 25° internally rotated TSTT was the smallest with 4.17 (2.4) cm, non-rotated TSTT was in the middle with 4.33 (2.4) cm, and 25° externally rotated TSTT was 4.69 (2.5) cm (Figure 3-8). Pairwise comparisons with Bonferroni corrections demonstrate significant differences between 25° internal rotation and no rotation (-0.158 cm, $p = 0.023$), between 25° internal rotation and 25° external rotation (-0.522 cm, $p < 0.001$), and between no rotation and 25° external rotation (-0.364 cm, $p < 0.001$) .

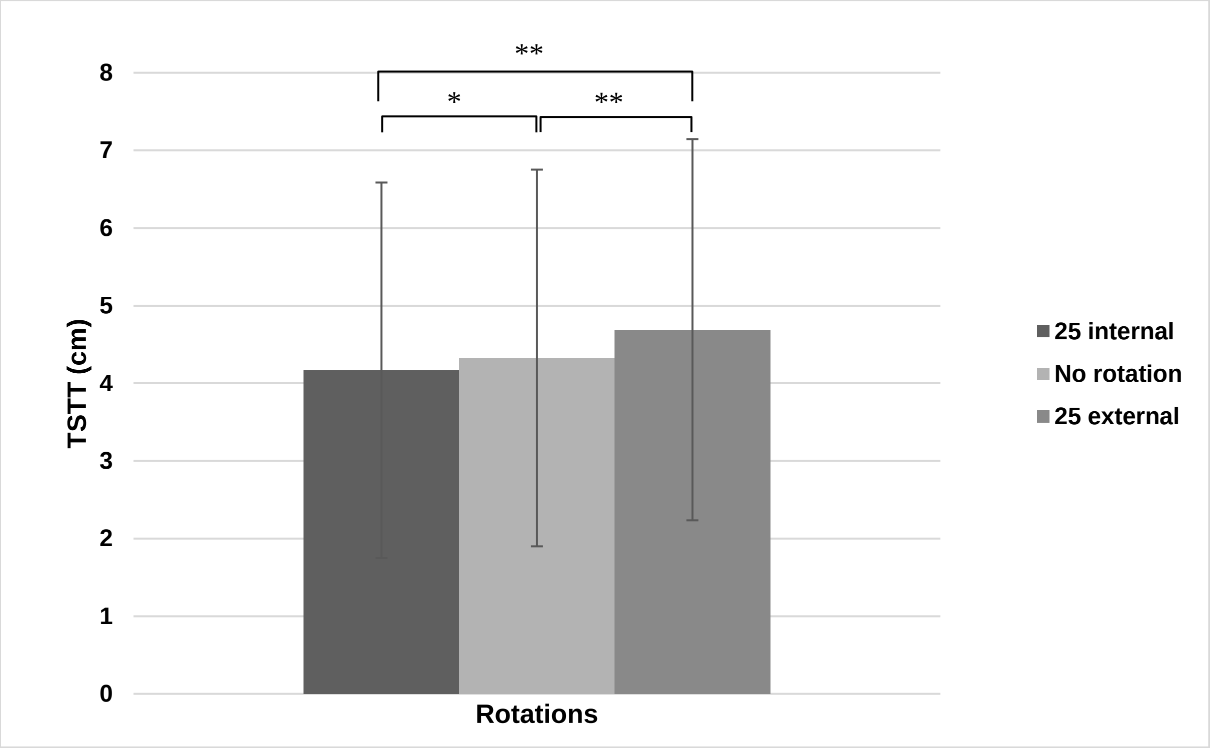


Figure 3-8. Mean (SD) TSTT in each hip rotation (* denotes $p < 0.05$, ** denotes $p < 0.001$)

3.3.5 US Intra-Rater Reliability

The ICC within BL was 0.980 for the standing US measurements, 0.972 for the supine US measurements, and 0.977 for the side-lying US measurements. The mean bias of the differences was -0.07 cm for the standing US (Figure 3-9, Left), -0.04 cm for the supine US (Figure 3-9, Middle), and -0.02 cm for the side-lying US (Figure 3-9, Right). The CR for the standing US was 0.87 cm (Figure 3-9, Left), 1.32 cm for the supine US (Figure 3-9, Middle), and 0.69 cm for the side-lying US (Figure 3-9, Right). Lastly, 93.3% of the standing US differences were within the 0.96 cm bounds (Figure 3-9, Left), 80% for the supine US differences (Figure 3-9, Middle), and 97.8% for the side-lying measurements (Figure 3-9, Right). The mean (and range) CV of the standing measurements was 6.50 (0.3-47.1) %, the CV of the supine measurements was 6.35 (0.2-21.3) %, and the CV of the side-lying measurements was 5.81 (0-23.7) %.

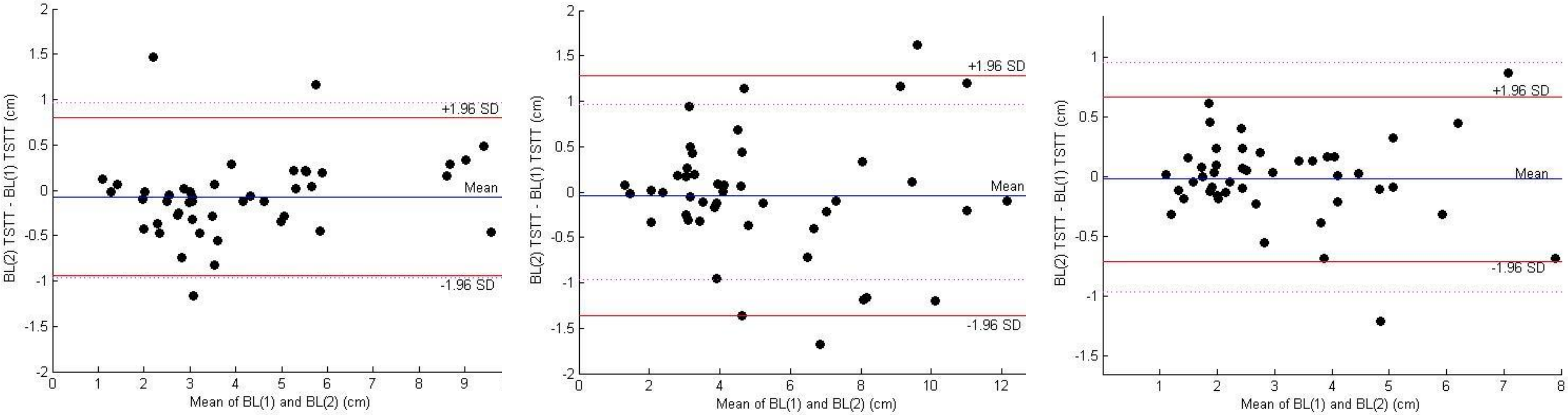


Figure 3-9. Left: Bland-Altman plot for standing US TSTT within the first two measurements of the principal investigator BL. The mean bias is -0.07 cm (blue line) and the CR is 0.87 cm (red lines). 93.3% of the standing US differences were within the 0.96 cm bounds (pink lines).

Middle: Bland-Altman plot for supine US TSTT within the first two measurements of the principal investigator BL. The mean bias is -0.04 cm (blue line) and the CR is 1.32 cm (red lines). 80% of the supine US differences were within the 0.96 cm bounds (pink lines).

Right: Bland-Altman plot for side-lying US TSTT within the first two measurements of the principal investigator BL. The mean bias is -0.02 cm (blue line) and the CR is 0.67 cm (red lines). 97.8% of the side-lying US differences were within the 0.96 cm bounds (pink lines).

3.3.6 US Inter-Rater Reliability

The ICC between BL and AT was 0.970 for the standing US measurements, 0.939 for the supine US measurements, and 0.948 for the side-lying US measurements. The mean bias of the differences was -0.08 cm for the standing US (Figure 3-10), -0.19 cm for the supine US (Figure 3-10), and 0.11 cm for the side-lying US (Figure 3-10). The CR for the standing US was 1.17 cm (Figure 3-10), 2.08 cm for the supine US (Figure 3-10), and 1.10 cm for the side-lying US (Figure 3-10). Lastly, 93.3% of the standing US differences were within the 0.96 cm bounds (Figure 3-10), 73.3% for the supine US differences (Figure 3-10), and 93.3% for the side-lying measurements (Figure 3-10). The mean (and range) CV of the standing position was 8.9 (0.26 – 41.1) %, the supine position CV was 9.9% (1.12 to 39.6) %, and the side-lying position CV was 10.5 (0.74 to 31.7) %

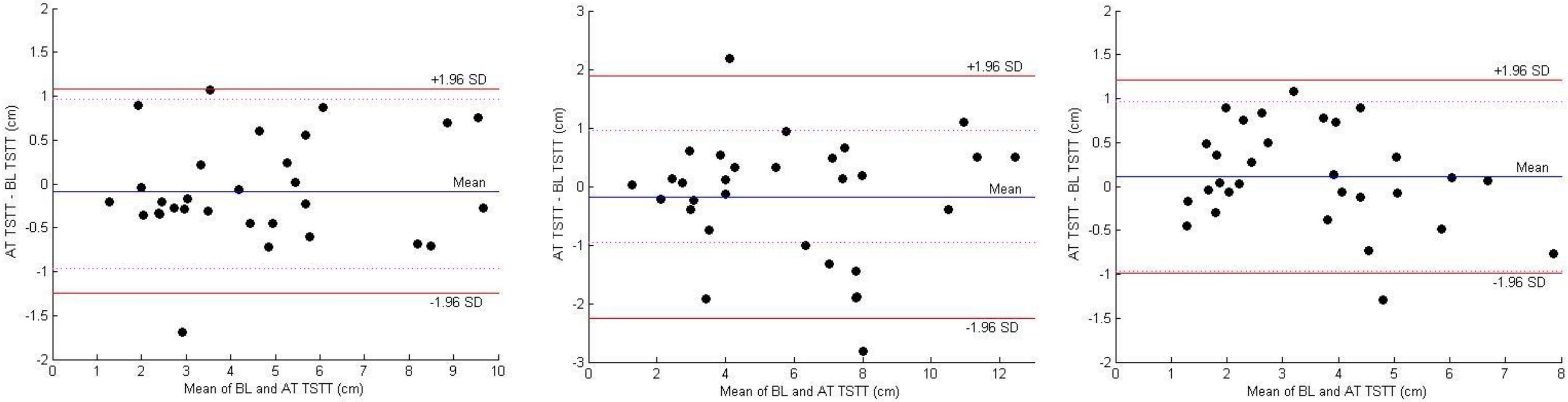


Figure 3-10. Left: Bland-Altman plot for standing US TSTT between investigators BL and AT. The mean bias is -0.08 cm (blue line) and the CR is 1.17 cm (red lines). 93.3% of the standing US differences were within the 0.96 cm bounds (pink lines).

Middle: Bland-Altman plot for supine US TSTT between investigators BL and AT. The mean bias is -0.19 cm (blue line) and the CR is 1.10 cm (red lines). 73.3% of the supine US differences were within the 0.96 cm bounds (pink lines).

Right: Bland-Altman plot for side-lying US TSTT between investigators BL and AT. The mean bias is 0.11 cm (blue line) and the CR is 1.10 cm (red lines). 93.3% of the side-lying US differences were within the 0.96 cm bounds (pink lines).

3.3.7 US Image Inter-Day Reliability

The agreement between the original and re-analysed (through ImageJ) first supine US image was extremely high. The ICC was 0.999, the bias was -0.04 cm, the CR was 0.25 cm, and all of the differences were within the 0.96 cm (Figure 3-11). The mean (and range) CV was 1.6 (0.01 and 20.1) %.

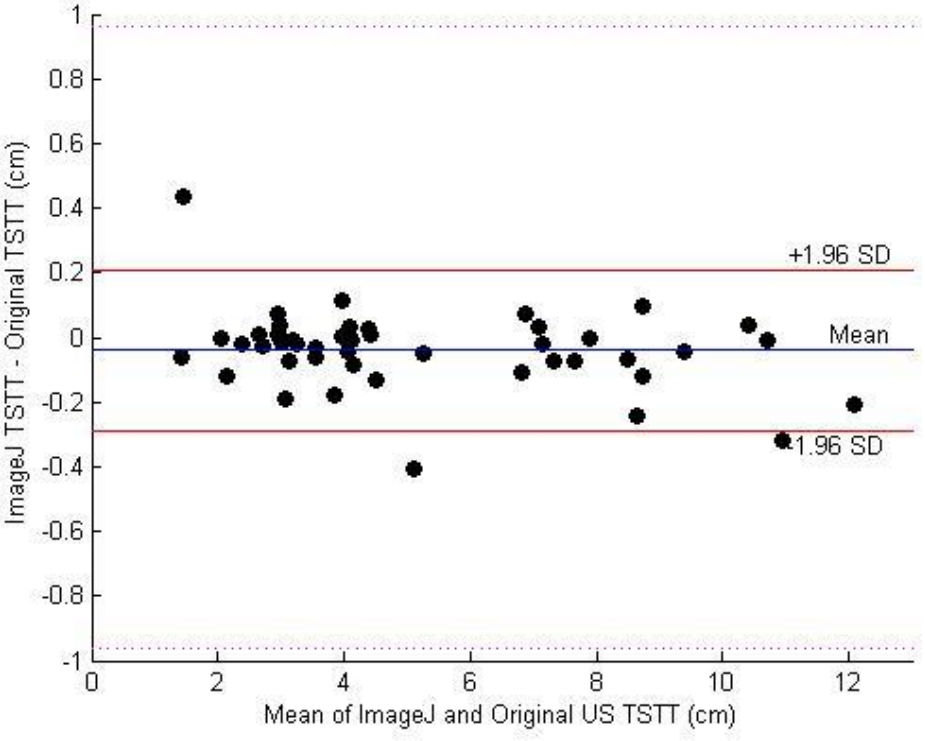


Figure 3-11. Bland-Altman plot for the re-analyzed supine US images. The mean bias is -0.04 cm (blue line) and the CR is 0.25 cm (red lines). 100% of the differences were within the 0.96 cm bounds (pink lines).

3.3.8 DXA Intra-Rater Reliability

For the DXA measurements, the ICC within BL was 0.995. The bias of the differences was -0.07 cm and the CR was 0.45 cm (Figure 3-12). All of the differences were within the 0.96 cm bounds, with the largest difference being 0.5 cm (Figure 3-12). The mean (and range) CV for the DXA measurements was 3.3 (0 – 16.6) %.

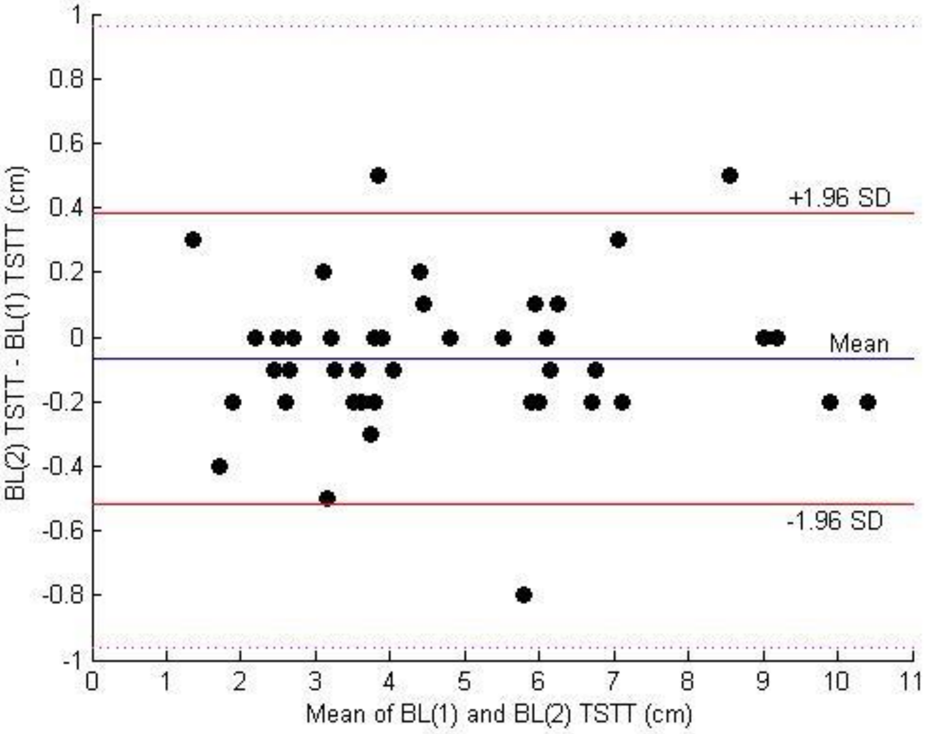
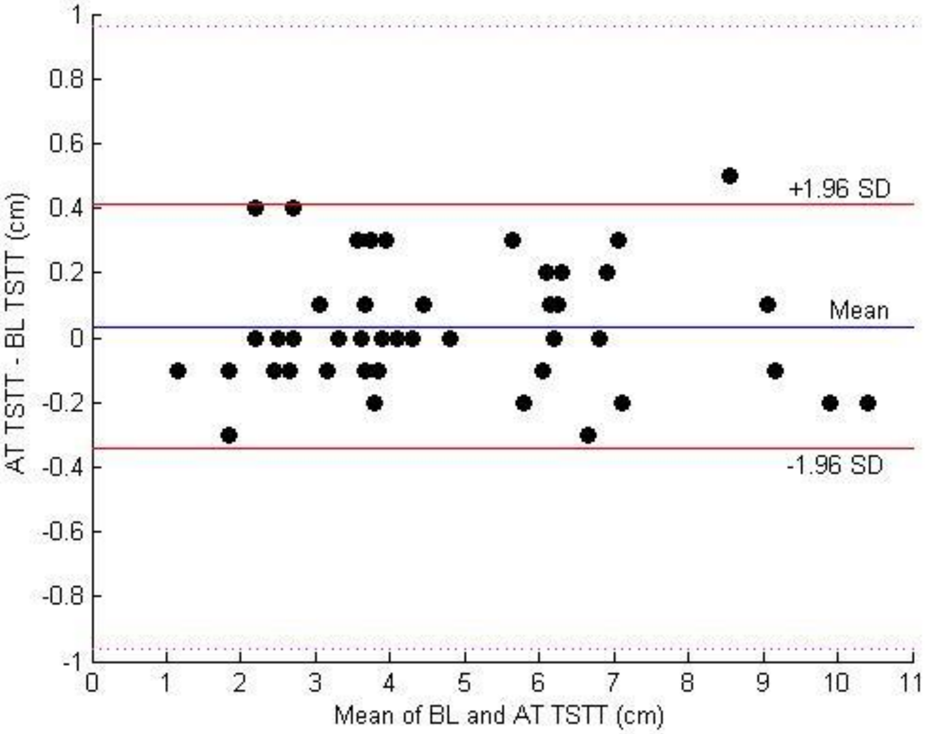


Figure 3-12. Bland-Altman plot for DXA TSTT within BL. The mean bias is -0.07 cm (blue line) and the CR is 0.45 cm (red lines). 100% of the DXA differences were within the 0.96 cm bounds (pink lines).

3.3.9 DXA Inter-Rater Reliability

For the DXA measurements, the ICC between BL and AT was 0.995. The bias of the differences was -0.03 cm and the CR was 0.38 cm (Figure 3-13). All of the differences were within the 0.96 cm bounds, with the largest difference being 0.5 cm (Figure 3-13). The mean (and range) CV was 2.6 (0 – 12.9) %.



12

Figure 3-13. Bland-Altman plot for DXA TSTT between investigators BL and AT. The mean bias is -0.03 cm (blue line) and the CR is 0.38 cm (red lines). 100% of the DXA differences were within the 0.96 cm bounds (pink lines).

3.3.10 Summary of Results

The following tables present a summary of the results found in this thesis. **Error! Reference source not found.** is a summary of the mean (SD) TSTT measurements used to test the thesis hypotheses, whereas **Error! Reference source not found.**, **Error! Reference source not found.**, and **Error! Reference source not found.** are a summary of the thesis objectives, hypotheses acceptance or rejection, and interpretation of the results. The green (lighter shade) squares indicate an accepted hypothesis, whereas the red (darker shade) squares indicate a rejected hypothesis.

Table 3-3. Mean (SD) of the TSTT measurements used to test the thesis hypotheses

STANDING TSTT (cm)	BL1, 25int	BL2, 25int	BL, 0rot	BL, 25ext	AT, 25int
Mean (SD)	4.16 (2.2)	4.08 (2.3)	4.22 (2.1)	4.61 (2.2)	4.03 (2.2)
Mean (SD) of BL measures	4.33 (2.1)				
SUPINE TSTT (cm)	BL1, 25int	BL2, 25int	BL, 0rot	BL, 25ext	AT, 25int
Mean (SD)	5.24 (2.8)	5.21 (2.8)	5.46 (2.8)	5.99 (2.7)	5.18 (2.7)
Mean (SD) of BL measures	5.57 (2.8)				
Side-lying TSTT (cm)	BL1, 25int	BL2, 25int	BL, 0rot	BL, 25ext	AT, 25int
Mean (SD)	3.10 (1.6)	3.08 (1.6)	2.84 (1.4)	3.30 (1.8)	3.18 (1.6)
Mean (SD) of BL measures	3.29 (1.7)				
DXA TSTT (cm)	BL1	BL2	BL3	AT	
Mean (SD)	4.78 (2.3)	4.71 (2.3)	4.77 (2.3)	4.81 (2.3)	
Mean (SD) of BL measures	4.75 (2.3)				

Table 3-4. Results of hypotheses testing and interpretation for the first objective. The green squares indicate an accepted hypothesis, whereas the red squares indicate a rejected hypothesis

First Objective	Determine the concordance validity of TSTT between the supine US and DXA measurements			
	Hypothesis	Result	Hypothesis	Result
First measure	i) = ICC > 0.8	ICC (2,1) = 0.898	ii) = CR < 0.96	CR = 2.15 cm
Interpretation	The concordance validity between the supine US and DXA TSTT is poor, especially in participants with more TSTT.			

Table 3-5. Results of hypotheses testing and interpretation for the second objective. The green (lighter shade) squares indicate an accepted hypothesis, whereas the red squares (darker shade) indicate a rejected hypothesis

Second Objective	Determine if there are any significant differences in TSTT between standing, supine, and sideline measurements, as well as between internal hip rotation of 25°, 0° rotation, external hip rotation of 25, and a possible interaction			
	Hypothesis	Result	Pairwise Comparisons	Result
Interaction of Body Position and Hip Rotation	iii) No interaction between body position and hip rotation	$p = 0.018$	N/A	N/A
Main Effect of Body Position	iv) $\text{TSTT}_{\text{Stand}} \neq \text{TSTT}_{\text{Side}} \neq \text{TSTT}_{\text{Sup}}$, $p<0.05$	$p < 0.001$	$\text{TSTT}_{\text{Stand}} \neq \text{TSTT}_{\text{Stand}}$ $\text{TSTT}_{\text{Stand}} \neq \text{TSTT}_{\text{Side}}$, $\text{TSTT}_{\text{Side}} \neq \text{TSTT}_{\text{Sup}}$	$p < 0.001$ $p < 0.001$ $p < 0.001$
Main Effect of Hip Position	v) $\text{TSTT}_{\text{25int}} \neq \text{TSTT}_{\text{0rot}} \neq \text{TSTT}_{\text{25ext}}$, $p<0.05$	$p < 0.001$	$\text{TSTT}_{\text{25int}} \neq \text{TSTT}_{\text{0rot}}$ $\text{TSTT}_{\text{0rot}} \neq \text{TSTT}_{\text{25ext}}$ $\text{TSTT}_{\text{25int}} \neq \text{TSTT}_{\text{25ext}}$	$p = 0.023$ $p < 0.001$ $p < 0.001$
<u>Interpretation</u>	Body position and hip rotation influence TSTT measurements, and these two factors combined will also influence the TSTT measurement			

Table 3-6 Results of hypotheses testing and interpretation for the third objective. The green squares indicate an accepted hypothesis, whereas the red squares indicate a rejected hypothesis

Third Objective	Determine the intra and inter-rater reliability across DXA and US measurements made in a standing, supine, and side-line position			
US Intra-Rater Reliability	Hypothesis	Result	Hypothesis	Result
Standing measures	vi) = ICC > 0.8	ICC (2,k) = 0.980	vii) = CR < 0.96	CR = 0.87 cm

Supine measures	vi) = ICC > 0.8	ICC (2,k) = 0.972	vii) = CR < 0.96	CR = 1.32 cm
Sideline measures	vi) = ICC > 0.8	ICC (2,k) = 0.977	vii) = CR < 0.96	CR = 0.69 cm
Interpretation	The US intra-rater reliability is good in the standing and sideline position, but mediocre in the supine position, especially in participants with more TSTT.			
US Inter-Rater Reliability	Hypothesis	Result	Hypothesis	Result
Standing measures	vi) = ICC > 0.8	ICC (2,k) = 0.970	vii) = CR < 0.96	CR = 1.17 cm
Supine measures	vi) = ICC > 0.8	ICC (2,k) = 0.939	vii) = CR < 0.96	CR = 2.08 cm
Sideline measures	vi) = ICC > 0.8	ICC (2,k) = 0.948	vii) = CR < 0.96	CR = 1.10 cm
Interpretation	The US inter-rater reliability is mediocre in the standing and sideline position, but poor in the supine position, especially in participants with more TSTT			
US Image Inter-Day Reliability	Hypothesis	Result	Hypothesis	Result
Supine	vi) = ICC > 0.8	ICC (2,k) = 0.999	vii) = CR < 0.96	CR = 0.25 cm
Interpretation	The US image inter-day reliability is very good			
DXA Intra-Rater Reliability	Hypothesis	Result	Hypothesis	Result
	vi) = ICC > 0.8	ICC (2,k) = 0.995	vii) = CR < 0.96	CR = 0.45 cm
Interpretation	The DXA intra-rater reliability is very good			
DXA Inter-Rater Reliability	Hypothesis	Result	Hypothesis	Result
	vi) = ICC > 0.8	ICC (2,k) = 0.995	vii) = CR < 0.96	CR = 0.38 cm
Interpretation	The DXA inter-rater reliability is very good			

3.4 Discussion

The first objective of this study was to determine the agreement between supine US and DXA measurements of TSTT, and it was found to be poor between the two imaging devices. On average, DXA underestimated TSTT when compared to ultrasound, and a positive proportional bias was found between the two. The second objective was to examine the potential influence of three postures and hip rotations, and both of these factors were found to have main effects on TSTT. These findings suggest there is value in developing a standardized approach for measuring TSTT in future research or clinical applications. The third objective was to determine the intra and inter-rater reliability of measuring TSTT in different positions and by different devices. It was found that the DXA reliability is superior to US, and the likely cause of the increased error with US is due to landmarking the greater trochanter on the participant. Overall, the findings in this study should help future researchers, clinicians, or policy-makers select the most appropriate approach for measuring TSTT; many of the advantages and disadvantages of current measurement approaches are discussed below.

3.4.1 Concordance Validity between US and DXA

The ICC, bias of the differences and CR were 0.896, 0.46 cm and 2.15 cm, respectively. Similar results were produced when conducting the same analyses on the average measures, thus only the first measure analyses were presented. Larger differences between the two types of analyses were expected, as taking an average of three measures could have eliminated random error, but this was not the case.

When examining the Bland-Altman plots between US and DXA (Figure 3-5), it is evident that the agreement is poor. One of the main causes of the disagreement is the proportional bias. To provide context, a positive difference means that the US provided a larger TSTT than DXA, and vice-versa for a negative difference. A good portion of the differences was above 0 cm, so the bias between the two devices can be considered a systematic error. After reviewing the literature, the likely cause of the systematic error is due to the fan beam technology of the DXA scanner used in this study. Since the X-Ray beam is projected in a fan shape, there is an unavoidable magnification or demagnification of the width of the object of interest (Griffiths et al., 1997). When the object of interest is too close to the X-ray source, magnification occurs, and when the object is too far, demagnification occurs (Figure 3-14). Due to the posteroanterior scanning method in Hologic QDR machines (Pocock et al., 1997), it is more likely that the scanning object is further away from the source, explaining the decrease in TSTT when compared to US. Griffiths et al. (1997) demonstrated this experimentally by measuring the width of an aluminum phantom at different table heights (it was adjustable with the Lunar Expert) – they found that the predicted width changed from 7.61% to -0.84% by varying the table height from 2 cm to 22 cm. Also, the DXA scanner used in this study was a Hologic QDR 4500 Discovery, which has been reported to use a wide beam angle of 30° (Oldroyd et al., 2003). With a wider beam, it is more likely that demagnification occurs than magnification, explaining the smaller TSTT measurements when compared to US.

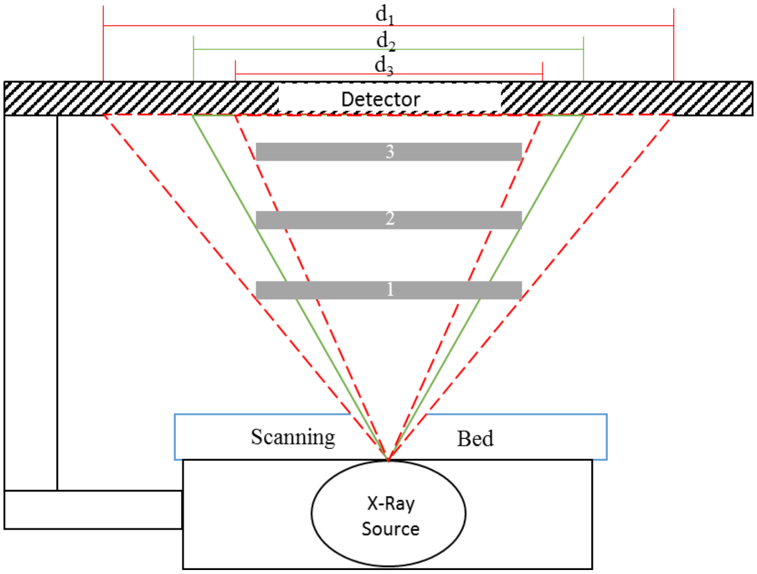


Figure 3-14. Examples of magnification errors with the fan beam DXA. Rectangle 1 is too close to the X-ray source, so the width is magnified (d_1). Rectangle 2 is in the optimal position, so the width is not magnified. Rectangle 3 is too far from the X-ray source, so the width is de-magnified.

The only study that measured width of an object with a fan beam DXA found that femoral neck axis length decreased by 11.4% (or a 1.9 SD) between the largest women and smallest women when using the fan beam DXA (Pocock et al., 1997). To determine the decrease in femoral neck axis length with an increase of distance from the X-ray source, the researchers compared the QDR 2000 fan beam DXA scanner to a Lunar DPXL pencil beam scanner (the pencil beam being the gold standard because of its linear measuring technique). The larger women had more soft tissue over their ischial tuberosity, explaining the decrease of femoral neck axis length as they were further away from the X-Ray source.

This is not the first study to make US and DXA TSTT measurements, as Maitland et al. (1993) found a strong correlation between the two imaging devices (r = 0.903). However, the researchers of this study compared US while standing with their DXA measurements. Therefore, I decided to determine the concordance validity between standing US and DXA (in case it was not clear, all of the analyses above were between the supine US and DXA). The standing+25 external rotation configuration was chosen, as the Maitland group had their participants stand with "uniform weight distribution", and people naturally stand externally rotated (see section 3.4.4). Interestingly, the agreement was better between the standing US and DXA than the supine US and DXA (Figure 3-15). The ICC was 0.92 versus 0.90, the CR was 1.75 versus 2.15, and most importantly, the bias was non-proportional, with the mean difference being -0.17 cm versus 0.46 cm. It seems as though the Hologic demagnification error balances out with the change of soft tissue spreading while standing. This finding is extremely important as DXA is believed to overestimate TSTT in the literature (Bouxsein et al., 2007; Levine et al., 2015; Maitland et al., 1993; Schacter and Leslie, 2014), but the results of this study challenge this belief. <u>However, this is only an applicable statement for Hologic fan beam DXAs.</u> For example, Lunar DPXL has been shown magnify the width of objects because of its anteroposterior scanning (Pocock et al., 1997). Accordingly, the bias of fan-beam DXAs will be dependent on the manufacturer, so TSTT could be underestimated or overestimated.

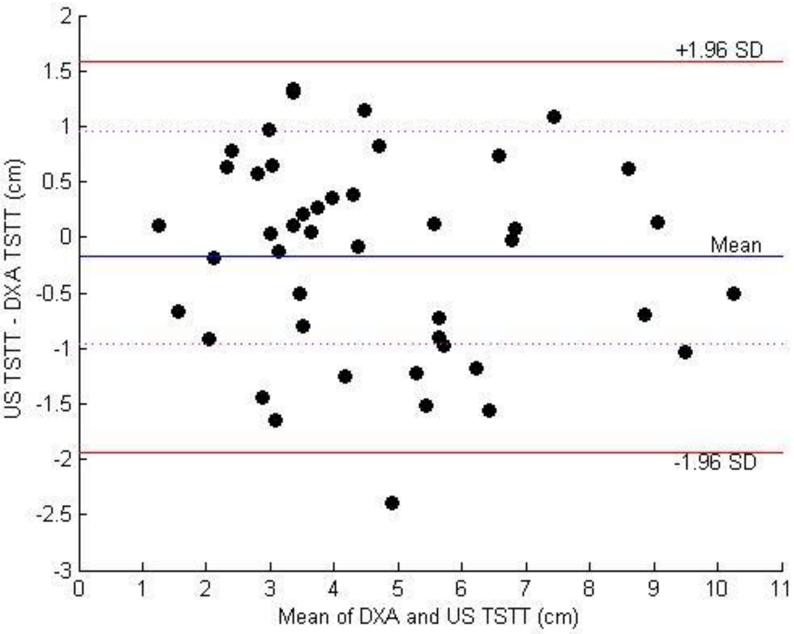


Figure 3-15. Bland Altman plot for the <u>standing</u> US and DXA measurements. The mean bias is -0.17 cm (blue line) and the CR is 1.75 cm. 68.9% of the differences were within the 0.96 cm bounds.

3.4.2 TSTT Interaction between Position and Rotation

Body position and hip rotation combined had an influence on the TSTT value. When the participant was in a position associated with a greater TSTT, externally rotating the hip had a much larger effect of increasing TSTT than in a position associated with a smaller TSTT. To quantify, the differences between 25° internal rotation and 25° external rotation between standing, supine, and side-lying were 0.45 cm, 0.75 cm, and 0.37 cm. This interaction can be explained by the anatomy of the hip. When the hip is internally rotated, the greater trochanter is pointed more anteriorly, and when the hip is externally rotated, the greater trochanter is pointed more posteriorly. During a standing or side-lying position, the gluteal muscles (specifically gluteus maximus) does not get compressed, but it does get compressed in a supine position (Bouxsein et al., 2007; Maitland et al., 1993). Accordingly, when the hip is externally rotated while lying on your back, the greater trochanter will be more tucked behind the "pushed-up" gluteus maximus, but this would not be the case while side-lying because the gluteus maximus is being pulled away medially by gravity. This explains the large rotation differences in the supine position and the equal differences in the side-lying position. In side-lying, the difference can be solely explained by tendon "slack", which will be discussed in the next section. In the standing position, the greater trochanter will point towards the gluteus maximus more than in a side-lying position, but not as much as the supine position. The results of this section demonstrate the importance of standardizing hip rotation in a supine position, like during a DXA or CT scan, when obtaining TSTT.

3.4.3 Main effect of Body Position on TSTT

The mean (SD) TSTT in the standing position was 4.16 (2.1) cm, which is much greater than what was found in the literature. For example, Robinovitch et al. (1991) reported TSTT values of 2.61 (1.27) cm, Minns et al. (2007) reported mean values of 1.81 cm and 2.79 cm for fracture cases and controls and Levine et al. (2015) reported TSTT of 2.81 (0.97). These differences can be explained by the different populations used: Minns et al. (2007) measured hip fracture or hip/knee replacement patients, whereas Robinovitch et al. (1991) and Levine et al. (2015) measured younger adults. Maitland et al. (1993) is the only study that measured community-dwelling older adults like this study, and although they do not report their mean TSTT, the range of 1.5 cm to 8.5cm is similar to this one (1.03 to 9.8 cm). The mean (SD) TSTT in the side-lying position was 3.29 (1.7) cm, which is equivalent to Choi et al. (2015) results of 3.21 (0.72) cm in younger adults and 3.04 (1.5) cm in older adults. To the best of my knowledge, no study has measured TSTT in the supine position with US, so there is no comparison that can be made.

Nevertheless, the supine US TSTT was 5.57 (2.8) cm, which is greater than any mean reported in the literature (Table 2-1).

It came to no surprise that there were significant differences of TSTT between positions, yet the magnitude of the differences was unexpected. Accordingly, the magnitude of the differences should warrant using a standardized position for measuring TSTT in future research or practice. The differences between positions can be explained by the effect of gravitational and normal forces. In the side-lying position, the soft tissue is being pulled medially, so it mostly parts at the greater trochanter. While supine, the gluteal soft tissue gets compressed by the normal force from the table, and it spreads laterally on top of the greater trochanter. Standing TSTT is in the middle because the soft tissue is not being parted nor compressed in this position. Although it's only speculation, it is likely that these effects are amplified by age. It is well known that soft tissue is less stiff in older adults (Choi et al., 2010), so these effects may not be present in younger adults.

Figure 3-16 demonstrates the relationship between different positions, and regression equations are provided to convert TSTT from one position to another. Note that the mean of every measurement made by BL was used for this figure, and not solely the first measurement. With reference to the standing (ST) position, supine (SUP) TSTT is approximately 1.3 times greater, whereas side-lying (SID) TSTT is about 0.75 times smaller. Side-lying TSTT is approximately 0.6 the thickness of supine TSTT. The regression equations are provided in Figure 3-16 for anyone who would like to convert TSTT with a good amount of accuracy (≈ 90% of the variance explained). The regression was set to have a (0,0) intercept – the models without this intercept constraint only explained 0.25% more variance than without. These equations can be used by future researchers if they ever need to convert TSTT across positions.

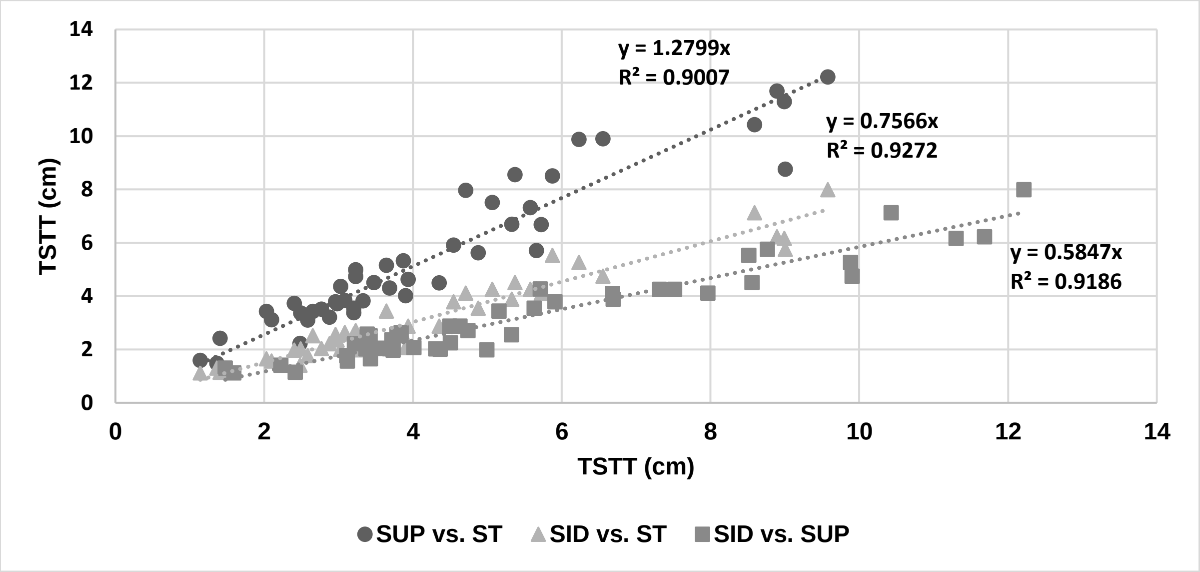


Figure 3-16. Scatter plot of all the TSTT measurements made in the different positions. The legend indicates the particular relationship.

3.4.4 Main effect of Hip Rotation on TSTT

While the effect was not as pronounced as body position, there was a main effect of hip rotation on TSTT. Although a direct comparison cannot be made, a previous study has demonstrated that postero-lateral falls (i.e. falls where the hip is impacted internally rotated) produce slightly higher impact forces than lateral falls (Nankaku et al., 2005). This coincides with the results of this study, as TSTT is lower when the hip is internally rotated.

The biggest difference was between 25° internal and external rotation, and the average difference was 0.522 cm. When examining the data, some differences exceeded the theoretical 0.96 cm threshold, which should be considered biologically significant (Appendix B). Levine et al. (2015) also demonstrated significant changes when different positions were used; compared to quiet standing, extending the hip by 30° increased TSTT by 0.671 cm, and flexing the hip by 60° increased TSTT by 0.385 cm. The results from our laboratory should warrant controlling the posture of the hip in the future. To do so, future researchers or clinicians should use an immobilization device like the foam triangle used in this study, or the triangle provided with DXA scanners. In this study, the mean (SD) natural hip rotation during standing was 19.2° (4.8°) externally, and it ranged from 9.5° to 29.5°. Accordingly, using one's natural posture is unwise for TSTT measurements as it varies from person to person. If a rotation had to be recommended, it would be to use internal rotation as it leads to a smaller TSTT (“worst case scenario”), and it is the position used when measuring hip BMD in a DXA scanner. However, participants in this study mentioned that it was the most uncomfortable out of the three rotations, so that can be an issue. For example, while standing internally rotated, the participant is less stable, so something like a chair or table should be provided to help maintain balance.

The TSTT differences between hip rotations can be explained by the lateral hip anatomy. The greater trochanter serves as attachment point for many muscles. They include gluteus medius, gluteus minimus, and the six external rotators of the hip: piriformis, obturator internus, obturator externus, superior gemelli, inferior gemelli, and the quadratus femoris (Guay, 2005; Marieb and Hoehn, 2010; Moore et al., 2010). The gluteus maximus and tensor of fascia lata attach to the iliotibial tract, which also rests on top of the greater trochanter. Judging by the images taken in this thesis, it seems as though the tendons of these muscles are closer to the greater trochanter when internally rotated, and deviate from the greater trochanter when externally rotated (Figure 3-17). In other words, the tendons are tight to the greater trochanter when internally rotated, and then the tendons have more slack when externally rotated (Figure 3-17). This “tendon slack” pushes all of the superficial structure outwards, creating a larger TSTT. Based on consultations with two senior anatomy demonstrators, I believe that the tendon is the gluteus medius or minimus tendon. When the participant was internally rotated, their internal rotators (i.e. gluteus medius and minimus) were activated. When the muscles are activated, tension is created at the tendon, thus being tight with the greater trochanter. When the muscles were inactive during external rotation, the tendon became slack, thus pushing superficial structures outwards. Videos of internal-external hip rotation were taken to also demonstrate this point. Hip rotation effects were also evident without the ultrasound. Some participants had an indent when internally rotated, and then this indent would be filled out once externally rotated (Figure 3-18). It is important to note that Figure 3-17 and Figure 3-18 are not the same participants.

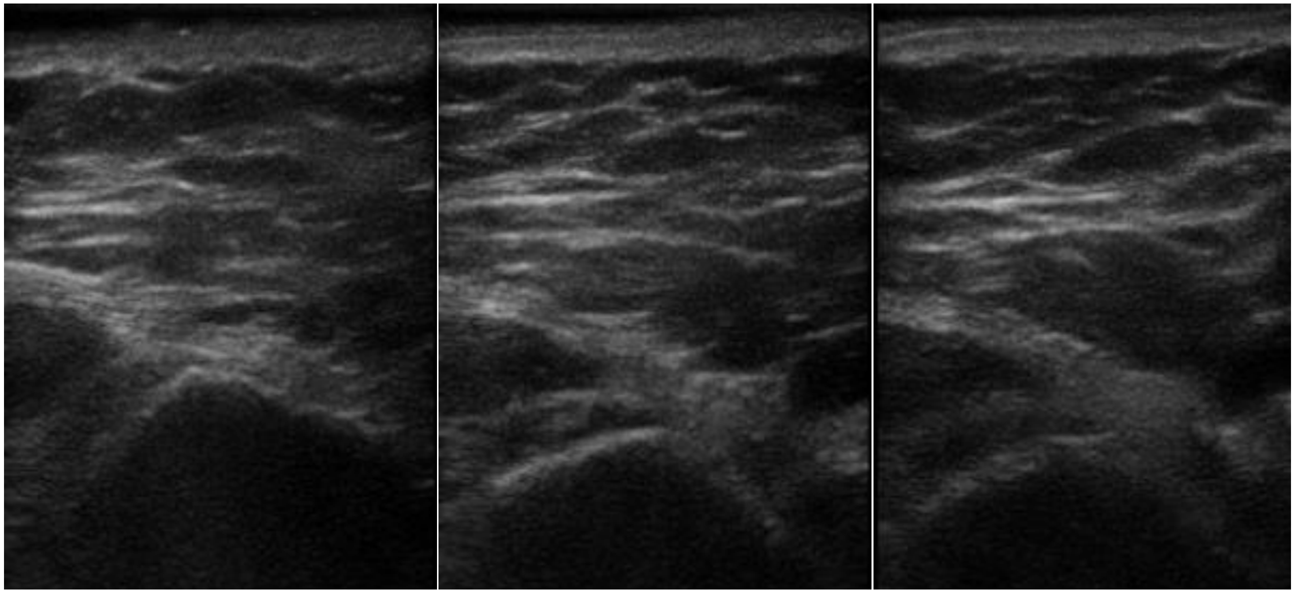


Figure 3-17. An ultrasound image of 25° internal rotation (left), no rotation (middle), and 25° external rotation (right). By looking at the left side of the images, the tendon deviates from the greater trochanter as the participant externally rotates. The left image TSTT is 3.02 cm, the middle image is 3.61 cm, and the right image is 3.70 cm.

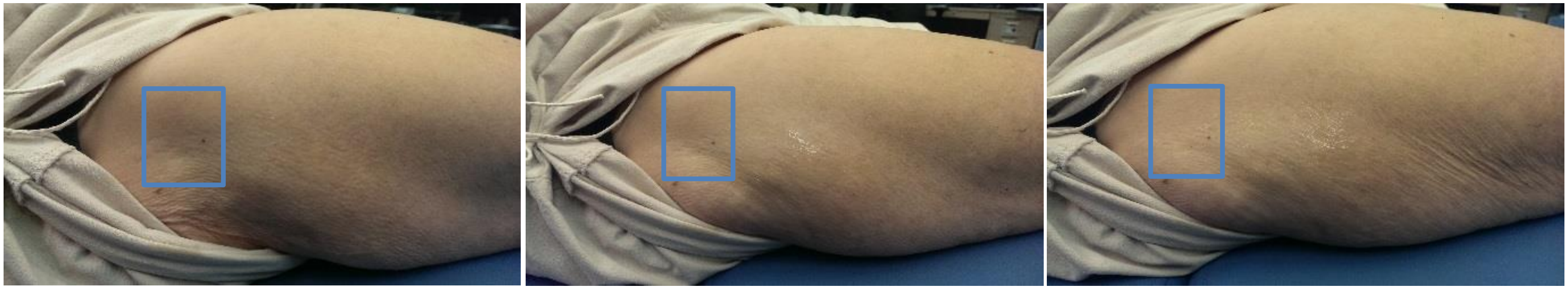


Figure 3-18. An example of the hip rotation effect on TSTT. In the rectangles is the approximate location of the greater trochanter. When the participant is internally rotated (left) there is an indent. This indent starts to fill out during a neutral rotation (middle) and disappears during external rotation (right).

3.4.5 US Intra-Rater Reliability

US intra-rater reliability was good in the standing and side-lying position, but mediocre in the supine position. The ICCs within the primary investigator (BL) were very strong: 0.980 for the standing position, 0.972 for the supine position, and 0.977 for the side-lying position. These values correspond to the only other study to report intra-rater reliability of ultrasound measurements of TSTT (ICC of 0.98 while standing, Levine et al., 2015). The Bland-Altman plot also shows relatively good agreement between the two measures taken by BL. The biases were all very small within BL (-0.07 cm for standing, -0.04 cm for supine, and -0.02 cm for side-lying), meaning that there was no systematic error. Also, a proportional bias was tested using a regression in the BA plot, and it was found to be non-significant. The CR of the standing and side-lying position are below 0.96 cm (0.87 cm and 0.69 cm respectively), but the CR of the supine position is greater than 0.96 cm with 1.32 cm. The reason that the supine position is less repeatable is likely because of the greater amounts of TSTT in this position. When there is a greater amount of TSTT, it is more difficult to get a good echo off the greater trochanter while ensuring that the soft tissue does not get compressed. Moreover, is it more difficult to palpate the greater trochanter when there is more TSTT, so the measurer has to rely heavily on the ultrasound image to confirm the greater trochanter landmark (and not the femoral shaft). Also, the resolution decreases when TSTT is above 10 cm with the ultrasound device: it changes to 0.1 cm instead of 0.01 cm. Lastly, it is more difficult to distinguish landmarks (e.g. non-circular bone, gluteus minimus or medius tendons), making it more difficult to find the greater trochanter in the image. Examples of these issues are presented in Figure 3-19.

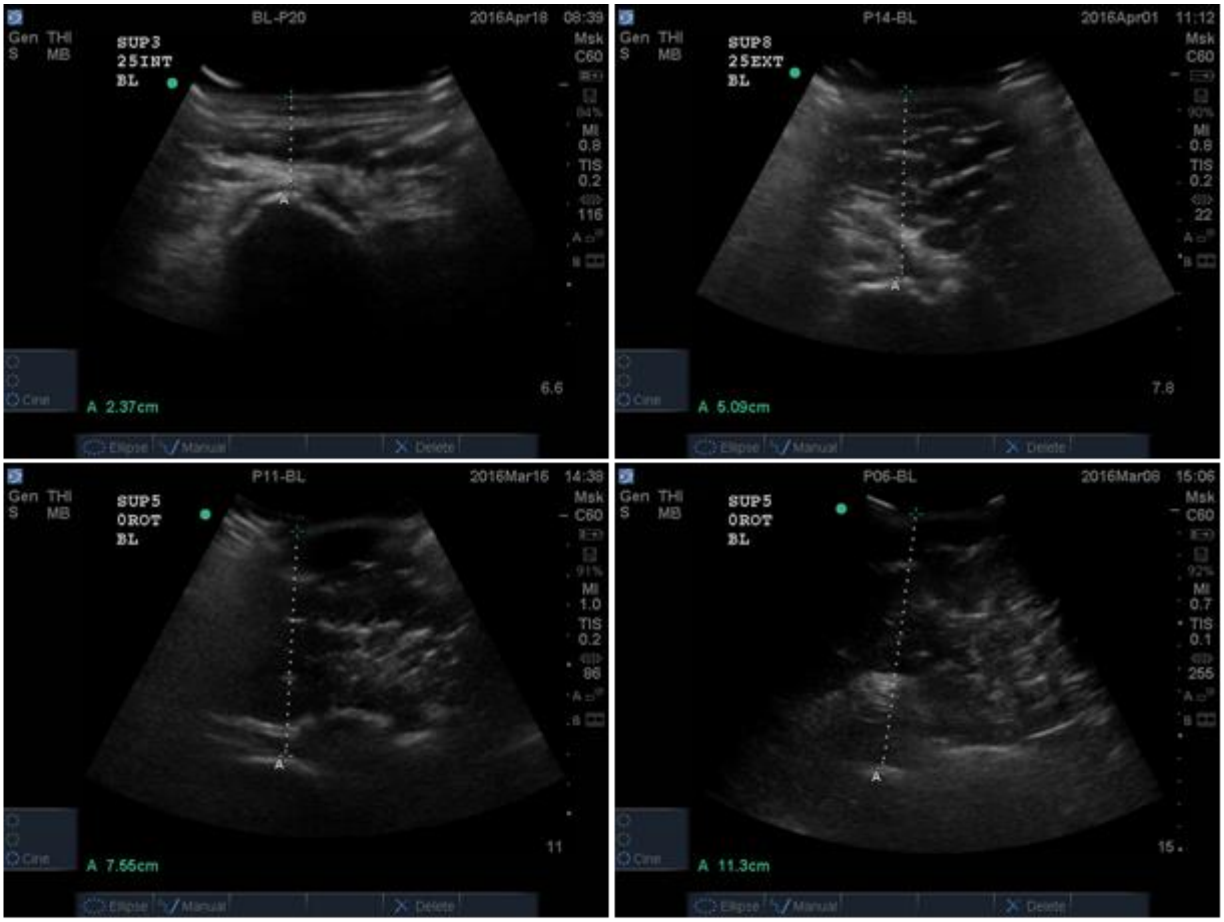


Figure 3-19. Random supine pictures of smaller TSTT (top left) to larger TSTT (bottom right). Notice that landmarks become increasingly difficult to identify as TSTT increases.

This decrease in reliability with an increase of TSTT can be easily identified in the Bland-Altman plot, as every measure above 4 cm typically has a larger difference than those below 4cm. To further this point, Figure 3-20 is a custom Bland-Altman plot that differentially illustrates the bias and CR measurements that fall below and above the median TSTT value of 3.9 cm. Specifically, for TSTT values below the median, the bias and CR are 0.1 cm and 0.71cm compared to bias and CR of -0.17 and 1.70 for TSTT values greater than the median. This indicates that the US protocol is much more reliable when TSTT is smaller, explaining why the standing and side-lying measurements had a better agreement. Although this custom Bland-Altman analysis was not presented for the standing or sideline results, the same trend occurred. This is an important finding because hip fracture patients usually have lower amounts of TSTT (approximately 3 cm for males, and approximately 4 cm for females; Table 2-1), so TSTT of this population could be more reliably measured by US.

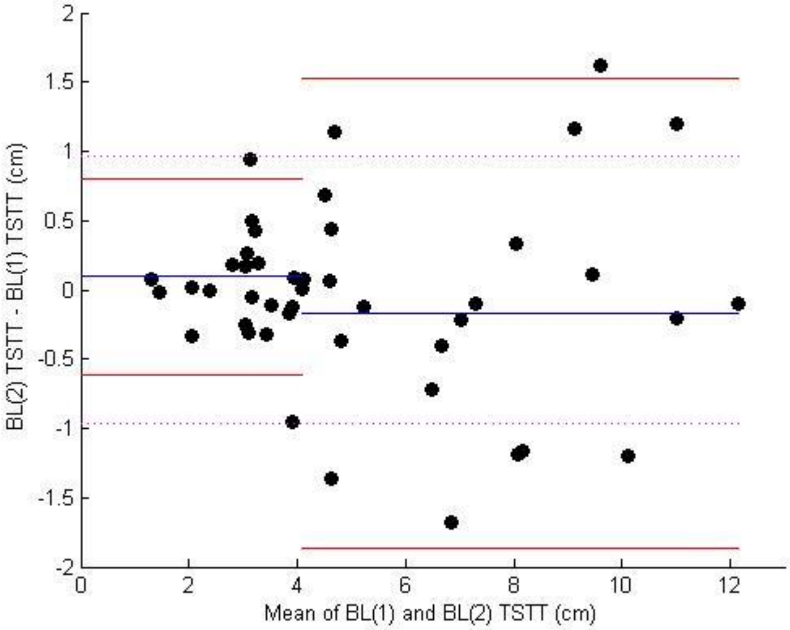


Figure 3-20. A customized Bland-Altman plot of the supine US TSTT within BL. The mean bias and CR for the lower half of TSTT was 0.1 cm and 0.71 cm, and the mean bias and CR for the upper half of TSTT was -0.17 cm (blue line) and 1.70 cm (red lines). For the first half, 100% of the measurements was within the 0.96cm bounds, whereas 61% of the measurements was within 0.96cm bounds for the second half. The split was based on the median TSTT value of 3.9 cm.

To provide another perspective, 93%, 80%, and 98% of the differences were within the 0.96 cm bounds for the within standing, supine, and side-lying US measurements. With the exception of the supine measurements, this indicates that it is highly unlikely to make a US measurement that will be mistaken as a potential non-fracture case versus a fracture case (as defined by the clinical threshold of 0.96 cm).

3.4.6 US Inter-Rater Reliability

The US inter-rater reliability is mediocre in the standing and sideline position, but poor in the supine position. The ICCs between BL and AT were 0.969, 0.939, and 0.948 for the standing, supine, and side-lying measurements, respectively. These values are quite high, meaning that there is a very strong agreement between the two raters. However, the Bland-Altman does not demonstrate the same type of agreement. There were minor biases of -0.08 cm (standing), -0.19 cm (supine), and 0.11 cm (side-lying) between the BL and AT. A proportional bias was tested using a regression in the BA plot, and it was found to be non-significant. To provide context, a negative bias means that AT had a smaller measurement than BL, whereas a positive bias means that AT had a larger measurement that BL. There is no explicit explanation for the bias of each position but is most likely explained by the different landmarking of the greater trochanter. The investigators were likely to always image the same point on the greater trochanter, which could be different between raters. The greater trochanter is a relatively large bony landmark, so it is likely to have different thicknesses even if it is being imaged. For example, one investigator may image the greater trochanter more superiorly (which would give a smaller TSTT), and the other more inferiorly (which would give a larger TSTT). Second, one of the investigators could have compressed the TSTT without noticing it on the ultrasound image. This is especially true with the curvilinear probe, as the skin-air interface can be difficult to identify with the cone-shaped field of view. Third, the cone-shaped field of view makes it difficult to determine what the smallest TSTT is, particularly when the greater trochanter is not in the middle. Although the curvilinear transducer is not ideal for TSTT measurements, it did demonstrate good concordance validity with the linear transducer (Appendix A). Therefore, the difference between BL and AT is likely explained by different landmarking on the participant. The bias did not exceed 0.2 cm, nor were the differences consistently positive or negative, so it is unlikely that there is a systematic error between the investigators. Yet, these possible sources of error can explain why the CR between BL and AT was above 0.96 cm; 1.07cm for the standing position, 2.08 cm for the supine position, and 1.10 cm for the side-lying position. As explained in section 3.4.1, the supine position is less repeatable because of the greater amount of TSTT produced by the position. Also similar to the intra-rater reliability, the agreement between the two raters decreased at TSTT increased. The aforementioned custom Bland-Altman analysis was executed for each position, and although the agreement was better in the first half of the measures than the second half, only the first-half CR for the side-lying measurement was under 0.96 cm. Still, this outcome indicates that a low-TSTT/high-fracture-risk population can be reliably measured between two people in the side-lying position.

There is some value in mentioning the percentage of differences that were within the 0.96 cm bounds: 93% for standing TSTT, 73% for supine TSTT, and 93% for the side-lying TSTT. With such high percentages for the standing and side-lying TSTT, it is somewhat surprising that the CR is not below 0.96cm. When examining their respective Bland-Altman plots (Figure 3-10), it can be seen that a few outliers are stretching out the 95% confidence intervals. Without these outliers, the CRs are likely to be under 0.96 cm, explaining the different interpretation of reliability between these two positions and the supine position. Nevertheless, future researchers or clinicians that would like to measure TSTT with ultrasound should undergo training in order to minimize measurement error. Expert sonographers would probably have better inter-rater reliability than BL and AT, but the results of this study can be generalized to those with basic ultrasound training and a good understanding of human anatomy.

3.4.7 US Image Inter-Day Reliability

The inter-day image reliability within the supine images of BL was excellent. The supine position was chosen because it was the position with the most difficult images to determine TSTT. Thusly, the inter-day image reliability results are expected to be the same with regards to the standing and side-lying position. As mentioned, the greater trochanter and skin-air interface are more difficult to landmark in participants with greater amounts of TSTT, but it was still identifiable as the bias does not seem to increase in the Bland-Altman plot (Figure 3-11). Knowing that the inter-day image reliability was very good, it can be determined that the landmarking of the greater trochanter on the person is the likely cause of the mediocre “full” US intra-rater reliability. It also reinforces the idea that the inter-rater reliability was reduced because of different landmarking on the person. It is important to note that the re-analysis was not possible with the US device, so the ImageJ program had to be used.

3.4.8 DXA Intra-Rater Reliability

The intra-rater reliability for the DXA analyses of TSTT was excellent. This is quite remarkable as the resolution is much lower in the DXA scans when compared to the ultrasound. The smallest increment of change in the DXA analyses is 0.1 cm, whereas the smallest increment of change is 0.01 cm (for TSTT below 10 cm; it is 0.1 cm for TSTT above 10 cm). Accordingly, there is less room for error in the DXA analyses, as being “one pixel” off will result in a greater change.

Technically speaking, it should be said that the inter-day image reliability is very good with the DXA analyses, as a *true* intra-rater reliability protocol would be to compare two DXA scans that were made on the same person at a different time. A comparison between these two types of reliability was only found in one study, and they had an intra-analyzer CV of 2.6%, and an inter-scan CV of 6.4% (Nielson et al., 2009). This type of protocol could provide different results, as any change in the positioning of the participant (which was done by the Medical Radiation Technologist), or any movement during the scan can influence the TSTT analysis. Older adults can lie down on their backs without moving, so this should not influence the reliability results heavily. Even if they did move, the movement artifact can usually be identified and avoided in the DXA scan. If there were to be differences between the reliability of this study and an inter-scan reliability, it would likely be explained by differences in positioning. As mentioned, rotation of the hip joint in any direction will affect TSTT, so it's important that the leg is always positioned the same way. However, this can be easily achieved by doing "quality checks" before starting the scan. Several studies have shown that the precision of DXA scanners is good for body composition analyses (Bonnick, 2004; Toombs et al., 2012), so the reliability results of this study should be similar to *true* reliability results.

The DXA TSTT reported in this study is similar to the literature. The mean (SD) DXA TSTT of the male participants was 3.16 (1.1) cm, which is nearly identical to Nielson et al. (2009) non-fracture group value of 3.10 (1.2) cm. DXA TSTT of the female participants was 6.08 (2.2) cm, which is slightly greater than Bouxsein et al. (2007) control group value of 4.98 (1.7) cm. Schacter and Leslie (2014) reported a DXA TSTT value of 4.9 (2.3) cm, which is very close to this study's grand mean of 4.78 (2.3) cm. Also, the mean (and range) CV for the DXA measurements was 3.3 (0 – 16.6) %. Although studies have used different methods to report CV (e.g. repeat the analysis nine times in one participant, or have three repeated measurements in twenty participants), the resulting values are similar to the one in this study. Specifically, Maitland et al. (1993) calculated a CV of 3.7%, Bouxsein et al. (2007) calculated a CV of 3.9%, and Nielson et al. (2009) calculated a CV of 2.6%. This resemblance of mean TSTT and CV with the other studies confirms the integrity of all of the DXA measurements.

The DXA TSTT protocol is highly reliable, but it is definitely not perfect as some issues arose during the analyses. Movement artifacts were the most common issue, and examples are demonstrated in Figure 3-21. Also, some participants were too big for the DXA table, so their hands had to be tucked too closely to their hips (Figure 3-21), or two half-scans had to be taken. The hand does not have to be near the hip or two scans does not have to be taken for the TSTT analysis, but it would be if the researcher is interested in body composition results (the whole-body scans were shared with another researcher, hence why this was done). Future researchers should be wary of these issues so that it does not affect their DXA TSTT analyses.

Figure 3-21. Examples of movement artifact are demonstrated in the three left images. The right image demonstrates how the hand can interfere with the TSTT analysis.

3.4.9 DXA Inter-Rater Reliability

As with the intra-rater reliability, the DXA inter-rater reliability was excellent. Oddly enough, the bias and CR were better between BL and AT than within BL, but the differences were minimal. The mean inter-observer CV of 2.6 % is similar to the 3.5% found by Schacter and Leslie (2014), thus demonstrating the integrity of the results.

An important aspect to mention about the DXA inter-rater reliability is that minimal training was performed before doing the analyses. Similar to the US image inter-day reliability (i.e. measuring TSTT with an existing image), anyone with basic knowledge of human anatomy should be able to measure TSTT from a DXA scan. Each rater analyzed 5 random scans as practice before conducting the actual analyses. This is useful information for future studies that may not necessarily have the time to do training with the ultrasound device, or for studies that would like the analysis done by several different people.

3.4.10 Limitations

Despite the care and novelty of the study presented, there are some important limitations that need to be acknowledged. First, the researchers who measured TSTT with the ultrasound device were not expert sonographers. AT had experience using an ultrasound device in a veterinary setting, but her expertise is still considered minimal compared to the training received by certified imaging professionals. The reliability of measuring TSTT is expected to be greater for expert sonographers.

Second, the linear US transducer should have been used instead of the curvilinear transducer. In hindsight, I believe that better measurements would have been taken with the linear transducer, as the image landmarks are easier to identify. This was the initial plan, but after collecting the first two participants with the linear transducer, it was clear that bigger participants would have a TSTT that exceeded the depth of the linear transducer. At the time it was not known if there would be systematic differences between the linear and curvilinear transducers, so it seemed unwise to use two different transducers in a reliability study. Therefore, the curvilinear transducer was used, as it can be used in all participants. After the collections, it was found that the agreement between the linear and curvilinear transducer was very good for TSTT measurements under 8 cm, so the transducers could have been used interchangeably without major repercussions. Since there was good agreement, it is unlikely that the results would be critically influenced if the linear transducer were to be used (Appendix A). Although it is not as relevant to this thesis, using the curvilinear transducer compromised the ability to differentiate tissue types in the image.

Third, there could be researcher bias in the US reliability results as the researchers were always in the same room with the participant. The one investigator could examine the technique used to locate the greater trochanter on the person, thus influencing their approach. This is especially true as the investigators knew that their reliability was being tested. Ideally, the investigator not conducting the measurement should not be in the room. However, they needed to be in the room to make the collection quicker, and to help with other tasks. Fortunately, these safeguards were taken to minimize this effect:

- A sticky note was always placed on the ultrasound screen, so the thickness could never be seen by any investigator during the collection.
- After obtaining TSTT, the gel was completely wiped off the thigh, and the depth of the US transducer was changed to the default setting (i.e. 9.2 cm), so it was like the participant was never measured
- For each position, the first three measures were taken by BL, then AT, then BL again. Accordingly, BL would walk away from the participant's right side, thus allowing a brief wash-out period.
- The investigator not taking a measurement would always tend to the participant by ensuring their comfort, conversing with them, and by holding their feet in the side-lying position. With these tasks, they were less likely to examine the other investigator.

Fourth, the foam triangle that was used to internally rotate the hip during the US analyses was not used during the DXA analyses. The whole body scans were being shared with another researcher, and the foam triangle would have compromised the body composition analyses. As a safeguard to this limitation, I checked that the hip was rotated 25° with a large protractor, and then taped the feet together to ensure the preservation of this joint rotation angle.

Fifth, there were no underweight (i.e. BMI < 20 kg/m^2) individuals recruited for this study. It would have been ideal to collect this population as they are at higher risk of hip fracture, and to increase the range of TSTT collected. However, there is potential value in improving hip fracture risk assessment in the medium risk population, who typically do not have low BMI. This group was well represented in the current study.

Chapter 4 Thesis Synthesis and Conclusion

4.1 Novel Contributions and Impact

There are several novel findings in this thesis study. This is the first study to determine the concordance validity between US and DXA measurements of TSTT. I found that the agreement between the two imaging devices is poor and that DXA underestimates TSTT by an average of 11.6% when compared to US. It was also found that the differences and bias were accentuated in people with a greater amount of TSTT (e.g. >4 cm). Interestingly, due to the DXA underestimation of TSTT, the US TSTT measured in the standing position had a better agreement than the US TSTT measured in the supine position. Therefore, it may be acceptable to compare studies that measured TSTT by using DXA or standing US. However, researchers and clinicians should still be aware of this notion, as the current consensus in the literature is that DXA overestimates TSTT because of the compressed gluteal soft tissue.

Second, this study found a main effect of body position and hip rotation on TSTT, along with a significant interaction. This emphasizes the importance of controlling for posture as TSTT can easily be changed. For example, in some of the bigger participants, TSTT doubled between the side-lying and supine positions. Between 25° of internal and external rotation, the TSTT in some participants changed by more than 1 cm. This type of dramatic change can severely influence the estimation of force applied to the hip, or the assessment of hip fracture risk. For example, when calculating the Factor of Risk of the eleventh participant (see section 2.3.3), she was considered not at risk in the supine position ($\Phi = 0.66$), almost at risk in the standing position ($\Phi = 0.93$), and at risk in the side-lying position ($\Phi = 1.07$). Accordingly, future practice should use a standardized approach when measuring TSTT, similar to the DXA hip scan where the leg is immobilized. Lastly, it would be unwise to compare studies that measured TSTT in different position (with the exception of DXA and standing TSTT, see above). For those who would like to make these comparisons, equations are provided in this thesis (see section 3.4.3) that can convert between TSTT values in different positions.

Third, the reliability of several TSTT measurement approaches was assessed extensively in this study, and DXA was found to be the best. In contrast, the intra-rater reliability for the US analysis was good in the standing and side-lying position, but mediocre in the supine position. The US inter-rater reliability was mediocre in the side-lying position, and the standing position, but poor in the supine position. Knowing that the image inter-day reliability was very good in the supine position, it can be presumed that landmarking of the greater trochanter was the likely cause of the sub-par reliability results. It is also important to note that all types of reliability were much better in participants with less than 4cm of TSTT. This is a positive finding from a clinical perspective as those at the highest risk of a hip fracture are usually lean and frail with low TSTT. It is also important to mention that the reliability assessment was novel in itself. According to a 2012 systematic review of ultrasound reliability, none of the 24 studies reviewed included re-landmarking as a potential source of error in their protocols / analyses (English et al., 2012). Furthermore, these studies only reported ICCs or Bland-Altman plots without the use of pre-defined clinically acceptable difference thresholds. If this thesis study had employed the approaches commonly used in literature, the reliability results would have appeared substantially more promising. The current results suggest that a more robust picture of reliability can be ascertained by utilizing the assessment approaches used in this thesis.

In the literature review, it was mentioned that TSTT is a promising variable that can improve hip fracture risk assessment, or more specifically, improve the assessment in the “medium risk” population. To provide an example, two participants similar in anthropometrics were chosen: P07 is a healthy community-dwelling woman, 67 years old, has a BMI of 26.6, and a femoral neck BMD T-Score of -1.1; P14 is a healthy community-dwelling woman, 67 years old, has a BMI of 25.6, and a femoral neck BMD T-Score of -0.2. When their information was inputted into FRAX, their 10-year major osteoporotic and hip fracture risk was similar - 7.4% and 0.7% for P07, and 6.3% and 0.3% for P14. With these percentages, they are both nearing the medium risk category according to the clinical guidelines, and an intervention is not required. When the Factor of Risk is calculated using DXA TSTT for both of these participants (see section 2.3.3), P07 is not considered at risk with a ratio of 0.79, whereas P14 is considered at risk with a ratio of 1.06. This is because TSTT was vastly different between the two participants (6.8 cm for P07 versus 2.7 cm for P14). TSTT was different because of their anthropometrics: P07 has a waist to hip ratio of 0.78, whereas P14 has a waist to hip ratio of 0.91. This type of information is not captured in the assessment made by FRAX even though it can be extremely useful for the clinician or individual. Much more research is required, but I would hypothesize that adding TSTT to the current clinical risk factors could significantly improve hip fracture risk assessment. If it is eventually shown to improve hip fracture risk, this thesis will be very impactful as it provides important information and recommendations with regards to the measurement of TSTT.

4.2 Future Research

There is much more research that needs to be done with regards to TSTT. From a basic science standpoint, the attenuation properties of TSTT are still not well defined. Currently, it is very difficult to estimate the attenuation that occurs when the hip impacts the ground. The composition of TSTT can probably provide insight in force attenuation (e.g. fat/muscle ratio), but this has never been explored. Additionally, the composition of TSTT has never been characterized in different populations like men versus women, or controls versus fracture patients. There may be more than the thickness that can be used to estimate soft tissue attenuation in TSTT measurements. Although TSTT is usually associated with impact force, it may provide insight into balance control. For example, Addison et al. (2014) demonstrated that older adults with greater intramuscular adipose tissue in their gluteus medius and minimus had greater variability in their gait and poorer balance. This type of information could get captured while measuring TSTT with ultrasound, thus improving the risk of hip fracture assessment. There may be a similar type of relationship of balance and TSTT with body composition measures of the proximal thigh.

In this thesis, a “gold standard” approach of measuring TSTT was not determined as it is not known what TSTT would best assist in predicting hip fractures. Accordingly, a longitudinal or cohort study could test which measurement approach of TSTT (i.e. ultrasound in what position, or DXA) best predicts fractures. The best way to measure TSTT would probably be in the position taken right before a hip impact, but that would be very difficult to do in older adults. Still, an interesting future study should characterize the TSTT right before impact, and compare it to the TSTT of current approaches used in this thesis. The TSTT that has the best correlation with the “pre-impact” TSTT would probably be considered the “gold standard” as it is the best representative of the hip fracture scenario.

4.3 Conclusion

The findings of this thesis demonstrate the importance of using a standardized approach for measuring TSTT, as factors including imaging device, body position and hip rotation significantly influenced the measured value. It is difficult to choose what method to standardize as it is not known what TSTT best predicts hip fractures. As demonstrated by Table 4-1, many approaches are reliable, quick, and comfortable for participants, so it is up to the measurer to determine what set of advantages and disadvantages best suits their needs. However, I would choose the US side-lying approach for basic science studies, and the DXA approach for clinical studies. My reasoning for side-lying ultrasound is that it has better resolution, and it can provide better composition information needed to fully understand the attenuation properties, thus being ideal for studies focused on the impact mechanics. My reasoning for DXA is that it can be used to collect BMD simultaneously with TSTT, making it a very useful tool for clinical studies. Overall, information on reliability, time required to measure TSTT, feasibility, and participant comfort are presented in Table 4-1. This information may assist researchers, clinicians, policy-makers and those who play a role in hip fracture assessment in deciding which measurement approach is most appropriate for future applications.

Table 4-1. Information of different TSTT measurement approaches

Measurement Approach	Reliability	Total Duration (from entering room to obtaining TSTT)	Feasibility	Comfort	Other
Standing US	Good intra-rater and mediocre inter-rater reliability.	10 to 20 minutes	Older adults may have difficulties preserving their balance. Would need a chair or table in front	Comfortable if the participant can tolerate prolonged standing	Produces a TSTT that is smaller than supine, but larger than side-lying
Supine US	Mediocre intra-rater and poor inter-rater reliability. The least reliable.	10 to 25 minutes	Very feasible	Most comfortable	Produces the largest TSTT
Side-lying US	Good intra-rater and mediocre inter-rater reliability	8 to 15 minutes. Side-lying typically takes the least amount of time	With a triangle between the feet, it is difficult for the individual to stay upright, Would need a second person to hold the individual	Uncomfortable with the triangle between the feet	Produces the smallest TSTT
DXA	Excellent intra and inter-rater reliability. The most reliable, even with a decreased resolution	10 to 25 minutes. DXA typically takes the most amount of time	DXA is not portable like ultrasound, so the scan would have to be conducted in a specific setting	Very comfortable	BMD can be simultaneously collected with TSTT, but it exposes the individual to low-dose radiation

Appendix A - Concordance Validity between an Ultrasound Linear Transducer and Curvilinear Transducer while Measuring Trochanteric Soft Tissue Thickness

Introduction: To examine the lateral hip using ultrasound, the literature recommends the use of a 9 to 15 MHz for most patients, and a 5 MHz transducer for obese or muscular patients (Hill and Leiszler, 2013; O'Neill, 2008). However, a previous study that measured trochanteric soft tissue thickness (TSTT) in obese participants have found thicknesses that exceed the 5MHz penetration depth of 9 cm (Schacter and Leslie, 2014). Accordingly, a curvilinear transducer with a greater penetration depth of 9 cm would be needed to determine TSTT in these individuals. Although it is possible to measure a small amount of TSTT with a curvilinear transducer, it's convex shape can cause major distortion of superficial structures (e.g. the skin-air interface), making it less accurate than its linear counterpart. To the best of my knowledge, no study has determined the agreement between a linear transducer and a curvilinear transducer when measuring TSTT. It is not known whether the curvilinear transducer is an acceptable device to measure TSTT in the general population.

Purpose and Hypothesis: Determine the concordance validity between a 5–10 MHz linear transducer (Sonosite L38) and a 2–5 MHz curvilinear transducer (Sonosite C60x). The associated hypotheses are that the two-way random intraclass correlation will be strong (ICC (2,1) > 0.8), and that the coefficient of repeatability (CR) will be below a clinically significant value of 0.96 cm (which is the difference between hip fracture cases and non-fracture cases).

Methods: Forty-one participants were recruited for this study. Participants were between the age of 40 to 86 years, the BMI ranged from 20 to 40 kg/m^2, and the TSTT ranged from 1.09 to 8.24 cm. TSTT was measured once by both transducers while the participant was in the same side-lying position. The TSTT measurement was blinded by the researcher during the collection, and the greater trochanter had to be re-landmarked for each image. This type of protocol was used to simulate clinical practice.

Results: The ICC was 0.96 and the CR was 0.78 cm. The mean bias of the differences was -0.14 cm and 95.6% of the differences were within the 0.96 cm bounds (Figure A-1).

Discussion: The two hypotheses of this study have been accepted: the ICC 0.96 > 0.8 and the CR 0.78 cm < 0.96 cm. By examining the Bland-Altman plot, it is clear that the differences become larger as TSTT becomes larger. This is due to the difficulty of getting a good echo off the greater trochanter at greater depths. Also, there was a minimal bias of -0.14cm, meaning that the linear transducer usually measured a smaller TSTT than the curvilinear transducer. With the linear transducer, the skin-air interface is more clear, whereas there is more speckle pattern for the curvilinear transducer (i.e. scattering of the ultrasound echo). This speckle pattern is the likely cause of the greater TSTT measure in the curvilinear transducer. Approximately 96% of the differences were within the 0.96 cm bounds, meaning that only 4% of the differences were deemed unacceptable.

Conclusion: Although image quality is not as good as a linear transducer, the curvilinear transducer can be used to reliably measure small amounts of TSTT.

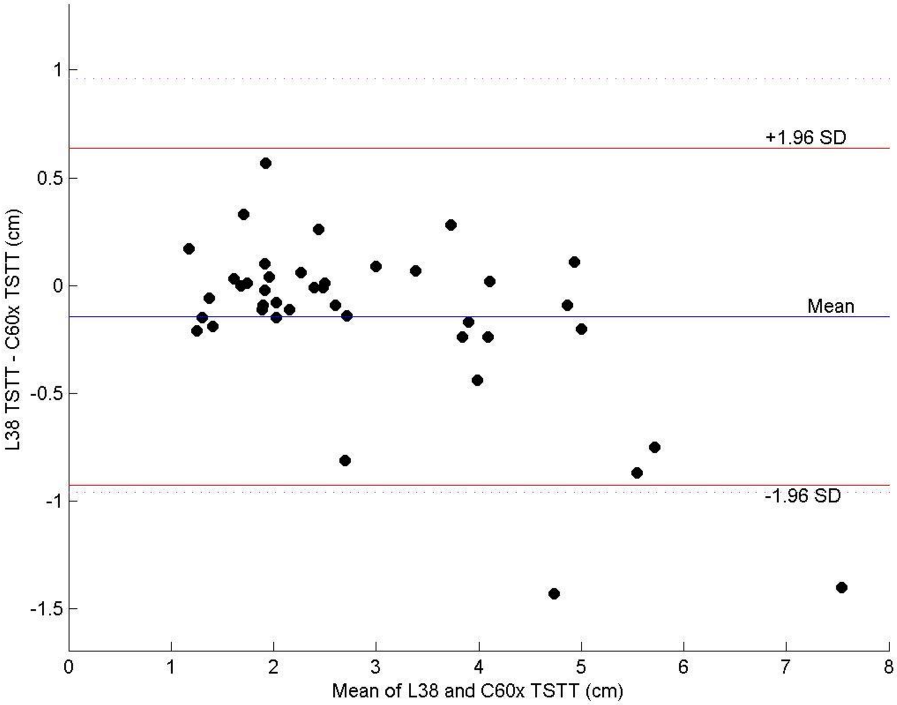


Figure A-1. Bland-Altman plot for TSTT measurements between L38 and C60x. The mean bias is -0.14 cm (blue line) and the CR is 0.78 cm (red lines). 95.6% of the differences were within the 0.96 cm bounds (pink lines).

Appendix B – Grand Table of Data

Table B-0-1. All of the standing and supine TSTT measurements. At the bottom of the table are the mean and SD. Yellow (light shading) represents females, whereas blue (dark shading) represents males. Orange text represents MP measurements

	MP					MP						
Standing TSTT (cm)					Supine TSTTT (cm)							
ST1, 25int, BL	ST2, 25int, AT	ST3, 25int, BL	ST4, 0rot, BL	ST5, 25ext, BL	SUP1, 25int, BL	SUP2, 25int, AT	SUP3, 25int, BL	SUP4, 25int, BL	SUP5, 0rot, BL	SUP6, 0rot, BL	SUP7, 25ext, BL	SUP8, 25ext, BL
1.47	2.37	2.94	3.12	3.08	3.17	2.77	2.92	3.02	3.61	3.7	3.78	3.76
6.07	5.47	5.62	5.58	5.62	6.87	7.36	6.47	6.44	6.14	6.35	7.17	7.34
8.53	7.85	8.69	8.18	8.96	10.7	10.3	9.5	10.7	10.5	10.6	10.5	10.5
5.15	5.39	5.37	5.33	5.64	7.89	8.07	8.22	8.24	9.35	8.22	9.14	8.85
9.8	9.53	9.34	9.18	9.99	12.2	12.7	12.1	12.4	11.9	11.8	12.2	12.9
9.16	9.92	9.65	8.65	8.5	11.1	11.6	10.9	11.4	11.3	11.7	11.4	11.3
5.41	5.96	5.63	5.65	5.62	7.14	7.8	6.92	7.12	7.16	7.25	7.6	8.01
2.87	2.6	2.62	2.91	3.1	3.18	2.95	3.13	3.09	3.12	3.25	3.38	3.33
4.22	4.16	4.1	4.6	4.49	5.32	5.64	3.96	3.93	4.24	4.31	4.94	4.76
5.63	6.5	5.67	5.43	6.77	8.75	6.85	7.59	8.42	8.14	8.69	8.96	9.05
5.2	4.48	4.91	4.94	5.22	7.35	7.49	7.25	6.93	7.55	7.09	8.09	8.35
1.38	1.17	1.44	1.33	1.31	1.25	1.27	1.33	1.44	1.52	1.44	1.64	1.72
3.64	3.33	3.35	3.6	4.15	3.94	4.05	3.77	4.29	4.7	3.94	4.78	4.72
3.01	4.08	2.99	3.87	4.03	4.07	3.93	4.15	4.19	4.24	4.57	5.22	5.09
2.21	1.85	1.78	2.16	2.25	2.72	2.78	2.9	2.99	2.99	3.41	3.43	3.38
2.57	2.36	2.1	2.62	2.64	2.21	1.99	1.88	1.99	1.99	2.08	2.6	2.8
3.1	2.81	3.02	3.38	3.41	4.11	4.43	5.25	4.03	4.37	4.34	5.73	5.35
3.11	2.94	2.99	4.13	3.08	3.9	3.15	3.99	3.99	3.49	3.46	4	3.9
8.85	8.14	9.18	9.07	8.92	8.53	7.08	9.69	9.15	9.18	8.28	8.1	8.43
2.02	1.98	1.92	2.92	3.47	2.38	2.51	2.37	2.47	2.87	3.01	4.18	4.47
5.79	5.56	5.98	6.46	7.98	9.41	6.6	9.52	9.87	9.86	9.87	10.3	10.5
4.67	4.22	4.55	4.29	4.68	5.29	6.22	5.17	5.99	5.95	5.39	6.79	6.78
2.56	2.23	2.44	2.18	2.78	2.65	3.25	3.59	2.86	3.61	3.63	3.68	3.56
2.58	2.23	2.53	2.35	2.16	3.01	5.2	3.44	3.44	3.94	3.58	4.44	4.27
8.52	9.22	8.81	9.1	9.14	10.4	11.5	11.6	12	11.7	11.9	12.1	12.1
5.44	5.46	5.64	5.04	5.2	7.68	6.36	6	5.68	6.67	6.44	6.9	7.48
5.16	4.71	6.32	6.52	6.93	8.78	6.9	10.4	9.72	10.1	10.1	10.6	9.47
3.75	2.06	4.04	4.21	3.61	4.39	2.47	3.44	3.58	3.87	4.01	4.58	4.22
3.51	3.46	3.58	3.61	5.04	4.07	4.47	4.08	4.1	4.47	4.57	5.62	5.51
3.96	2.99	3.13	3.27	5.12	4.18	4.2	4.86	4.39	5.75	4.7	6.83	6.57
3.65	2.39	2.49	2.66	4.01	2.92	4.25	3.42	3.13	3.64	3.58	4	4.18
5.16	4.19	4.82	4.18	4.68	8.67	6.62	7.48	7.52	7.56	9.02	7.94	7.56
2.47	2.58	2.1	2.85	3.65	2.94	3.24	3.2	3.32	3.3	3.56	4.16	4.11
2.87	3.08	2.6	3.12	4.32	4.56	4.22	4.63	4.22	5.25	5.3	5.42	5.56
1.03	1.17	1.15	1.15	1.23	1.46	1.49	1.44	1.33	1.36	1.64	1.88	2.02
5.3	4.73	5.32	5.14	6.88	4.99	5.31	4.62	4.71	5.54	6.32	6.85	6.94
2.02	2.18	2	2.08	2.02	3.25	3.37	2.94	2.99	3.53	3.53	3.87	3.88
4.34	4.94	4.28	5.61	5.27	6.85	5.85	6.13	5.62	4.7	5.46	5.21	5.42
3.21	3.43	2.89	2.89	3.35	3.59	4.13	3.27	3.58	3.8	3.58	4.46	4.6
2.87	2.97	2.89	2.87	3.2	3.57	3.43	3.46	3.44	3.65	3.77	4.6	3.99
3.89	3.45	3.34	3.63	3.71	4.42	5.54	4.86	4.24	5.31	5.17	6.21	5.87
3.19	3.76	2.45	2.6	3.87	3.96	3.95	3.84	3.82	4.22	4.29	5.43	4.99
3.44	2.96	2.97	2.85	3.55	2.97	3.37	3.14	3.37	3.27	3.6	3.77	3.56
1.29	1.7	1.27	1.48	1.58	2.04	2.8	2.06	2.1	2.45	2.39	2.99	2.89
3.05	2.88	2.91	2.93	3.03	3.18	3.47	3.37	3.3	3.91	3.72	4.16	4.37
4.16	**4.03**	**4.08**	**4.22**	**4.61**	**5.24**	**5.18**	**5.21**	**5.21**	**5.46**	**5.48**	**5.99**	**5.96**
2.15	2.19	2.26	2.11	2.19	2.85	2.68	2.84	2.93	2.82	2.82	2.71	2.74

Table B-0-2. All of the side-lying and DXA TSTT measurements. At the bottom of the table are the mean and SD. Yellow (light shading) represents females, whereas blue (dark shading) represents males. Orange text represents MP measurements

Sideline TSTT (cm)						DXA TSTT (cm)			
SID1, 25int, BL	SID2, 25int, AT	SID3, 25int, BL	SIDX, 25int, BL	SID4, 0rot, BL	SID5, 25ext, BL	TSTT1	TSTT2	TSTT3	AT TSTT
2.21	2.24	2.62		2.79	2.45	2.5	2.4	2.4	2.9
3.84	3.98	4.01		3.7	4.85	5.5	5.5	5.7	5.8
6.64	6.71	7.51		7.59	6.75	10	9.8	9.9	9.8
4.45	4.33	4.48		4.52	4.58	7.2	7	7.2	7
8.24	7.47	7.56	6.84	7.81	8.36	10.5	10.3	10.4	10.3
6.09	5.61	5.78	5.34	6.54	6.25	9.2	9.2	9.2	9.1
3.96	4.85	4.13	3.72	4.1	4.86	6.8	6.7	6.9	6.5
2.07	2	1.92	1.99	2.11	2.75	3.9	3.9	3.9	3.9
2.65	3.73	2.86	2.56	2.75	3.25	4.1	4	4.1	4.1
5.1	5.02	5.02	4.9	5.72	6.28	6.8	6.6	6.7	6.8
4.1	4.03	4.11	4.12	4.25	4.56	6.2	6.1	6.2	6.3
1.5	1.05	1.32	1.31	1.4	0.97	1.2	1.5	1.4	1.1
1.94	1.64	1.82	1.98	2.11	2.23	3.8	3.8	3.5	4.1
1.86	1.9	2.1	1.96	2.45	2.61	2.7	2.6	2.7	2.7
1.4	1.88	1.56	1.34	1.5	1.8	3.9	3.6	3.9	3.8
1.38	1.21	1.27	1.23	1.33	1.64	2	1.8	1.9	1.7
2.49	2.99	2.4	2.48	3.19	2.76	3.3	3.2	3.4	3.3
2.21	3.05	2.08	2.1	3.26	2.54	3.2	3.2	3.3	3.1
4.88	5.21	4.78	4.99	6.8	6.59	8.3	8.8	8.3	8.8
1.68	1.64	1.76	1.68	1.82	1.84	2.5	2.4	2.5	2.5
5.45	4.16	4.24	4.02	4.96	4.37	6.9	7.2	7.2	7.2
3.99	3.61	3.61	3.82	3.7	3.87	6.2	6.3	6.4	6.4
1.54	2.43	2.16	1.87	2.38	2.03	2	1.8	2	2.4
1.92	2.67	1.96	1.9	1.94	2.08	3.6	4.1	3.9	3.9
5.98	6.08	6.43	5.11	6.68	5.82	9	9	8.8	9.1
3.59	4.32	3.73	3.87	3.76	4.48	6.1	6.1	6.2	6.2
4.91	4.18	5.24	4.82	5.51	5.39	6.2	5.4	6	6.2
1.64	2	2.1	2.21	2.06	2.47	3.4	2.9	2.7	3.7
2.78	2.67	2.56	2.64	2.8	3.37	3.9	3.7	3.9	3.7
2.49	2.42	2.55	2.5	2.47	2.71	4.3	4.5	4.6	4.3
1.74	1.68	1.74	1.75	1.94	2.76	2.7	2.5	2.7	2.6
4.21	4.03	3.53	3.77	3.9	4.79	5.9	6	6	5.7
2.1	1.87	1.92	1.95	1.92	2.2	3.6	3.4	3.6	3.7
1.94	2.03	1.86	1.85	1.96	2.21	4.4	4.5	4.5	4.5
1.36	1.06	1.05	1.15	1.01	1.09	1.9	1.5	1.5	1.8
4.21	4.33	4	3.97	4.29	4.56	6.8	6.6	6.5	7
1.6	1.58	1.56	1.63	1.62	1.84	2.2	2.2	2.2	2.2
3.35	4.13	3.49	3.42	3.44	3.89	6	5.8	6	6.2
2.31	2.58	2.55	2.57	2.57	3.09	2.7	2.7	2.8	2.7
3.1	2.27	2.55	2.29	2.23	2.35	3.7	3.5	3.7	3.6
2.95	4.21	2.99	3.04	3.61	4.23	6.1	5.9	6	6
1.94	2.56	2.04	1.83	1.99	2.06	3.6	3.5	3.6	3.6
2.4	2.27	2.47	2.39	2.47	2.95	4.8	4.8	4.8	4.8
1.09	1.01	1.11	1.26	1.27	1.09	2.5	2.5	2.5	2.4
2.24	2.31	2.2	2.3	2.29	2.69	3	3.2	3.2	3.1
3.10	**3.18**	**3.08**	**2.84**	**3.30**	**3.47**	**4.78**	**4.71**	**4.77**	**4.81**
1.65	1.57	1.62	1.36	1.76	1.71	2.32	2.33	2.32	2.32

Appendix C - Main Effect of Sex, and an Interaction between Sex and Position on TSTT

TSTT was calculated for each sex in each position as it may be useful for comparisons with other studies (Figure C-1). Subsequently, it was decided to include sex as a factor to the position/rotation ANOVA, and there was a main effect of sex ($p < 0.001$), and a significant interaction between sex and position ($p = 0.002$). There was no significant interaction between sex and rotation ($p = 0.766$), and there was no significant three-way interaction between sex, position, and rotation ($p = 0.403$). The p-value of the previous main effects and interaction were still well below 0.05.

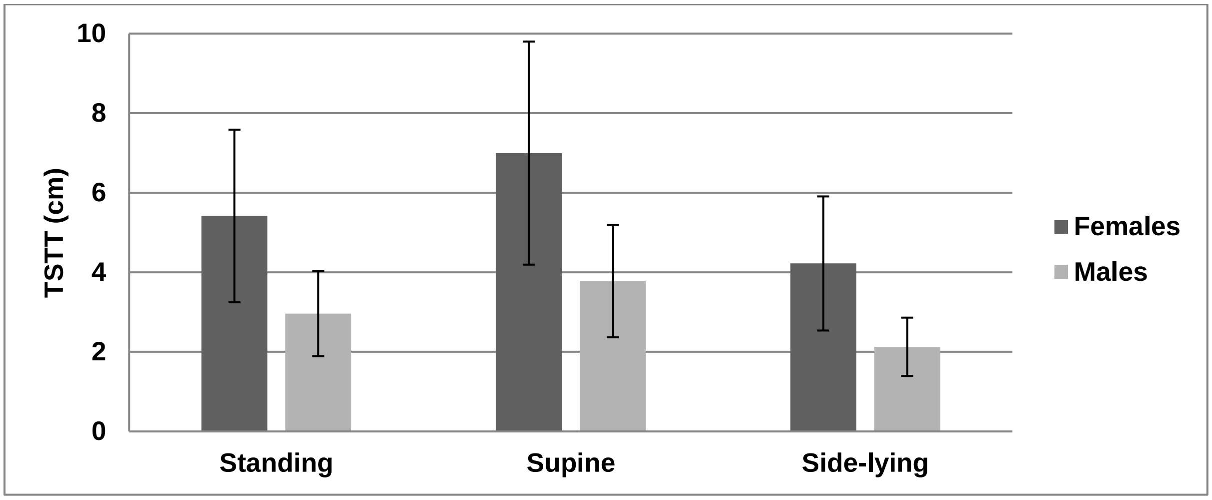


Figure C-1. Mean (SD) TSTT of female and male participants while standing, supine, and side-lying.

The main effect of sex on TSTT was expected as females tend to have a gynoid distribution of adipose tissue (i.e. in the hip/buttocks region), whereas males tend to have an android distribution of adipose tissue (i.e. in the abdomen). Since there is a greater amount of TSTT in women than men, it is more influenced by gravitational changes, thus explaining the interaction between sex and position. It could also be due to differences in soft tissue stiffness between males and females. As mentioned, TSTT is less reliably measured in those with greater amounts of TSTT. Accordingly, it can be assumed that TSTT measurements in women would be less reliable than men. This is unfortunate as more women suffer hip fractures than men (Leslie et al., 2009).

Appendix D - Additional Details of Methodology

Questionnaires

Participants completed two questionnaires. The Medical History Questionnaire collected the following information: age, sex, ethnicity, previous fracture, parental history of fracture, smoking history, alcohol intake, glucocorticoid use, secondary osteoporosis, rheumatoid arthritis, hormone replacement therapy, menopausal symptoms, endocrine disorders, asthma, chronic obstructive pulmonary disease, dementia, cancer, cardiovascular disease, chronic liver disease, chronic kidney disease, systemic lupus erythematosus, epilepsy/anticonvulsants, antidepressant use, diabetes, nursing/home care resident, and history of falls. These answers are inputs needed for the FRAX, QFracture, and Garvan model, which are models that calculate the probability of a hip fracture. The Occupational Activity Questionnaire queried each participant about the type of work they did in the past, and how many hours they spent sitting/standing/walking for each occupation. All of the answers to these questionnaires will be used for future research purposes.

Anthropometric measurements

A total of nine anthropometric measurements were acquired. The first was waist circumference, where waist was defined as the midpoint between the lowest rib and the top of the iliac crest. I followed the Rudolf et al. (2007) recommendations by always taking the measurement 4 cm above the umbilicus. Next was the hip circumference, which is at the level of the greatest protrusion of the gluteal muscles (Reiman and Manske, 2009). Thigh circumference was collected by starting the measurement 2 cm below the gluteal line (Reiman and Manske, 2009). The distance between the greater trochanter and the lateral condyle, and to the floor, which I called thigh length and leg length respectively, was also measured in each participant. Lastly, I measured standing width and standing hip angle. Standing width is the distance between the two heels, whereas standing hip angle is the angle between the vertical and an imaginary line between the vertical and the lateral border of the foot. These measurements were taken to make a multi-variate regression equation of TSTT, but this will not be addressed in this thesis.

DXA Hip Scan

After the whole body DXA scan, a right hip scan was performed by the MRT to acquire BMD of the hip. The regions of interest of the hip scan are the femoral neck, the trochanter, the intertrochanteric area, Ward’s triangle, and the total hip area. These BMD values will be used to calculate the Factor of Risk, but it will not be addressed in this thesis.

Second Investigator for Inter-Rater Reliability (MP)

Two secondary investigators were used in this study, and they had different types of training. AT trained with the main investigator BL, where both investigators practiced the protocol of this thesis on several graduate students. MP did another type of training with the ultrasound device. He is accustomed to measuring muscle thicknesses throughout the body, but this protocol did not include ultrasounding the greater trochanter. However, MP has more experience palpating the greater trochanter and using the ultrasound device than BL and AT. MP completed 15 participants with BL, whereas AT completed 30 participants with BL. For this thesis, only the results between BL and AT was presented

Ultrasound Checklist

Table D-0-1. Ultrasound protocol used for this thesis study. This table was always used for the collection sessions

Position	Rotation	Investigator	File Name	Complete?
Standing	25° internal rotation	BL	ST1, 25int, BL	
Standing	25° internal rotation	AT / MP	ST2, 25int, AT	
Standing	25° internal rotation	BL	ST3, 25int, BL	
Standing	0 rotation	BL	ST4, 0rot, BL	
Standing	25° external rotation	BL	ST5, 25ext, BL	
CHECK TO MAKE SURE ALL IMAGES ARE TAKEN				
Supine	25° internal rotation	BL	SUP1, 25int, BL	
Supine	25° internal rotation	AT / MP	SUP2, 25int, AT	
Supine	25° internal rotation	BL	SUP3, 25int, BL	
Supine	25° internal rotation	BL	SUP4, 25int, BL	
Supine	0 rotation	BL	SUP5, 0rot, BL	
Supine	0 rotation	BL	SUP6, 0rot, BL	
Supine	25° external rotation	BL	SUP7, 25ext, BL	
Supine	25° external rotation	BL	SUP8, 25ext, BL	
CHECK TO MAKE SURE ALL IMAGES ARE TAKEN				
Side-lying	25° internal rotation	BL	SID1, 25int, BL	
Side-lying	25° internal rotation	AT / MP	SID2, 25int, AT	
Side-lying	25° internal rotation	BL	SID3, 25int, BL	
Side-lying	25° internal rotation	BL with linear	SIDX, 25int, BL	
Side-lying	0 rotation	BL	SID4, 0rot, BL	
Side-lying	25° external rotation	BL	SID5, 25ext, BL	
CHECK TO MAKE SURE ALL IMAGES ARE TAKEN				

Bibliography

Addison, O., Young, P., Inacio, M., Bair, W.-N., Prettyman, M.G., Beamer, B. a, Ryan, A.S., Rogers, M.W., 2014. Hip but not thigh intramuscular adipose tissue is associated with poor balance and increased temporal gait variability in older adults. Curr. Aging Sci. 7, 137–43.

Balta, P.J., Ward, M.W., Tomkins, A.M., 1981. Ultrasound for measurement of subcutaneous fat. Lancet 1, 504–505.

Behari, J., 2009. Biophysical Bone Behavior. Wiley.

Bhan, S., Levine, I.C., Laing, A.C., 2014. Energy absorption during impact on the proximal femur is affected by body mass index and flooring surface. J. Biomech. 47, 2391–7.

Bonnick, S. Lou, 2004. Bone Densitometry in Clinical Practice - Application and Interpretation, Second. ed. Humana Press, Totowa.

Bouxsein, M.L., Szulc, P., Munoz, F., Thrall, E., Sornay-Rendu, E., Delmas, P.D., 2007. Contribution of trochanteric soft tissues to fall force estimates, the factor of risk, and prediction of hip fracture risk. J. Bone Miner. Res. 22, 825–831.

Braithwaite, R.S., Col, N.F., Wong, J.B., 2003. Estimating hip fracture morbidity, mortality and costs. J. Am. Geriatr. Soc. 51, 364–370.

Choi, W.J., Hoffer, J.A., Robinovitch, S.N., 2010. The effect of positioning on the biomechanical performance of soft shell hip protectors. J. Biomech. 43, 818–825.

Choi, W.J., Russell, C.M., Tsai, C.M., Arzanpour, S., Robinovitch, S.N., 2015. Age-related changes in dynamic compressive properties of trochanteric soft tissues over the hip. J. Biomech. 48, 695–700.

Cooper, C., Campion, G., Melton III, L.J., 1992. International Original Article Hip Fractures in the Elderly : A World-Wide Projection. Osteoporos. Int. 285–289.

Cosman, F., Lindsay, R., LeBoff, M.S., Jan de Beur, S., Tanner, B., 2014. 2014 Clinician's Guide to Prevention and Treatment of Osteoporosis. Natl. Osteoporos. Found. 1, 55.

Cotter, P.E., Timmons, S., O'Connor, M., Twomey, C., O'Mahony, D., 2005. The financial implications of falls in older people for an acute hospital. Ir. J. Med. Sci. 175, 11–13.

Dufour, A.B., Roberts, B., Broe, K.E., Kiel, D.P., Bouxsein, M.L., Hannan, M.T., 2012. The factor-of-risk biomechanical approach predicts hip fracture in men and women: The Framingham Study. Osteoporos. Int. 23, 513–520.

Earthman, C.P., 2015. Body Composition Tools for Assessment of Adult Malnutrition at the Bedside: A Tutorial on Research Considerations and Clinical Applications Introduction to Lean Tissue Terminology, Journal of Parenteral and Enteral Nutrition.

English, C.K., Thoirs, K.A., 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, 368–376.

Etheridge, B.S., Beason, D.P., Lopez, R.R., Alonso, J.E., McGwin, G., Eberhardt, A.W., 2005. Effects of trochanteric soft tissues and bone density on fracture of the female pelvis in experimental side impacts. Ann. Biomed. Eng. 33, 248–254.

Fairhead, A.C., Wittingham, T.A., 2012. Diagnostic Ultrasound. In: Dendy, P.P., Heaton, B. (Eds.), Physics for Diagnostic Radiology. CRC Press, pp. 489–562.

Feldman, F., Robinovitch, S.N., 2007. Reducing hip fracture risk during sideways falls: Evidence in young adults of the protective effects of impact to the hands and stepping. J. Biomech. 40, 2612–2618.

Goh, J.C.H., Shah, K.M., Bose, K., 1995. Biomechanical study on femoral neck fracture fixation in relation to bone mineral density. Clin. Biomech. (Bristol, Avon) 10, 304–308.

Griffiths, M.R., Noakes, K. a, Pocock, N. a, 1997. Correcting the magnification error of fan beam densitometers. J. Bone Miner. Res. 12, 119–123.

Guay, M., 2005. Anatomie fonctionnelle de l'appareil locomoteur. Les Presses de l'Universite de Montreal.

Hayes, W.C., Myers, E.R., Robinovitch, S.N., Van Den Kroonenberg, A., Courtney, A.C., McMahon, T.A., 1996. Etiology and prevention of age-related hip fractures. Bone 18, 77S–86S.

Hayes, W.C., Piazza, S.J., Zysset, P.K., 1991. Biomechanics of fracture risk prediction of the hip and spine by quantitative computed tomography. Radiol. Clin. North Am. 29, 1–18.

Hibbeler, R.C., 2011. Mechanics of Materials, 8th ed. Pearson.

Hill, J.C., Leiszler, M.S., 2013. Hip. In: Daniels, J., Dexter, W. (Eds.), Basics of Musculoskeletal Ultrasound. pp. 87–91.

Ioannidis, G., Papaioannou, A., Hopman, W.M., Akhtar-Danesh, N., Anastassiades, T., Pickard, L., Kennedy, C.C., Prior, J.C., Olszynski, W.P., Davison, K.S., Goltzman, D., Thabane, L., Gafni, A., Papadimitropoulos, E. a., Brown, J.P., Josse, R.G., Hanley, D. a., Adachi, J.D., 2009.

Relation between fractures and mortality: Results from the Canadian Multicentre Osteoporosis Study. Cmaj 181, 265–271.

Johnell, O., Kanis, J. a., 2004. An estimate of the worldwide prevalence, mortality and disability associated with hip fracture. Osteoporos. Int. 15, 897–902.

Kanis, J. a., Johnell, O., Oden, A., Johansson, H., McCloskey, E., 2008. FRAX and the assessment of fracture probability in men and women from the UK. Osteoporos. Int. 19, 385–397.

Kanis, J. a., McCloskey, E., Johansson, H., Oden, A., Leslie, W.D., 2012. FRAX with and without bone mineral density. Calcif. Tissue Int. 90, 1–13.

Katch, F.I., 1983. Individual differences of ultrasound assessment of subcutaneous fat: effects of body position. Hum. Biol. an Int. Rec. Res. 55, 789–795.

Korhonen, N., Niemi, S., Parkkari, J., Sievänen, H., Palvanen, M., Kannus, P., 2013. Continuous decline in incidence of hip fracture: nationwide statistics from Finland between 1970 and 2010. Osteoporos. Int. 24, 1599–603.

Laing, A.C., Robinovitch, S.N., 2010. Characterizing the effective stiffness of the pelvis during sideways falls on the hip. J. Biomech. 43, 1898–1904.

Lekamwasam, S., Lenora, J., 2003. Effect of Leg Rotation on Hip Bone Mineral Density Measurements. J. Clin. Densitom. 6, 331–336.

Leslie, W.D., Donnell, S.O., Jean, S., Walsh, P., Bancej, C., Hanley, D. a, 2009. Trends in Hip Fracture Rates in Canada 302, 883–889.

Leslie, W.D., O'Donnell, S., Lagacé, C., Walsh, P., Bancej, C., Jean, S., Siminoski, K., Kaiser, S., Kendler, D.L., Jaglal, S., 2010. Population-based Canadian hip fracture rates with international comparisons. Osteoporos. Int. 21, 1317–1322.

Levine, I.C., Bhan, S., Laing, A.C., 2013. The effects of body mass index and sex on impact force and effective pelvic stiffness during simulated lateral falls. Clin. Biomech. 28, 1026–1033.

Levine, I.C., Minty, L.E., Laing, A.C., 2015. Factors that influence soft tissue thickness over the greater trochanter: Application to understanding hip fractures. Clin. Anat. 28, 253–261.

Luo, Y., 2015. A biomechanical sorting of clinical risk factors affecting osteoporotic hip fracture. Osteoporos. Int.

Maitland, L. a, Myers, E.R., Hipp, J. a, Hayes, W.C., Greenspan, S.L., 1993. Read my hips: measuring trochanteric soft tissue thickness. Calcif. Tissue Int. 52, 85–89.

Majumder, S., Roychowdhury, A., Pal, S., 2008. Effects of trochanteric soft tissue thickness and hip impact velocity on hip fracture in sideways fall through 3D finite element simulations. J. Biomech. 41, 2834–2842.

Majumder, S., Roychowdhury, A., Pal, S., 2013. Hip fracture and anthropometric variations: Dominance among trochanteric soft tissue thickness, body height and body weight during sideways fall. Clin. Biomech. 28, 1034–1040.

Marieb, E.N., Hoehn, K., 2010. Anatomie et Physiologie Humaines, Quatrième. ed. ERPI.

Martin Bland, J., Altman, D., 1986. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. Lancet 327, 307–310.

Martinoli, C., Bianchi, S., 2007. Hip. In: Ultrasound of the Muscoskeletal System. pp. 551–610.

McCloskey, E., Johansson, H., Oden, A., Kanis, J. a., 2012. Fracture risk assessment. Clin. Biochem. 45, 887–893.

McGlasson, R., Zellermeyer, V., MacDonald, V., Lo, N., Spafford, D., Legge McMullan, J., Beaupre, L., Sims Gould, J., Saryeddine, T., Scott, V., 2011. National Hip Fracture Toolkit.

Minns, R.J., Marsh, a. M., Chuck, a., Todd, J., 2007. Are hip protectors correctly positioned in use? Age Ageing 36, 140–144.

Moore, K.L., Dalley, A.F., Agur, A.M.R., 2010. Clinically Oriented Anatomy. Lippincott Willams & Wilkins.

Nankaku, M., Kanzaki, H., Tsuboyama, T., Nakamura, T., 2005. Evaluation of hip fracture risk in relation to fall direction. Osteoporos. Int. 16, 1315–1320.

Nielson, C.M., Bouxsein, M.L., Freitas, S.S., Ensrud, K.E., Orwoll, E.S., 2009. Trochanteric soft tissue thickness and hip fracture in older men. J. Clin. Endocrinol. Metab. 94, 491–496.

Nigg, B.M., Herzog, W., 2007. Biomechanics of the Musculo-skeletal System, Third. ed. Wiley.

Njeh, C.F., Shepard, J.., 2004. Absorptiometric measurement. In: Njeh, C.F., Langton, C.M. (Eds.), The Physical Measurement of Bone. Institute of Physics Publishing, pp. 267–307.

O’Neill, J., 2008. Musculoskeletal ultrasound: Anatomy and technique, Musculoskeletal Ultrasound: Anatomy and Technique.

Oldroyd, B., Smith, a H., Truscott, J.G., 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.

Papaioannou, a., Kennedy, C.C., Ioannidis, G., Sawka, a., Hopman, W.M., Pickard, L., Brown, J.P., Josse, R.G., Kaiser, S., Anastassiades, T., Goltzman, D., Papadimitropoulos, M., Tenenhouse, a., Prior, J.C., Olszynski, W.P., Adachi, J.D., 2009. The impact of incident fractures on health-related quality of life: 5 years of data from the Canadian Multicentre Osteoporosis Study. Osteoporos. Int. 20, 703–714.

Papaioannou, A., Morin, S., Cheung, A.M., Atkinson, S., Brown, J.P., Feldman, S., Hanley, D. a., Hodsman, A., Jamal, S. a., Kaiser, S.M., Kvern, B., Siminoski, K., Leslie, W.D., 2010. 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: Summary. Cmaj 182, 1864–1873.

Pocock, N. a., Noakes, K. a., Majerovic, Y., Griffiths, M.R., 1997. Magnification error of femoral geometry using fan beam densitometers. Calcif. Tissue Int. 60, 8–10.

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

Reiman, M.P., Manske, R.C., 2009. Functional Testing in Human Performance. Human Kinetics.

Roberts, B.J., Thrall, E., Muller, J.A., Bouxsein, M.L., 2010. Comparison of hip fracture risk prediction by femoral aBMD to experimentally measured factor of risk. Bone 46, 742–746.

Robinovitch, S.N., Hayes, W.C., McMahon, T.A., 1991. Prediction of femoral impact forces in falls on the hip. J. Biomech. Eng. 113, 366–374.

Robinovitch, S.N., Hayes, W.C., McMahon, T.A., 1997. Distribution of contact force during impact to the hip. Ann. Biomed. Eng. 25, 499–508.

Robinovitch, S.N., McMahon, T.A., Hayes, W.C., 1995. Force attenuation in trochanteric soft tissues during impact from a fall. J. Orthop. Res. 13, 956–962.

Roudsari, B.S., Ebel, B.E., Corso, P.S., Molinari, N.-A.M., Koepsell, T.D., 2005. The acute medical care costs of fall-related injuries among the U.S. older adults. Injury 36, 1316–1322.

Rudolf, M.C.J., Walker, J., Cole, T.J., 2007. What is the best way to measure waist circumference? Int. J. Pediatr. Obes. 2, 58–61.

Schacter, I., Leslie, W.D., 2014. Estimation of Trochanteric Soft Tissue Thickness From Dual-Energy X-ray Absorptiometry. J. Clin. Densitom. 17, 54–59.

Statistics Canada, 2009. Canadian Health Measures Survey : Cycle 1 Data Tables – 2007 to 2009 Distribution of the household population aged 18 to 79 , by body mass index norms based on measured inputs , by age and sex , Canada , 2007 to 2009 Statistics Canada – Catalogue no . 82-. Heal. Meas. Surv. 2009.

Tarride, J.E., Hopkins, R.B., Leslie, W.D., Morin, S., Adachi, J.D., Papaioannou, a., Bessette, L., Brown, J.P., Goeree, R., 2012. The burden of illness of osteoporosis in Canada. Osteoporos. Int. 23, 2591–2600.

Thayalan, K., 2014. The Physics of Radiology and Imaging. Jaypee Brothers Medical Publishers.

Toombs, R.J., Ducher, G., Shepherd, J. a, De Souza, M.J., 2012. The impact of recent technological advances on the trueness and precision of DXA to assess body composition. Obesity (Silver Spring). 20, 30–9.

Wiktorowicz, M.E., Goeree, R., Papaioannou, a, Adachi, J.D., Papadimitropoulos, E., 2001. Economic implications of hip fracture: health service use, institutional care and cost in Canada. Osteoporos. Int. 12, 271–278.

Wilmore, J.H., Costill, D.L., Kenney, W.L., 2009. Physiologie du sport et de l’exercice. De Boeck.