Investigation of a Double Calibration Technique to Reduce Soft Tissue Artifact Error in High Flexion

by

Madalyn Natalia Tworzyanski

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, 2022

Author's Declaration

I declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Abstract

High knee flexion tasks (knee flexion angle > 120 degrees) are performed frequently in both daily living activities (gardening, religious practice, exercise, etc.) and occupational settings (childcare, roofing, construction, floor laying, etc.). These tasks are associated with an increased risk of knee osteoarthritis development, which can alter movement patterns. These types of movement differences can be captured when analyzing high flexion postures using optical motion capture; however, we do not know how accurate reported kinematic outcomes are because of an inherent source of error known as soft tissue artifact (STA). This error is defined as the movement of skin markers relative to the underlying bone and affects the thigh markers more than the shank. It cannot be filtered out of data after processing because it has a similar frequency content to the movements themselves (~5-10 Hz). Therefore, all reported measures of knee kinematics obtained using optical motion capture include an unknown level of error and can affect clinical and biomechanical interpretations of knee pathologies. This thesis investigated the use of a double calibration technique to improve the accuracy of landmarks tracked in high flexion postures with an anterolaterally located marker cluster in the mid-thigh region. Thirty-three participants performed flatfoot squatting, heels-up squatting, dorsiflexed kneeling and plantarflexed kneeling movements. The position of the functional hip joint center was defined in a standing reference position using a functional calibration trial. This landmark was then simultaneously tracked with the thigh cluster and pelvis cluster during high flexion movements. The landmark tracking with the thigh cluster was referred to as the femoral head center and was obtained in two ways: using either single or double calibration techniques. The landmark tracking with the pelvis, referred to as the hip joint center, was considered the gold standard and thus was the position to which the single and double calibration-based femoral head center

positions were compared. Root-mean squared (RMS) error was calculated between the hip joint center and femoral head centers in the global x, y, and z directions. Resultant error (distance between the hip joint center and the femoral head centers) was also determined. The bias and limits of agreement on the resultant error were used to evaluate the accuracy of locating the femoral head center using each calibration technique relative to the location of the hip joint center. Paired t-tests revealed RMS error and resultant error were not significantly lower using the double calibration technique, and the limits of agreement were wider in the double calibration. Data were then separated into percentile groups to evaluate the double calibration technique after controlling for mid-thigh circumference, a subject-specific variable. RMS errors in the global y and z directions and resultant error were significantly greater in the double calibration for the 25^{th} and 50^{th} percentile groups, while there were no significant differences in RMS error nor resultant error between calibration techniques in the 15^{th} and 75^{th} percentile group. Resultant error also increased from 15^{th} to 50^{th} percentile groups and limits of agreement increased with higher percentiles. Due to poor performance of the double calibration despite separating data into mid-thigh percentile groups, the double calibration method was evaluated using a different landmark to predict the position of the femoral head center in the seated calibration posture (as the functional hip joint center trial was only able to be applied to the standing posture). Double calibration using the lateral femoral condyle performed only slightly superior than using the greater trochanter, the landmark initially used in the double calibration method; however single calibration continued to perform best. Therefore, the double calibration was not recommended for use with the studied high flexion movements. The conclusions of this thesis direct future research to evaluate the composition of the thigh (adipose tissue vs muscular tissue) and its relation to STA, and to assess ways to improve palpation of landmarks in high

flexion postures (e.g., using fluoroscopy measures). Once accurate palpation is achieved in high flexion, the double calibration technique could be revisited to evaluate its effectiveness to reduce STA.

Acknowledgement

This Masters thesis was completed with the guidance and assistance of various researchers that I worked closely with at the University of Waterloo. I would like to acknowledge Dr. Stacey Acker, my Masters supervisor, for their expertise and creativity in guiding me throughout this thesis and my Masters career. We shared amazing moments of "research excitement" when discussing different ideas and research projects, and I am very thankful to have worked with you. I would also like to acknowledge my thesis committee members, Dr. Clark Dickerson and Dr. Monica Maly, for their input and guidance throughout my thesis proposal and defence. You both challenged my critical thinking skills and provided extremely valuable input that further developed this thesis document.

My lab members provided continued support and I would like to acknowledge Annemarie Laudanski, Jessa Buchman-Pearle, Terri Weeks, Michelle Loo, Kim Peckett and Dave Varandas. You all made my Masters experience something that I will never forget. I would especially like to thank Annemarie and Jessa, who mentored me from the day I started my Masters all the way until I completed my thesis defence. Thank you both for your insightful knowledge and guidance throughout this process.

Lastly, I would like to acknowledge my family for their love and support as I pursued my Masters career.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Rotations and Translations

FL/EX	Flexion/extension
AB/AD	Abduction/adduction
INT/EXT	Internal/external
A/P	Anterior/posterior
M/L	Medial/lateral
PROX/DIS	Proximal/distal

Primary Variables in Equations for the Double Calibration

A(t)	The reconstructed position of the femoral head center in the thigh cluster coordinate system at each frame of motion trials.
A_1	Position of the femoral head center in the standing calibration posture in the thigh cluster coordinate system.
A_2	Position of the femoral head center at maximal flexion in the thigh cluster coordinate system.
f(t)	Knee flexion angle at each frame of the motion trials determined from the single calibration.
f_1	Knee flexion angle in the standing calibration posture.
f_2	Knee flexion angle at maximal flexion in each motion trial.
FHC_L	Functional hip joint center in the first frame of quiet standing in the thigh cluster coordinate system. Also defined as A_1.
TCS_R	Rotation matrix from the global coordinate system to the thigh cluster coordinate system in the first frame of quiet standing.
$FHJC_G$	Functional hip joint center in the global coordinate system in the first frame of quiet standing.
TCS_o	Origin of the thigh cluster coordinate system in the first frame of quiet standing.
GT_L	Greater trochanter in the thigh cluster coordinate system at the first frame of the quiet standing trial.
GT_{LSit}	Greater trochanter in the cluster coordinate system at the first frame of the seated calibration trial.
$FHC(t)_R$	Reconstructed femoral head center position in the global coordinate system at each frame in the motion trials (t).
$TCS(t)_R$	Local to global rotation matrix of the thigh cluster coordinate system at each frame in the high flexion motion trials.
$TCS(t)_o$	Origin of the thigh cluster coordinate system at each frame in the high flexion motion trials.
HJC	Hip joint center.
FHC	Femoral head center.

Other

OA	Osteoarthritis
ACL	Anterior cruciate ligament
STA	Soft tissue artifact
ROM	Range of motion
BMI	Body mass index
RMS	Root-mean squared
GT-based	Greater trochanter-based
LFC-based	Lateral femoral condyle-based
HJC_L	Hip joint center in the pelvis cluster coordinate syst

Chapter 1 – Introduction

1.1 Introduction

Knee pathologies are commonly analyzed through differences in knee mechanics. Pathological knee mechanics have been compared against healthy typical mechanics to identify severity, and monitor progression of rehabilitation in osteoarthritis (OA) (Hamai et al., 2009), total knee replacements (McClelland et al., 2011), obesity (Clément et al., 2018; Li et al., 2019), and anterior cruciate ligament (ACL) deficiencies (Hamai et al., 2009; Shybut et al., 2015). Previous research found that knee motion in the anterior/posterior (A/P) direction and axial rotation was most affected by medial knee compartment OA (Hamai et al., 2009), maximum flexion angle was most affected by total knee replacements (McClelland et al., 2011), and obesity significantly altered the flexion/extension (FL/EX) and abduction/adduction (AB/AD) range in gait (Li et al., 2019). ACL deficiencies with/without a lateral meniscal tear also had significantly greater anterior translation than healthy ACLs in translational testing (Shybut et al., 2015). Since differences in kinematics can be important indicators of severity and progression, they could also play a role in elucidating mechanisms of pathology development. Thus, these measures should be as accurate as possible to ensure correct clinical and biomechanical interpretations are made.

Using optical motion capture systems, marker clusters on the thigh and shank are tracked and their coordinates are used to create segment coordinate systems and calculate knee joint kinematics and kinetics. These clusters, however, are affected by soft tissue artifact (STA); a source of error caused by the movement of skin markers relative to the underlying bone (Andersen et al., 2012; Barré et al., 2015; Cappello et al., 2005). STA is unable to be filtered out of data because it has a similar frequency content to dynamic motions (Camomilla et al., 2015;

Cappello et al., 2005) and is also larger than inherent errors from the motion capture system itself (Cappozzo et al., 1996). Traditionally, most motion capture data does not account for these errors in reporting joint kinematics or kinetics, leaving collected data contaminated with an unknown error.

Activities that result in more soft tissue deformation may be susceptible to greater STA. High knee flexion activities, where knee flexion exceeds 120 degrees, can cause significant deformation of soft tissue, particularly on the thigh, due to thigh-calf contact. These postures, such as squatting and kneeling, are adopted in daily life activities such as gardening, religious practice, and exercise (Li et al., 2004) and occupational work such as childcare, roofing, construction (Cooper et al., 1994), and floor laying (Jensen et al., 2000). In cases where these postures are very repetitive, as seen in these occupational and daily tasks, there is an increased risk of knee OA development (Amin et al., 2008; Coggon et al., 2000; Cooper et al., 1994; Dulay et al., 2015). It is important that we know, and take the best possible approach to ensure, the accuracy of the kinematics obtained using optical motion capture in high knee flexion activities because: high knee flexion is associated with pathology, and, while the literature has established that changes in joint kinematics can provide us with critical information regarding the pathology, it is also clear that high knee flexion could result in greater uncertainty in these kinematic measures due to increased soft tissue deformation.

One method commonly explored to reduce STA in motion capture data is to consider the effects of marker placement. Previous research evaluated different marker cluster locations on the thigh to determine the most appropriate placement (Akbarshahi et al., 2010; Buchman-Pearle & Acker, 2021; Cockcroft et al., 2016; Gao & Zheng, 2008) because it is subject to larger STA than the shank (Barré et al., 2013; Benoit et al., 2015; Clément et al., 2018; Dumas et al., 2014;

Gao & Zheng, 2008; Kuo et al., 2011; Stagni et al., 2005). Different placements, however, are specific to the task performed and the primary plane of motion, meaning the most appropriate placement would not be similar amongst a variety of tasks. This limits the generalizability of marker placement for various movements and requires evaluation for specific tasks.

Only one study has evaluated the performance of thigh marker clusters in high flexion, using Bland-Altman analysis to determine the agreement of a thigh-cluster tracked functional hip joint center position with a pelvis cluster tracked functional hip joint center position. The study concluded the most accurate thigh cluster placement to track digitized landmarks was the mid-anterolateral thigh; however, this placement did not reduce all STA (Buchman-Pearle & Acker, 2021). A promising potential method for further reducing STA is double calibration. This method digitizes landmarks relative to a marker cluster in two different postures (usually at the extreme ranges of motion of the activity being analysed) to create a 2-point linear calibration curve that estimates how the position of a landmark changes throughout the full range of motion (ROM) (Cappello et al., 2005). This method has been found to improve knee joint rotations and translations as a result of STA compensation (Cappello et al., 2005).

1.2 Rationale

High flexion postures are commonly adopted in occupational and daily work but risk the development of knee joint disorders. STA is inherent in measurement of these movements using motion capture systems, caused by the sliding of skin markers relative to the underlying tissue. It greatly affects kinematic outcomes of motion capture data obtained from the thigh, especially in high flexion postures, and has only been mitigated thus far by controlling for cluster location on the thigh.

Drawing conclusions on knee pathology-induced movement variations should be made with full confidence. Knee OA (Hamai et al., 2009; Yang et al., 2010; Zhang et al., 2004), total knee replacements (McClelland et al., 2011), obesity (Clément et al., 2018; Li et al., 2019), and ACL deficiencies (Hamai et al., 2009; Shybut et al., 2015) all cause knee motions to deviate from healthy individuals. In high flexion especially, values cannot be reported with full confidence because STA remains despite previous efforts to reduce the effects of thigh cluster location. Therefore, there is a need to investigate other methods to reduce the effect of STA in optical motion capture data in high flexion postures and provide recommendations to ensure data is most accurately collected and reported.

1.3 Objectives

Primary Objective: To investigate the use of a double calibration technique to improve the accuracy of prediction of the femoral head center tracked by the mid-anterolateral thigh cluster in dynamic kneeling and squatting (Chapter 3). This study recruited a healthy student population. Landmarks were digitized in two calibration postures, one standing and one seated, to create a two-point linear calibration curve for the double calibration technique that predicted the position of the femoral head center relative to the mid-anterolateral thigh cluster throughout the full range of knee flexion. Errors in locating the femoral head center using the thigh cluster in the single and double calibration techniques were calculated by comparing to the location of the hip joint center, which tracked with the pelvis cluster.

Secondary Objective 1: To determine the effectiveness of the double calibration technique within participant groups (Chapter 4). Participants were divided into four groups based on the mid-thigh circumference. This objective emerged from results and limitations in Chapter 3, which addressed the primary objective.

Secondary Objective 2: To evaluate whether choosing a different palpated anatomical landmark from which to predict the femoral head center in the seated calibration posture influenced double calibration results. The femoral head center location was predicted using a functional hip joint center trial, which was unable to be applied to the seated calibration posture. Therefore, a palpated landmark (greater trochanter) was chosen, and a vector was created from the greater trochanter to the femoral head center in the standing calibration posture. This vector was then applied the greater trochanter in the seated calibration posture to predict the femoral head center location. In this secondary objective, the lateral femoral condyle replaced the greater trochanter to predict the femoral head center location. This objective emerged from results emerged in Chapters 3 and 4, and acknowledging palpation was more difficult in a seated posture compared to standing.

Chapter 2 – Literature Review

2.1 Soft Tissue Artifact

2.1.1 Overview of STA

STA is a source of error in motion capture data. It is defined as the movement of skin markers on segments, which can be caused by wobbling mass (Andersen et al., 2012; Barré et al., 2015), muscular contractions (Andersen et al., 2012; Barré et al., 2015; Cappello et al., 2005), and skin viscoelasticity (Andersen et al., 2012; Barré et al., 2015; Cappello et al., 2005). This can lead to errors in skin marker positions that are higher than inherent errors from the motion capture system (Cappozzo et al., 1996) and affect reported kinematics (Barré et al., 2015; Benoit et al., 2006) and kinetics (Kuo et al., 2011). This decreases the accuracy and reliability to measure true bone motion. (Andersen et al., 2010, 2012; Barré et al., 2015; Benoit et al., 2006; Cappello et al., 1997, 2005).

STA has a very similar frequency content as bone movement (Cappello et al., 2005; Cappozzo et al., 1996). It occurs between 5-10 Hz as in gait trials (Camomilla et al., 2015), with normal movement occurring around 6 Hz; therefore, STA is unable to be filtered out of data (Cappello et al., 2005). STA acts as noise in data, and procedures should be applied to minimize this error because STA causes skin markers to undergo deformation and rigid displacement because of the underlying soft tissue (Gao & Zheng, 2008). Previous research has defined deformation as non-rigid changes in shape variation (Cappello et al., 1997, 2005; Dumas et al., 2014; Grimpampi et al., 2014), meaning changes occur due to the behaviour of soft tissues relative to the bone (Gao & Zheng, 2008). Displacement has been defined as a rotation and translation (Cappello et al., 1997, 2005; Grimpampi et al., 2014) and is rigidly associated with the segments, meaning it occurs together with the bone. Although traditional motion capture data

collections assume skin marker clusters are rigidly to the segments, the movement of the markers does not represent true bone motion (Benoit et al., 2015; Cappozzo et al., 1996).

An increase in joint ROM can propagate errors in skin marker position due to STA. This is problematic because activities such as gait, sit-to-stand, kneeling and squatting move through wide variations of joint ROM, and some parts of the movement will be subjected to higher STA than others, as seen in the swing phase of gait (Andriacchi et al., 1998). Previous research found that STA rigid transformation of skin markers increased with knee flexion angle (Cappozzo et al., 1996; Clément et al., 2018), leading to errors in skin marker readings because of the rigid motion (Andersen et al., 2012; Grimpampi et al., 2014). Of the two segments neighbouring the knee joint, the thigh is subject to higher STA than the shank (Barré et al., 2013; Benoit et al., 2015; Clément et al., 2018; Dumas et al., 2014; Gao & Zheng, 2008; Kuo et al., 2011) and propagated error can vary based on the location on the thigh segment. Although no developed method can eliminate STA contamination in motion capture data, understanding ways to mitigate this error is advantageous for data collections.

2.1.2 Subject, Task and Location specificity

STA is subject-, task- and location-specific, meaning it is variable between individuals, tasks performed, and locations between and within segments. It is important to highlight these characteristics because STA is not generalizable, meaning a simple offset in data cannot account for the STA experienced in various conditions. The importance of these characteristics will be relevant to the proposed thesis.

STA is supported in the literature to be a subject-specific variable (Benoit et al., 2015; Camomilla et al., 2015; Camomilla et al., 2013; Fiorentino et al., 2017). This means that errors associated with STA is not comparable between subjects. Previous research looking at STA in

gait tasks found that root-mean square (RMS) errors of knee joint angles calculated from x-ray fluoroscopy and motion capture data varied between subjects (Akbarshahi et al., 2010). Similarly, another study using bone-pins and motion capture data found that kinematic profiles and errors of the knee joint in cutting and gait tasks also varied between subjects; however, within-subject error was similar across the two movements (Benoit et al., 2006). This highlights that there is/are some individual characteristics that influence the effect of STA. Other research found body mass index (BMI) correlates with differences in femoral vertical axis rotation (Barré et al., 2013) while another study found mixed results with STA and BMI but was limited to the obese subjects having knee OA (Clément et al., 2018). Higher thigh mass and shorter leg length was also correlated to greater STA (Grimpampi et al., 2014) while few gender differences were found in rotation and translation of thigh skin markers (Gao & Zheng, 2008).

STA is a task-specific variable (Akbarshahi et al., 2010; Camomilla et al., 2015; Fiorentino et al., 2017). Previous research found STA to not be comparable between subjects completing walking trials; and attributed differences in STA to differences in walking speed (Barré et al., 2013). Another study found STA in cutting tasks to be larger than walking, due to sudden changes in motion and higher impact, causing the skin markers to be less sensitive to the underlying bone motion (Benoit et al., 2015; Benoit et al., 2006). High flexion postures are also affected by STA, specifically at the thigh and shank, due to increased soft tissue deformation, especially when thigh-calf contact occurs (Kingston & Acker, 2018). Differences in STA in the few studies that have measured it during different tasks indicated that STA cannot be generalized to a variety of different movements but would depend on the ROM and speed at which they are completed.

STA is a location-specific variable, meaning it varies between different segments and locations within segments. In terms of segment type, the thigh is subjected to higher STA than the shank and has been widely supported in the literature (Barré et al., 2013; Benoit et al., 2015; Clément et al., 2018; Dumas et al., 2014; Gao & Zheng, 2008; Kuo et al., 2011). In terms of marker location on the thigh segment, previous research found RMS error between skin marker and fluoroscopy (Akbarshahi et al., 2010; Barré et al., 2015) and differences in knee angles relative to marker position (Cockcroft et al., 2016; Gao & Zheng, 2008) varied based on location. Recommended positions will be elaborated on in Section 2.1.4.

2.1.3 Assessment of Skin Marker Accuracy

There are many ways to assess STA. Some techniques include modelling, while others use motion capture data and compare that to data obtained through bone images or bone-pins to allow for a gold standard comparison between skin marker readings and true bone motion. These methods determine the accuracy of skin marker readings in representing bone motion (Benoit et al., 2006). In the case that true bone motion cannot be estimated using fluoroscopy, other proposed gold-standard methods compared a landmark tracked by two different marker cluster coordinate systems (Buchman-Pearle & Acker, 2021; Schache et al., 2008).

STA has been quantified using simultaneous motion capture and imaging techniques. This approach compared motion capture data with dual fluoroscopy to determine differences in measured and reconstructed skin marker locations (Barré et al., 2015; Fiorentino et al., 2016; Kuo et al., 2011). This allows for a gold standard comparison to true bone motion (Fiorentino et al., 2016; Lin et al., 2016; Stagni et al., 2009; Tsai et al., 2011). When using this technique, previous research reported kinematic data as RMS error between fluoroscopy-based, and skin marker-based locations (Akbarshahi et al., 2010) and joint angles (Cappello et al., 2005;

Fiorentino et al., 2017; Stagni et al., 2009). Previous research also used a paired t-test between motion capture- and fluoroscopy-based joint angles/translations and moments at 10 % increments of a sit-to-stand task (Kuo et al., 2011), which allowed for the comparison of the data set across the movement. The inclusion of a gold standard fluoroscopy is very valuable; however, some limitations of using imaging techniques include the quality of the image during dynamic tasks (Akbarshahi et al., 2010), radiation exposure, lower sample size (Fiorentino et al., 2016, 2017, 2020), lower number of trials to be collected (Kapron et al., 2014) and lack of accessibility for many researchers. A wide range of subject BMI's should also be included when using this technique in research so errors in STA are not biased towards individuals of one particular size (Fiorentino et al., 2016, 2017).

Another method used to assess STA combines the use of motion capture data using skin markers and bone-pins. This is achieved by implanting pins into the bones, for example, the femur and tibia (Benoit et al., 2015) or the pelvis (Grimpampi et al., 2014) for lower extremity testing, where the exposed portion of the bone pin is instrumented with reflective markers to be captured by motion capture systems (Benoit et al., 2015). This method similarly calculates RMS error in kinematics derived from both skin and bone-pin markers (Andersen et al., 2010; Benoit et al., 2015). Bone-pin markers have been considered a gold standard (Peters et al., 2010); however, it may restrict the movement of soft tissues around the pin which could propagate to an error in measuring STA in skin markers (Blache et al., 2017; Fiorentino et al., 2017). Additionally, this is an invasive procedure with increased risk of complications.

A non-invasive approach to measure STA for a particular segment focused on using marker clusters of adjacent segments as a comparator, such as measuring thigh STA by comparing measures to that obtained from the pelvis or shank (Buchman-Pearle & Acker, 2021;

Schache et al., 2008). Previous research sought to evaluate thigh STA by defining a particular anatomical landmark in a reference posture and track its movement throughout motion trials from two coordinate systems (either anatomical or cluster-based) on the thigh and shank (Lucchetti et al., 1998; Ryu et al., 2009; Schache et al., 2008). This procedure then quantified STA through the difference in anatomical landmark location (Lucchetti et al., 1998; Schache et al., 2008) or through differences in calculated knee joint translations between the two coordinate systems throughout motion (Schache et al., 2008). To reduce any artifact from the comparator marker cluster, participants can complete activities where the joint between the comparator and segment of interest remains stiff, while the other joint adjacent to the segment of interest rotates in motion. For example, if the shank was defined as the comparator for evaluating thigh STA, a hip rotation task where the knee joint remained extended would be appropriate for measuring thigh STA (Ryu, 2012), while it is assumed that shank STA is very low (Schache et al., 2008). This method is not feasible in high flexion postures because the ankles, knees and hips flex in squatting and kneeling tasks, and there is increased deformation of the shank and thigh due to thigh-calf contact in high flexion (Kingston & Acker, 2018). Additionally, if evaluating thigh STA using the shank cluster for knee flexion movements, additional error may be introduced due to knee joint translations. For example, if a vector was created from a shank and thigh marker to the functional hip joint center, knee joint translation would cause additional movement of the vector and lead to uncertainty in the correct position of the functional hip joint center; negatively affecting estimation of STA. Therefore, a more appropriate segment to use as the comparator would be the pelvis. One study to date assessed thigh STA in high flexion movements using the pelvis cluster as the comparator. Buchman-Pearle and Acker (2021) reconstructed the functional hip joint center in the thigh cluster coordinate system and the pelvis cluster coordinate system.

This method also was founded on the assumption that pelvis STA was minimal (Buchman-Pearle & Acker, 2021), thus any error in the reconstruction of the femoral head center attributed to thigh STA (Buchman-Pearle & Acker, 2021). Although this method did not represent a true gold-standard, in that the hip joint center position is still indirectly located using a skin-mounted clusters, it is an accessible alternative to assess STA when radiographic methods or bone-pins are not readily available or feasible in data collections.

2.1.4 Compensation Methods for STA

The location specificity of STA logically indicates that there are better and worse locations to place markers/clusters and considering this in data collection can help reduce STA. Previous research identified thigh skin marker translational error in walking, step-up, and knee flexion tasks (Figure 2.1) by comparing skin marker position to the relative position as represented in the bone-embedded frame defined by fluoroscopy (Akbarshahi et al., 2010). RMS differences are reported in Table 2.1. Across all tasks, most markers located at the distal-anterior and distal-lateral thigh had the lowest translational errors while the mid-proximal-anterior thigh and distal-anterior thigh were the highest. However, there were mixed results for the distal-anterior placement of markers as, in some planes, this marker had the lowest translation error, while for other planes, it had the highest. No single marker location accounted for all the highest or lowest translational errors. Instead, the location with the lowest error was specific to the task and plane of motion (Akbarshahi et al., 2010).

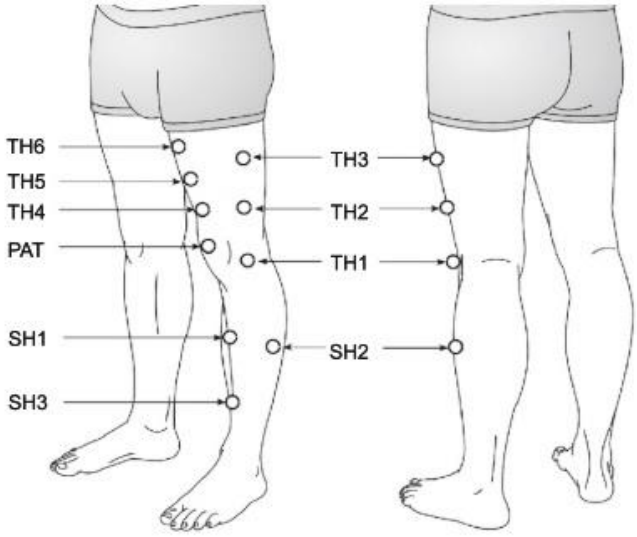


Figure 2.1: Thigh marker placements where TH6, TH5 and TH4 are located at the mid-proximal, mid-, and distal-anterior thigh, and TH3 and TH2 are located at the mid- and distal-lateral thigh, respectively (Akbarshahi et al., 2010).

Table 2.1: RMS error (mm) of skin marker position relative to the anatomical reference frame defined by fluoroscopy for walking, step-up and knee flexion tasks. Errors were reported in anterior/poster (A/P), proximal/distal (PROX/DIS), and medial/lateral (M/L) directions. Highest [1] and lowest [2] values are marked in the table. Adapted from Akbarshahi et al. (2010).

Marker	Location	Activities								
		Walking			Step-up			Knee flexion		
		A/P	PROX /DIS	M/L	A/P	PROX /DIS	M/L	A/P	PROX /DIS	M/L
TH2	Distal lateral	5.6 (1.6)	4.9 (3.4)[2]	9.2 (6.1)	5.3 (2.2)	9.3 (5.9)[2]	6.4 (1.9)[2]	4.1 (2.8)	3.0 (1.0)[2]	7.1 (2.2)[2]
TH3	Mid-lateral	5.1 (1.9)	5.4 (3.1)	8.9 (5.1)[2]	8.1 (2.1)[1]	11.3 (4.2)	8.2 (2.5)	7.5 (3.0)[1]	3.2 (1.5)	7.5 (2.1)
TH4	Distal anterior	4.6 (1.4)[2]	10.9 (3.1)[1]	10.9 (6.5)[1]	2.7 (1.5)[2]	17.5 (6.5)[1]	10.4 (6.1)	3.4 (1.6)[2]	19.5 (6.4)[1]	8.8 (2.9)
TH5	Mid-anterior	5.7 (2.1)	10.2 (2.6)	10.6 (6.3)	3.5 (1.6)	15.1 (4.2)	14.6 (10.8)	3.7 (1.7)	13.2 (2.4)	8.6 (3.6)
TH6	Mid-proximal anterior	6.5 (2.6)[1]	9.9 (3.0)	10.0 (4.4)	5.1 (2.3)	16.8 (2.8)	19.4 (14)[1]	5.0 (1.7)	9.2 (0.5)	10.7 (4.5)[1]

Locations are specific to the acknowledged study and do not represent locations for this thesis.

Other research found similar translational and rotational patterns among three (anterior, anterolateral, and lateral) vertical columns of markers during walking tasks by tracking the movement of markers/clusters to a reference point (T7) (Figure 2.2) (Gao & Zheng, 2008); These results agree with gait data presented by Akbarshahi et al. (2010) for A/P translation, where the mid-proximal-anterior marker had the greatest error, but did not agree in other directions. In terms of the lowest translational error, both studies support the use of mid- and distal- clusters; however, there were mixed results on anterior, anterolateral and lateral placement except for proximal/distal (PROX/DIS) error which supported the use of a distal-lateral cluster (Akbarshahi et al., 2010; Gao & Zheng, 2008).

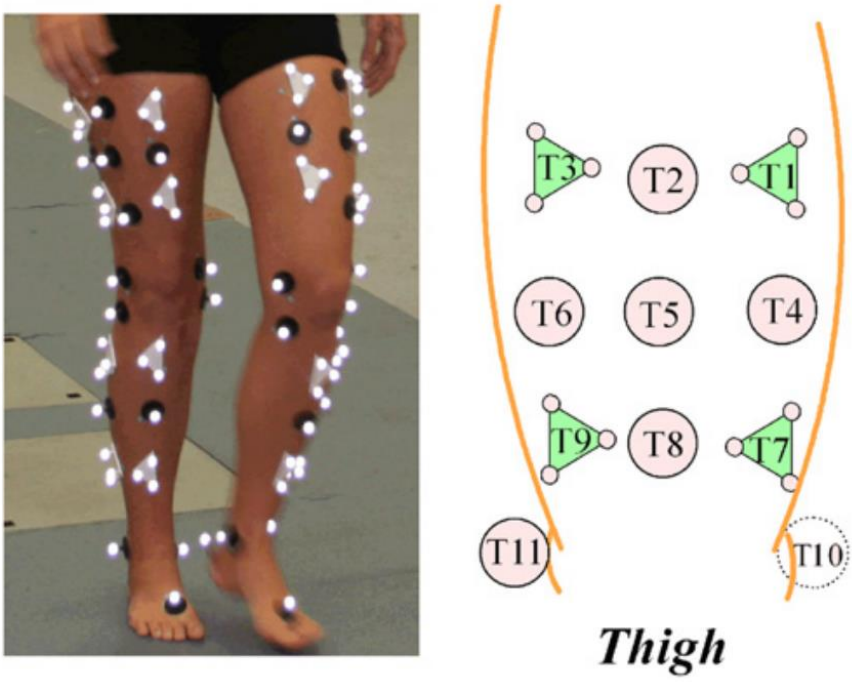


Figure 2.2: Thigh skin marker (T2, T4, T5, T6, and T8) and cluster (T1, T3, and T9) locations to describe greatest translational and rotational error from a reference point (T7) (Gao & Zheng, 2008).

STA can also be described in terms of error in joint angles. Previous research quantified error in hip and knee angles due to STA across markers located at the proximal and distal third of the lateral thigh during walking trials, and found less STA, reduced valgus-varus range, and increased reliability in hip joint angles with the proximal markers (Cockcroft et al., 2016). This would oppose the previous studies that demonstrated the mid- and distal- clusters had the lowest translational errors (Akbarshahi et al., 2010; Gao & Zheng, 2008); however, two of these studies did not include a gold standard fluoroscopy comparison (Cockcroft et al., 2016; Gao & Zheng, 2008). Another recent study reconstructed the hip joint center in high flexion movements using different thigh marker cluster positions and compared that against the hip joint center determined from the pelvis cluster (Buchman-Pearle & Acker, 2021). Although this study did not analyze proximal thigh cluster placements, clusters located on the mid-anterolateral and mid-anterior thigh demonstrated the best reconstruction of the hip joint center, as determined using a Bland-Altman analysis (Buchman-Pearle & Acker, 2021). Additionally, STA of the thigh was noticeably altered at flexion angles at the initiation of thigh-calf contact, particularly affecting the lateral thigh clusters (Buchman-Pearle & Acker, 2021). Cluster placement influenced the accuracy of data and can be accounted for in data collections by choosing the most appropriate location.

The use of multiple calibrations is another method used to minimize errors in STA where calibration trials in two extremes of movements are collected. This method improved the estimation of anatomical landmarks, so that they more closely matched true landmark position; however, it violates the rigid body assumption because the cluster movement defining anatomical landmarks in both extremes of postures were accounted for throughout the calibration technique (Cappello et al., 1997, 2005). One main advantage of this method is the use of joint

angles and time to determine the appropriate reconstruction of skin markers and anatomical landmarks because STA increases with joint angle and will have different weightings of error throughout the motion (Cappozzo et al., 1996; Clément et al., 2018).

While post-processing methods to compensate for STA are outside the scope of the proposed work, such methods include the use of joint constraints and modelling. Joint constraints have been implemented in data processing methods of the lower extremity to minimize errors in skin marker data; however, there are mixed results with this method. Previous research implemented spherical and revolute knee joint constraints but found this processing method did not improve kinematic measures and in some cases, worsened them (Andersen et al., 2010). Constraining the hip joint was also not supported and generated greater errors in the prediction of the hip joint center (Fiorentino et al., 2016). However, one study that found joint constraints improved kinematic RMS error, used joint constraints at the hip, knee, and ankle, and were designed to be subject-specific. This means the femur and tibia were constructed based on subject-radiographs rather than standardized joint geometry from the literature before going through optimization procedures that reduced errors in marker-based coordinates (Clément et al., 2015). This advanced technique, however, would complicate data processing methods.

Modelling techniques can isolate STA and successfully reduce the error in data as if using other pre-defined methods. One study developed a linear model to estimate the movement of skin markers in various tasks and used principal component analysis to determine what components were most important to consider in the model. Opposed to the traditional assumption that skin markers move rigidly with segments, the model allowed for the non-rigid movements of the clusters and largely reduced thigh artifact that was experimentally quantified vs. modelled in walking (27 to 5.1 mm), cutting (22.7 to 1.9 mm), and hopping tasks (16.2 to 3.5 mm) (Andersen

et al., 2012). Another study linearly modelled the rigid component of STA (e.g. translation and rotation as defined above) and subtracted this from collected data to improve the accuracy of marker readings. They found this improved the accuracy of measured kinematics, emphasizing the rigid component contributed to most differences in the outcome measures (Camomilla et al., 2015; Cappello et al., 2005). Both these models considered the rigid component of STA, which has a greater contribution to error as opposed to non-rigid components (Andersen et al., 2012; Benoit et al., 2015). A limitation, however, is the assumption of linearity as the behavior of STA vibrations during different activities is unknown (Camomilla et al., 2015).

2.1.5 Quantification of Kinematic Errors

By assessing the accuracy of skin markers in motion capture systems, researchers can quantify the sensitivity of measures and account for any differences in future work. Table 2.2 compares kinematic outcomes of skin marker-based motion capture against gold standard measures for knee flexion, treadmill walking, step-up, quasi-static squatting and sit-to-stand tasks. The outcomes between motion capture and a gold standard measure are represented by RMS error, absolute/mean difference, and/or paired t-test results to indicate to the effect of STA. RMS error between motion capture and fluoroscopy gold standard measures across all tasks for knee joint rotations were as high as 8.3 degrees, 8.1 degrees and 7.2 degrees in FL/EX, internal/external (INT/EXT) and AB/AD, respectively, while translational errors were as high as 13.1 mm, 13.2 mm and 13.1 mm in A/P, medial/lateral (M/L) and PROX/DIS. Absolute difference between skin marker and bone-pin knee joint rotations were as high as 2.8 degrees, 5.4 degrees, and 6.7 degrees in FL/EX, INT/EXT, and AB/AD, respectively, while translational differences were 12.0 mm, 8.3 mm, and 5.9 mm in A/P, M/L and PROX/DIS. Additionally, significant differences in all rotations and translations along the axes were found amongst all

activities in different studies. The quantification of error can indicate how data are affected by STA.

Table 2.3: Summary table of study participant characteristics, methodology (experimental design and statistics), kinematic and kinetic outcomes of soft tissue artifact (STA), and limitations. Experimental design highlights equipment used to compare skin marker motion capture data to a gold standard and activities performed in the study.

Author(s)	Participant Characteristics	Experimental Design	Statistics	STA Outcome(s)	Limitation(s)
(Akbarshahi et al., 2010)	4 males **Age:** 30 (3) years **BMI:** 22.4 (1.7) kg/m^2	**Equipment**: skin marker motion capture and x-ray fluoroscopy **Activity**: open-chain knee flexion, treadmill walking and step-up tasks	Average RMS error of knee joint rotations (deg) between x-ray and skin markers.	**Open-chain knee flexion:** 8.3 FL/EX, 7.2 AB/AD, 6.4 INT/EXT **Treadmill walking:** 4.5 FL/EX, 4.45 AB/AD, 5.91 INT/EXT **Step-up:** 4.3 FL/EX, 4.9 AB/AD, 7.2 INT/EXT	Low sample size.
(Barré et al., 2015)	11 females, 8 males **Age:** 70 (6) years **BMI:** 28.4 (4.1) kg/m^2	**Equipment**: skin marker motion capture and biplane fluoroscopy **Activity**: treadmill walking (evaluated clusters on areas of thigh that were most and least affected by STA)	Average RMS error of knee joint rotations (deg) and translations (mm) between biplane fluoroscopy and various skin marker clusters.	**Most affected:** 3.5 (1.7) FL/EX, 4.4 (2.2) AB/AD, 6.1 (1.7) INT/EXT, 13.1 (4.0) A/P, 13.2 (8.5) M/L, 13.1 (4.7) PROX/DIS **Least affected: (used in comparisons between studies)** 3.5 (0.8) FL/EX, 3.5 (1.2) AB/AD, 8.1 (2.2) INT/EXT, 5.1 (2.8) A/P, 4.8 (2.2) M/L, 6.1 (2.8) PROX/DIS	Elderly population with prosthetic knees.
(Barré et al., 2013)	11 females, 8 males **Age:** 70 (6) years **BMI:** 28.4 (4.1) kg/m^2	**Equipment**: skin marker motion capture and biplane fluoroscopy **Activity**: treadmill walking	Median difference of translation amplitude (mm) of the thigh cluster between bi-plane fluoroscopy and skin marker location as % completion of gait cycle. Broke STA component into rigid movement of cluster (rotation and	**0-60% gait cycle:** 5.9 STA, 5.4 rigid STA, 2.7 local STA **60-100% gait cycle:** 13.0 STA, 13.0 rigid STA, 4.4 local STA	Elderly population with prosthetic knees. Slow walking speed compared to normal.

			translation error) and local deformation.		
(Benoit et al., 2006)	8 males **Age:** 26 years **BMI:** 24.6 kg/m^2	**Equipment:** skin marker and bone pin motion capture. Digitized anatomical landmarks through biplanar calibration box **Activity:** walking and lateral cutting	Absolute difference between knee kinematics derived from skin markers and pin markers at heel strike (HS), mid-stance (MS), and toe-off (TO). Only significant rotation (deg) and translation (mm) values were reported ($p < 0.05$). Standard error of estimate of predicting knee kinematics using skin markers.	**Walking** **Foot strike:** 2.8 (2.6) FL/EX, 2.8 (2.0) INT/EXT, 7.7 (4.4) A/P, 5.0 (2.9) PROX/DIS **Mid Stance:** 2.4 (2.0) FL/EX, 5.5 (3.1) M/L, 3.3 (2.4) PROX/DIS **Toe off:** 4.4 (3.2) AB/AD, 13.0 (5.0) A/P, 5.0 (2.5) PROX/DIS **Standard error:** 2.5 FL/EX, 3.6 AB/AD, 2.9 INT/EXT, 6.8 A/P, 5.9 M/L, 2.7 PROX/DIS **Cutting** **Foot strike:** 6.7 (5.4) AB/AD, 5.4 (4.2) INT/EXT, 5.6 (5.1) A/P, 6.3 (4.0) PROX/DIS **Mid Stance:** 5.9 (3.1) AB/AD, 5.4 (4.0) INT/EXT, 6.7 (4.4) A/P, 5.9 (4.5) M/L, 5.6 (3.8) PROX/DIS **Toe off:** 3.3 (1.8), INT/EXT, 8.3 (6.2) PROX/DIS **Standard error:** 6.3 FL/EX, 4.5 AB/AD, 3.0 INT/EXT,5.5 A/P, 8.0 M/L, 7.1 PROX/DIS	Low sample size.
(Clément et al., 2018)	**Group 1:** 5 females, 4 males **Age:** 54.2 (9.6) years **BMI:** 24.8 (2.3) kg/m^2 **Group 2:** 7 females, 1 male **Age:** 59.0 (5.9) years	**Equipment:** biplane radiography of bone and exoskeleton markers fitted to lower extremity **Activity:** quasi-static squatting at 0, 30, 40, 50, and 60 deg knee flexion	Mean error of exoskeleton marker knee kinematics (deg) compared to bone.	**Group 1:** 1.0 FL/EX, 4.9 AB/AD, 5.5 INT/EXT **Group 2:** 0.7 FL/EX, 2.3 AB/AD, 4.4 INT/EXT	Markers adhered to an exoskeleton on the lower extremity.

	BMI: 34.3 (2.7) kg/m^2				
(Kuo et al., 2011)	10 subjects (M/F not specified) **Age:** 77.7 (6.5) years **BMI:** 27.8 (1.6) kg/m^2	**Equipment:** skin marker motion capture and fluoroscopy imaging system **Activity:** sit-to-stand	Paired t-test between skin marker and fluoroscopy knee kinematic (deg and mm) measures obtained through bottom-up inverse dynamics at 10% intervals of sit-to-stand cycle. Cycles of significance ($p < 0.05$) will only be reported. Differences in moments are reported as a percentage of the max moment calculated using fluoroscopy.	**Kinematics** **0-70%:** FL/EX, AB/AD, INT/EXT, A/P, PROX/DIS **80-90%:** A/P, PROX/DIS FL/EX, AB/AD and INT/EXT rotations via skin markers were underestimated compared to fluoroscopy. A/P and PROX/DIS knee joint center translations were overestimated.	Low sample size and participants had total knee replacement on tested limb.
(Cappello et al., 2005)	2 subjects (sex not specified) **Age:** 67 and 64 years **BMI:** 24 and 22 kg/m^2	**Equipment:** skin marker motion capture and fluoroscopy imaging system. Use of single and double calibration **Activity:** sit-to-stand	Mean RMSE between gold standard fluoroscopy and skin marker knee joint rotations (deg) and translations (mm) using single and double calibrations.	**Single Calibration** 6.4 FL/EX, 3.7 AB/D, 3.7 INT/EXT, 12.9 A/P, 6.3 M/L, 11.9 PROX/DIS **Double Calibration** 1.5 FL/EX, 1.4 AB/AD, 1.6 INT/EXT, 2.0 A/P, 2.1 M/L, 2.8 PROX/DIS	Low sample size, sex not specified, and participants had total knee replacement on tested limb.
(Lucchetti et al., 1998)	3 males **Ages:** 30, 27 and 45 years **Body Mass:** 67, 78 and 75 kg	**Equipment:** skin marker motion capture. Use of marker clusters on thigh and shank, where the shank was the comparator for thigh STA **Activity:** counterclockwise leg swing from hip extension to hip flexion, with knee	RMSE in angular (deg) and linear (mm) displacement of knee joint kinematics obtained from the thigh and shank marker clusters. Femoral landmarks were rigidly associated with the shank markers, to be able to evaluate thigh STA.	**Subject 1:** 10.9 A/P, 3.9 M/L, 4.9 PROX/DIS, 3.7 FL/EX, 1.3 AB/AD, 4.4 I/E **Subject 2:** 5.8 A/P, 5.8 M/L, 3.3 PROX/DIS, 1.9 FL/EX, 1.2 AB/AD, 4.7 I/E **Subject 3:**	Low sample size, one subject had a knee prosthesis, and no true gold-standard measure.

		extended (resembled swing phase in gait)		13.2 A/P, 10.1 M/L, 4.2 PROX/DIS, 5.0 FL/EX, 2.5 AB/AD, 1.1 I/E	
(Schache et al., 2008)	20 subjects (5M/15F) **Age:** 20.8 (4.1) years **BMI:** 22.2 kg/m^2	**Equipment:** skin marker motion capture. Use of marker clusters on thigh and shank, where shank was the comparator for thigh STA **Activity:** gait	Mean absolute difference (mm) in knee joint center translations calculated from the shank and thigh coordinate systems, where the shank was the comparator segment, and the thigh was of interest to measure STA error.	**Distal thigh rigid cluster:** 8.27 A/P, 4.29 M/L, 3.90 PROX/DIS	No true gold-standard and low subject BMI.

Flexion/extension (FL/EX), abduction/adduction (AB/AD), internal/external (INT/EXT), anterior/posterior (A/P), medial/lateral (M/L) and proximal/distal (PROX/DIS)

As previously discussed in STA modelling, the rigid component contributed most to differences in kinematic outcomes (Andersen et al., 2012; Benoit et al., 2015; Camomilla et al., 2015; Cappello et al., 2005). Previous research used skin marker motion capture and biplane fluoroscopy to separate STA into rigid and non-rigid vector components in gait trials (Figure 2.3) and similarly found the rigid component dominated over local deformation (non-rigid component) (Barré et al., 2013). Rigid components represent rotations and translations and was addressed in the subsequent papers to define kinematic and kinetic errors.

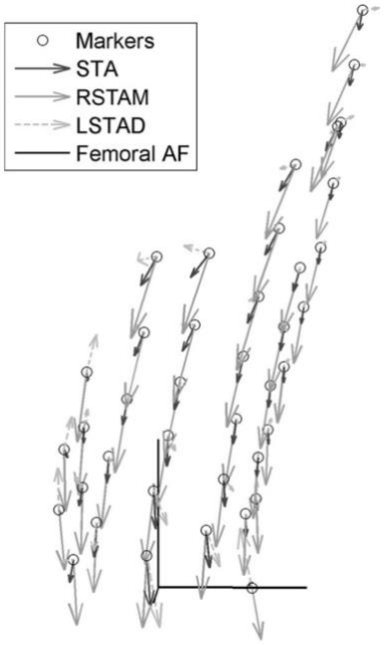


Figure 2.3: Rigid (RSTAM) and local deformation (LSTAD) vector components of soft tissue artifact on the thigh for a representative subject at mid-swing of gait. Marker translations depicted from gold standard fluoroscopy measures Barré et al. (2013).

Walking was the most frequent task assessed and both rotational and translational errors in knee kinematics were reported; however, results from fluoroscopy and bone-pin studies did not agree. Previous research that quantified RMS errors in walking using skin marker and fluoroscopy data found INT/EXT rotational errors to be highest, while FL/EX or AB/AD results

were the least susceptible depending on the study and marker location (Akbarshahi et al., 2010; Barré et al., 2015). Using the same methodology, PROX/DIS translational error was highest, and M/L was lowest but was only determined in one study (Barré et al., 2015). Other research using bone-pins reported standard error in estimating rotational and translational kinematics throughout most of the gait cycle was largest in AB/AD and A/P, and the lowest in FL/EX and PROX/DIS (Benoit et al., 2006). Two different methods of measuring STA yielded inconsistent results in which rotation and translation had the greatest and least STA error; therefore, differences in the methodology could potentially explain differences in outcome measures. Bone pins may also restrict the movement of soft tissues around the pin which could propagate to an error in measuring STA in skin markers (Blache et al., 2017; Fiorentino et al., 2017).

Other studies using fluoroscopy measured joint kinematics in open-chain knee flexion, step-up, quasi-static squatting, and sit-to-stand tasks. These activities are also subject to rotational and translational errors due to STA, and presented unique variance due to the task specificity piece of STA. Results can be similar for tasks with comparable flexion angles and this was seen with rotational errors in quasi-static squatting tasks (Clément et al., 2018) and a representative subject completing a step-up task (Akbarshahi et al., 2010). Both reached 60 degrees knee flexion. A representative subject in the cutting task reached close to 50 deg flexion (Benoit et al., 2006) while one in open-chain knee flexion reached close to 100 deg flexion (Akbarshahi et al., 2010). Although different tasks, the highest and lowest rotational errors were similar; however, this could be due to random chance.

2.2 High Knee Flexion

2.2.1 Overview of High Flexion and OA

In high knee flexion postures, the flexion knee angle exceeds 120 degrees (Zelle et al., 2009) and experiences increased loading (Nagura et al., 2002). These postures are adopted in occupational work such as childcare, roofing, construction (Cooper et al., 1994), floor laying (Jensen et al., 2000), or daily activities such as gardening, religious practices, and exercise (Li et al., 2004) in the form of kneeling and squatting. In adopting these postures, previous research identified sex differences in peak knee adduction and internal rotation for both squatting and kneeling, while differences in range were found for adduction and internal rotation for squatting, and adduction for kneeling (Han et al., 2014). Other studies have used kinematics to classify differences between two groups; however, the accuracy of kinematic measures have not yet been quantified in high flexion.

There is a need to study high flexion activities and kinematics associated with these movements, to best understand knee pathology (Hamai et al., 2013). The adoption of high flexion postures is associated with increased risk of knee OA development (Amin et al., 2008), especially in exposures with a combination of kneeling, squatting and lifting, and prolonged postures (Cooper et al., 1994; McMillan & Nichols, 2005; Zhang et al., 2004). The development of knee OA and responses to loading depend on age (Felson & Zhang, 1998), sex, genetics (Jordan et al., 2000), knee alignment (Besier et al., 2005; Sun, 2010), joint geometry (Besier et al., 2005), tissue properties, and cartilage thickness (Besier et al., 2005). Knee valgus and varus can shift tibial contact forces from lateral to medial where a varus alignment, or greater adduction moment, is associated with the increased risk (Yang et al., 2010).

In addition to the increased risk factors, movements of those with knee OA differs from the movement of healthy individuals. Previous research found knee OA affected movements of AB/AD and A/P displacement, and INT/EXT rotation (Hamai et al., 2009). Knowing that

kinematics change with disease status, and that high flexion increases OA risk, it is important that we quantify the accuracy of these kinematics in high flexion.

2.2.2 Knee Joint Motion

When the knee moves from extension to high flexion, the tibia will internally rotate because the lateral femoral condyle translates posteriorly along the lateral meniscus while the medial side remains constrained (Yao et al., 2008). These geometric properties cause the medial meniscus to undergo higher strain at the anterior horn and impingement of the posterior horn compared to the lateral side (Yao et al., 2008). The tibia also undergoes anterior translation; however, previous research that tested tibial motion using cadaver specimens found that from 120 degrees to 150 degrees flexion, the tibia had much less translation and rotation at/near angles of 150 degrees (Li et al., 2004). The knee joint is more constrained beyond 120 degrees which leads to impingement of the tissues and increased stress on the posterior compartment of the knee (Li et al., 2004; Nagura et al., 2001; Williams & Logan, 2004) to greater extents compared to gait activities because this is in high flexion (Nagura et al., 2001). Cumulatively, damage to the knee joint cartilage can reach a point where it becomes more susceptible to knee OA (Felson & Zhang, 1998).

Chapter 3 – Investigation of a Double Calibration Approach

3.1 Objective and Hypothesis

To investigate the effectiveness of a double calibration technique applied to the thigh segment to reduce error in femoral head center reconstruction in high flexion motion trials in a healthy population. The double calibration technique was evaluated with respect to the femoral head center reconstructed using a mid-anterolateral thigh cluster and the hip joint center location reconstructed using the pelvis cluster. The hip joint center was the gold standard comparator, determined using the star-arc pattern which best predicted the hip joint center location (Camomilla et al., 2006; Ehrig et al., 2006), and was defined in the upright standing calibration posture. The femoral head center (from each of the single and double calibration techniques) was compared to hip joint center locations using two analyses: RMS error between the femoral head center and hip joint center locations calculated in the global x, y, and z directions, and resultant error (distance between the femoral head center and hip joint center locations). Bland-Altman analyses were also carried out on the resultant error and provided a bias and limits of agreement between the two locations. Paired t-tests were conducted between the single and double calibration, on the mean RMS errors and resultant error from the Bland-Altman analysis between the femoral head center and hip joint center. Paired t-test outcomes determined the calibration method least susceptible to rigid STA.

It was hypothesized that the double calibration method would have significantly lower average RMS error between the femoral head center and hip joint center locations across all high flexion trials. It was also hypothesized that the bias from the double calibration would be significantly lower and would have narrower limits of agreement. In previous work, double calibrations in the two extreme postures of a movement significantly improved knee joint

translations and rotations when measured against that obtained through fluoroscopy, in step-up step-down, sit-to-stand/stand-to-sit and knee flexion tasks (Cappello et al., 2005).

3.2 Methods

A secondary analysis was conducted for this thesis using previously collected lower extremity data for high flexion motion trials and only the procedures relevant to this analysis were presented in the methods (Buchman-Pearle & Acker, 2021). Participants were screened for exclusion criteria and completed an informed consent document if they chose to proceed with the study. Demographic and anthropometric information on participants (age, sex, height, body mass, thigh length, and thigh circumference at mid-thigh) were collected, and the researcher provided verbal instructions and demonstrations of the squatting and kneeling tasks included in the study protocol. Participants had an opportunity to practice all tasks and ask questions. Squatting (flatfoot and heels-up) and kneeling (dorsiflexed and plantarflexed) movements were then recorded throughout data collection, and the investigation of a double calibration method was completed by comparing the femoral head center location reconstructed from the double calibration and single calibration using the mid-anterolateral cluster, against the hip joint center location reconstructed using the pelvis cluster. This thigh cluster location reconstructed the hip head center best, as determined using a Bland-Altman analysis (Buchman-Pearle & Acker, 2021).

An overview of all methodological procedures was represented in Figure 3.1. Motion capture was used in the data collection procedures along with a variety of software programs to process all data.

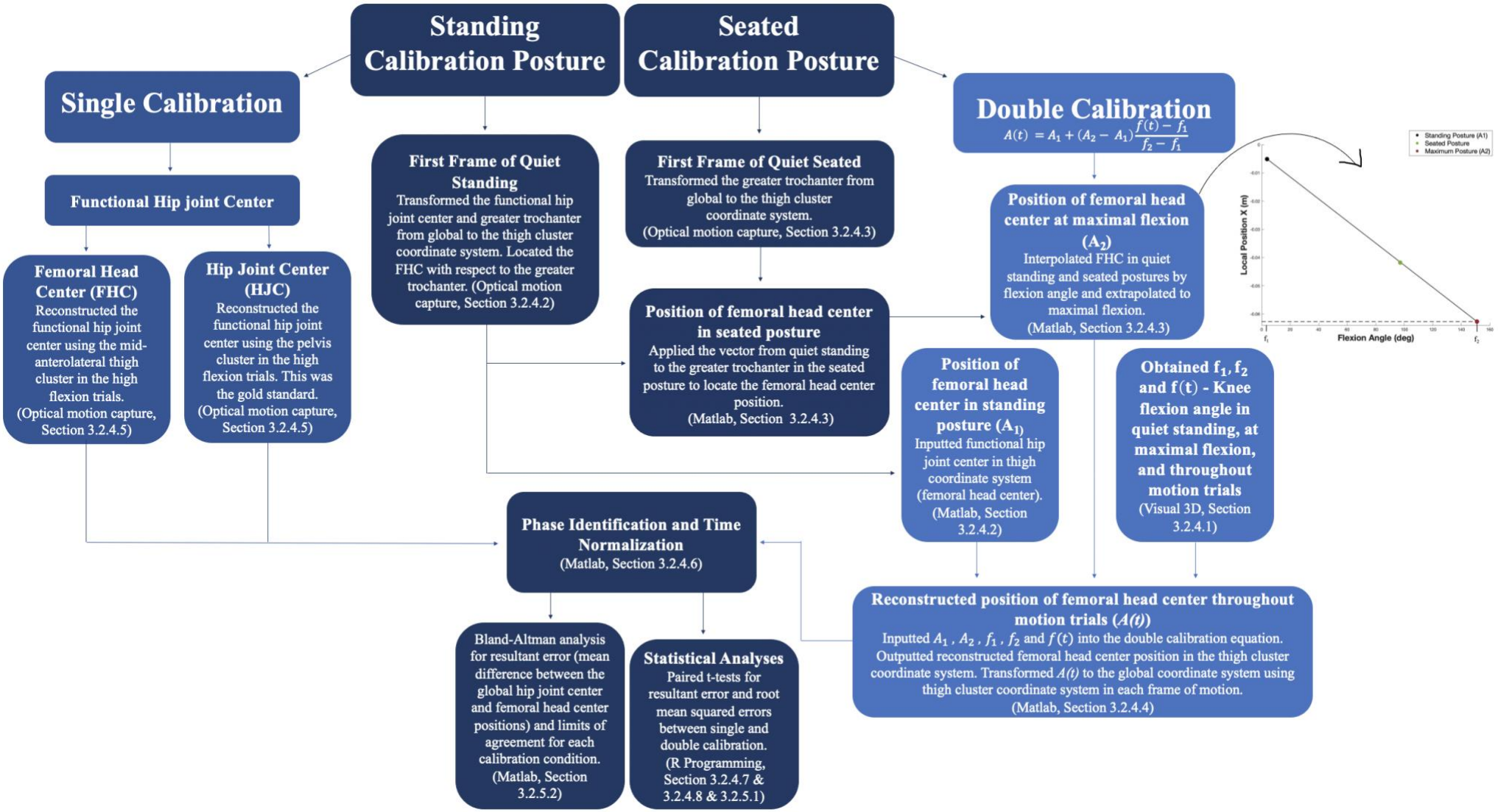


Figure 3.1: Overview of methodological procedures for data collection, data processing, data analysis, and statistical analysis.

3.2.1 Participants

Lower limb kinematic data for thirty-three participants were used for this study (Table 3.1). Using the G*Power 3 program (Faul et al., 2007), a one-tailed compromise power analysis for paired t-tests was calculated using the sample size of 33, projected high effect size of 0.8 (Cohen, 1992), and β to α ratio of 4 (Faul et al., 2007). The outputted power for this test ($1 - \beta$) was 0.97, with an alpha level of 0.007, and beta of 0.03.

All participants in this study met the inclusion criteria, which included no allergies to adhesives, no treatment for any medical condition/injury to the lower back or lower extremities, and no pain/discomfort in squatting and kneeling tasks. After participants were deemed eligible for participation, they provided their informed consent which was approved by the University of Waterloo's Research Ethics Board.

Table 3.1:Anthropometric data for participants included in the study. Data were presented as mean (+/- standard deviation).

Parameter	Male (n = 17)	Female (n = 16)	All (n = 33)
Age (yrs)	21.2 (2.5)	20.6 (2.0)	20.9 (2.3)
Body Mass (kg)	77.8 (10.6)	62.8 (11.1)	70.5 (13.1)
Height (m)	1.79 (0.06)	1.65 (0.06)	1.72 (0.11)
Thigh length (cm)	44.7 (2.2)	43.0 (3.0)	43.9 (2.7)
Mid-thigh circumference (cm)	53.4 (4.1)	50.6 (4.9)	52.0 (4.7)

3.2.2 Instrumentation

Kinematic data were collected at 50 Hz using a 6 tri-camera optical motion capture system (Optotrak, Certus/3020, NDI, Waterloo, ON) (Buchman-Pearle & Acker, 2021). Participants were instrumented with rigid marker clusters located at the pelvis, and bilaterally on

the thighs, distal-lateral shanks, and lateral feet using Velcro straps and medical grade tape to limit any movement. The leg to which the mid-anterolateral cluster was affixed was used for this analysis (Figure 3.2), because this cluster location had the lowest difference in limits of agreement when reconstructing the hip joint center against the pelvis cluster (Buchman-Pearle & Acker, 2021). This cluster was placed slightly proximal to the mid-thigh, which was calculated at 60% of the length from the lateral femoral condyle to the greater trochanter. The appropriate anterolateral placement was approximately halfway between two imaginary vertical lines, one that connected the greater trochanter to the lateral femoral condyle and the second that connected the anterior superior iliac spine to the patella (Buchman-Pearle, 2020). After marker clusters were administered, participants completed an upright standing and seated calibration trial, followed by functional hip (Camomilla et al., 2006) and knee trials. The functional hip joint center trial included FL/EX (~60 deg) and AB/AD (~30 deg) movements in a star-arc pattern (Camomilla et al., 2006; Kainz et al., 2015; Sangeux et al., 2011). This trial was demonstrated by the researcher to ensure participants understood the appropriate star-arc movements of this trial. Participants were also corrected if their ROM did not approximate ~60 degrees FL/EX and ~30 deg AB/AD. A maximal flexion posture was originally considered as the second calibration posture to follow procedures of previous double calibration work (Cappello et al., 1997, 2005);

however, palpating landmarks in full flexion proved unreliable and inaccurate so the seated calibration posture was used instead.

Figure 3.2: Position of mid-anterolateral (red) thigh cluster location. Adapted from Buchman-Pearle and Acker (2021).

3.2.3 Experimental Protocol

After informed consent was received, anthropometric data were collected, which included age, sex, height, body mass, thigh length, and thigh circumference at the mid-thigh (Section 3.2.1, Table 3.1). Mid-thigh circumference was measured at 60% of thigh length from the lateral femoral condyle to the greater trochanter. An overview of the experimental protocol was then provided to participants, and they had an opportunity to ask questions or clarify any information on the study.

Participants were then instrumented with the rigid marker clusters and the researcher demonstrated all the high flexion tasks in data collection and provided instruction. Instruction was given such that participants were aware they had to reach maximal flexion after descending into the high flexion tasks, had to hold the static phase of the task for approximately 5 seconds, and then raise to ascend. All phases of the tasks were performed at a comfortable pace. Participants were given an opportunity to practice all movements until they were comfortable in

performing the tasks. Four tasks from the data collection were extracted for this thesis, which included flatfoot squatting, heels-up squatting, dorsiflexed kneeling and plantarflexed kneeling. The squatting tasks began in upright standing, where participants descended into maximal flexion held the static phase, and then ascended. The kneeling tasks began with participants on their knees (approximately 90 degrees knee flexion), where they sat back onto their heels to reach maximal flexion, held the static posture, and then returned to 90 degrees knee flexion (Figure 3.3). Participants performed 5 repetitions of each high flexion task in a randomized order.

Figure 3.3*: Movements in experimental protocol: (A) flatfoot squat, (B) heels-up squat, (C) plantarflexed kneel, and (D) dorsiflexed kneel. For both squatting movements, A.1 and B.1 represented the frontal plane view at full flexion, while A.2 and B.2 represented a sagittal view at full flexion. For the kneeling movements, C.1 and D.1 represented the start and end position of the kneel, while C.2 and D.2 represented the kneel at full flexion.*

3.2.4 Data Processing

Motion capture kinematic data were processed in Visual 3D (C-Motion, Inc., Germantown, MD). The upright standing calibration model was applied to the dynamic motion trials and knee joint angles were calculated for the sagittal plane, following a Z-X-Y (M/L, A/P, INT/EXT) rotation sequence for the knee joint (Grood & Suntay, 1983). To process the seated calibration posture, the model template from the upright standing posture was applied to the seated posture. Kinematic data were then interpolated to fill gaps up to 10 frames in length (200 ms) using a spline technique of 3rd order polynomial (Howarth & Callaghan, 2010), and padded with 50 points (1 second) using the reflection method (Howarth & Callaghan, 2009). Data were also filtered using a low pass 2nd order Butterworth filter with a cutoff frequency of 6 Hz (Howarth & Callaghan, 2010; Winter, 2009). The greater trochanter, lateral femoral condyle, five pelvis markers, four mid-anterolateral thigh markers, and knee FL/EX angles from both calibration trials and motion trials were exported into Matlab (The Mathworks, Natick, MA, R2019b) for the remainder of processing. The functional hip joint center applied to the quiet standing trial and motion trials were also exported into Matlab (The Mathworks, Natick, MA, R2019b). The functional hip joint center trial was not able to be applied to the seated calibration posture; however, its position was determined using the procedures described in Section 3.2.4.3.

3.2.4.1 Double Calibration

Data processing for the double calibration was completed in Matlab (The Mathworks, Natick, MA, R2019b) using equation 1:

$$A(t) = A_1 + (A_2 - A_1)\frac{f(t) - f_1}{f_2 - f_1} \quad (1)$$

where $A(t)$ is the reconstructed position of the landmark of interest in the cluster coordinate system at each frame of the motion trial, A_1 is the position of the landmark of interest

in the first calibration posture in the cluster coordinate system, A_2 is the position of the landmarks of interest at maximal flexion in the cluster coordinate system, $f(t)$ is the knee flexion angle at each frame of the motion trials determined from the single calibration, f_1 is the knee flexion angle in the first calibration posture, and f_2 is the knee flexion angle at maximal flexion (Cappello et al., 2005). The FL/EX angle was chosen as the weighting factor in the double calibration equation because it was most robust against STA (Stagni et al., 2005). For this thesis, the first calibration posture was taken in upright standing and the second was taken in a seated posture, and the landmark of interest for reconstruction was the femoral head center. In knowing that the seated calibration posture was not at maximum flexion, A_2 was predicted through linear interpolation of the femoral head center from the standing and seated calibration posture and was extrapolated to f_2, the maximal flexion angle in each motion trial. f_1, f_2 and $f(t)$ were calculated in the above step in Visual 3D (C-Motion, Inc., Germantown, MD) (Section 3.2.4).

Each step described in the remainder of the data processing section are broken down into determining A_1, A_2 and A(t) in the double calibration equation. Data processing procedures are also described as if being completed for a single participant.

3.2.4.2 Calculation of A_1

For this thesis, A_1 is the position of the femoral head center in the thigh cluster coordinate system during the standing calibration trial (this is where the femoral head center would remain fixed when using a single calibration approach). Using the first frame of the quiet standing trial, a transformation matrix was created from the global coordinate system to the thigh cluster coordinate system (Appendix A: Thigh Cluster Coordinate System). The first frame in quiet standing was chosen because, to compare the single and double calibration methods in the thigh cluster coordinate system, the position of the femoral head centers must be coincident so that any

divergence would not be due to differences in start position. Participants were also as still as possible in the quiet standing trial; therefore, it was not necessary to take the average transformation matrix over the quiet standing trial. The functional hip joint center in the first frame of quiet standing was then transformed from the global coordinate system to the thigh cluster coordinate system (FHC_L) using equation 2:

$$FHC_L = TCS_R * (FHJC_G - TCS_o) \quad (2)$$

where TCS_R is the rotation matrix from the global coordinate system to the thigh cluster coordinate system in the first frame of quiet standing, $FHJC_G$ is the functional hip joint center in the global coordinate system in the first frame of quiet standing, and TCS_o is the origin of the thigh cluster coordinate system in the first frame of quiet standing. FHC_L was assigned as variable *A1* in the double calibration equation. The construction of the coordinate systems are elaborated on in Appendix A.

3.2.4.3 Calculation of A_2

A2 is the position of the femoral head center in the thigh cluster coordinate system when the participant is at maximal flexion. Before locating this landmark, some additional processing of the standing and seated calibration trial was needed because the femoral head center cannot be palpated. The goal was to have a best estimate of the femoral head center location in the seated calibration posture. Therefore, the location of the femoral head center in the thigh cluster coordinate system with respect to the greater trochanter (a palpable landmark) was first determined in the standing calibration posture, for later use in the seated calibration posture. To palpate the greater trochanter, participants were instructed to axially rotate their hip with their heel on the ground (as if squishing a bug). The bone feature that was most prominent at the location of the greater trochanter was palpated.

Prediction of the femoral head center in the quiet seated trial is illustrated in Figure 3.4. Using the same process as equation 2, the greater trochanter in the first frame of the quiet standing trial was transformed from the global coordinate system to the thigh cluster coordinate system (resulting in the coordinates GT_L). The greater trochanter was palpated in this posture by first locating its position in standing posture using the same instructions as above. Then, the researcher placed their finger on its position and held it while the participant assumed the seated position where it was digitized. Then, a 1x3 vector with x, y and z pairs was created from the greater trochanter to the femoral head center ($\overrightarrow{GT:FHC}$) , expressed in the thigh cluster coordinate system, using equation 3:

$$\overrightarrow{GT:FHC} = FHC_L - GT_L \quad (3)$$

where FHC_L is the femoral head center in the thigh cluster coordinate system calculated from equation 2 and GT_L is the greater trochanter in the thigh cluster coordinate system at the first frame of the quiet standing trial.

In the quiet seated calibration trial, a transformation matrix was created at the first frame to transform the greater trochanter from the global coordinate system to the thigh cluster coordinate system, resulting in coordinates GT_{LSit}. The $\overrightarrow{GT:FHC}$ vector calculated in equation 3 was applied to the greater trochanter to determine the position of the femoral head center (Figure 3.4), in the thigh cluster coordinate system, in the seated calibration trial (FHC_{LSit}) using equation 4:

$$FHC_{LSit} = \overrightarrow{GT:FHC} + GT_{LSit} \quad (4)$$

where GT_{LSit} is the location of the greater trochanter in the cluster coordinate system at the first frame of the seated calibration trial. A visual example of this process and applying the vector from standing calibration to seated calibration is represented in Figure 3.4.

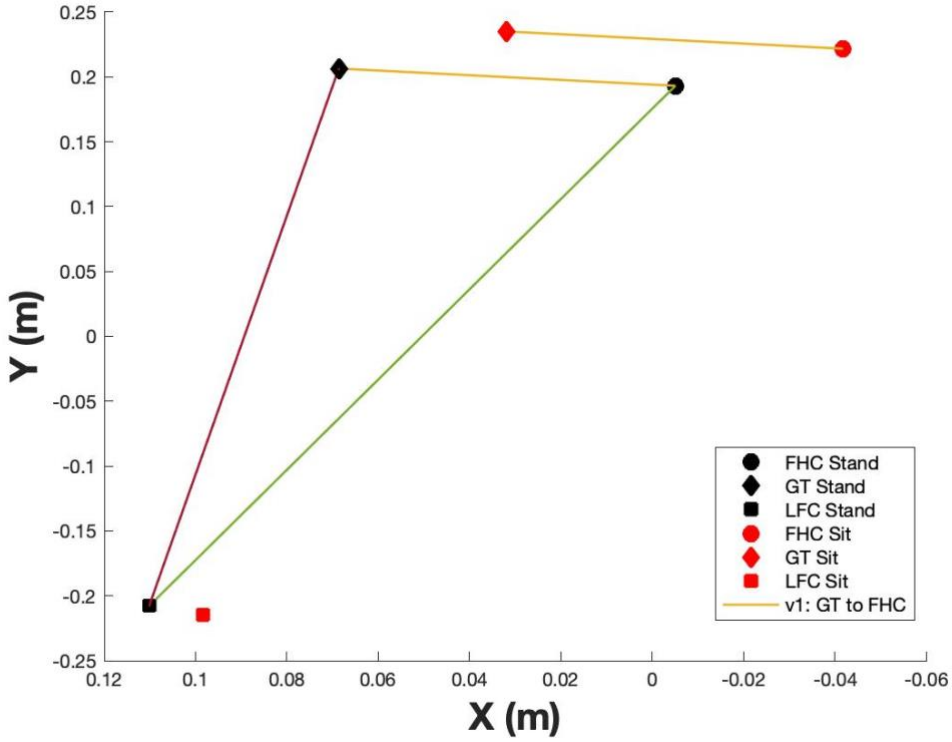


Figure 3.4: Visual schematic in the thigh cluster coordinate system of the greater trochanter, GT, (diamond), lateral femoral condyle, LFC, (square), and femoral head center, FHC, (circle) in the standing (black) and seated (red) calibration postures for Subject 1. V_1 represented the vector calculated in the standing posture from the GT to FHC, which was applied to the seated calibration GT digitized landmark to predict the FHC location.

Using the two locations of the femoral head center with respect to the thigh cluster coordinate system determined from the two calibration postures, and the associated knee joint flexion angles in each posture, three two-point linear calibration curves (one for each of the x, y, and z coordinates) were created to predict the femoral head center position in the thigh cluster coordinate system as a function of knee joint angle, and were extrapolated to maximal flexion angle in each high flexion motion trial (Figure 3.5). This femoral head center location in the thigh cluster coordinate system at maximum flexion became A_2 in the double calibration equation. While this was unique to each high flexion trial performed by each participant, the

slope of the calibration curve did not change since the curve was linear and based on the two calibration postures.

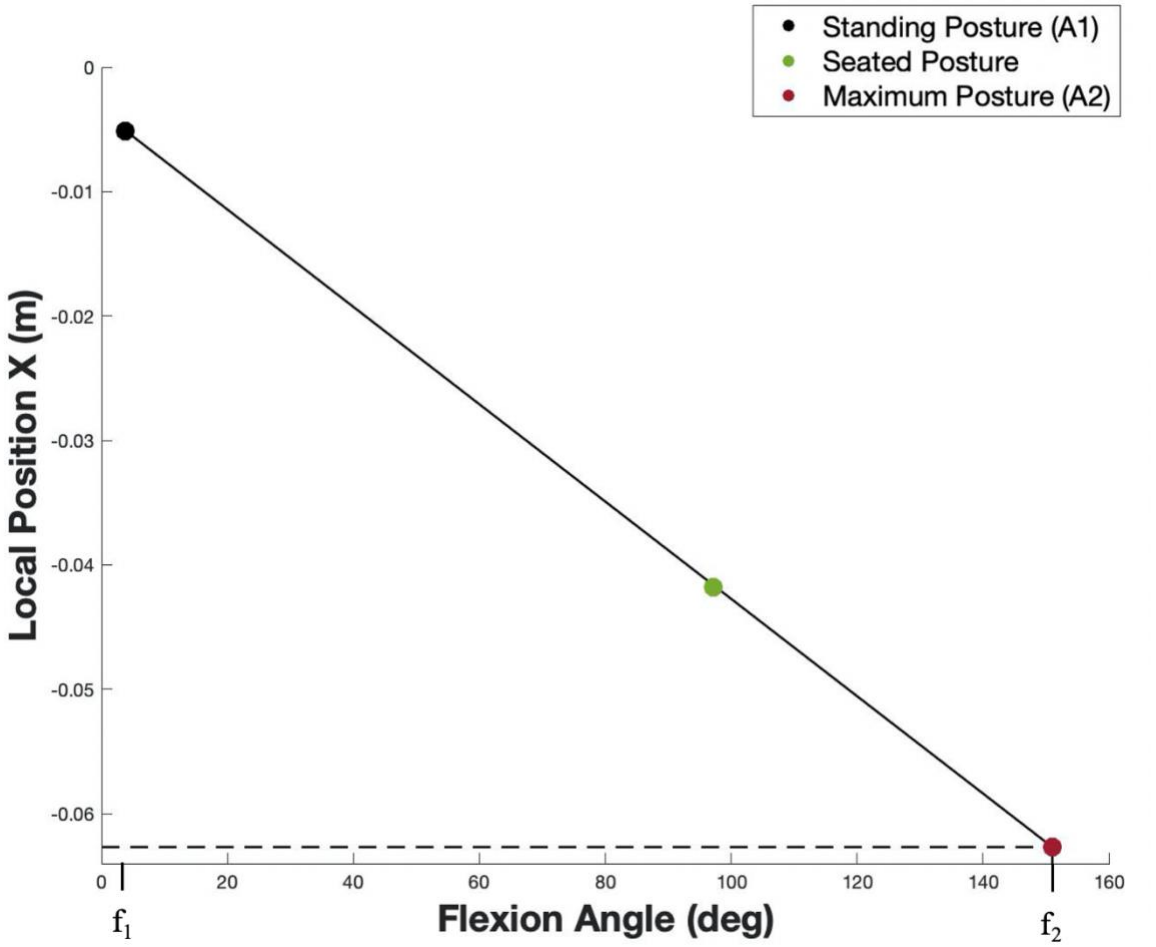


Figure 3.5: Three-point calibration curve to predict the x-coordinate of the femoral head center at maximal flexion for a Subject 1. The position of the femoral head center in the standing posture (A_1 – black circle) and seated posture (green circle) were interpolated from the standing (f_1) to seated flexion angle. This calibration curve was then extrapolated to maximum flexion angle (f_2) to obtain the position of the femoral head center at maximal flexion (A_2).

3.2.4.4 Calculation of A(t)

A(t) is the reconstructed position of the femoral head center in the thigh cluster coordinate system at a given time, t, during the high flexion trial. These positions were calculated using the double calibration equation (Equation 1, Section 3.2.4.1). A transformation matrix was then created at each frame of the high flexion motion trials and the position of A(t) was transformed from the thigh cluster coordinate system back to the global coordinate system using equation 5:

$$FHC(t)_R = (TCS(t)_R * A(t)) + TCS(t)_o \quad (5)$$

where $FHC(t)_R$ is the reconstructed femoral head center position in the global coordinate system at each frame in the motion trials, *t*, $TCS(t)_R$ is the local to global rotation matrix of the thigh cluster coordinate system at each frame in the high flexion motion trials, and $TCS(t)_o$ is the origin of the thigh cluster coordinate system at each frame in the high flexion motion trials. The construction of the transformation matrices are further explained in Appendix A.

3.2.4.5 Single Calibration

The functional hip joint center was reconstructed in single calibration using the mid-anterolateral thigh cluster. Transformation matrices were calculated at each frame of the high flexion motion trials and FHC_L from Section 3.2.4.2 was transformed from the thigh cluster coordinate system to the global coordinate system. This defined the position of the femoral head center in the single calibration. FHC_L, the position of the femoral head center in the first frame of quiet standing in the thigh cluster coordinate system, was used to reconstruct the global positions of the femoral head center in single calibration so there were no differences in position of the femoral head center at the start of motion for each calibration technique.

The functional hip joint center was reconstructed using traditional single calibration techniques where the landmark is fixed in the cluster coordinate system, using the pelvis cluster. This position was used as the comparator for the single and double calibration techniques, to determine which one predicted the position of the femoral head center best. The functional hip joint center was transformed from the global coordinate system to the pelvis cluster coordinate system in the first frame of quiet standing, following equation 2 in Section 3.2.4.2, to obtain the fixed position hip joint center in the local coordinate system (HJC_L). Local-to-global transformation matrices for the pelvis were then calculated at each frame of the high flexion motion trials and HJC_L was transformed to the global coordinate system to define the "gold standard" comparator.

3.2.4.6 Phase Identification and Time Normalization

Femoral head center positions in single and double calibration, and the hip joint center positions, were truncated from the start to end of the motion trials using the vertical trajectory of the pelvis marker. Start of descent was marked as the first instance the pelvis marker fell below 1 cm of the mean marker position in the first 10 frames of each motion trial. End of ascent was similarly marked except the final 10 frames were used to calculate the mean position of the pelvis marker and the last instance the pelvis marker fell below the 1 cm of the mean was used. There were however some exceptions to the truncating method. In some cases, participants fidgeted at the start or end of the trial so the mean from the stable end (either first or last 10 frames) was used as the threshold to determine start of descent or end of ascent. In other cases where the first and last instance of the pelvis marker below the 1 cm threshold did not adequately mark the start or end points, each successive frame after or before the incorrectly identified frame was analyzed until there was a downwards or upwards trend in the movement depending

on whether the start or end needed correcting. These newly identified start or end points were verified and/or corrected in Visual 3D if there was uncertainty. Data were then time normalized to 101 frames and trials with missing position data and/or blips/spikes in knee joint angle were dropped. These abnormalities in joint angle were speculated to be due to larger gaps in data that were not filled.

3.2.4.7 RMS Error

For each individual trial, RMS error was calculated along each of the three global directions, and for each of the two calibration types between the femoral head center (from either the single or double calibration) and hip joint center using equation 6, resulting in six RMS errors per trial:

$$RMS\ error_{i\,C} = \sqrt{\frac{\Sigma_{r=1}^{n}\ (HJC_r - FHC_{r\,C})^2{}_i}{n}} \quad (6)$$

where $_i$ indicates either the x, y, or z coordinate, $_C$ is the calibration type (single or double), $_r$ is the frame number, *HJC* is the appropriate global coordinate of the hip joint center, *FHC* is the appropriate global coordinate of the femoral head center, and *n* is the number of truncated and time-normalized frames of data (101 points). Overall means of RMS error were calculated across all trials for a given participant, all high flexion conditions and all high flexion movements combined, resulting in one mean RMS error along each of the global axes across all high flexion postures for each participant.

3.2.4.8 Resultant Error

The distance ($\overrightarrow{Diff}$) is the magnitude of the vector between the femoral head center (from either the single or double calibration) and the hip joint center position. $\overrightarrow{Diff}$ was calculated using equation 7:

$$\overrightarrow{Diff_C} = \sqrt{(HJC_x - FHC_{xC})^2 + \left(HJC_y - FHC_{yC}\right)^2 + (HJC_z - FHC_{zC})^2} \quad (7)$$

where C is the calibration type of the femoral head center (single or double calibration), HJC_x, HJC_y, and HJC_z are the global coordinates of the gold standard hip joint center, and FHC_x, FHC_y, and FHC_z are the global coordinates of the femoral head center. The $\overrightarrow{Diff}$ thus expresses the resultant error in locating the femoral head center (the thigh cluster-based approximation of the hip joint center.) The mean resultant error, also known as the bias in Bland-Altman analyses (Section 3.2.5.2), were used to compare the overall error in global position using both calibration techniques. The resultant errors (one for each of the two calibration techniques) were calculated for each frame of a motion trial, then averaged into one mean resultant error value per calibration technique for the entire trial. These mean trial resultant error values were then averaged for each high flexion movement, and then for all high flexion conditions combined.

3.2.5 Statistical Analyses

3.2.5.1 Paired t-tests

One-tailed paired t-tests (α= 0.05, R-Programming, R Core Team, 2020) provided within-participant comparisons of the mean RMS errors and resultant error between the single and double calibrations. In these tests, it was hypothesized the mean RMS errors and resultant error would be lower in the double calibration. It was also hypothesized the limits of agreement would be narrower for the double calibration. Data were tested for normality using a Shapiro-Wilk Test and if significant, a Wilcox Signed Rank Test was used as the non-parametric equivalent.

3.2.5.2 Bland-Altman Analysis

Bland-Altman analyses were conducted to determine the agreement between the femoral head center location, obtained through single or double calibration, and the hip joint center

location. To use this method in analyses, the bias (mean difference between the two points) and standard deviation of the difference were needed. For this thesis, the mean resultant error was reported as the bias. To calculate the standard deviation, additional steps were required because each participant had multiple trials, referred to as multiple observations per participant, and the within and between subject variances had to be accounted for (Bland & Altman, 2007). First, 2 *one-way ANOVAs* were conducted: One on the resultant error between the double calibration-based femoral head center and the hip joint center, and one on the resultant error between the single calibration-based femoral head center and the hip joint center. Table 3.2 represents an ANOVA table that measured within subject differences for single calibration resultant error. These outputs were used to calculate the standard deviation for Bland-Altman analyses with multiple observations.

Table 3.2: ANOVA table for the single calibration demonstrating group mean-squared (MS) and error MS values used in the Bland-Altman multiple observations method to calculate the standard deviation for the calibration techniques. Other values include sums of squares (SS), degrees of freedom (df), and F-statistic (F).

Source	SS	df	MS	F	Prob>F
Groups	0.05945	32	0.00186	25.92	1.8101e-90
Error	0.04093	571	0.00007		
Total	0.10038	603			

The *variability* in the resultant error across all participants was calculated from the residual mean squares (Error MS) subtracted from the mean squares of subjects (Groups MS). A value based on the number of observations (*Obs*) for each subject included in the analysis was then calculated using equation 8:

$$Obs = \left(\frac{(\sum m_i)^2 - \sum m_i^2}{(n-1)\sum m_i}\right) \quad (8)$$

where m_i is the total number of observations, or trials, per subject and n is the total number of subjects (Bland & Altman, 2007). The *variability* (Groups MS – Error MS) was then divided by *Obs* and returned a value that represented the *heterogeneity* of variance of the one-way ANOVA. The *heterogeneity* was then added to the residual mean squares (Error MS) to determine the *total variance* of the difference between subjects, which was then square rooted to determine the *standard deviation* (Bland & Altman, 2007).

Using the standard deviations, the limits of agreement were calculated. These limits were defined as the interval of 2 standard deviations above and below the resultant error (to span 95% of all observations) and were calculated for each calibration method between the femoral head center and hip joint center (Bland & Altman, 2007). Bland & Altman (2007) acknowledged that this method using multiple observations per individual resulted in narrower limits of agreement, rather than treating each individual trial as a separate observation; however, were not concerned as the limits of agreement were only smaller by a slight bit.

To interpret the Bland-Altman analyses, visual observation of the bias and size of the limits of agreement was used. For the bias, the calibration method with bias closest to zero represented the method that predicted location of the femoral head center (from the thigh coordinate system) with the most similarity to the location of the hip joint center (using the pelvis coordinate system). For the limits of agreement, the calibration method with the narrowest limits of agreement was considered to have the least variation in resultant error (reported as the bias) across the participants.

3.3 Results

Thirty-three participants were included in the results. Outcomes were averaged such that there was one value per participant that represented each of the four outcomes of interest: one

resultant error and three single-axis RMS errors. These values were for all high flexion movements combined. Results were presented this way because the purpose of the double calibration technique was to evaluate its effectiveness across all high flexion movements. Mean RMS error and resultant error separated by high flexion movement are presented in Appendix B.

A one tailed paired t-test that compared RMS errors between the single and double calibration approach revealed that the double calibration error was not significantly lower than the single calibration error along the global x, y, and z directions (Table 3.3). The observed relationship was in fact the opposite, where the double calibration RMS errors were greater than the single calibration errors. This would explain why all p-values were at or close to 1 because the one-tailed paired t-test null hypothesis was not rejected. Figures 3.6 and 3.7 show the position of the hip joint center and femoral head centers for the first two subjects completing a heels-up squat and a plantarflexed kneel from start of descent to end of ascent, respectively. These figures demonstrate the increased error in double calibration.

Table 3.3: Results from one-tailed paired t-tests (α=0.05) on the mean RMS error in the global x, y and z directions between single and double calibration. The femoral head center was compared against the hip joint center in RMS calculations. RMS errors were presented as a mean (+/- standard deviation) in meters (m).

Condition		Double Calibration	Single Calibration	p-value	Effect Size
All movements	x	0.03 (0.02)	0.021 (0.01)	0.983	0.429
	y	0.04 (0.026)	0.015 (0.01)	1	0.952
	z	0.019 (0.011)	0.01 (0.004)	1	0.765

* $p < 0.05$

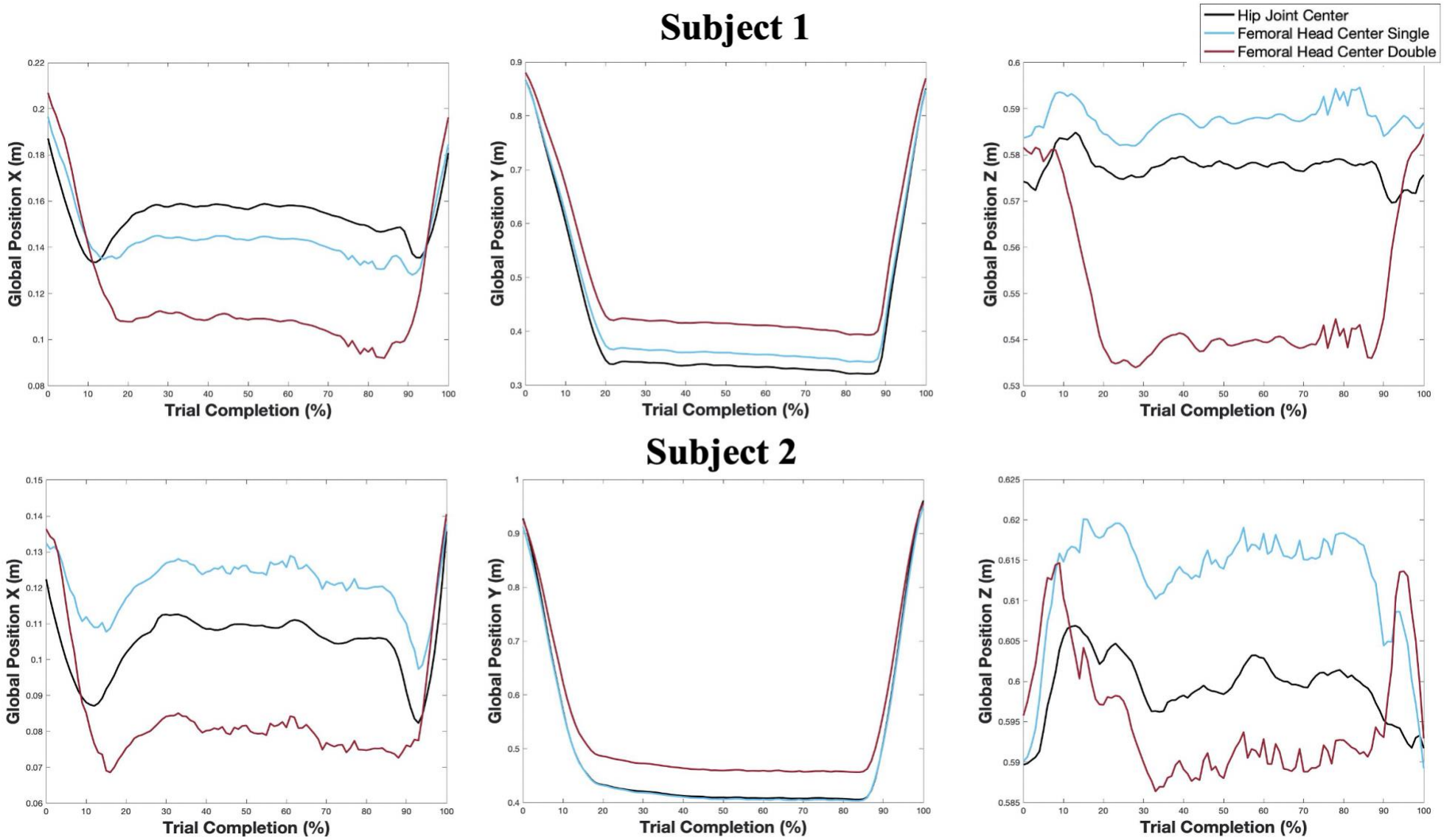


Figure 3.6: Global x, y and z coordinates of the hip joint center and femoral head center reconstructed from single and double calibrations during a heels-up squatting task. The upper row represented data from Subject 1 and the lower row represented Subject 2.

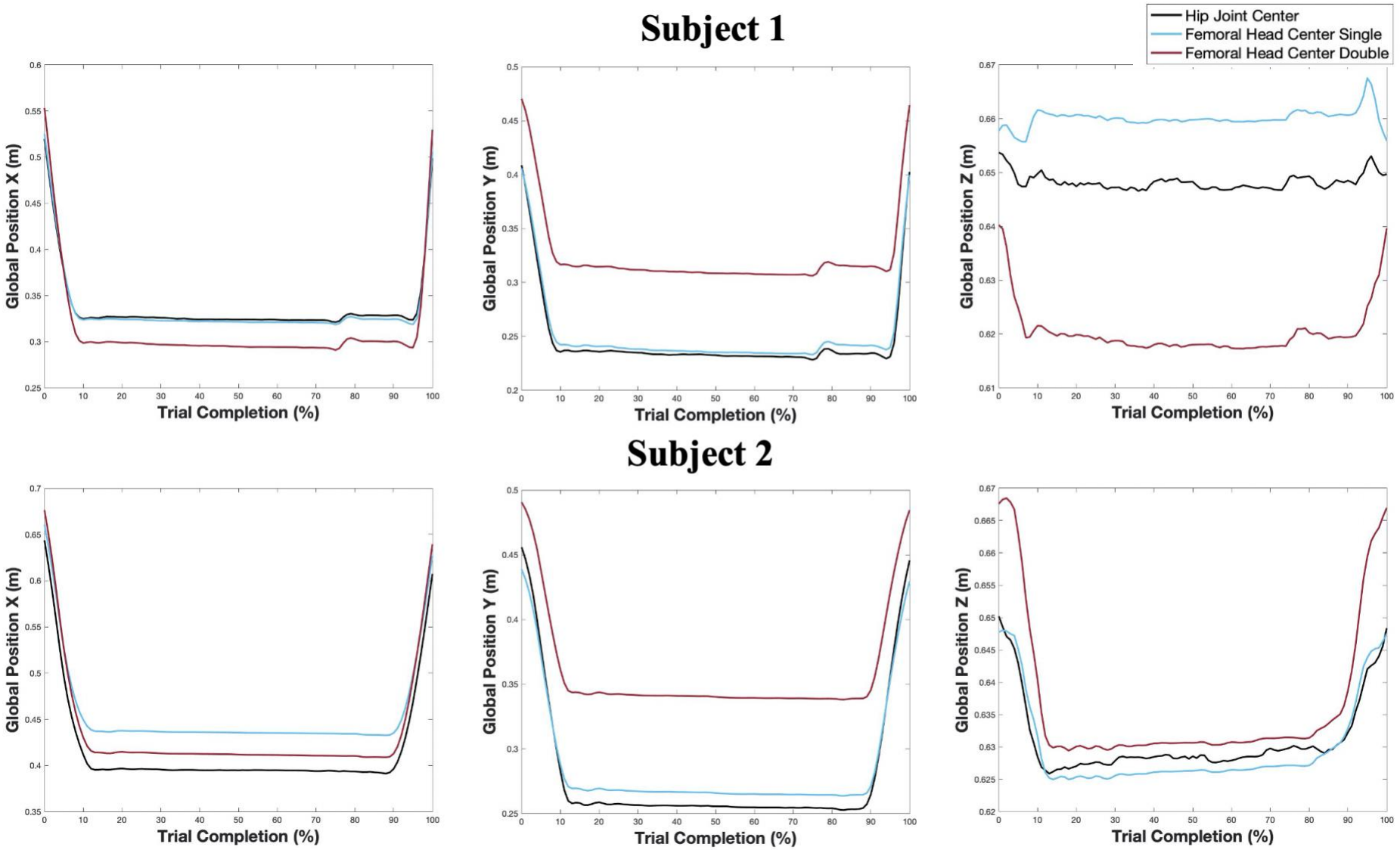


Figure 3.7: Global x, y and z coordinates of the hip joint center and femoral head center reconstructed from single and double calibrations during a plantarflexed kneeling task. The upper row represented data from Subject 1 and the lower row represented Subject 2.

For the Bland-Altman analysis (Bland & Altman, 2007), the resultant error and limits of agreement were plotted for the two calibration conditions (Figure 3.8). Plots for each calibration condition with the mean of the femoral head center and hip joint center magnitude and difference in mean magnitude (mean distance) for each participant are presented in Appendix C. The double calibration resultant error was 5.8 cm while the single calibration error was 3.1 cm. The double calibration resultant error was further away from the zero line, indicating less agreement between the two methods. A one-tailed paired t-test comparing resultant error between the two calibration methods also revealed that the double calibration resultant error was not significantly lower than the single calibration (Table 3.4). The limits of agreement were also wider in the double calibration compared to the single calibration. The wider limits of agreement in the double calibration method indicated that the spread of errors were even greater than that in the single calibration. Upper and lower limits of agreement reached 11.7 cm and -0.20 cm in the double calibration, respectively, while the single calibration only reached 5.6 cm and 0.5 cm. The range of error between the upper and lower limits of agreement was 11.9 cm and 6.1 cm in double and single calibrations, respectively.

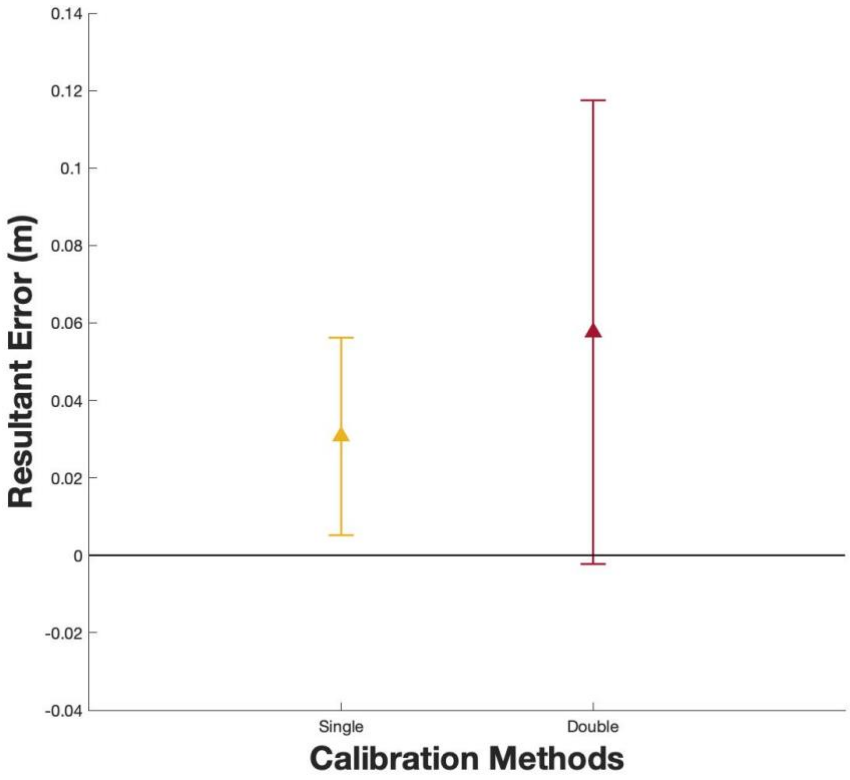


Figure 3.8: Resultant error plotted for the single and double calibration methods. The error bars extend to the upper and lower limits of agreement.

Table 3 4: Results from one-tailed paired t-tests (α=0.05) between single and double calibration on the resultant error. Resultant errors were presented as a mean (+/- standard deviation) in meters (m).

Condition	Double Calibration	Single Calibration	p-value	Effect Size
All movements	0.058 (0.028)	0.031 (0.01)	1	1.02

* $p < 0.05$

3.4 Discussion

RMS error, resultant error and limits of agreement evaluated the effectiveness of the double calibration technique. These analyses concluded that the double calibration performed worse than the single calibration on locating the femoral head center, when compared to the hip

joint center, using both individual x, y, and z coordinates and the distance from the hip joint center. Similarly, the limits of agreement were wider for the double calibration, indicating greater variability amongst participant resultant errors. Therefore, the double calibration method was not more effective at reducing thigh STA compared to traditional single calibration when using a mid-anterolateral thigh cluster in high flexion.

The double calibration equation used in this thesis was adapted from Cappello et al. (2005). The published method sought to reduce error in kinematic outcomes by calibrating anatomical landmarks in two extremes of a movement (at minimum and maximum ROM) to account for rigid displacement and non-rigid deformation of a marker cluster in response to soft tissue movement, that would have introduced error in the anatomical landmark location throughout a motion trial and calculated joint angles (Cappello et al., 2005). Results from previous research were favorable towards a double calibration, where knee joint rotation and translation RMS errors were lower for double calibration than for single calibration, when compared to measures using fluoroscopy (Cappello et al., 2005; Stagni et al., 2006). This thesis found opposing results, in that RMS error was 3 cm, 4 cm and 1.9 cm in x, y and z positions for the femoral head center location using the double calibration, compared to 2.1 cm, 1.5 cm and 1 cm using the single calibration (Table 3.3). Although this thesis compared the position of a non-palpable anatomical landmark (the femoral head center) – predicted from a palpable anatomical landmark (the greater trochanter) – and previous research used knee joint rotations and translations to evaluate the effectiveness of calibration techniques, it was not believed that these differences in reported kinematic outcomes would have affected the conclusion drawn because error in anatomical landmark position translates to error in joint angles. When constructing anatomical coordinate systems with anatomical landmarks to calculate knee joint angles, the

location of the femoral head center alone affected hip and knee joint angles and moments – the hip moments to a larger degree where anterior mislocation attributed error of -22 %, lateral mislocation of -15%, and posterior mislocation of 25% of underestimation (-) and overestimation (+) of the moment curves – after introducing 3 cm of error to the location of the hip joint center position (Stagni et al., 2000). Errors in knee joint FL/EX and AB/AD angles were within 1.5 degrees before and after 3 cm of error was introduced to the hip joint center (Stagni et al., 2000); however, if knee joint angles were to be calculated from coordinate systems constructed from anatomical landmarks in this thesis, the double calibration equation would have been applied to all anatomical landmarks tracked from the thigh cluster and would likely risk greater errors in calculated knee joint angles. One traditional method to construct the vertical axis of the thigh anatomical coordinate system axis uses the intersection of the midpoint between the medial and lateral femoral epicondyle and the hip joint center (i.e. femoral head center) and proceed to calculate knee joint angles as the rotation of the shank local coordinate system relative to the thigh coordinate system (Grood & Suntay, 1983). Although knee joint angles were not investigated in this study, it is expected that the increased error from the double calibration would mean that knee joint angles calculated from the double calibration-based landmarks would be less accurate than those using single calibration-based landmarks.

The mean range of FL/EX angle for the motor tasks in previous research using the double calibration was 95 degrees knee flexion (Cappello et al., 2005; Stagni et al., 2006); however, in this thesis, knee flexion angle exceeded 120 degrees. It was increasingly difficult to palpate bony landmarks at maximal flexion since they were much less prominent than at 95 degrees; therefore, a novel step to the double calibration was added such that the second calibration posture was palpated at ~90 degrees knee flexion in a seated position, and then the landmarks at maximal

flexion were predicted through linear extrapolation of its position in the thigh cluster coordinate system up to maximal flexion. Previous research used assumptions of linearity to predict the behaviour of STA (Andersen et al., 2012; Camomilla et al., 2015; Cappello et al., 2005); however, based on the poor performance of the double calibration in high flexion activities, this may not have been an appropriate assumption for modeling the movement of anatomical landmarks in response to STA in high flexion. Kingston & Acker (2018) found that after thigh-calf contact was initiated between 119.7 and 128.1 degrees in high flexion squatting and kneeling, the total thigh-calf contact force (%BW) increased non-linearly in squatting and exponentially in kneeling right before reaching maximal flexion. Zelle et al. (2007) also modeled the total force of thigh-calf contact in deep squatting and kneeling and found exponential behaviour from thigh-calf contact onset to maximal flexion angle. Since soft tissue deformation increased in high flexion movements (Kingston & Acker, 2018), perhaps it might have been appropriate to follow similar behaviour to that of thigh-calf contact force, modeling the movement of anatomical landmarks non-linearly or exponentially from its position in quiet seated onwards. Although it was not possible to model these relationships in this thesis, as we did not have the position of the femoral head center at maximal flexion, this is an area of future investigation.

The resultant error calculated in this study was greater in the double calibration compared to the single calibration, and associated limits of agreement were wider (Figure 3.8, Table 3.4). If the double calibration were to affect all participant data the same, the limits of agreement should have been the same size for the single and double calibration; however, this was not the case. There was increased variability in the spread of overall error after performing the double calibration, which indicated that the effect of double calibration varied between participants.

Higher thigh mass and shorter leg length has been correlated to greater STA (Grimpampi et al., 2014), and perhaps greater STA could relate to greater error in the prediction of the femoral head center. This would need to be further investigated before any conclusions are drawn for a subject-specific metric.

3.5 Limitations

The are some limitations in the current study that may have contributed to the observed error in double calibration results. It was not possible to accurately palpate bony landmarks on the femur at maximum flexion, so a seated position was chosen as the second calibration posture where the bony landmarks were more prominent. This captured the digitized position of the greater trochanter in a seated posture, rather than relying on the mid-anterolateral cluster to predict its location which would have included STA. The natural movement of the mid-anterolateral thigh cluster and associated change in greater trochanter position in the thigh cluster coordinate system (used to predict femoral head center position) was still captured from a standing to seated posture (approx. 0 to 90 degrees flexion). The second limitation referred to the assumption of linearity in the double calibration equation. Cappello et al. (2005) reconstructed the position of a given anatomical landmark in the corresponding cluster coordinate system based on its position from A_1 to A_2 and linearly weighted this change in position based on flexion angle. Flexion angle was chosen as the weighting factor because the behaviour of STA depends on the ROM of a task (Cappello et al., 1997) and was the most appropriate measure to provide this insight. This angle was also least affected by STA (Stagni et al., 2005). A true gold-standard femoral head center location in the thigh cluster coordinate system would also be needed to investigate varying relationships of the double calibration equation. A third limitation is the use of the hip joint center reconstructed from the pelvis cluster as the comparator to evaluate thigh

STA, as opposed to a true gold-standard. The chosen algorithm (Schwartz & Rozumalski, 2005) and FL/EX, AB/AD, and circumduction functional hip joint movements (Camomilla et al., 2006) used to calculate the functional hip joint center were most accurately cited methods with 2.5 (2.1 to 2.9) mm of pelvis STA, calculated as the mean absolute deviation in the functional joint center trial (Fiorentino et al., 2016). This error was much smaller compared to the femoral head center location error measured in this thesis which reached 3.1 and 5.8 cm of error in the single and double calibration resultant error, respectively.

Chapter 4 – Investigation of Subject Specificity and Double Calibration Results

4.1 Objectives and Hypothesis

The objective of this investigation was to determine the effectiveness of the double calibration technique within participant groups based on the mid-thigh circumference. This objective emerged from Section 3.4 after the limits of agreement were wider in the double calibration technique and led to some speculation that subject-specific characteristics may have influenced the success of the double calibration. Mid-thigh circumference groupings were determined using standardized data from the 2003-2006 National Health Statistics Report from the United States (McDowell et al., 2008). This data category was excluded from further reports. For this thesis, four percentile categories were identified: 0-24th percentile, 25-49th percentile, 50-74th percentile, and 75+ percentile. These will be referred to as the 15th, 25th, 50th and 75th percentiles. Paired t-tests were conducted for the mean RMS error and resultant error within each percentile group between the femoral head center location in the single and double calibration and the hip joint center. Comparing the t-test results across percentile groups indicated if there was a difference in the performance of the double calibration versus the single calibration in those with different thigh circumferences. Paired t-tests were chosen over a two-way ANOVA because of the unequal sample sizes in the percentile groups, and between-percentile group comparisons were not of interest to answer the research question. Bland-Altman analyses were also completed to determine the mean bias and limits of agreement for the resultant error between both calibration methods.

It was hypothesized that there would be significant differences in mean RMS error, resultant error, and limits of agreement for the calibration techniques in the 25th and 50th percentile groups. These groups were chosen because larger thigh mass correlated to greater STA

(Grimpampi et al., 2014) and it was assumed that higher thigh circumference would relate to larger thigh mass. It was also hypothesized there would be no differences in mean RMS error, resultant error and limits of agreement between the calibration techniques in the 15th percentile group because thigh circumference was smallest in this group and assumed to have the least amount of soft tissue. Two-tailed statistical tests were conducted in this chapter because the results in Chapter 3 concluded the double calibration error was not lower than the single calibration error. Hypotheses were not included for the 75th percentile group because it had 3 participants and would therefore require a larger t-statistic to reject the null hypothesis.

4.2 Methods

4.2.1 Participants

Participant hip joint center and femoral head center position from the single and double calibration in Chapter 3 were separated into percentile groups based on mid-thigh circumference (Table 4.1). Each percentile category included male and female data but were categorized using sex-specific standardized measurements. Data were not separated by sex, in addition to percentile group, because the objective of this section was to investigate a subject-specific metric that was localized to the thigh, only.

Table 4.1: Mid-thigh circumference measures for participants after separating data into percentile groups. Data were presented as mean (+/- standard deviation) in centimeters (cm).

Parameter	15th percentile (n=6)	25th percentile (n=12)	50th percentile (n=12)	75th percentile (n=3)
Mid-thigh circumference (cm)	45.35 (2.73)	50.84 (2.11)	54.78 (2.13)	59.23 (3.65)

4.2.2 Data Processing

Processed data from Chapter 3 were used to complete this analysis.

4.2.3 Data Analysis

Using the calculated resultant error from Section 3.2.4.8 and mean RMS Error from Section 3.2.4.7, participant data were separated into the four percentile groups.

4.2.4 Statistical Analysis

Two-tailed paired t-tests ($\alpha = 0.05$) were conducted between single and double calibration on the mean RMS error and resultant error within each percentile group. Resultant error (bias) and limits of agreement were also calculated for each percentile group using the same procedures in Section 3.2.5.2. This included separate calculations for the *one-way ANOVA*, *variability*, *Obs*, *heterogeneity*, *total variance*, *standard deviation*, and *limits of agreement* for each percentile group. The resultant error and limits of agreement were plotted to visually compare the bias and limits of agreement between the single and double calibration within each percentile group.

4.3 Results

There were six participants in the 15th percentile group, twelve in the 25th and 50th percentile groups, and three in the 75th percentile group. Results from two-tailed paired t-tests on mean RMS error between the single and double calibration in each percentile group revealed that the double calibration mean RMS error was significantly greater in the y and z position of the femoral head center for the 25th and 50th percentile groups ($p < 0.05$) (Table 4.2). There were no significant differences between the double and single calibration in the 15th and 75th percentile groups, despite the large differences and effect sizes in single-axis RMS error. One potential explanation for lack of significance despite the large differences and effect sizes is the low sample size in the 75th percentile group (n=3), resulting a high critical value to reject the null hypothesis and low power in the 15th group.

Table 4.2: Results from two-tailed paired t-tests (α=0.05) between single and double calibration mean RMS error in x, y, and z positions of the femoral head center in each percentile group. Femoral head center data were compared against the hip joint center data to obtain RMS error. RMS error was reported in meters (m).

Percentile		Double Calibration	Single Calibration	p-value	Effect Size
15th (n=6)	x	0.025 (0.01)	0.022 (0.009)	0.491	0.303
	y	0.022 (0.019)	0.015 (0.011)	0.321	0.450
	z	0.018 (0.013)	0.008 (0.006)	0.156	0.767
25th (n=11)	x	0.029 (0.021)	0.018 (0.008)	0.129	0.513
	y	0.038 (0.029)	0.012 (0.006)	0.000488*	0.890
	z	0.018 (0.012)	0.009 (0.004)	0.0122*	0.743
50th (n=12)	x	0.033 (0.025)	0.025 (0.012)	0.569	0.317
	y	0.049 (0.026)	0.018 (0.013)	0.00488*	1.18
	z	0.018 (0.01)	0.011 (0.004)	0.0413*	0.667
75th (n=3)	x	0.03 (0.006)	0.018 (0.011)	0.0758	1.98
	y	0.049 (0.024)	0.018 (0.006)	0.142	1.36
	z	0.022 (0.004)	0.014 (0.002)	0.151	1.31

* $p < 0.05$ indicates a significant difference between the single and double calibrations for a specific outcome measure.

The resultant error and limits of agreement were plotted for each percentile group in the single and double calibration (Figure 4.1). Plots for each calibration condition with the mean of the femoral head center and hip joint center magnitude and difference in mean magnitude (mean distance) for each participant are presented in Appendix D. The double calibration resultant error was farther away from the zero line in all percentile groups compared to the single calibration. Two-tailed paired t-tests between the resultant error for single and double calibration in each percentile group also revealed that the double calibration had significantly greater resultant errors in the 25th and 50th percentile groups ($p < 0.05$) (Table 4.3). The double calibration resultant error was 5.6 cm and 6.6 cm in the 25th and 50th percentile groups, while these two groups in the single calibration had a bias of 2.5 cm and 3.6 cm (Table 4.4). The limits of agreement in the double calibration were also wider for all percentile groups, and spanned 18.2 cm, 20.1 cm, 21.2

cm, and 26.6 cm of error for the 15th, 25th, 50th and 75th percentiles, respectively, compared to 6.3 cm, 5.9 cm, 8.5 and 15.3 cm. The difference in the size of the limits of agreement from the double calibration to the single calibration were 11.8 cm, 14.1 cm, 12.8 cm, and 11.3 cm, showing no relationship in the mid-thigh circumference percentile group.

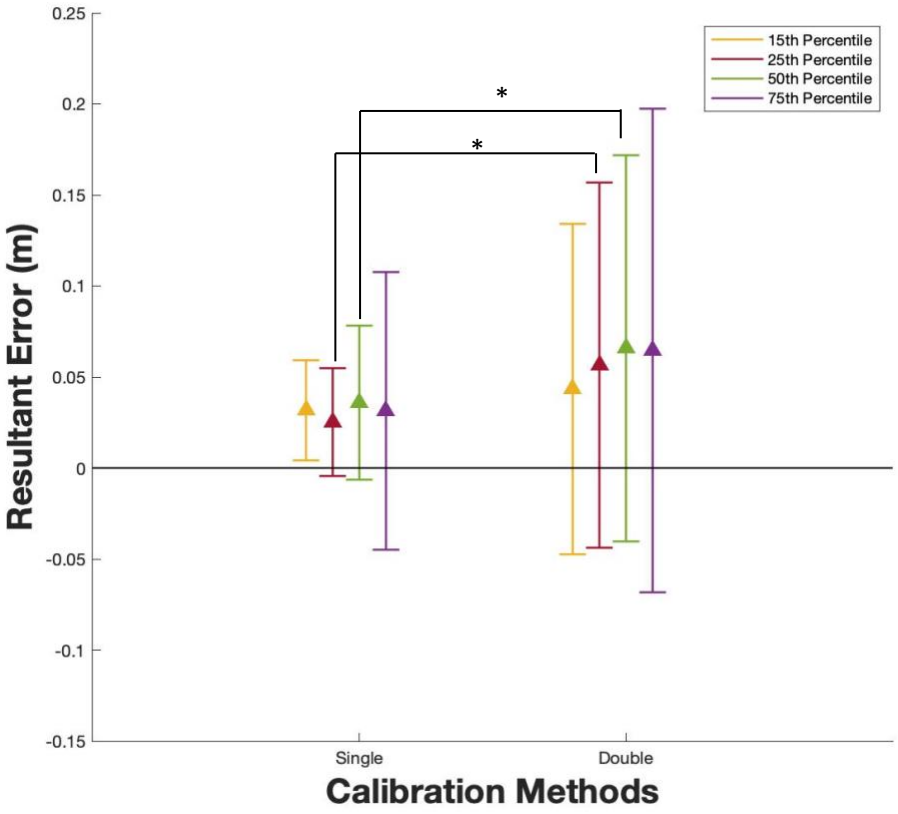


Figure 4.1: Resultant error plotted for the single and double calibration methods, grouped by mid-thigh circumference percentiles. The error bars represent the upper and lower limits of agreement. * $p < 0.05$ indicates a significant difference between the single and double calibrations for a specific outcome measure within that percentile.

Table 4.3: Results from two-tailed paired t-tests (α=0.05) between single and double calibration resultant error of the femoral head center compared against the hip joint center in each percentile group. Resultant errors were presented as a mean (+/- standard deviation) in meters (m).

Percentile	Double Calibration	Single Calibration	p-value	Effect Size
15th	0.041 (0.022)	0.031 (0.006)	0.217	0.576
25th	0.057 (0.029)	0.025 (0.008)	0.00251*	1.12
50th	0.066 (0.031)	0.036 (0.011)	0.00448*	1.03
75th	0.065 (0.017)	0.031 (0.01)	0.119	1.52

* $p < 0.05$ indicates a significant difference between the single and double calibrations for a specific outcome measure.

Table 4.4: Parameters from Figure 4.1 representing the resultant error (mean bias) and the limits of agreement in meters (m).

Percentile	Parameter	Double Calibration	Single Calibration
15th (n=6)	Upper limit (m)	0.134	0.059
	Mean bias (m)	0.043	0.031
	Lower limit (m)	-0.047	0.004
25th (n=11)	Upper limit (m)	0.159	0.055
	Mean bias (m)*	0.057	0.025
	Lower limit (m)	-0.044	-0.004
50th (n=12)	Upper limit (m)	0.172	0.078
	Mean bias (m)*	0.066	0.036
	Lower limit (m)	-0.040	-0.006
75th (n=3)	Upper limit (m)	0.197	0.108
	Mean bias (m)	0.065	0.031
	Lower limit (m)	-0.068	-0.045

* $p < 0.05$ indicates a significant difference between the single and double calibrations on the mean bias.

4.4 Discussion

RMS error, resultant error and Bland-Altman agreements were used to evaluate the effectiveness of the double calibration technique after separating data based on mid-thigh circumference percentiles. This analysis provided insight into whether the performance of the double calibration differed based on a subject-specific characteristic. These analyses concluded that when separating data using mid-thigh circumference, the double calibration resultant error and RMS error in y and z directions were significantly different than the single calibration in the 25^{th} and 50^{th} percentile groups, where the double calibration had larger error. There were no differences between the single and double calibration RMS error in the x-direction for the 25^{th} and 50^{th} percentiles, nor for RMS error (in all directions) and resultant error in the 15^{th} and 75^{th} percentile groups. Despite no differences in the 15^{th} and 75^{th} percentiles, the limits of agreement were wider in all percentile groups compared to the single calibration. This similar effect was observed in Chapter 3, which continued to support conclusions that the double calibration introduced more variability in the data; however, we observed in this section that the variability increased as the grouping percentile increased. The effect of the double calibration resultant error was also larger with increasing percentile groups, with a medium effect size in the 15^{th} percentile group and high effect sizes in the 25^{th}, 50^{th} and 75^{th} percentile groups. We speculated this result may have been due to greater inaccuracies in palpation of the landmarks in the seated posture for those with larger thigh circumferences or greater deformation of the thigh in the higher percentile groups, at flexion angles beyond the seated posture, that would not have been captured using the double calibration equation.

In the single calibration, resultant error did not increase with percentile groups; however, the range between the upper and lower limits of agreement increased when moving from the 25^{th}

to 75th percentiles. Previous research has evaluated STA between individuals of varying BMI (Barré et al., 2013; Clément et al., 2018); however, found mixed results of STA in calculated knee joint angles where the highest BMI group did not always have greatest STA. Although we did not statistically compare the performance of the double calibration between different percentile groups (due to varying sample sizes in each group), the results from this thesis present similar results where there wasn't a definitive relationship between resultant error and thigh circumference measures for the single calibration technique. The work in this thesis also builds on previous findings by determining the limits of agreement, where there was increasing trend in variability with grouping percentile from the 25th to 75th percentile. Subject-specific characteristics that include effects of thigh composition may need to be investigated to draw more accurate conclusions on subject-specific variability of STA, such as wobbling mass (Andersen et al., 2012; Barré et al., 2015), muscular contractions (Andersen et al., 2012; Barré et al., 2015; Cappello et al., 2005), and skin viscoelasticity (Andersen et al., 2012; Barré et al., 2015; Cappello et al., 2005).

4.5 Limitations

There are two limitations to be addressed in this chapter and both are related to sample size. The 15th and 75th percentile sample sizes were quite small and had a medium to large effect size in resultant error differences, producing a power level of 0.54 and 0.69, respectively. Although the power level is acceptable in the 75th percentile group with three participants, the critical value to reject the null hypothesis would have been quite high, making it more likely to return a false negative for the paired t-test. Therefore, appropriate conclusions could not be drawn for the 75th percentile groups. The 15th percentile group may have also returned a false negative due to a low power value. Increased sample size would have been needed to address

these limitations.; however, data were only collected for 33 participants. The second limitation referred to the method used to categorize percentile groups. Due to low sample size and dividing data in to four groups, male and female thigh circumference percentiles were grouped. This resulted in different threshold values for determining percentile groups. For example, the 25th percentile female mid-thigh circumference was 47.3 cm while the male circumference was 50.8 cm. This may have misrepresented thigh circumference RMS error and resultant error because males and females had different ranges and when grouped together, included higher and lower mid-thigh circumference values in each percentile group. To mitigate potential effects of these groupings, paired t-tests were conducted between single and double calibration RMS error and resultant error, resulting in within participant comparisons for each percentile group.

Chapter 5 – Investigation of Digitization Error and Use of Alternative Landmark to Predict the Femoral Head Center in Double Calibration

5.1 Objective and Hypothesis

The objective of this sub-investigation was to evaluate whether choosing a different palpated anatomical landmark to construct the vector that predicted the femoral head center in the seated calibration posture influenced double calibration results. This objective emerged from the results presented in Chapters 3 and 4, which indicated that the double calibration was not successful in reconstructing the femoral head center position most similar to the gold standard hip joint center, despite controlling for subject-specific characteristics. As described in Chapter 3, the femoral head center could not be directly palpated in the standing and seated reference positions, so its position relative to the greater trochanter was assumed to be fixed in the thigh cluster coordinate system. In the double calibration, the greater trochanter was modeled to move within the thigh cluster coordinate system, and so the femoral head center moved with it as well. In this chapter, the results of the double calibration when the lateral femoral condyle was used to locate the femoral head center will be compared to the double calibration results when the greater trochanter was used (from Chapter 3), in order to determine if one method results in improved double calibration performance over the other.

The choice of anatomical landmark on which to base the location of the femoral head centre matters, since any mislocation of the anatomical landmarks in the seated posture would cause the predicted femoral head center location to not match its position in quiet standing. The femur is a rigid segment and the distance between any two landmarks in the thigh-cluster coordinate system should not change if they are accurately palpated in standing and seated postures. However, as mentioned previously, the palpation of these landmarks is more challenging in the seated posture, resulting in greater uncertainty in the accuracy of their

palpation. The evaluation of the inter-landmark distances between the standing and seated postures in the thigh cluster coordinate system provide insight into the digitization error, assuming the standing calibration inter-landmark distance was most correct (i.e., gold standard). A one-way ANOVA between the single calibration, greater trochanter-based (GT-based) double calibration and lateral femoral condyle-based (LFC-based) double calibration were conducted for each single-axis RMS error and resultant error. Bland-Altman analyses were also conducted to evaluate the resultant error and limits of agreement between the single calibration and the LFC-based double calibration methods. A two-tailed paired t-test evaluated the greater trochanter and lateral femoral condyle inter-landmark distances in the thigh cluster coordinate system between the standing and seated postures.

It was hypothesized that there would be significant differences in femoral head center RMS error and resultant error from the single calibration, GT-based double calibration and LFC-based double calibration. It was also hypothesized there would be differences in the size of limits of agreement of the single and double calibration. Significant differences between the single and GT-based double calibrations were previously reported in Chapters 3 and 4. It was also reported after palpating the seated posture that the lateral femoral condyle was less difficult to palpate compared to the greater trochanter, which may influence the GT-based and LFC-based double calibration error and inter-landmark distance in the standing and seated postures. It was hypothesized there would be differences in inter-landmark distance between standing and seated postures.

5.2 Methods

The same participant data from Chapter 3 were used for this analysis. The double calibration was run a second time using the lateral femoral condyle to predict the femoral head center. An overview of the methodological procedures were presented in Figure 5.1.

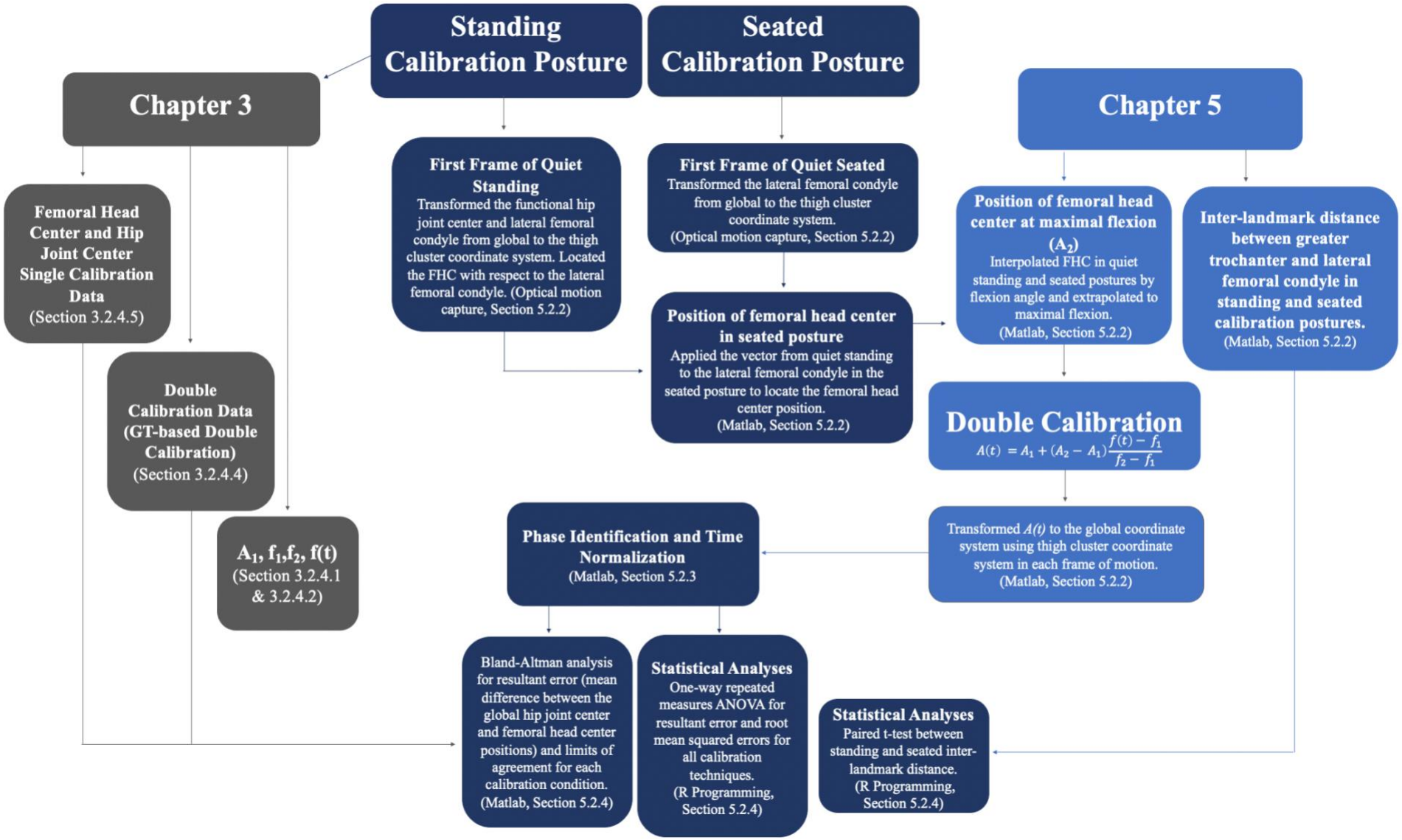


Figure 5.1: Overview of methodological procedures where data from Chapter 3 (grey boxes) were referenced and associated with Chapter 5 analyses.

5.2.1 Participants

The same participants were used as described in Chapter 3.

5.2.2 Data Processing

Processed single calibration femoral head center and hip joint center data from Section 3.2.4.5 were used for this analysis. For the double calibration, A_1, f_1, f_2, and $f(t)$ defined from Section 3.2.4.1 and 3.2.4.2 were used; however, the position of A_2 was modified. Using the position of the lateral femoral condyle in the first frame of the quiet standing trial, a transformation matrix was applied such that the lateral femoral condyle was transformed to the thigh cluster coordinate system using equation 2 from Section 3.2.4.2. Then, a vector was created between the location of the lateral femoral condyle and femoral head center in the thigh cluster coordinate system using equation 3 from Section 3.2.4.3. In the quiet seated calibration trial, a transformation matrix was created in the first frame of motion to transform the lateral femoral condyle from the global coordinate system to the thigh cluster coordinate system. The vector, created from the first frame of the standing trial as described above, from the lateral femoral condyle to the femoral head center was added to the position of the lateral femoral condyle in the thigh cluster coordinate system using equation 4 in Section 3.2.4.3 (Figure 5.2). To palpate the lateral femoral condyle in the standing posture, the researcher had participants slightly bend their knees and then located the most lateral point of the femur above the joint space between the tibia and femur. Participants straightened their knee for digitization, while the researcher continually followed the movement of the identified landmark via palpation with their finger. To digitize the lateral femoral condyle in the seated posture, the researcher similarly palpated the lateral femoral condyle when participants had their knees slightly flexed. Once located, the research held their

finger on the landmark as participants flexed their knee back into the flexed position in the seated posture.

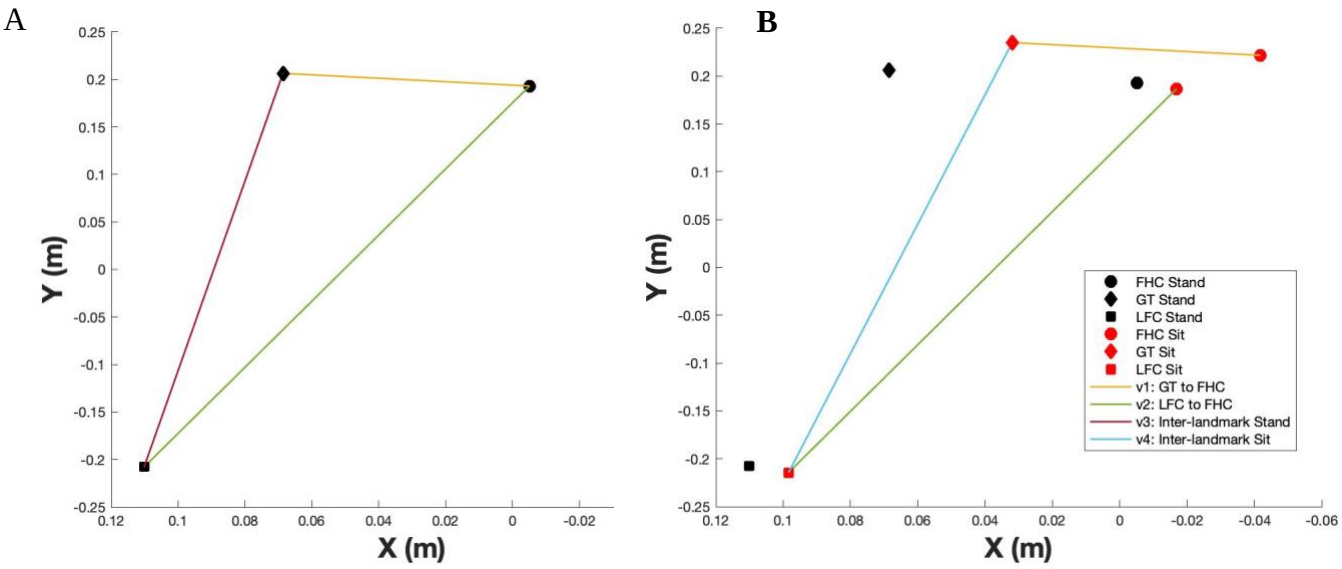


Figure 5.2: Position of the digitized greater trochanter (GT), lateral femoral condyle (LFC), and predicted femoral head center (FHC) position in the thigh cluster coordinate system for Subject 1. A 2D representation of a 3D landmark position was chosen to simplify visualization of the landmarks as if viewing from the anterolateral thigh. The orientation of the thigh coordinate system is provided in Appendix A. Plot A represented the standing calibration and three associated vectors: v_1, v_2 and v_3. v_1 was used to predict the location of the FHC in the seated calibration posture from the GT digitized landmark, while v_2 was used to predict the location of the FHC from the LFC digitized landmark. v_3 represented the inter-landmark distance between the GT and LFC. Plot B represented the digitized GT and LFC landmarks with the associated predictions of the FHC using v_1 and v_2. v_4 represented the inter-landmark distance between the GT and LFC in the seated calibration.

Using the position of the femoral head center in the standing and seated calibration trial, data points were linearly interpolated and extrapolated to the maximal knee flexion angle for each motion trial, as in Section 3.2.4.3 to calculate A_2, and then inputted into the double calibration equation presented in Section 3.2.4.1 and transformed to the global coordinate system following the same procedure in Section 3.2.4.4. This process was repeated for x, y, and z coordinates. An example figure of the new femoral head center position predicted from the

lateral femoral condyle is presented in Figure 5.3. The original prediction of the femoral head center using the greater trochanter was also included in the figure for visual observation of femoral head center location differences.

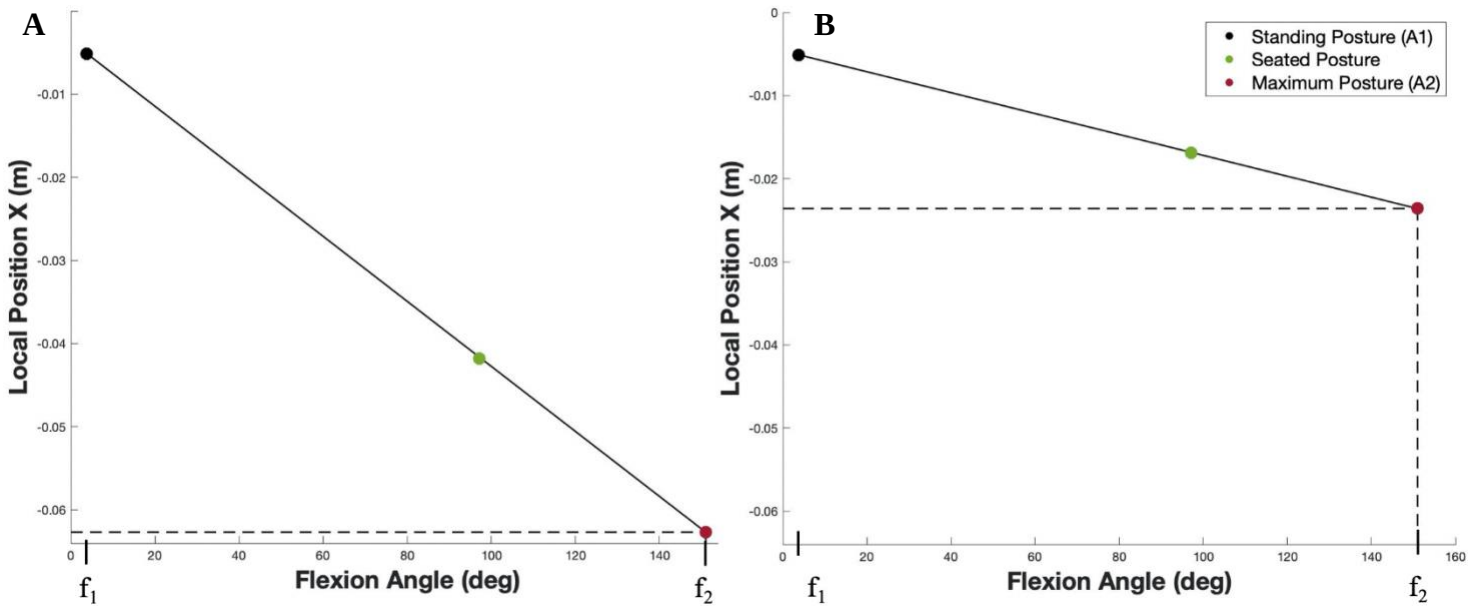


Figure 5.3: Three-point calibration curves in the thigh cluster coordinate system for the x-coordinate Subject 1. The femoral head center locations (red) at the maximum flexion posture (A2) were predicted using two methods to predict the femoral head center location in the seated calibration posture (green). The left plot (A) is from Chapter 3 (Figure 3.5) which predicted the femoral head center in the seated posture using the greater trochanter landmark. The right plot (B) predicted the femoral head center in the seated posture using the lateral femoral condyle. The femoral head center position in a standing posture (black) is the same in both calibration methods, but then diverge when creating a three-point calibration curve while passing through the femoral head center position in the seated posture. Flexion angle in the standing (f_1) and maximum posture (f_2), determined in a heels-up squat, are the same in both calibration conditions.

5.2.3 Data Analysis

The resultant error and RMS error between the femoral head center in the single calibration and hip joint center from Sections 3.2.4.7 and 3.2.4.8 were used for this analysis. Femoral head center LFC-based double calibration data were truncated using the start and end frames defined from Section 3.2.4.6 and were time normalized to 101 frames. RMS error and resultant errors were then calculated between the reconstructed femoral head center and hip joint

center positions using the same procedures in Sections 3.2.4.7 and 3.2.4.8. The mean RMS error and resultant error across all trials were calculated, followed by the mean of each trial within each high flexion condition, and the mean of all conditions for each participant. The inter-landmark distance between the digitized greater trochanter and digitized lateral femoral condyle in the thigh cluster coordinate system was then calculated for the standing and seated trials to quantify digitization error between standing and seated postures.

5.2.4 Statistical Analysis

A one-way repeated measures ANOVA was conducted to compare single-axis RMS error and resultant error across each of the three calibration conditions: 1) single calibration, 2) LFC-based double calibration, and 3) GT-based double calibration. For significant results after the Greenhouse-Geisser correction (since comparing 3 groups in the repeated measures ANOVA and correcting for sphericity), data were broken down through pairwise comparisons with a Bonferroni correction. If the assumption of normality was not met, the non-parametric equivalent, a Friedman-ANOVA, was conducted and a Wilcox Signed Rank Test broke down the data for pairwise-comparisons and was interpreted with a Bonferroni correction. A Wilcox-Signed Rank Test of the greater trochanter and lateral femoral condyle inter-landmark distance (v_3 and v_4 from Figure 5.2) was then conducted between the standing and seated calibration postures and was reported with a Bonferroni correction Bland-Altman analyses were also conducted using the same procedures as described in Section 3.2.5.2. The resultant error and limits of agreement were plotted for the single calibration, GT-based double calibration and LFC-based double calibration data for visual comparison of all evaluated calibration techniques.

5.3 Results

Results from the one-way ANOVAs for single-axis RMS error between single, GT-based and LFC-based double calibrations are presented in Table 5.1. RMS error along the global y and z axes were significantly different, where post-hoc pairwise comparisons revealed that the GT-based double calibration and LFC-based double calibration both had significantly greater RMS error compared to the single calibration. Comparisons also revealed that the GT-based double calibration had significantly greater RMS error in the y direction and significantly lower RMS-error in the z direction, compared to the LFC-based double calibration. The x-axis returned non-significant results.

Table 5.1: Results from one-way ANOVAs (α=0.05) for single-axis mean RMS error between the single calibration, greater trochanter-based (GT-based) double calibration and lateral femoral condyle-based (LFC-based) double calibration. RMS errors are calculated based on differences between the coordinates of the femoral head center (predicted from the thigh cluster and landmarks) and the hip joint center (gold standard, predicted from the pelvis cluster). RMS errors were presented as a mean (+/- standard deviation) in meters (m), for each global axis separately.

Condition		Single Calibration	GT-based Double Calibration	LFC-based Double Calibration	p-value
All Movements	x	0.021 (0.01)[a]	0.03 (0.02)[a]	0.023 (0.014)[a]	0.0913
	y	0.015 (0.01)[a]	0.04 (0.026)[b]	0.025 (0.018)[c]	< 0.05*
	z	0.01 (0.004)[a]	0.019 (0.011)[b]	0.033 (0.015)[c]	< 0.05*

Values with the same letter within a given row were not significant from each other
* $p < 0.05$

The mean resultant error and limits of agreement were plotted for both GT-based and LFC-based double calibrations, and the single calibration. Between both double calibrations, the resultant error was closest to zero for the LFC-based double calibration (Figure 5.4); however, it was still further away from zero compared to the single calibration. A one-way ANOVA revealed significant differences between the calibration techniques. Pairwise comparisons revealed the GT-based and LFC-based double calibration resultant error was significantly greater than the single calibration, but there were no significant differences between the GT-based and

LFC-based double calibration resultant error (Table 5.2). The limits of agreement were also narrower for the LFC-based double calibration (spanned 7.91 cm) compared to the GT-based double calibration (spanned 11.9 cm) (Table 5.3).

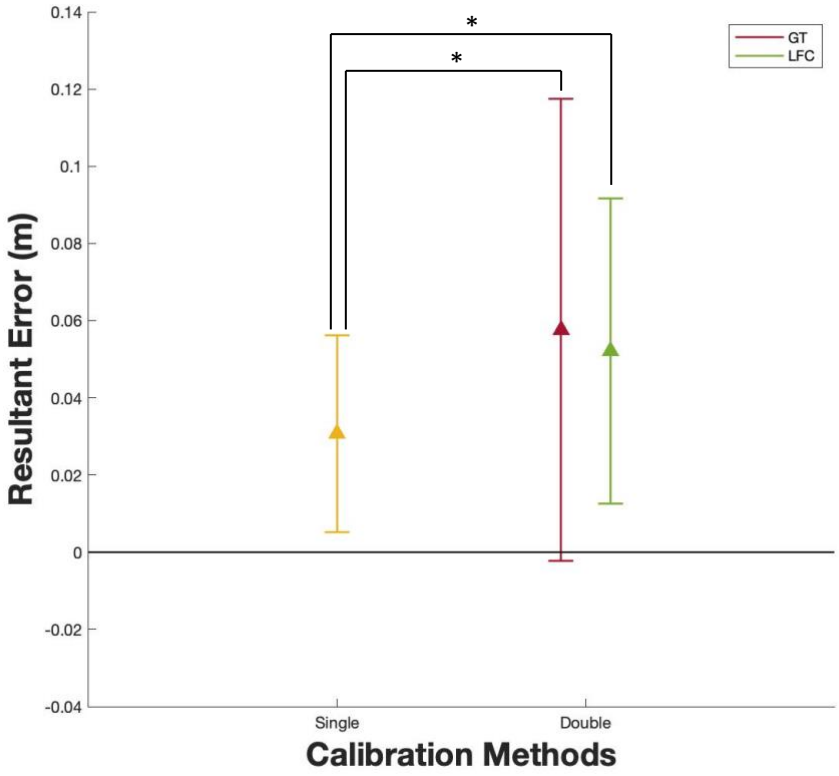


Figure 5.4: Resultant error plotted for the single, greater trochanter (GT) based and lateral femoral condyle (LFC) based double calibration methods. The standard deviation lines represent the limits of agreement, with a marked upper and lower threshold.

Table 5.2: Results from one-way ANOVAs (α=0.05) for resultant error between the single calibration, greater trochanter-based (GT-based) double calibration and lateral femoral condyle-based (LFC-based) double calibration. Resultant errors are calculated based on differences between the coordinates of the femoral head center (predicted from the thigh cluster and landmarks) and the hip joint center (gold standard, predicted from the pelvis cluster). Resultant errors were presented as a mean (+/- standard deviation) in meters (m), for each global axis separately.

Condition	Single Calibration	GT-based Double Calibration	LFC-based Double Calibration	p-value
All movements	0.031 (0.01)	0.058 (0.028)[a]	0.052 (0.018)[a]	<0.05*

Values with the same letter within a given row were not significant from each other
* $p < 0.05$

Table 5.3: Parameters from Figure 5.1 representing the single, greater trochanter (GT)-based, and lateral femoral condyle (LFC)-based double calibration resultant error and the limits of agreement. Values are presented in meters (m).

Parameter	Single Calibration	GT-based Double Calibration	LFC-based Double Calibration
Upper limit (m)	0.056	0.117	0.092
Mean bias (m)	0.031	0.058	0.052
Lower limit (m)	0.005	-0.002	0.013

A two-tailed paired t-test for mean inter-landmark distance between standing and seated calibration revealed no significant differences (Table 5.4). The mean inter-landmark distance (which can be interpreted as an approximate measure of the length of the thigh) in the seated calibration posture exceeded the mean inter-landmark distance in the standing posture by 0.2 cm (Table 5.4). However, the absolute mean within-participant difference in the inter-landmark distance between the standing and seated postures was 1.9 cm (+/- 1.3 cm). Within-participant thigh length (inter-landmark distance) measures are presented in Appendix E.

Table 5.4: Results from two-tailed paired t-tests (α=0.05) between the mean inter-distance of the greater trochanter to lateral femoral condyle in standing and seated calibration in the thigh cluster coordinate system. Inter landmark distance was calculated using the x, y and z position of the greater trochanter and lateral-femoral condyle. Inter-landmark distances were presented as a mean (+/- standard deviation) in meters (m).

	Standing Calibration	**Seated Calibration**	**p-value**	**Effect Size**
Inter-landmark Distance	0.417 (0.028)	0.419 (0.033)	0.711	-0.0894

* $p < 0.05$

5.4 Discussion

Like the GT-based double calibration, the LFC-based double calibration performed worse than the single calibration. When comparing both double calibration methods, the GT-based method had lower and higher RMS error in y and z positions, respectively, with no difference in x position. There were also no differences between GT-based and LFC-based resultant errors; however, the limits of agreement were narrower for the LFC-based method, indicating less variability amongst participants. Based on the above information, the LFC-based double calibration performed better than the GT-based double calibration; however, would still not be recommended as a superior technique to the single calibration. In a paired t-test, there was no significant difference in inter-landmark distance between the standing and seated calibration postures, which would indicate that the relative location of the greater trochanter and lateral femoral condyle digitization were similar. However, the mean within-participant difference in inter-landmark distance between standing and seated postures was 1.9 cm (+/- 1.3 cm), which would indicate meaningful uncertainty in the accuracy of landmark palpation. This finding cannot directly indicate palpation inaccuracy in the seated calibration posture, but one could reasonably assume the standing palpation was more accurate due to ease of palpation. Therefore, this meaningful uncertainty can be interpreted as evidence that palpation inaccuracy, more in the

seated position, likely contributed to the overall poor performance of the double calibration. Although no previous literature evaluated palpation accuracy of the thigh landmarks in a seated posture, the results from this thesis showed no differences between mean GT-based and LFC-based double calibration resultant error.

The investigation of two-different landmarks in the double calibration technique stemmed from feedback received during data collection, that in the seated calibration posture, the lateral femoral condyle seemed to be easier to palpate than the greater trochanter. This speculation could not be verified based on data collected; however, differences in palpation accuracy between landmarks would influence double calibration accuracy because the landmarks were used to predict the position of the femoral head center in the seated calibration posture. The seated calibration posture was also used by Cappello et al. (2005), and no difficulties in palpation were noted. Although in this thesis there were some differences in RMS error when different landmarks were used, they were not all favorable towards one double calibration method. The narrower limits of agreement in the LFC-based double calibration indicated less variability in resultant error amongst participants compared to the GT-based double calibration; however, it still would not be recommended over the single calibration.

5.5 Limitations

The conclusions from this Chapter were limited by true assessment of palpation accuracy, which could have been achieved by radiographic imaging. Instead, the inter-landmark distance between the greater trochanter and lateral femoral condyle was used to approximate palpation accuracy by observing if the landmarks were digitized approximately equal distances from each other between the two calibration postures.

Chapter 6– Conclusions and Future Directions

The results of this thesis concluded that the double calibration, which was effective in reducing STA in previous work (Cappello et al., 1997, 2005; Stagni et al., 2006), was not effective at reducing the effects of STA compared to single calibration using the mid-anterolateral thigh cluster in high flexion. The purpose of this thesis was to not replicate the methodology of Cappello et al (2005), but to use the double calibration equation and to apply it to high flexion postures. This thesis evaluated the double calibration technique in high flexion squatting and kneeling, used a rigid marker cluster to track thigh motion throughout high flexion movements, predicted the location of the femoral head center at maximum flexion (A_2) using the two-point calibration curve, and used the hip joint center as the gold standard. This methodology differed from that of Cappello et al. (2005), in that the mean range of the movements were ~95 degrees knee flexion (therefore not capturing the thigh-calf contact phase), the thigh anatomical landmarks were palpated at standing and maximum flexion postures, a non-rigid marker cluster was used to track thigh motion, and fluoroscopy was the gold standard. Based on the results of this thesis, we speculated that the prediction of A_2 in high flexion using the methodology in this thesis did not accurately capture the relative movement of the anatomical landmarks in high flexion movements, and thus was the main contributor to double calibration inaccuracies.

Chapter 3 revealed the linearity assumption in the double calibration was not appropriate for modelling high flexion tasks. Despite having palpated the greater trochanter in a seated posture and allowing the natural movement of the thigh cluster, extrapolation of the calibration curve likely did not represent true anatomical landmark movement. Thigh-calf contact force showed non-linear and exponential relationships in previous literature (Kingston & Acker, 2018; Zelle et al., 2007), and this intersegmental contact will influence soft tissue behaviour in high

flexion movements, which would not have been captured using this double calibration technique. Future investigations should consider modelling the change in anatomical landmark location from a standing position to full ROM that includes thigh-calf contact. This work can then be applied to the double calibration approach in hope to improve its performance for high flexion squatting and kneeling. Chapter 3 also revealed increased variability in the double calibration data, leading to future investigations in Chapter 4, that investigated how mid-thigh circumference might affect the performance of double calibration.

In Chapter 4, resultant error increased from 15th to 50th percentile groups while the limits of agreement increased from 15th to 75th percentile groups, indicating larger error and greater variability with increasing mid-thigh circumferences. Larger thigh circumference may indicate greater soft tissue volume, which could make accurate palpation of bony landmarks more challenging. This line of reasoning, in addition to wider limits of agreement across the entire study sample for the double calibration in Chapter 3, led to preliminary investigation of palpation accuracy in Chapter 5. The single calibration data revealed no observed relationship between mid-thigh circumference and resultant error; however, limits of agreement increased from 25th to 75th percentile groups. Future investigations should consider soft tissue composition and its effect on STA, as mid-thigh circumference did not capture a relationship of increasing resultant error with percentile groups in the single calibration.

Chapter 5 revealed the LFC-based double calibration did not present a significant improvement in double calibration, although the limits of agreement were narrower for the LFC-based double calibration indicating less variability in data when using the lateral femoral condyle to predict the femoral head center, opposed to the greater trochanter. In addition to future directions discussed for palpation accuracy, further investigations should also consider

developing guidelines for palpation in seated and high flexion calibration postures for the purpose of performing double calibrations. Previous research found that beyond 120 degrees flexion, the femur had greater M/L translation compared to between 30 and 120 degrees (Qi et al., 2013) and slight internal rotation compared to a flexion angle below 120 degrees (Pinskerova et al., 2009). Without guidelines for palpation in these postures, researchers may not capture the true movement of the underlying landmarks for double calibrations, which will directly affect reported results and the accuracy of this calibration method. The more trust one has in results, the better performance and confidence the results are for drawing both clinically and biomechanically meaningful contributions. Once accurate palpation is achieved and the behaviour of STA error in high flexion is modelled, the double calibration technique could be revisited to evaluate its effectiveness in reducing STA in high flexion.

References

Akbarshahi, M., Schache, A. G., Fernandez, J. W., Baker, R., Banks, S., & Pandy, M. G. (2010). Non-invasive assessment of soft-tissue artifact and its effect on knee joint kinematics during functional activity. *Journal of Biomechanics*, *43*(7), 1292–1301. https://doi.org/10.1016/j.jbiomech.2010.01.002

Amin, S., Goggins, J., Niu, J., Guermazi, A., Grigoryan, M., Hunter, D. J., Genant, H. K., & Felson, D. T. (2008). Occupation-related squatting, kneeling, and heavy lifting and the knee joint: A magnetic resonance imaging-based study in men. *Journal of Rheumatology*, *35*(8), 1645–1649.

Andersen, M. S., Benoit, D. L., Damsgaard, M., Ramsey, D. K., & Rasmussen, J. (2010). Do kinematic models reduce the effects of soft tissue artefacts in skin marker-based motion analysis? An in vivo study of knee kinematics. *Journal of Biomechanics*, *43*(2), 268–273. https://doi.org/10.1016/j.jbiomech.2009.08.034

Andersen, M. S., Damsgaard, M., Rasmussen, J., Ramsey, D. K., & Benoit, D. L. (2012). A linear soft tissue artefact model for human movement analysis: Proof of concept using in vivo data. *Gait and Posture*, *35*(4), 606–611. https://doi.org/10.1016/j.gaitpost.2011.11.032

Andriacchi, T. P., Alexander, E. J., Toney, M. K., Dyrby, C., & Sum, J. (1998). A point cluster method for in vivo motion analysis: Applied to a study of knee kinematics. *Journal of Biomechanical Engineering*, *120*(6), 743–749. https://doi.org/10.1115/1.2834888

Barré, A., Jolles, B. M., Theumann, N., & Aminian, K. (2015). Soft tissue artifact distribution on lower limbs during treadmill gait: Influence of skin markers' location on cluster design. *Journal of Biomechanics*, *48*(10), 1965–1971. https://doi.org/10.1016/j.jbiomech.2015.04.007

Barré, A., Thiran, J., Jolles, B. M., Theumann, N., & Amianim, K. (2013). Walking in Subjects With Total Knee Arthroplasty. *IEEE Transactions on Biomedical Engineering*, *60*(11), 3131–3140.

Benoit, D.L., Damsgaard, M., & Andersen, M. S. (2015). Surface marker cluster translation, rotation, scaling and deformation: Their contribution to soft tissue artefact and impact on knee joint kinematics. *Journal of Biomechanics*, *48*(10), 2124–2129. https://doi.org/10.1016/j.jbiomech.2015.02.050

Benoit, Daniel L., Ramsey, D. K., Lamontagne, M., Xu, L., Wretenberg, P., & Renström, P. (2006). Effect of skin movement artifact on knee kinematics during gait and cutting motions measured in vivo. *Gait and Posture*, *24*(2), 152–164. https://doi.org/10.1016/j.gaitpost.2005.04.012

Besier, T. F., Gold, G. E., Beaupré, G. S., & Delp, S. L. (2005). A modeling framework to estimate patellofemoral joint cartilage stress in vivo. *Medicine and Science in Sports and Exercise*, *37*(11), 1924–1930. https://doi.org/10.1249/01.mss.0000176686.18683.64

Blache, Y., Dumas, R., Lundberg, A., & Begon, M. (2017). Main component of soft tissue artifact of the upper-limbs with respect to different functional, daily life and sports movements. *Journal of Biomechanics*, *62*, 39–46. https://doi.org/10.1016/j.jbiomech.2016.10.019

Bland, J. M., & Altman, D. G. (2007). Agreement between methods of measurement with multiple observations per individual. *Journal of Biopharmaceutical Statistics*, *17*(4), 571–582. https://doi.org/10.1080/10543400701329422

Buchman-Pearle, J.M. (2020). *An Investigation of Methods to Attenuate Soft Tissue Artifact of the Thigh in High Knee Flexion* [Unpublished Masters thesis]. University of Waterloo.

Buchman-Pearle, J. M., & Acker, S. M. (2021). Estimating soft tissue artifact of the thigh in high knee flexion tasks using optical motion Capture: Implications for marker cluster placement. *Journal of Biomechanics*, *127*(July), 110659. https://doi.org/10.1016/j.jbiomech.2021.110659

Camomilla, V., Bonci, T., Dumas, R., Chèze, L., & Cappozzo, A. (2015). A model of the soft tissue artefact rigid component. *Journal of Biomechanics*, *48*(10), 1752–1759. https://doi.org/10.1016/j.jbiomech.2015.05.007

Camomilla, Valentina., Cereatti, A., Chèze, L., & Cappozzo, A. (2013). A hip joint kinematics driven model for the generation of realistic thigh soft tissue artefacts. *Journal of Biomechanics*, *46*(3), 625–630. https://doi.org/10.1016/j.jbiomech.2012.09.018

Camomilla, Valentina, Cereatti, A., Vannozzi, G., & Cappozzo, A. (2006). An optimized protocol for hip joint centre determination using the functional method. *Journal of Biomechanics*, *39*(6), 1096–1106. https://doi.org/10.1016/j.jbiomech.2005.02.008

Cappello, A., Cappozzo, A., La Palombara, P. F., Lucchetti, L., & Leardini, A. (1997). Multiple anatomical landmark calibration for optimal bone pose estimation. *Human Movement Science*, *16*(2–3), 259–274. https://doi.org/10.1016/S0167-9457(96)00055-3

Cappello, A., Stagni, R., Fantozzi, S., & Leardini, A. (2005). Soft tissue artifact compensation in knee kinematics by double anatomical landmark calibration: Performance of a novel method during selected motor tasks. *IEEE Transactions on Biomedical Engineering*, *52*(6), 992–998. https://doi.org/10.1109/TBME.2005.846728

Cappozzo, A., Catani, F., Leardini, A., Benedetti, M. G., & Della Croce, U. (1996). Position and orientation in space of bones during movement: Experimental artefacts. *Clinical Biomechanics*, *11*(2), 90–100. https://doi.org/10.1016/0268-0033(95)00046-1

Clément, J., de Guise, J. A., Fuentes, A., & Hagemeister, N. (2018). Comparison of soft tissue artifact and its effects on knee kinematics between non-obese and obese subjects performing a squatting activity recorded using an exoskeleton. *Gait and Posture*, *61*(January), 197–203. https://doi.org/10.1016/j.gaitpost.2018.01.009

Clément, J., Dumas, R., Hagemeister, N., & de Guise, J. A. (2015). Soft tissue artifact compensation in knee kinematics by multi-body optimization: Performance of subject-specific knee joint models. *Journal of Biomechanics*, *48*(14), 3796–3802. https://doi.org/10.1016/j.jbiomech.2015.09.040

Cockcroft, J., Louw, Q., & Baker, R. (2016). Proximal placement of lateral thigh skin markers reduces soft tissue artefact during normal gait using the Conventional Gait Model. *Computer Methods in Biomechanics and Biomedical Engineering*, *19*(14), 1497–1504. https://doi.org/10.1080/10255842.2016.1157865

Coggon, D., Croft, P., Kellingray, S., Barrett, D., McLaren, M., & Cooper, C. (2000). Occupational activities and osteoarthritis of the knee. *Arthritis & Rheumatism*, *43*(7), 1443–1449. https://doi.org/10.1093/bmb/lds012

Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science*, *1*(3), 98–101.

Cooper, C., McAlindon, T., Coggon, D., Egger, P., & Dieppe, P. (1994). Occupational activity and osteoarthritis of the knee. *Annals of the Rheumatic Diseases*, *53*(2), 90–93. https://doi.org/10.1136/ard.53.2.90

Dulay, G. S., Cooper, C., & Dennison, E. M. (2015). Knee pain, knee injury, knee osteoarthritis

& work. *Best Practice and Research: Clinical Rheumatology*, *29*(3), 454–461. https://doi.org/10.1016/j.berh.2015.05.005

Dumas, R., Camomilla, V., Bonci, T., Cheze, L., & Cappozzo, A. (2014). A qualitative analysis of soft tissue artefact during running. *Computer Methods in Biomechanics and Biomedical Engineering*, *17*(SUPP1), 124–125. https://doi.org/10.1080/10255842.2014.931518

Ehrig, R. M., Taylor, W. R., Duda, G. N., & Heller, M. O. (2006). A survey of formal methods for determining the centre of rotation of ball joints. *Journal of Biomechanics*, *39*(15), 2798–2809. https://doi.org/10.1016/j.jbiomech.2005.10.002

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behaviour Research Methods*, *39*(2), 175–191.

Felson, D. T., & Zhang, Y. (1998). An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. *Arthritis and Rheumatism*, *41*(8), 1343–1355.

Fiorentino, N. M., Atkins, P. R., Kutschke, M. J., Bo Foreman, K., & Anderson, A. E. (2020). Soft tissue artifact causes underestimation of hip joint kinematics and kinetics in a rigid-body musculoskeletal model. *Journal of Biomechanics*, *108*, 109890. https://doi.org/10.1016/j.jbiomech.2020.109890

Fiorentino, N. M., Atkins, P. R., Kutschke, M. J., Foreman, K. B., & Anderson, A. E. (2016). In-vivo quantification of dynamic hip joint center errors and soft tissue artifact. *Gait and Posture*, *50*, 246–251. https://doi.org/10.1016/j.gaitpost.2016.09.011

Fiorentino, N. M., Atkins, P. R., Kutschke, M. J., Goebel, J. M., Foreman, K. B., & Anderson, A. E. (2017). Soft tissue artifact causes significant errors in the calculation of joint angles and range of motion at the hip. *Gait and Posture*, *55*(December 2016), 184–190. https://doi.org/10.1016/j.gaitpost.2017.03.033

Gao, B., & Zheng, N. (Nigel). (2008). Investigation of soft tissue movement during level walking: Translations and rotations of skin markers. *Journal of Biomechanics*, *41*(15), 3189–3195. https://doi.org/10.1016/j.jbiomech.2008.08.028

Grimpampi, E., Camomilla, V., Cereatti, A., De Leva, P., & Cappozzo, A. (2014). Metrics for describing soft-tissue artefact and its effect on pose, size, and shape of marker clusters. *IEEE Transactions on Biomedical Engineering*, *61*(2), 362–367. https://doi.org/10.1109/TBME.2013.2279636

Grood, E. S., & Suntay, W. J. (1983). A joint coordinate system for the clinical description of three-dimensional motions: Application to the knee. *Journal of Biomechanical Engineering*, *105*(2), 136–144. https://doi.org/10.1115/1.3138397

Hamai, S., Dunbar, N. J., Moro-Oka, T. A., Miura, H., Iwamoto, Y., & Banks, S. A. (2013). Physiological sagittal plane patellar kinematics during dynamic deep knee flexion. *International Orthopaedics*, *37*(8), 1477–1482. https://doi.org/10.1007/s00264-013-1958-6

Hamai, S., Moro-oka, T. A., Miura, H., Shimoto, T., Higaki, H., Fregly, B. J., Iwamoto, Y., & Banks, S. A. (2009). Knee kinematics in medial osteoarthritis during in vivo weight-bearing activities. *Journal of Orthopaedic Research*, *27*(12), 1555–1561. https://doi.org/10.1002/jor.20928

Han, S., Ge, S., & Liu, H. (2014). Gender differences in lower extremity kinematics during high range of motion activities. *Journal of Medical Imaging and Health Informatics*, *4*(2), 272–276. https://doi.org/10.1166/jmihi.2014.1236

Howarth, S. J., & Callaghan, J. P. (2009). The rule of 1 s for padding kinematic data prior to digital filtering: Influence of sampling and filter cutoff frequencies. *Journal of*

Electromyography and Kinesiology, *19*(5), 875–881. https://doi.org/10.1016/j.jelekin.2008.03.010

Howarth, S. J., & Callaghan, J. P. (2010). Quantitative assessment of the accuracy for three interpolation techniques in kinematic analysis of human movement. *Computer Methods in Biomechanics and Biomedical Engineering*, *13*(6), 847–855. https://doi.org/10.1080/10255841003664701

Jensen, K., Mikkelsen, S., Loft, I. P., & Eenberg, W. (2000). *Radiographic knee osteoarthritis in floorlayers and carpenters Author (s): Lilli Kirkeskov Jensen , Sigurd Mikkelsen , Inger P Loft , Winnie Eenberg , Inge Bergmann and Vibeke Løgager Source : Scandinavian Journal of Work , Environment & Health , June 2*. *26*(3), 257–262.

Jordan, J., Kingston, R., Lane, N., Nevitt, M., Zhang, Y., & Showers, M. (2000). NIH Conference Osteoarthritis : New Insights. *Annals of Internal Medicine*, *133*(8), 637–639. https://doi.org/10.7326/0003-4819-133-8-200010170-00016

Kainz, H., Carty, C. P., Modenese, L., Boyd, R. N., & Lloyd, D. G. (2015). Estimation of the hip joint centre in human motion analysis: A systematic review. *Clinical Biomechanics*, *30*(4), 319–329. https://doi.org/10.1016/j.clinbiomech.2015.02.005

Kapron, A. L., Aoki, S. K., Peters, C. L., Maas, S. A., Bey, M. J., Zauel, R., & Anderson, A. E. (2014). Accuracy and feasibility of dual fluoroscopy and model-based tracking to quantify in vivo hip kinematics during clinical exams. *Journal of Applied Biomechanics*, *30*(3), 461–470. https://doi.org/10.1123/jab.2013-0112

Kingston, D. C., & Acker, S. M. (2018). Thigh-calf contact parameters for six high knee flexion postures: Onset, maximum angle, total force, contact area, and center of force. *Journal of Biomechanics*, *67*, 46–54. https://doi.org/10.1016/j.jbiomech.2017.11.022

Kuo, M. Y., Tsai, T. Y., Lin, C. C., Lu, T. W., Hsu, H. C., & Shen, W. C. (2011). Influence of soft tissue artifacts on the calculated kinematics and kinetics of total knee replacements during sit-to-stand. *Gait and Posture*, *33*(3), 379–384. https://doi.org/10.1016/j.gaitpost.2010.12.007

Li, G., Zayontz, S., DeFrate, L. E., Most, E., Suggs, J. F., & Rubash, H. E. (2004). Kinematics of the knee at high flexion angles: An in vitro investigation. *Journal of Orthopaedic Research*, *22*(1), 90–95. https://doi.org/10.1016/S0736-0266(03)00118-9

Li, J., Tsai, T., Clancy, M. M., Li, G., Lewis, C. L., & Felson, D. T. (2019). Weight loss changed gait kinematics in individuals with obesity and knee pain. *Gait and Posture*, *68*(December 2018), 461–465. https://doi.org/10.1016/j.gaitpost.2018.12.031

Lin, C. C., Lu, T. W., Lu, H. L., Kuo, M. Y., & Hsu, H. C. (2016). Effects of soft tissue artifacts on differentiating kinematic differences between natural and replaced knee joints during functional activity. *Gait and Posture*, *46*, 154–160. https://doi.org/10.1016/j.gaitpost.2016.03.006

Lucchetti, L., Cappozzo, A., Cappello, A., & Della Croce, U. (1998). Skin movement artefact assessment and compensation in the estimation of knee-joint kinematics. *Journal of Biomechanics*, *31*(11), 977–984. https://doi.org/10.1016/S0021-9290(98)00083-9

McClelland, J. A., Webster, K. E., Feller, J. A., & Menz, H. B. (2011). Knee kinematics during walking at different speeds in people who have undergone total knee replacement. *Knee*, *18*(3), 151–155. https://doi.org/10.1016/j.knee.2010.04.005

McDowell, M. A., Fryar, C. D., Ogden, C. L., & Flegal, K. M. (2008). Anthropometric reference data for children and adults: United States, 2003-2006. *National Health Statistics Reports*, *10*, 2003–2006.

McMillan, G., & Nichols, L. (2005). Osteoarthritis and meniscus disorders of the knee as occupational diseases of miners. *Occupational and Environmental Medicine*, *62*(8), 567–575. https://doi.org/10.1136/oem.2004.017137

Nagura, T., Dyrby, C. O., Alexander, E. J., & Andriacchi, T. P. (2001). Mechanical loads on the knee joint during deep flexion. *American Society of Mechanical Engineers, Bioengineering Division (Publication) BED*, *50*, 393–394.

Peters, A., Galna, B., Sangeux, M., Morris, M., & Baker, R. (2010). Quantification of soft tissue artifact in lower limb human motion analysis: A systematic review. *Gait and Posture*, *31*(1), 1–8. https://doi.org/10.1016/j.gaitpost.2009.09.004

Pinskerova, V., Samuelson, K. M., Stammers, J., Maruthainar, K., Sosna, A., & Freeman, M. A. R. (2009). The knee in full flexion: An anatomical study. *Journal of Bone and Joint Surgery - Series B*, *91*(6), 830–834. https://doi.org/10.1302/0301-620X.91B6.22319

Qi, W., Hosseini, A., Tsai, T. Y., Li, J. S., Rubash, H. E., & Li, G. (2013). In vivo kinematics of the knee during weight bearing high flexion. *Journal of Biomechanics*, *46*(9), 1576–1582. https://doi.org/10.1016/j.jbiomech.2013.03.014

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ryu, T. (2012). Application of Soft Tissue Artifact Compensation Using Displacement Dependency between Anatomical Landmarks and Skin Markers. *Anatomy Research International*, *2012*, 1–10. https://doi.org/10.1155/2012/123713

Ryu, T., Choi, H. S., & Chung, M. K. (2009). Soft tissue artifact compensation using displacement dependency between anatomical landmarks and skin markers - a preliminary study. *International Journal of Industrial Ergonomics*, *39*(1), 152–158. https://doi.org/10.1016/j.ergon.2008.05.005

Sangeux, M., Peters, A., & Baker, R. (2011). Hip joint centre localization: Evaluation on normal subjects in the context of gait analysis. *Gait and Posture*, *34*(3), 324–328. https://doi.org/10.1016/j.gaitpost.2011.05.019

Schache, A. G., Baker, R., & Lamoreux, L. W. (2008). Influence of thigh cluster configuration on the estimation of hip axial rotation. *Gait and Posture*, *27*(1), 60–69. https://doi.org/10.1016/j.gaitpost.2007.01.002

Schwartz, M. H., & Rozumalski, A. (2005). A new method for estimating joint parameters from motion data. *Journal of Biomechanics*, *38*(1), 107–116. https://doi.org/10.1016/j.jbiomech.2004.03.009

Shybut, T. B., Vega, C. E., Haddad, J., Alexander, J. W., Gold, J. E., Noble, P. C., & Lowe, W. R. (2015). Effect of lateral meniscal root tear on the stability of the anterior cruciate ligament-deficient knee. *American Journal of Sports Medicine*, *43*(4), 905–911. https://doi.org/10.1177/0363546514563910

Stagni, R., Fantozzi, S., & Cappello, A. (2006). Propagation of anatomical landmark misplacement to knee kinematics: Performance of single and double calibration. *Gait and Posture*, *24*(2), 137–141. https://doi.org/10.1016/j.gaitpost.2006.08.001

Stagni, R., Fantozzi, S., & Cappello, A. (2009). Double calibration vs. global optimisation: Performance and effectiveness for clinical application. *Gait and Posture*, *29*(1), 119–122. https://doi.org/10.1016/j.gaitpost.2008.07.008

Stagni, R., Fantozzi, S., Cappello, A., & Leardini, A. (2005). Quantification of soft tissue artefact in motion analysis by combiningStagni, R., Fantozzi, S., Cappello, A., & Leardini, A. (2005). Quantification of soft tissue artefact in motion analysis by combining 3D

fluoroscopy and stereophotogrammetry: A study on. *Clinical Biomechanics*, *20*(3), 320–329. https://doi.org/10.1016/j.clinbiomech.2004.11.012

Stagni, R., Leardini, A., Cappozzo, A., Grazia Benedetti, M., & Cappello, A. (2000). Effects of hip joint centre mislocation on gait analysis results. *Journal of Biomechanics*, *33*(11), 1479–1487. https://doi.org/10.1016/S0021-9290(00)00093-2

Sun, H. B. (2010). Mechanical loading, cartilage degradation, and arthritis. *Annals of the New York Academy of Sciences*, *1211*, 37–50. https://doi.org/10.1111/j.1749-6632.2010.05808.x

Tsai, T. Y., Lu, T. W., Kuo, M. Y., & Lin, C. C. (2011). Effects of soft tissue artifacts on the calculated kinematics and kinetics of the knee during stair-ascent. *Journal of Biomechanics*, *44*(6), 1182–1188. https://doi.org/10.1016/j.jbiomech.2011.01.009

Williams, A., & Logan, M. (2004). Understanding Tibio-Femoral motion. *Knee*, *11*(2), 81–88. https://doi.org/10.1016/j.knee.2003.12.004

Winter, D. a D. a. D. a. (2009). Biomechanics and motor control of human body. In *Motor Control* (Vol. 2nd). https://doi.org/10.1002/9780470549148

Yang, N. H., Canavan, P. K., & Nayeb-Hashemi, H. (2010). The effect of the frontal plane tibiofemoral angle and varus knee moment on the contact stress and strain at the knee cartilage. *Journal of Applied Biomechanics*, *26*(4), 432–443. https://doi.org/10.1123/jab.26.4.432

Yao, J., Lancianese, S. L., Hovinga, K. R., Lee, J., & Lerner, A. L. (2008). Magnetic resonance image analysis of meniscal translation and tibio-menisco-femoral contact in deep knee flexion. *Journal of Orthopaedic Research*, *26*(5), 673–684. https://doi.org/10.1002/jor.20553

Zelle, J., Barink, M., De Waal Malefijt, M., & Verdonschot, N. (2009). Thigh-calf contact: Does it affect the loading of the knee in the high-flexion range? *Journal of Biomechanics*, *42*(5), 587–593. https://doi.org/10.1016/j.jbiomech.2008.12.015

Zelle, J., Barink, M., Loeffen, R., De Waal Malefijt, M., & Verdonschot, N. (2007). Thigh-calf contact force measurements in deep knee flexion. *Clinical Biomechanics*, *22*(7), 821–826. https://doi.org/10.1016/j.clinbiomech.2007.03.009

Zhang, Y., Hunter, D. J., Nevitt, M. C., Xu, L., Niu, J., Lui, L. Y., Yu, W., Aliabadi, P., & Felson, D. T. (2004). Association of Squatting With Increased Prevalence of Radiographic Tibiofemoral Knee Osteoarthritis: The Beijing Osteoarthritis Study. *Arthritis and Rheumatism*, *50*(4), 1187–1192. https://doi.org/10.1002/art.20127

Appendices

Appendix A: Thigh Cluster Coordinate System

Mid-anterolateral thigh cluster with marker labels

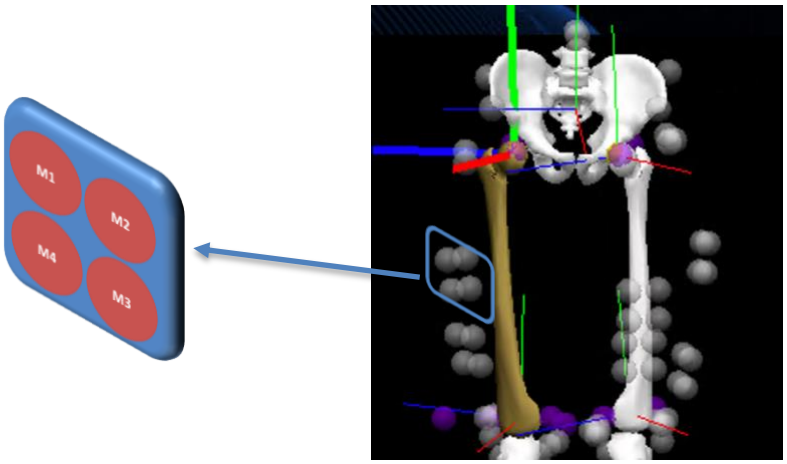


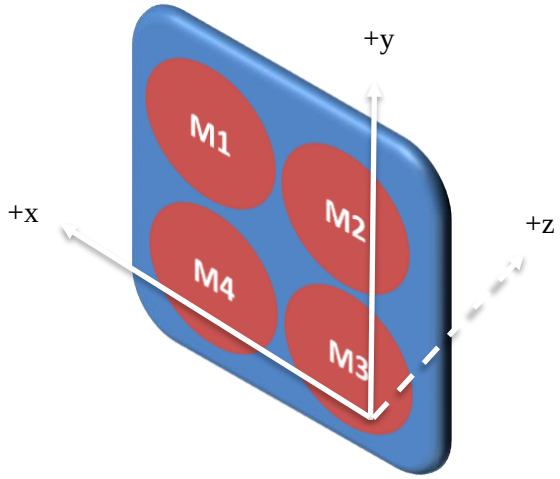
Figure 8.1: Image of the mid-anterolateral thigh marker cluster (outlined in blue) from Visual 3D and the corresponding marker labels to create the thigh cluster coordinate system.

Thigh Cluster Coordinate System

Table 8.1: Thigh cluster coordinate system with corresponding marker labels.

origin	M3
y-axis	Vector connecting M2 and M3
temporary x-axis	Vector connecting M4 and M3
z-axis	Cross product of temporary x-axis and y-axis, to produce z-axis
x-axis	Cross product of y-axis and z-axis

Orientation of Thigh Cluster Coordinate System



Appendix B: Statistical Analyses from Chapter 3 Separated by High Flexion Movement

Table 8 2: Results from one-tailed paired t-tests (α=0.05) for mean RMS error in x, y and z calculated between single and double calibrations. The femoral head center position in both calibration conditions were compared against the hip joint center position in each frame of high flexion movement. RMS errors were presented as a mean (+/- standard deviation) in meters (m).

Condition		Double Calibration	Single Calibration	p-value	Effect Size
Flatfoot squatting	x	0.027 (0.018)	0.025 (0.016)	0.683	0.139
	y	0.033 (0.026)	0.017 (0.014)	1	0.662
	z	0.017 (0.011)	0.011 (0.007)	0.995	0.496
Heels-up squatting	x	0.031 (0.021)	0.021 (0.012)	0.996	0.491
	y	0.039 (0.029)	0.017 (0.014)	1	0.808
	z	0.021 (0.013)	0.012 (0.006)	1	0.776
Dorsiflexed kneeling	x	0.03 (0.025)	0.018 (0.01)	0.984	0.437
	y	0.042 (0.031)	0.014 (0.008)	1	0.906
	z	0.018 (0.012)	0.009 (0.005)	1	0.666
Plantarflexed kneeling	x	0.031 (0.025)	0.022 (0.01)	0.962	0.375
	y	0.046 (0.031)	0.013 (0.009)	1	1.05
	z	0.018 (0.012)	0.009 (0.005)	1	0.682

Table 8.3: Results from one-tailed paired t-tests (α=0.05) for mean resultant error between single and double calibrations. Resultant error was calculated using the x, y and z position of the femoral head center reconstructed from the single and double calibration and the hip joint center for each high flexion movement. Resultant errors were presented as a mean (+/- standard deviation) in meters (m).

Condition	Double Calibration	Single Calibration	p-value	Effect Size
Flatfoot squatting	0.05 (0.026)	0.035 (0.016)	0.999	0.630
Heels-up squatting	0.058 (0.029)	0.032 (0.014)	1	0.996
Dorsiflexed kneeling	0.061 (0.031)	0.027 (0.009)	1	1.14
Plantarflexed kneeling	0.063 (0.033)	0.029 (0.009)	1	1.07

* $p < 0.05$

Appendix C: Participant Mean of Femoral Head Center and Hip Joint Center Magnitude, and the Difference in Mean Magnitude (mean distance) For Each Calibration Condition.

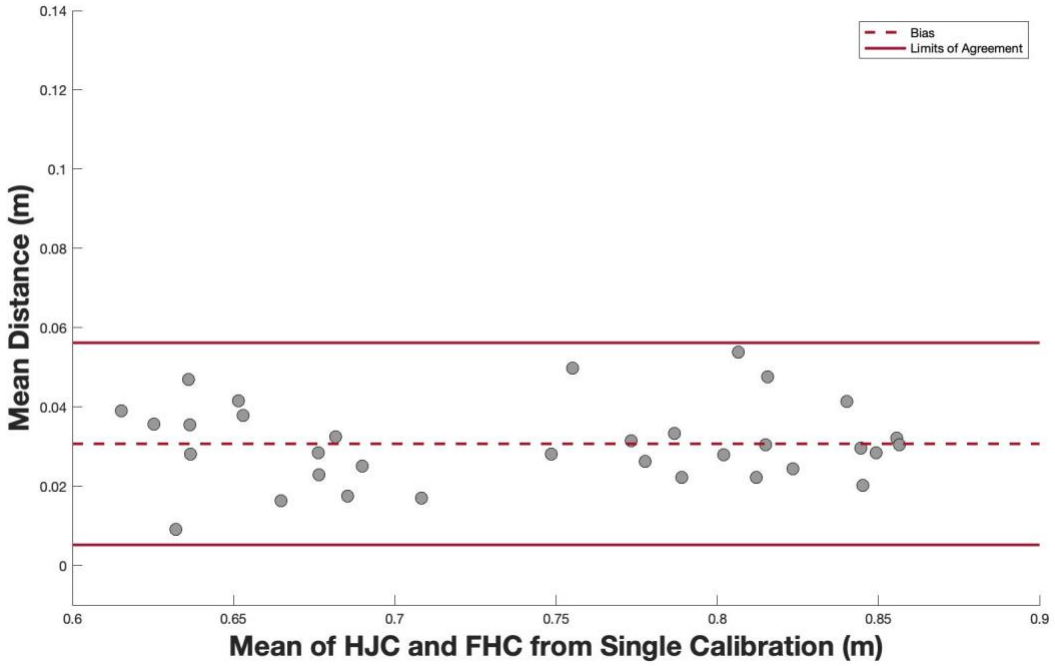


Figure 8.2: Mean of the magnitude of the x, y, and z coordinates of the hip joint center (HJC) and femoral head center (FHC) plotted against the mean distance between the HJC and FHC for each participant in the single calibration. The magnitude of the x, y and z coordinates were calculated as the vector norm for the HJC and FHC, while the mean distance was calculated as the vector norm difference between the HJC and FHC.

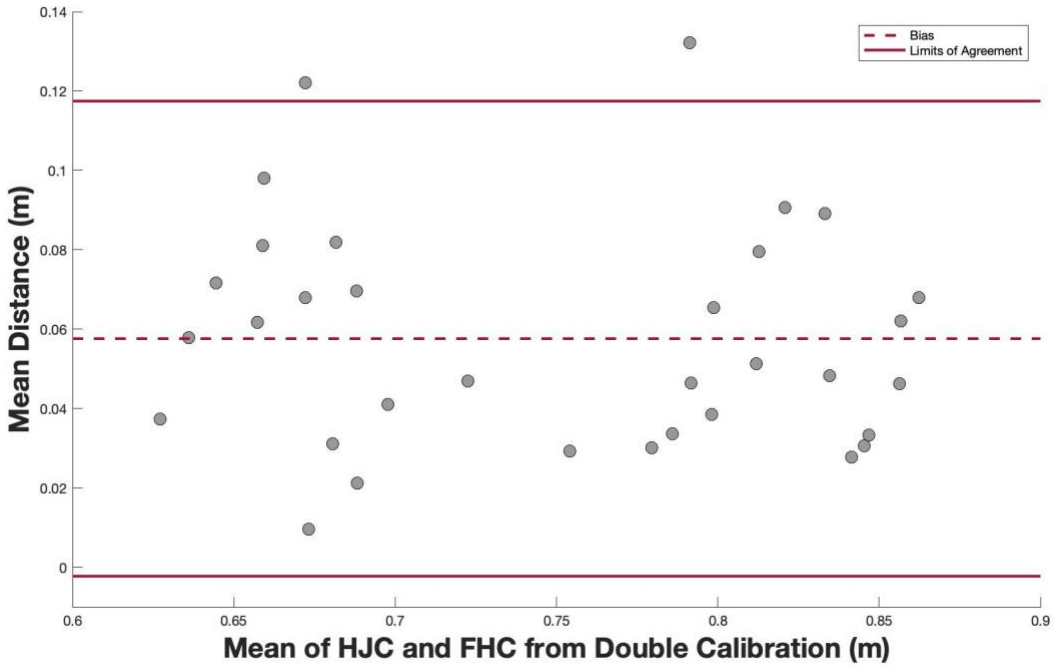


Figure 8.3: Mean of the magnitude of the x, y, and z coordinates of the hip joint center (HJC) and femoral head center (FHC) plotted against the mean distance between the HJC and FHC for each participant in the double calibration. The magnitude of the x, y and z coordinates were calculated as the vector norm for the HJC and FHC, while the mean distance was calculated as the vector norm difference between the HJC and FHC.

Appendix D: Percentile-Divided Participant Mean of Femoral Head Center and Hip Joint Center Magnitude, and the Difference in Mean Magnitude (mean distance) For Each Calibration Condition.

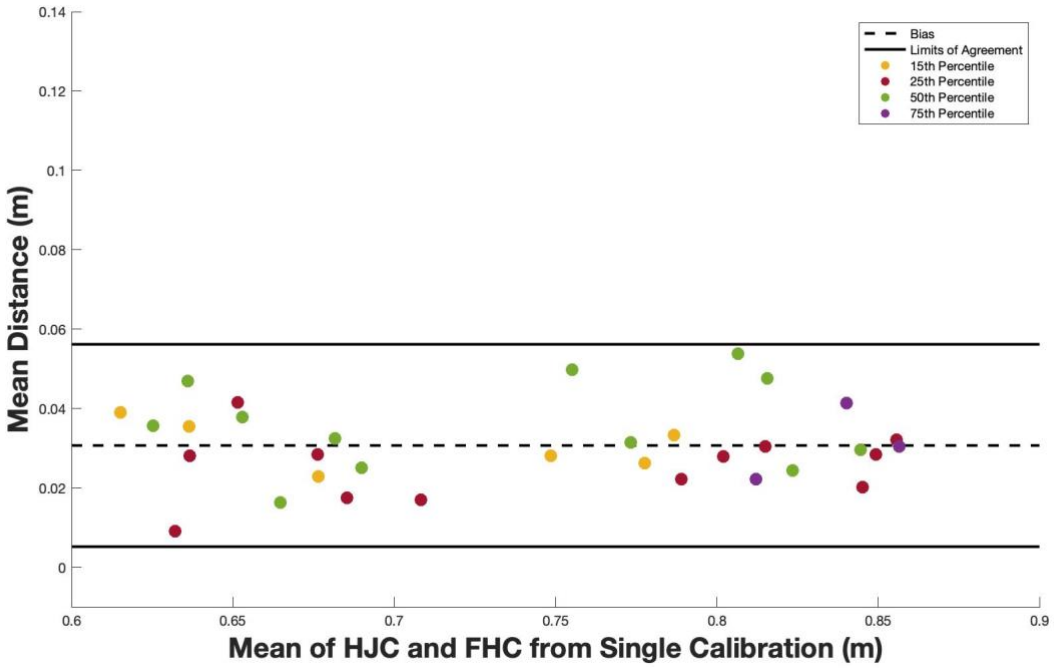


Figure 8.4: Mean of the magnitude of the x, y, and z coordinates of the hip joint center (HJC) and femoral head center (FHC) plotted against the mean distance between the HJC and FHC for each participant in the single calibration classified into 15th (yellow), 25th (red), 50th (green) and 75th (purple) percentiles. The magnitude of the x, y and z coordinates were calculated as the vector norm for the HJC and FHC, while the mean distance was calculated as the vector norm difference between the HJC and FHC.

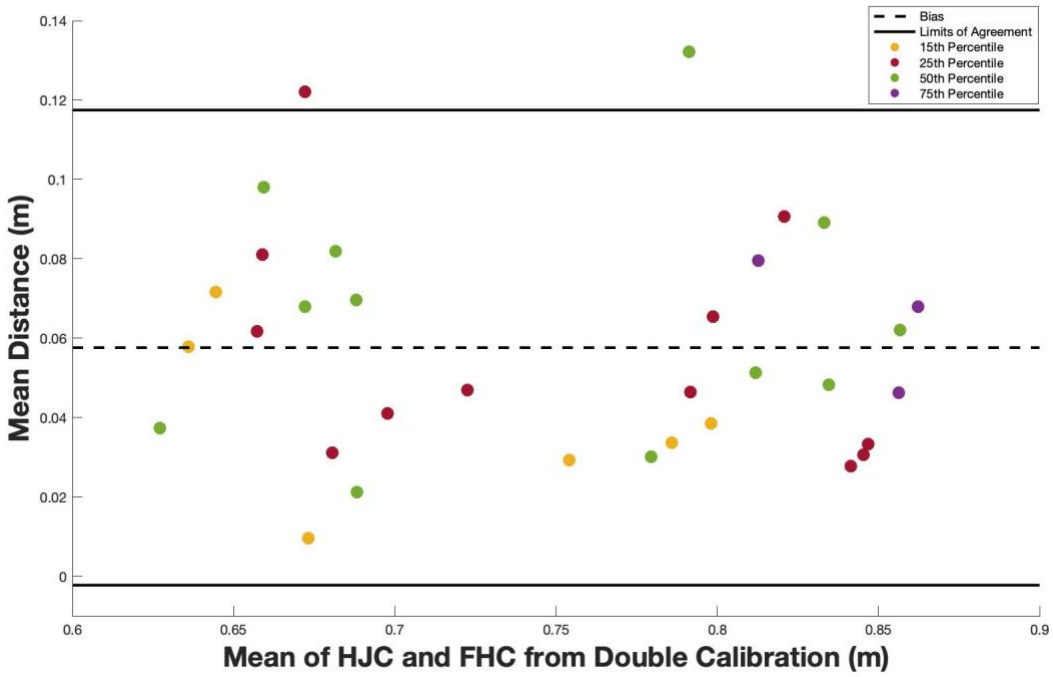


Figure 8.5: Mean of the magnitude of the x, y, and z coordinates of the hip joint center (HJC) and femoral head center (FHC) plotted against the mean distance between the HJC and FHC for each participant in the double calibration classified into 15th (yellow), 25th (red), 50th (green) and 75th (purple) percentiles. The magnitude of the x, y and z coordinates were calculated as the vector norm for the HJC and FHC, while the mean distance was calculated as the vector norm difference between the HJC and FHC.

Appendix E: Inter-Landmark Differences

Table 8.4: Absolute mean difference of inter-landmark distance between standing and seated calibration postures for each individual participant. The difference is presented as mean (+/- standard deviation) in centimeters (cm).

Participant	Absolute Difference
1	3.8
2	2.0
3	0.6
4	3.5
5	4.1
6	2.7
7	0.9
8	3.7
9	3.6
10	3.2
11	1.4
12	2.8
13	1.9
14	1.7
15	0.6
16	1.5
17	1.3
18	3.6
19	3.3
20	2.8
21	0.3
22	0.3
23	1.5
24	0.6
25	3.6
26	1.2
27	0.3
28	0.3
29	0.9
30	0.9
31	0.7
32	2.4
33	1.0