

Femoral bone quantitative CT analyses with
internal tissue-based phantomless densitometric CT calibration

by
Carla Winsor

A dissertation submitted in partial fulfilment of
the requirements for the degree of

Doctor of Philosophy
(Mechanical Engineering)

at the
UNIVERSITY OF WISCONSIN–MADISON

2022

Date of final oral examination:

This dissertation is approved by the following members of the Final Oral Committee:
Heidi-Lynn Ploeg, Professor, Mechanical and Materials Engineering
Kate Fu, Associate Professor, Mechanical Engineering
Elizabeth Meyerand, Professor, Biomedical Engineering and Medical Physics
Josh Roth, Assistant Professor, Mechanical Engineering
Shiva Rudraraju, Assistant Professor, Mechanical Engineering
Sylvana Garcia-Rodriguez, Research Associate, Radiology

© Copyright by Carla Winsor 2022

All rights reserved

Abstract

This dissertation develops a public-domain approach for evaluating and selecting patient tissues to use for deriving densitometric computed tomography calibration (DCTC). This method enables the evaluation of patient computed tomography (CT) scans captured without a densitometric calibration phantom in the scan field of view. Unlike other methods for estimating density from CT scans, this method can be applied in the context of CT-based patient-specific finite element (CTPSFE) models and analyses. CTPSFE analyses have been shown useful in a variety of applications including identifying patients at risk of imminent femoral fragility fracture. This dissertation aims to demonstrate, verify, and validate an approach to selecting patient tissues to use as the basis for deriving a phantomless DCTC equation. My analysis shows the demonstrated phantomless method was comparable with current clinical and orthopaedic research gold standard phantom-based calibration methods. The developed method shows promise as a public domain DCTC method capable of enabling further development of CTPSFE methods and broadening the clinical accessibility to quantitative CT analyses.

List of Publications

Journal Articles

C.L. Brunnquell, **C. Winsor**, H.L. Aaltonen, S. Telfer, 2021. "Sources of error in bone mineral density estimates from quantitative CT." EJR. 1-12. doi.org/10.1016/j.ejrad.2021.110001

C. Winsor, X. Li, M. Qasim, J. Zhang, C.R. Henak, P.J. Pickhardt, H. Ploeg, M. Viceconti, 2021. "Evaluation of patient tissue selection methods for deriving equivalent density calibration for femoral bone quantitative CT analyses." Bone. 1-12. doi.org/10.1016/j.bone.2020.115759

Conference Papers

D.C. Appel, **C. Winsor**, R.C. Tillinghast, M. Mansouri, 2022. "Understanding Obstacles in the STEM Career Pipeline through System Dynamics Modeling." 2022 11th IEEE Integrated STEM Education Conference, ISEC 2022.

D.C. Appel, R.C. Tillinghast, **C. Winsor**, M. Mansouri, 2020. "STEM Outreach: A Stakeholder Analysis." 2020 9th IEEE Integrated STEM Education Conference, ISEC 2020. Doi: 10.1109/ISEC49744.2020.9280723

R.C. Tillinghast, D.C. Appel, **C. Winsor**, M. Mansouri, 2020. "STEM Outreach: A Literature Review and Definition." 2020 9th IEEE Integrated STEM Education Conference, ISEC 2020. Doi: 10.1109/ISEC49744.2020.9280745

List of Presentations and Conference Abstracts

C. Winsor, A.N. Pernsteiner, P.J. Pickhardt, C.R. Henak, H. Ploeg, “Comparison of patient femoral quantitative computed tomography measures calibrated at different distances to isocenter,” Oral Presentation, Computer Methods in Biomechanics and Biomedical Engineering Symposium 2021, Virtual.

A.N. Pernsteiner, **C. Winsor**, P.J. Pickhardt, C.R. Henak, H. Ploeg, “Impact of Automatic versus Manual Calibration Reference Segmentation on CT-based Finite Element Analyses,” Oral Presentation, Computer Methods in Biomechanics and Biomedical Engineering Symposium 2021, Virtual.

A.N. Pernsteiner, **C. Winsor**, J. Ploeg, C.R. Henak, “Machine Learning Segmentation for Densitometric Phantom CT Data,” Poster Presentation, Summer Biomechanics, Bioengineering, and Biotransport Conference 2021, Virtual.

C. Winsor, H. Ploeg, P.J. Pickhardt, C.R. Henak, “Evaluation of distance and background medium effects on CT-density to CT Number calibration and Young’s Modulus,” Oral Presentation, Virtual Physiological Human Conference 2020, Paris, France.

A.N. Pernsteiner, E.G. Cameron, **C. Winsor**, C.R. Henak, and H. Ploeg, “Evaluation of the Effects of CT Scanner Model and Reconstruction Kernel on the CT Number Calibration,” Poster Presentation, Orthopedic Research Society 2020 Annual Meeting, Phoenix, Arizona, USA.

E.G. Cameron, A.N. Pernsteiner, **C. Winsor**, H. Ploeg, “Effects of CT X-ray Voltage and Current on Hydroxyapatite Phantom Density Plugs in Scans Reconstructed with Standard and BonePlus Kernels,” Poster Presentation, Orthopedic Research Society 2020 Annual Meeting, Phoenix, Arizona, USA.

C. Winsor, X. Li, J. Zhang, C.R. Henak, H. Ploeg, “Investigating sex-specific accuracy of proximal femur coordinate systems derived from statistical shape models,” **MS Paper Competition finalist**, Summer Biomechanics, Bioengineering, and Biotransport Conference 2019, Seven Springs, Pennsylvania, USA.

C. Winsor, L. Campello, M. Qasim, J. Zhang, X. Li, C.R. Henak, P.J. Pickhardt, H. Ploeg, M. Viceconti, “A phantom-less calibration technique for use in femur strength prediction,” Oral Presentation, 22nd International Workshop on Quantitative MSK Imaging 2019, Chateau Lake Louise, Canada.

C. Winsor, “A phantom-less calibration technique for use in femur strength prediction,” Mechanics Seminar, Madison, Wisconsin, 2018.

C. Winsor, L. Campello, M. Qasim, J. Zhang, X. Li, C.R. Henak, P.J. Pickhardt, H. Ploeg, M. Viceconti, “A phantom-less calibration technique for use in femur strength prediction,” Oral Presentation, Virtual Physiological Human Conference 2018, Zaragoza, Spain.

C. Winsor, M. Qasim, J. Zhang, X. Li, T. Szczykutowicz, C.R. Henak, P.J. Pickhardt, H. Ploeg, M. Viceconti, "Sensitivity analysis of density calibration phantom on finite element-predicted femoral strength," Poster Presentation, World Congress of Biomechanics 2018, Dublin, Ireland.

C. Winsor, "A phantom-less calibration technique for use in femur strength prediction," Invited talk, Whitaker International Bioengineering Symposium, Galway, Ireland, 2018.

C. Winsor, "Hip Fracture Risk as a Mechanics Problem," Integrated Musculoskeletal Biomechanics Group Seminar, Sheffield, England, 2018

C. Winsor, L. Campello, M. Qasim, J. Zhang, X. Li, C.R. Henak, P.J. Pickhardt, H. Ploeg, M. Viceconti, "A phantom-less calibration technique for use in femur strength prediction," Poster Presentation, INSIGNEO Showcase 2018, Sheffield, England.

C. Winsor, "Developing Clinical Diagnostics: Fracture Prediction in Patients with Osteopenia via Computational Methods," Whitaker Spring Enrichment Seminar, Budapest, Hungary, 2018.

Acknowledgements

I am forever grateful to Heidi-Lynn Ploeg for providing me with a second chance at conducting graduate research, supporting me in a successful application to the Whitaker International Program, and providing me with research support throughout my graduate studies. I am grateful to Corinne Henak for providing a research home after Heidi moved to Queen's University in Kingston, Ontario. I am honoured that Kate Fu made time to help me wrap up my PhD and created an environment where I could rediscover the fun in research.

There are no words to explain the amazing support and opportunity provided by Xinshan Li and Marco Viceconti. In addition to supporting me and providing feedback as I wrote my Whitaker International Program research project proposal, they provided me with foundational research training, an understanding of professional research courtesies, and an opportunity to experience first-hand the impact of a research vision carried out. Marco's ability to create a research vision is awe-inspiring and helped me to grow with respect to envisioning projects.

The year I spent at the University of Sheffield is the source of my happiest memories from graduate school. The University of Sheffield has provided access to a network of researchers that continually inspire, impress, and encourage me; and access to an alumni network of professionals I enjoy interacting with. I will always consider the faculty of the University of Sheffield and the researchers of Insigneo Institute for *in silico* Medicine as part of my research family.



Figure 0-1: The Integrated Musculo Skeletal Biomechanics (IMSB) group at the INSIGNEO Institute for *in silico* Medicine based at the Department of Mechanical Engineering of the University of Sheffield. This group picture was captured upon completion of our spring of 2018 team-building exercises.

I remain grateful for the interactions I had with my PhD committee members. Perry Pickhardt provided the clinical data that made this PhD possible, was been a consistent source of supportive discussion and clinical access to capture calibration data. His research insights have changed the way I see research and the way I think about setting up experiments. I have always found Shiva Rudraraju to be enthusiastic, knowledgeable, and encouraging. In under 90 minutes, he answered five years of built-up technical and mentoring questions in one of my first interactions with him. Josh Roth has been a wealth of information both on and off Twitter. We have had good conversations over cheese curds, and I appreciate his efforts towards the cadaveric experiments infeasible due to the pandemic. In addition to being a co-author on one of the papers my work builds upon, Sylvana Garcia-Rodriguez has been a role model through her role as an alumnus of

Heidi's lab, my interactions with her at the Summer Biomechanics, Bioengineering, and Biotransport Conference in 2019, and her work as a research associate.

I have also benefitted from excellent research conversations and collaborations with a variety of students, and postdoctoral researchers. Qasim Muhammad opened my eyes to CT-based patient-specific finite element analyses. I'm grateful to Pinaki Bhattacharya for mentoring, technical discussions, and making research fun. Zainab Altai demonstrated how awesome researchers can be and shared ANSYS macros. Soroush Irandoust has made the 2021-2022 academic year unforgettable and brought many good CT calibration and finite element analysis questions. One of my joys has been training, mentoring, and supporting undergraduate researchers. I thoroughly enjoyed early phantomless calibration discussions with Leticia Campello. Bara Kubanova's investment of time in segmenting CT images for the 80 patients in our fracture patient cohort provides a strong geometric foundation for my research. Conducting research with Ashley Pernsetiner, in the 2019-2022 academic years, have been some awesome memories in graduate school — from the late nights of trying to learn to write a research proposal in a weekend, to learning to make shakshuka, to road-tripping to Queen's University, to challenging me to provide guidance outside of my research experience.



Figure 0-2: Carla Winsor with Ashley Pernsteiner, Heidi Ploeg, Baixuan Yang, Lawrence Torkan, and Mahsa Zojaji during a January 2020 visit to Queen's University and the Centre for Health Innovation (formerly the Human Mobility Research Centre).

I have been incredibly blessed by two external collaborations. I am grateful that Professor Timothy Szczukutowicz connected me with Christina Brunquell. Christina and I have enjoyed collaborating on her literature review, and good technical and career discussions. I am grateful for her influence and enrichment of my graduate research experience. I am also grateful to Bryan Kirking and Nathan Pyle from DJO Global. It is not often that a PhD student gets to work with industrial researchers to reverse-engineer the methods of another PhD student. The technical discussions we have engaged in have mutually broadened our understanding of densitometric CT calibration in ways that will

enrich our work for years to come. Nathan and I are grateful to Andrew Michalski for including underlying densitometric relationships in his appendix that made his work more accessible than the published research manuscripts.

I am grateful to Daniel and Joan Appel. A friendly debate on whether or not writing could ever be "done" has turned into a series of writing collaborations that has enriched all three of our careers. Daniel has taught me a new precision in engineering language that will continue to enrich my writing. Joan has been a purveyor of medical insights. I cannot thank them enough for their help with my Whitaker International Fellowship application. Further, both have been instrumental in helping me find peace in the challenges that women engineers face and in providing me with the opportunity to improve the career selection process of future STEM professionals through our collaborations in Daniel's research. I have also enjoyed watching their family grow and interacting with their three daughters—who I hope will one day become STEAM-H



Figure 0-3: Abby learning about rigid-body dynamics by driving a hot wheels school bus on Carla's face under Joan Appel's supervision.

professionals. Their daughters may be the first people to provide me with a functional nickname, I gained the nickname of the "Pizza Princess" by showing up with pizza. Abby has brought us many laughs by learning about hot wheels by driving a school bus across my forehead, which Joan kindly captured and Daniel has immensely enjoyed. Joining us for many of our collaboration adventures, including my Whitaker International Program

Fellowship Application, Annette Dean taught me how to write a successful research proposal and has brought much clarity and guidance to my graduate school path. Annette Dean has taught me much about mental health and the importance of positivity through moving and assembling a bed to understanding academia. I'm grateful to Tucker Meyers and Anne Broadbridge for assistance with editing this dissertation.

My introduction to orthopaedic biomechanics and some of my best memories at the Illinois Institute of Technology came from my Inter-Professional Projects class.



Figure 0-4: Illinois Institute of Technology's Fall 2012 Inter-Professional Projects class led by Professor Kevin Meade (right) and industry mentor Michael Morley (left).

In that class, Kevin Meade used his rehabilitation journey following a haemorrhagic stroke to simultaneously teach us about orthopaedic biomechanics, patient-centred design, and the power of conquering disability as an inspirational and educational experience. As a Whitaker International Program Alumni and industry mentor, Michael Morley provided my first exposure to biomedical engineering, research, the Whitaker International Program, and how biomedical engineering can improve life for people in developing nations. Daniya

Kamran-Morley taught me how to inspire people with a vision, how to find and chase my passion for improving people's quality of life, and how to rise to the challenge in time for a deadline. I'm grateful Daniya, Michael, and Mauricio Palacio were willing to get up early on Friday mornings for our 3:00 am work sessions. Mauricio's enthusiasm, leadership, and willingness to ask for support opened my eyes to new design challenges, showed me how engineering design can support the digital inclusion of people with partial finger amputations, and helped me envision what I wanted to learn in graduate school.

Sue Collins has been one of the most influential mentors in my journey to becoming a mechanical engineer. She was the first female mechanical engineer I met, and we readily found ways to support each other. As a subject matter expert on groundwater and a mechanical engineer, Sue taught me the power of communicating concepts through visual representations as my Girl Scout Gold Award mentor. As a military wife, Sue taught me to see needs and support others by meeting their needs. She taught me intergenerational mentoring as I learned to mentor her daughters. She modeled generosity and grace under pressure. In supporting her, I discovered my own independence and strength. On visits home from the United States Coast Guard Academy (USCGA), I found myself drawn to Sue's house every time my plans were canceled. She and her husband stood up for me and mentored me throughout my time at the USCGA. More importantly, Sue was there to pick me up after the military with the Congressional Leadership Award Program. She encouraged me to give Mechanical Engineering a second try and helped me with my resume when applying for undergraduate internships at Sandia National Laboratories.

I am honored to also have received support from Sandia National Laboratories through both undergraduate and graduate student internships. I am grateful to Mar McCormack for demonstrating how proactive and supportive managers can be. Kate Rivera has shown me how effective, empowering, and intellectually stimulating a good manager can be. I am grateful to Laura Salguero for seeing the best in me, encouraging me to follow my dreams, and introducing me to model-based systems engineering. Sue Collins and Elsa Galloway have provided a good work environment and helped me learn how to create a work culture that feels safe. My colleagues at Sandia National Laboratories have helped me feel connected during the remote working days of the pandemic, and taught me that engineers can have fun at work. While interning at Sandia National Laboratories has provided good on-the-job training and opportunities for professional development, the most fulfilling aspects for me have been the opportunity to witness Sue's vision—serving in one area of the lab to gain professional development and bringing that knowledge back as a vision for how to grow her environmental program. I am honored that I could give back to Sue by laying the groundwork for preserving her legacy and impact at Sandia National Laboratories. I am grateful my schedule worked out such that I could walk Tallie during my lunch breaks (Figure 0-5).



Figure 0-5: I am grateful to Tallie for facilitating mental health breaks, inspiring playfulness, and encouraging physical fitness. Additional thanks go to Margaret Jenks and Randy Swiggum for thinking of me when Tallie needed walks during the day.

Dedication

This work is dedicated to:

The leaders who inspire us to speak up, initiate and facilitate change;
the survivors of domestic, academic, and workplace violence;
the allies who understand, support, and stand up for others;
And the advocates who create safe spaces, encourage rebuilding, and
promote access to healing.

Table of Contents

Abstract	i
List of Publications	ii
Journal Articles	ii
Conference Papers	ii
List of Presentations and Conference Abstracts	iii
Acknowledgements	v
List of Abbreviations and Symbols	xvii
List of Tables	xix
List of Figures	xx
1 Executive summary and motivation	1
2 Dissertation navigation	5
2.1 Objectives	6
2.2 Summary of original contributions	7
3 Clinical X-ray imaging: Definitions, assumptions, and limitations	8
3.1 Factors of X-ray based clinical imaging	8
3.1.1 X-ray-based method for the measurement of <i>in vivo</i> bone mineral	10
3.1.2 Computed Tomography	13
3.1.3 Acquiring the measurements and creating a CT image	16
3.1.4 Image reconstruction methods	17
3.2 Densitometric calibration of X-ray CT systems and measurement of material densities from CT images	18
3.2.1 Existing densitometric CT calibration standardized terms	19
3.2.2 Existing databases of X-ray mass attenuation coefficient tables	19
3.2.3 Determining theoretical effective energy	20
3.2.4 Phantom-based density approach to DCTC for equivalent BMD	21
3.2.5 Example phantom-based density approach to DCTC	22
3.2.6 Phantomless density approach to DCTC for equivalent BMD	24
3.2.7 Attenuation approach to DCTC method for equivalent BMD	25
3.2.8 Comparison of the density and attenuation DCTC methods	26
3.2.9 Mass-fraction-model approach to DCTC for integral BMD	27
3.2.10 Comparison of DCTC methods	30

3.3	Phantomless calibration facilitates QCT analyses	35
4	CTPSFE analysis	37
4.1	Geometry	39
4.1.1	Anatomical planes	39
4.1.2	Image segmentation	40
4.1.3	Mesh generation	42
4.2	Material properties	43
4.2.1	Deriving phantom-based effective BMD	43
4.2.2	Empirically derived density-elasticity relationships	46
4.2.3	Assignment of apparent elastic moduli onto the FE mesh	51
4.3	Boundary and loading conditions	52
4.3.1	Stance phase of walking loading	52
4.3.2	Sideways fall loading	53
4.3.3	Experimental validation of sideways fall loading	54
4.3.4	Specimen-specific numerical simulation of a sideways fall	55
4.3.5	Patient-specific numerical simulations of sideways fall loads	57
4.4	Post-processing	60
4.4.1	Estimating the femoral fracture load and minimum fall strength	60
4.5	CTPSFE limitations	62
4.6	CTPSFE shows promise for diagnosing patient femoral fracture risk	64
5	Evaluation of patient tissue selection methods for deriving equivalent density calibration for femoral bone quantitative CT analyses	65
5.1	Introduction	65
5.2	Materials and methods	71
5.2.1	Method of selecting patient scans for analysis	72
5.2.2	CT scanning protocol	74
5.2.3	Inline quantitative equivalent density calibration using the Mindways Model 3 BMD calibration phantom	74
5.2.4	Offline equivalent density calibration using a custom BMD phantom	75
5.2.5	Identify most consistent reference densities across patients	76
5.2.6	Experimental design to test repeatability of tissue identification	78
5.2.7	Finite element model BMD and Femur Strength	79
5.2.8	Statistical analysis	80
5.3	Results	80
5.4	Discussion	89

5.5	Acknowledgements	95
5.6	Conflict of interest	96
6	Comparison of internal tissue-based phantomless densitometric CT calibrations and error assessment	97
7	Discussion	103
8	Scientific knowledge generated	111
9	Future work	113
	References	116
	Appendix – Clinical handout for phantomless calibration	138
	Appendix – Matlab function for calibration equation derivation	140

List of Abbreviations and Symbols

CT: computed tomography BCT: biomechanical computed tomography

BMD: bone mineral density

aBMD: areal bone mineral density

DXA: dual-energy X-ray absorptiometry

CTXA: computed tomography X-ray absorptiometry

DCTC: densitometric CT calibration

QCT: quantitative computed tomography

HU: Hounsfield Units

FEA: finite element analysis

CTPSFE: computed tomography-based patient-specific finite element analysis

ACM: attenuation calibration method

DCM: density calibration method

MFM: mass-fraction-model calibration method

FDP: femoral density phantom

$Ca_{10}(PO_4)_6(OH)_2$: calcium hydroxyapatite

HA: calcium hydroxyapatite

K_2HPO_4 : dipotassium phosphate

ρ_{CT} , equivalent, CT or radiological density;

ρ_{QCT} , quantitative equivalent density;

ρ_{app} , apparent density, wet;

ρ_{ash} , ash density;

ρ_{eff} , effective density;

ρ_{nom} , nominal density;

ρ_{HA} , effective calcium hydroxyapatite density;

$\rho_{K_2HPO_4}$, effective dipotassium phosphate density;

GCF, general calibration factor;

FS, femoral strength;

MFS, minimum fall strength;

SSE, sum of squared error;

RMSE, root mean square error

ASTM: American Society for the Testing of Materials

ICRU: International Commission on Radiation Units

NIST: National Institute of Standards and Technology

List of Tables

Table 2-1: Summary of original contributions throughout this dissertation.....	7
Table 3-1: Example segmentation measurements from a CT examination of the femoral density phantom by densitometric standard.	23
Table 3-2: Example CT density measurement comparison of the density and attenuation methods for polyamide and polycarbonate from ASTM E1935-97. Observe that the results for the attenuation method could be rounded to the same values resulting from the density method, indicating that the extra work of the attenuation method may not be worth the increase in accuracy. Additionally, the attenuation method is precise to within 1%, and the density method is precise to within 5% relative to the published density. Note that (1) polyamide and polycarbonate have densities that fall within the range of human cortical femoral bone which makes this example relevant to this dissertation, and (2) this example was derived on an industrial CT scanner, which may differ from clinical CT scanners, particularly in terms of resolution.....	26
Table 3-3: Representative region-of-interest tissue segmentations from a CT scan of a 79-year-old patient at the proximal femur level.	32
Table 4-1: Example segmentation measurements for each densitometric standard in a CT examination of the CIRS Model 004 CT Simulator for Bone Mineral Analyses, a vertebral density phantom.	44
Table 4-2: Representative segmentation mean CT Number [HU] measurements for a 71-year-old patient with accompanying phantom-specific water and K ₂ HPO ₄ density tabulated values. The far-right column presents the CT Number [HU] minus water as a prerequisite for plotting.....	45
Table 5-1: The analysed cohort considered 211 distinct patients (bottom), with up to 258 scans (top). These scans were captured on nine different GE CT scanner models at the University of Wisconsin—Madison hospital.....	73
Table 5-2: BMD [mg/cm ³] from equivalent density [mg/cm ³] equations for the five patients scanned with the inline K ₂ HPO ₄ Mindways Model 3 BMD calibration phantom.....	75
Table 5-3: Nominal density values from NIST ²⁸	77
Table 5-4: Intra- and Inter-operator reanalysis precision error (root-mean-square) for FS, BMD, calibration equation, and tissue segmentations at the femur for n = 5. Coefficients of variation (CV _{RMS} , in %) and standard deviations (SD _{RMS} , in absolute units) are presented.	89
Table 5-5: Comparison of precision errors between the current study and literature.....	93
Table 9-1: Breakdown of prefracture and control patients segmented from the Madison fracture cohort for later CTPSFE analyses.....	115

List of Figures

Figure 3-1: Representative transverse or axial CT slice showing a 71-year-old patient at the proximal femur level..	14
Figure 3-2: Comparison of measured CT density for densitometric standards scanned with (Blue) and without (Red) 50% overlap between CT slices..	15
Figure 3-3: Example of a calibration curve for a calcium hydroxyapatite phantom.....	24
Figure 3-4: Representative slice effective energy calculations for tissue segmentations from the CT-scan of a 79-year-old patient..	33
Figure 3-5: Representative density [mg/cm^3] measurements for bracketing trabecular and cortical femoral bone standards for multiple calibration methods: femoral density phantom (FDP), phantomless density (ASTM DCM), phantomless attenuation (ASTM ACM–80, 97), and phantomless mass fraction model (MFM–80, 97)..	34
Figure 4-1: A visualization of computed tomography-based patient-specific finite element modelling pipeline steps..	38
Figure 4-2: Representative example of the anatomical planes applied to a human.	39
Figure 4-3: Representative axial CT slice and segmented proximal femur for a 71-year-old patient.....	40
Figure 4-4: Examples of representative tetrahedral elements, tet-4 on the left, tet-10 in the middle and a patient femur with an applied tetrahedral mesh on the right.	42
Figure 4-5: Example of a calibration curve linear regression for a calcium hydroxyapatite densitometric phantom	44
Figure 4-6: Example plots demonstrating why water CT Number [HU] measurements must be subtracted in order to regain linearity to find the linear calibration curve for the inline Mindways Model 3 BMD phantom.	46
Figure 4-7: Representative stance (left) and sideways fall (right) loading conditions for a 71-year-old patient.	53
Figure 4-8: Representative images demonstrating the method for estimating the full femur using the statistical shape modelling software, MapClient.	59
Figure 4-9: Representative image demonstrating the process of deriving a coordinate system using anatomical mapping.	60
Figure 5-1: Representative axial slice of a CT scan of an 85-year-old patient at the proximal femur level.	77
Figure 5-2: FS results derived using phantomless calibration displayed the least bias when compared against results derived using the average of the patient and scan-specific K_2HPO_4 calibration as shown by Bland-Altman analyses.....	85

Figure 5-3: BMD results derived using phantomless calibration displayed the least bias when compared against results derived using the average of the patient and scan specific K_2HPO_4 calibration as shown by Bland-Altman analyses.....	86
Figure 5-4: Phantomless calibration slopes derived from air, aortic blood and skeletal muscle segmentations displayed less bias than those derived from air and adipose when compared with patient and scan specific K_2HPO_4 calibration as shown by Bland-Altman analyses.....	87
Figure 5-5: These plots compare the ten best combinations of tissues in terms of calibration slopes and intercepts.	88
Figure 6-1: Slice effective energy calculations for tissue segmentations from the CT-scan of a 79-year-old patient.	98
Figure 6-2: Representative density [mg/cm^3] measurements for bracketing trabecular and cortical femoral bone standards for multiple calibration references: femoral density phantom (FDP); density and attenuation calibration methods for internal tissue-based phantomless calibration derived from air, adipose, aortic blood, and skeletal muscle; density and attenuation calibration methods for internal tissue-based phantomless calibration derived from air, aortic blood, and skeletal muscle; and density and attenuation calibration methods for internal tissue-based phantomless calibration derived from air, and adipose.....	100
Figure 6-3: Percent difference calculations between the femoral density phantom and phantomless density measurements for bracketing trabecular and cortical femoral densities.	101
Figure 6-4: Representative densitometric CT measurements for trabecular and cortical human femoral bone bracketing densities.....	102
Figure 7-1: Percent difference calculations for phantomless density measurements with femoral density phantom (FDP) based density measurements as the gold standard..	109
Figure 9-1: Representative CT slice of internal tissue segmentations for a 93-year-old patient with a surgical implant	114

1 Executive summary and motivation

Osteoporosis, a metabolic disease, is a known leading cause of femoral fragility fracture and has been widely studied. Femoral fragility fracture widely impacts the quality of life of people from a diverse range of backgrounds at-risk for reasons that may include aging, metabolic disease, traumatic injury, cancer, or space travel. Bone researchers with engineering backgrounds have proposed computed tomography-based patient-specific finite element (CTPSFE) modelling and analyses to identify patients at risk of femoral fracture. The first clinically implemented biomechanical computed tomography (BCT) product VirtuOst software (O.N. Diagnostics, Berkeley, CA) is newly available in the clinic in the USA as of 2020 after approval from the FDA. Despite this, universally identifying patients at risk of femoral fragility fracture and preventing fractures remain clinical challenges.

Mechanically, femoral fragility fracture is known to be a function of macro- and micro-architecture, density, loading conditions and the interactions of the bone with the environment during impact. Input variables for CTPSFE analyses include geometry, material properties, and loading and boundary conditions. Geometry and loading and boundary conditions can be defined using free-body diagrams and available instrumentation. Through medical image segmentation, digital patient-specific geometries are created from computed tomography (CT) scans. Loading and boundary conditions are commonly defined from either the stance phase of walking or a side fall scenario. For loading that represents the stance phase of walking, joint forces from instrumented hip implants provide the best currently available data on loading angles and forces. The

stress gradient in the combined loading that exists in either the walking or falling loading cases makes CTPSFE analyses sensitive to material properties. Defining bone material properties for CTPSFE analyses remains challenging.

Bone material properties are the most challenging aspect of CTPSFE analyses. Bone mineral density, a bone material property, can be measured using CT scanners. CT scanners measure X-ray attenuation and density after densitometric CT calibration (DCTC) is performed. With calibration and empirical material mapping equations, bone mineral density can define material properties for CTPSFE analyses. The empirical material mapping equations amplify the impacts of different methods of DCTC, due to a power-law relationship. Therefore, the CTPSFE analyses are sensitive to the DCTC.

Capturing bone specific DCTC data is not a universal clinical practice, and retrospectively may not be feasible due to clinics routinely upgrading CT scanners. VirtuOst overcame this limitation by developing and applying internal tissue-based (phantomless) DCTC. Their proprietary approach leverages deep learning to segment the ischioanal fossa and derive a DCTC equation from the ischioanal fossa and theoretical air, as defined by the Hounsfield Scale. The continued development of CTPSFE methods and broad clinical accessibility of quantitative CT analyses require a public domain method for deriving a DCTC equation.

This dissertation develops a public-domain approach for evaluating and selecting patient tissues to use for deriving DCTC. This method enables the evaluation of patient CT scans captured without a densitometric calibration phantom in the scan field of view. Unlike other methods for estimating density from CT scans, this method can be applied in the context of CTPSFE models and analyses. This dissertation aims to demonstrate,

verify, and validate a method for selecting patient tissues to use as the basis for deriving a phantomless DCTC equation.

The method for selecting tissues was applied to assess 258 CT scans of 211 patients. Tissue density assumptions were based on available data tables from the National Institute of Standards and Technology (NIST). Standardized DCTC and measurement methods were applied from the standard for calibrating and measuring CT density published by the American Standards for the Testing of Materials (ASTM). The phantomless calibration equation from air, aortic blood, and skeletal muscle demonstrated the least error across patients.

My analysis shows the demonstrated phantomless method was comparable with current clinical and orthopaedic research gold standard phantom-based calibration methods. Of the 211 patients, 5 were scanned with a clinical Mindways Model 3 BMD phantom in the scan field of view. The derived phantomless calibration equations for the 258 CT scans all fell within the range of the 5 available phantom-based calibration equations. A femoral density phantom (FDP) designed in accordance with the ASTM standard served as our orthopaedic research gold standard phantom. Notably, density measurements from the phantomless DCTCs for the 258 CT scans showed low overshoot when compared against density measurements from the FDP.

Together these results showed phantomless calibration is valid to serve as a basis for defining bone mineral density in the context of CTPSFE. The included analysis verified and validated the air, aortic blood, and skeletal muscle combination as the basis for phantomless calibration. Bias across patients was minimal indicating these methods may be suitable for analysing patient CT scans without a phantom in the scan field of view.

The developed method shows promise as a public domain DCTC method capable of enabling further development of CTPSFE methods and broadening the clinical accessibility to quantitative CT (QCT) analyses.

2 Dissertation navigation

This dissertation describes a method for predicting femoral fracture risk (FFR) based on densitometric CT calibration (DCTC) and CT-based patient-specific finite element (CTPSFE) analysis. This framework is designed to be accessible to a technical lay audience, researchers who have taken a break from this area, and students pursuing this area of research. After this chapter, this dissertation includes seven chapters. [Chapter 3](#) defines terms and explains assumptions and limitations of X-ray-based clinical imaging. [Chapter 4](#) identifies and discusses definitions, assumptions, and limitations of CTPSFE modelling. [Chapter 5](#) demonstrates, verifies, and validates a method for selecting patient tissues from which to derive phantomless DCTC data for use in CTPSFE analyses of the femur. [Chapter 6](#) presents error measurement as a comparison of phantomless DCTC methods to a femoral density phantom (FDP). [Chapter 7](#) discusses research conclusions and limitations. [Chapter 8](#) outlines the scientific knowledge generated by this work. [Chapter 9](#) previews future research directions.

This dissertation is modular. Following the table of contents, readers may cherry-pick the sections they want to read to understand various terms from Chapters 3 and 4. Of note are the sections in Chapter 3 on determining the density calibration of a CT system and Chapter 4 on material mapping, as these sections present groundwork for understanding the rest of the dissertation. Further, the example in Chapter 3's section 3.2.10 includes an illustration that will be expanded in Chapter 6. This scaffolded approach builds on the concepts presented and allows readers to interact with simple applications of these concepts before encountering more complicated versions.

2.1 Objectives

Demonstrate, verify, and validate a method of estimating densitometric CT calibration (DCTC) data for quantitative CT-based analyses of the femur.

Objective 1: Identify and characterize relevant and available patient tissues with the potential to serve as the basis for estimating DCTC data.

Objective 2: Devise quality checks for tissue segmentation.

Objective 3: Determine reasonable assumptions for tissue densities.

Objective 4: Demonstrate a repeatable and objective method for tissue combination selection.

Objective 5: Devise guidance for calibration equation quality checks.

Objective 6: Verify that patient-specific phantomless DCTC equations fall within the boundaries of available representative inline phantom-based DCTC measurements.

Objective 7: Validate patient-specific phantomless DCTC equations for patients whose CT scans contain inline densitometric phantoms.

2.2 Summary of original contributions

To form a cohesive narrative, I have added some original contributions throughout my dissertation (Table 2-1). Although not yet published, these contributions will be iterated and expanded upon for future publications.

Table 2-1: Summary of original contributions throughout this dissertation.

Section Number	Section	Original Contribution
3.3.9	Illustrative comparison of densitometric CT calibration methods	Clinically relevant demonstration of whether an attenuation calibration approach is more accurate than a density calibration approach when including the femur as an internal tissue reference.
5	Evaluation of patient tissue selection methods for deriving equivalent density calibration for femoral bone quantitative CT analyses	Demonstrated a repeatable and unbiased method for selecting patient tissues to serve as the basis for internal tissue-based phantomless DCTC. Verified patient-specific internal tissue-based phantomless DCTC equations results for 258 CT scans and 211 patients against sparse inline phantom-based DCTC data as representative field boundary measurements. Validated resulting patient-specific clinically relevant outcomes derived from internal tissue-based phantomless DCTC and phantom-based DCTC.
6	Comparison of Internal Tissue-Based Phantomless DCTCs and Error Assessment	Clinically relevant demonstration of whether an attenuation calibration approach is more accurate than a density calibration approach when excluding the femur as an internal tissue reference.

3 Clinical X-ray imaging: Definitions, assumptions, and limitations

In this chapter, we begin by constructing a mental model of how X-ray-based CT scans are captured by describing relevant mathematical principles. New CT scanners are constantly being introduced to the clinic, but a mental model of the fundamentals of CT scanning equips researchers to identify potentially statistically significant factors before analysing patient data even as the details of CT technology in practice evolve. The remainder of the chapter surveys existing methods for CT density calibration and measurement. Historically, the details of CT density calibration and measurement have been kept proprietary or within specific lab groups. The recent interest in phantomless densitometric CT calibration (DCTC) has meant more details are readily available in dissertations, so a survey of differing approaches is newly possible.

3.1 Factors of X-ray based clinical imaging

X-rays have been used to evaluate bone in the clinic for over a century¹. Because their index of refraction is close to one, X-rays are neither bent nor reflected as they pass through matter. X-ray *intensity* is the product of power [kilovolt power, kVp] and current [milliAmp, mA]. X-ray *absorption* follows Beer's Law, equation $I = I_0 e^{-\alpha x}$ (3.1), with I as the final X-ray intensity [kW/m²], I_0 as the initial X-ray intensity [kW/m²], α is the absorption coefficient, and x is the thickness [m].

$$I = I_0 e^{-\alpha x} \tag{3.1}$$

In the context of X-ray-based clinical imaging, α is the product of the mass attenuation

coefficient [m^2/g] and density [g/m^3]. The energy-specific mass attenuation coefficient, $\frac{\mu}{\rho}$, is the ratio of the linear attenuation, μ [$1/\text{m}$], and density, ρ [g/m^3]. Energy-dependent linear attenuation, μ [$1/\text{m}$], after the X-ray passes through the examination object, is shown in equation (3.2). PE and CS are the basis functions of the photoelectric effect and the Compton scatter effect, respectively². Note that a_1 , a_2 , m_1 , and m_2 are material-specific constants, and μ_1 and μ_2 are the X-ray attenuations of any two independent materials.

$$\mu(E) = a_1 PE(E) + a_2 CS(E) = m_1 \mu_1(E) + m_2 \mu_2(E) \quad (3.3)$$

The definition of linear attenuation can be logarithmically transformed and rearranged to take the alternate form shown in equation (3.4). In this form, μ is the energy-specific linear attenuation [$1/\text{cm}$], $\frac{\mu}{\rho}$ is the energy-specific mass attenuation coefficient [cm^2/mg], and ρ is density [mg/cm^3].

$$\mu = \frac{\mu}{\rho} * \rho \quad (3.4)$$

One advantage of the formulation shown in equation (3.4) for linear attenuation is the ease of rearranging to solve for density, as shown in equation (3.5).

$$\rho = \frac{\mu}{\left(\frac{\mu}{\rho}\right)} \quad (3.5)$$

Together, these relationships demonstrate that X-ray-based density measurements are sensitive to energy [kVp], current [mA], and thickness [m].

In conventional radiography, X-rays travel from the X-ray source through the examination object and are recorded on film expressed as a 2D planar image. If the X-ray measurements are recorded on film, the power and current are independently controlled

variables that must be selected before taking the measurement. When capturing X-ray measurements of bone, too much power can result in a brighter image with qualitatively higher bone density measurements and some loss of clarity. Similarly, too little current can produce an overly bright image, or too much current can produce a darker and shadowed image.

3.1.1 X-ray-based method for the measurement of *in vivo* bone mineral

In the 1960s, an improved method for measuring *in vivo* bone mineral was proposed by John Cameron and James Sorenson from the University of Wisconsin–Madison Departments of Radiology and Physics³. Cameron and Sorenson demonstrated an approach for using X-rays to measure areal bone mineral density (aBMD) by using Beer's Law, reproduced here in equations (3.6)–(3.9).

Let:

I_0 = X-ray intensity of unobstructed photon beam

T_b = equivalent bone thickness of compact bone mineral at density ρ_b

T_m = thickness of soft tissue

$T = T_b + T_m$ = thickness of tissues

I_0^* = X-ray intensity after passage of the beam through a thickness of tissue

I = X-ray intensity after passage of the beam through an equal thickness of bone mineral plus tissue

μ_b = mass absorption coefficient of bone

μ_m = mass absorption coefficient of tissue

Then, let X-ray intensity after passage through a thickness of tissue be:

$$I_0^* = I_0 e^{-\mu_m \rho_m t} \quad (3.6)$$

Equation (3.7) is rewritten such that the X-ray is attenuated by two tissues, and rearranging X-ray intensity to allow for equation (3.6) to be substituted in:

$$\begin{aligned} I &= I_0 e^{-\mu_m \rho_m T_m - \mu_b \rho_b T_b} \\ &= I_0 e^{-\mu_m \rho_m (T - T_b) - \mu_b \rho_b T_b} \\ &= I_0 e^{-\mu_m \rho_m T} e^{-\mu_b \rho_b T_b + \mu_m \rho_m T_b} \end{aligned} \quad (3.7)$$

Then substituting equation (3.6) into the final form of equation (3.7) yields:

$$I = I_0^* e^{-\mu_b \rho_b T_b + \mu_m \rho_m T_b} \quad (3.8)$$

Rearranging equation (3.8) to solve for bone thickness yields:

$$\begin{aligned} \frac{I}{I_0^*} &= e^{-\mu_b \rho_b T_b + \mu_m \rho_m T_b} \\ \ln\left(\frac{I}{I_0^*}\right) &= -\mu_b \rho_b T_b + \mu_m \rho_m T_b \\ T_b &= \frac{\left[\ln\left(\frac{I}{I_0^*}\right)\right]}{\mu_b \rho_b - \mu_m \rho_m} \end{aligned} \quad (3.9)$$

The practical application is that X-ray intensity measurements can be used to determine the cross-sectional area of compact bone mineral in a volume of interest. An equally spaced grid of X-rays travel through the tissue and bone in the specimen and the reduction in X-ray measurements is recorded at the grid locations. The cross-sectional area can then be found using the series of measurements across the intervals. This proof assumes a standard composition of bone, and equal thicknesses of bone and tissue. It also assumes that all non-bone mineral substances absorb radiation in a similar manner

to muscle tissue. Their proof clearly demonstrates that X-ray-based density measurements are sensitive to object thickness. As an illustration, Cameron and Sorenson provided a comparison of aBMD measurements for a 28-year-old normal woman (74.1 cm^2) and for an osteoporotic woman (30 cm^2). This approach is limited to a 2D measurement, while CT-based techniques discussed later involve 3D measurements.

Dual-energy X-ray absorptiometry (DXA) scanners were introduced in 1987⁴, applying the theory proposed by Cameron and Sorenson, and entered clinical practice shortly thereafter. One limitation to DXA is its sensitivity to tissue thickness as just discussed. In 2012, Yu et al. demonstrated this limitation empirically by showing that simulated increases in body fat increased variations in DXA measurements⁵. Their study found that increasing layers of fat around a phantom increased the BMD measured while increasing layers of fat around patients decreased the BMD measured. The same study also used CT scans as the basis for DXA measurements and found that this approach was less variable. Building on these methods, another improved approach, now called computed tomography X-ray absorptiometry (CTXA), is available in the clinic as an off-the-shelf solution from Mindways Software, Inc⁶.

CTXA serves several important functions. First, patients who are receiving routine virtual colonoscopies by CT examination can be screened for osteoporosis based on the same CT, with no additional scans. Second, CTXA allows for continuity of care because it is comparable with DXA measurements. The comparability provides clinicians with a consistent patient assessment over time regardless of which measurement was captured. Clinicians are also provided with data at a higher standard of care without needing continuing education to interpret measurement results.

Further, CTXA works seamlessly with the frameworks already in place to gain clinical meaning from DXA, such as the Fracture Risk Assessment Tool (FRAX). CTXA is an incremental step towards a CT-based clinical densitometric assessment on the way to the clinical adoption of CTPSFE-analysis-based techniques. Notably, following a similar scaffolded approach, VirtuOst's Biomechanical Computed Tomography tool provides both a CT-based, DXA-style analysis and a CTPSFE analysis⁷. Therefore, CTXA plays a vital role as a bridge for incremental clinical technology updates and for continuity of patient care over time.

3.1.2 Computed Tomography

A clinical CT scanner directs X-rays through a patient, detects changes in energy measured as X-ray attenuation, and generates cross-sectional images of the patient for the region of interest. Each cross-sectional CT image (Figure 3-1) is called a *slice*, and the collective group of images captured in one scan is referred to as a *stack*. Clinical CT scans are represented in greyscale (CT Number) on the Hounsfield Scale [Hounsfield Units, HU], where μ is the measured X-ray attenuation, μ_{water} is the X-ray attenuation of distilled water at standard temperature and pressure, and μ_{air} is the X-ray attenuation of air:

$$CT\ Number\ [HU] = \left(\frac{\mu - \mu_{water}}{\mu_{water} - \mu_{air}} \right) * 1000. \quad (3.10)$$

As evident from this definition, CT Numbers [HU] are a relative quantitative measurement of X-ray attenuation, normalized against water. Notably, substances less dense than water will have negative CT Numbers [HU]. By convention, CT images are rendered such that denser tissues display closer to white and less-dense tissues display closer to black.

Importantly, CT Numbers [HU] are roughly linearly proportional to density, due to Compton scatter effects⁸.

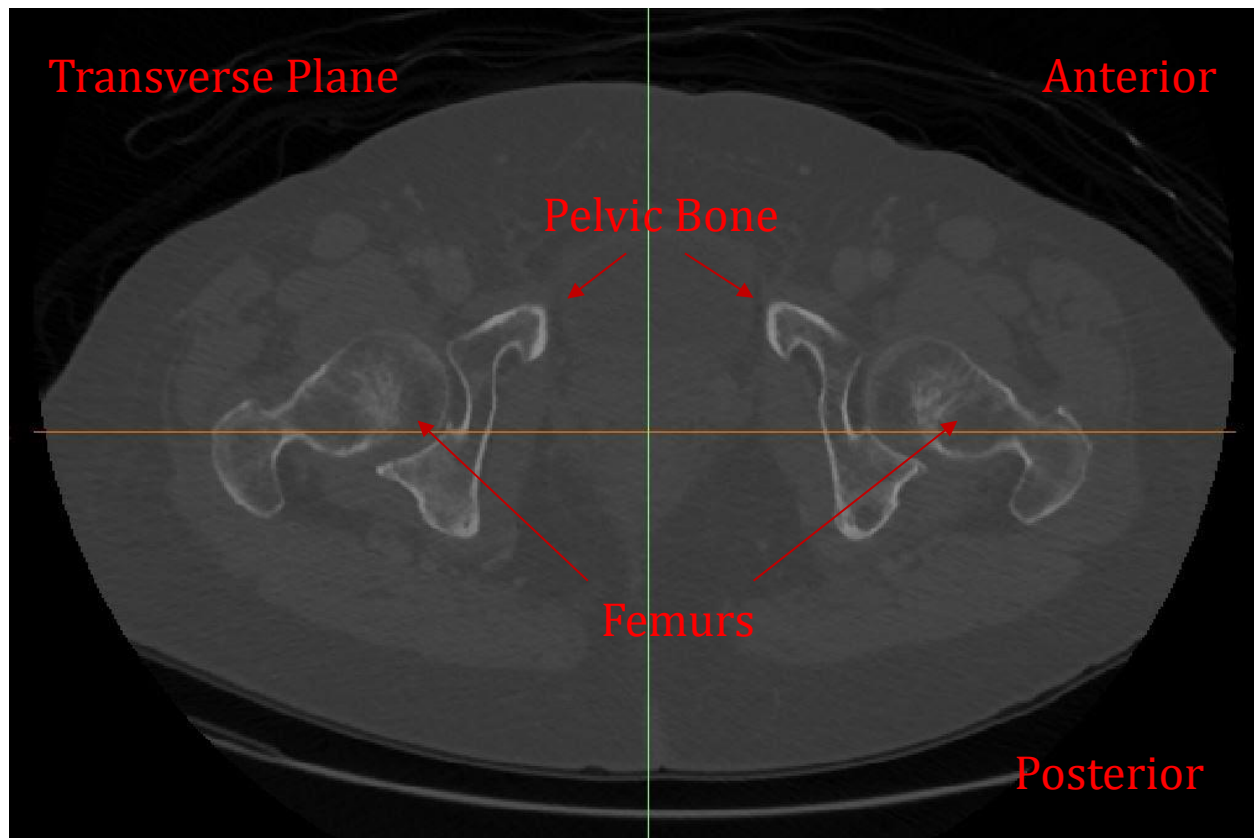


Figure 3-1: Representative transverse or axial CT slice showing a 71-year-old patient at the proximal femur level. The patient is lying supine.

Unlike conventional radiographic X-ray film-based images, CT images are digital, so image contrast and density can be adjusted at the display console after the image has been made. A CT image slice is composed of a finite collection of elements called *pixels*, each of which is assigned one of 4000 different CT Numbers [HU] based on their measured X-ray attenuation and the Hounsfield Scale. Since the X-rays only pass through the plane being imaged, each 2D pixel represents a 3D volume element called a *voxel*. The voxel's three dimensions are the length and width of the pixel and the depth of the slice. *Slice thickness*, or depth, is a variable prescribed at time of acquisition and is

constrained by the CT scanner's hardware capabilities. Example slice thicknesses include 1.25 mm, 2.5 mm, or 5 mm. *Slice increment* is the distance moved between image acquisitions. The minimum slice increment for a GE CT scanner is typically 0.3125 mm, for example.

Slice increment and slice thickness are not necessarily equal. When the slice increment and slice thickness are equal, the CT Number [HU] measurement may be around 106% of the actual density value (Figure 3-2). When the slice increment is less than the thickness so that slices overlap, it is possible to compute volumetric averages, which reduces CT Number [HU] measurement to around 107% of the actual density value (Figure 3-2). When it is desirable to limit patient exposure to radiation, larger slice increments can be used, leaving gaps between successive images.

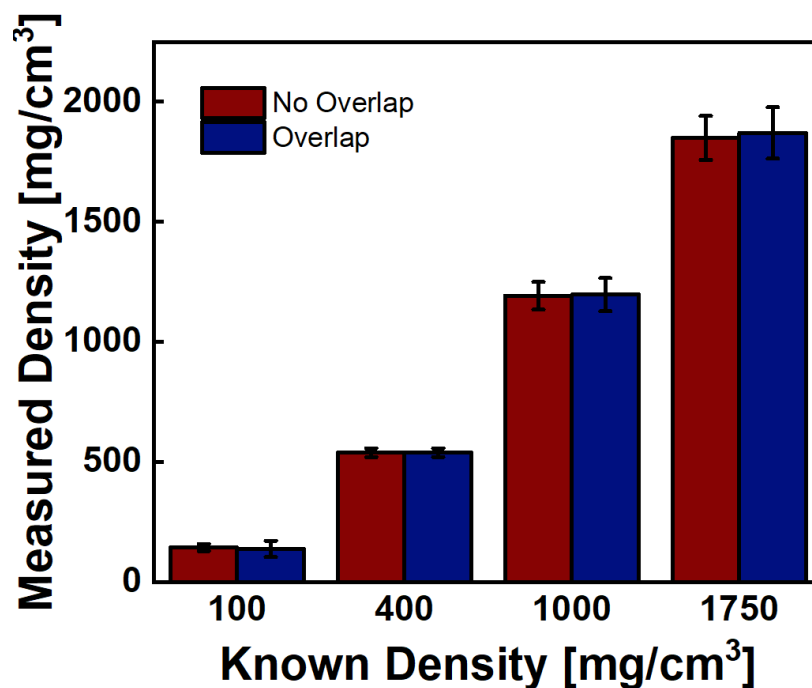


Figure 3-2: Comparison of measured CT density for densitometric standards scanned with (Blue) and without (Red) 50% overlap between CT slices. Note the overlap results in a lower CT density measurement for lower densities and a higher CT density for higher densities.

Once images have been constructed, image stacks can be *resliced* to a lower resolution by removing some slices. For CT images, reslicing cannot result in higher resolution, unlike magnetic resonance images (MRI). Depending on the size of voxels and the position of the patient or subject, some voxels may contain materials of very different densities and the resulting measured CT Number [HU] is an average. The averaged value may introduce error into the analysis and is referred to as a *partial volume artifact*. These artifacts frequently occur on the edges between bones and surrounding softer tissues.

3.1.3 Acquiring the measurements and creating a CT image

A CT measurement passes X-rays through objects. Depending on the CT scanner, different geometries may be applied, such as *parallel beam* or *fan beam*. In parallel beam geometry, the X-rays travel in parallel beams and detectors are arranged in a plane. In a fan beam geometry, the X-rays travel radially outward from a point to detectors arranged in an arc. This dissertation is limited to data collected on multidetector CT scanners (MDCT) which have multiple rows of parallel beam X-ray detectors.

A CT image is a visualization constructed from a matrix of X-ray attenuation measurements. Four terms are relevant to creating this matrix: ray, ray sum, view, and projection. A *ray* is a linear path through the examination object which may be represented as a vector. A *ray sum* is the sum of two or more rays (i.e., vector addition). In the context of CT, an X-ray attenuation measurement, or CT Number [HU], is a sum of linear attenuation coefficients along a ray through the examination object (e.g., tissues, in the case of a patient). A *view* is a complete set of rays captured for one departure of X-rays

from the X-ray source. A *projection* is a mathematical function describing the ray sums in a view.

3.1.4 Image reconstruction methods

Image reconstruction, performed by an algorithm, computes attenuation coefficients from different ray sums obtained as a projection. The algorithm, or *kernel*, manages image sharpness and noise, and can be used to create a sharper image of specific anatomical features. There are three kinds of image reconstruction algorithms used in clinical practice: (1) an iterative algorithm without statistical modelling, (2) an iterative algorithm with statistical modelling, and (3) a filtered back-projection. In an iterative algorithm, the algorithm assumes data, compares the assumption with measured data, and iterates until the two agree. When statistical modelling is included, the algorithm also considers several variables such as the X-ray source, image voxels, the detector, noise, data acquisition, and radiation attenuation. Since these variables vary by manufacturer, reconstruction algorithms tend to be proprietary.

Modern CT scanners reconstruct CT images using filtered back-projection, which was discovered by Cormack in 1963^{9–11} and is the most widely used clinical reconstruction algorithm today. In this method, each projection undergoes filtering, which includes adding extra negative numbers at the surrounding points. To achieve the required filtering effect, these negative numbers are proportional to the value of the projection and inversely proportional to the distance from the point. Once filtered, the projection values are projected back onto the reconstructed image. The resultant value at any point in the

image is the sum of the values from one point of the filtered and reversed projection. Repeating this for all projections results in a theoretically perfect reconstruction.

A filtered back-projection may also be called the convolution method. In modern CT scanners, the kernel may be labelled as reconstruction or convolution. Comparing results across patient CT scan images captured with different reconstruction kernels may not be straightforward because reconstruction kernels may generate statistically significantly different results^{12–14}. Examples of reconstruction kernel names relevant to GE scanners and this dissertation include “Standard”, “Bone”, “BonePlus”, and “BonePlus2”.

3.2 Densitometric calibration of X-ray CT systems and measurement of material densities from CT images

An exact *in vivo* measurement of bone mineral density from a clinical X-ray CT system may not be possible; but, there are several approaches that can provide good approximations. Notably, there is not yet a consensus among research labs or in the published literature on one method for DCTC which is necessary for approximation. One barrier to arriving at a consensus is the lack of publicly available details on the methods applied in the literature. This section explains different approaches based on the information available. Chapter 5 demonstrates ways to modify existing methods for different applications.

Before discussing densitometric CT calibration, equation (3.11) presents an alternate formulation of CT Numbers [HU] which more clearly demonstrates the dependence on energy, density, and chemical composition¹⁵. Here $\mu(E)$ is linear attenuation as a function of energy, $\frac{\mu(E)}{\rho}$ represents the mass attenuation coefficient, and c represents the

concentration.

$$CT\ Number(E) = 1000 * \frac{\left(\frac{\mu(E)}{\rho}\right)_{mineral} * c_{mineral} + \left(\frac{\mu(E)}{\rho}\right)_{H_2O} * c_{H_2O} - \mu(E)_{H_2O}}{\mu(E)_{H_2O}} \quad (3.11)$$

3.2.1 Existing densitometric CT calibration standardized terms

DCTC approximations depend on having either specific scans for DCTC or tables of energy-specific mass attenuation coefficients. The clinical gold standard for DCTC is to scan a BMD-specific densitometric reference^{16–24}. A *densitometric reference* is an object of known density or density that can be measured. Frequently the densitometric reference is made up of individual density references, called *standards*. A *phantom* contains multiple density references to quantify a range of densities. ASTM Standard E1935-97 recommends that phantoms include density standards bracketing the densities of interest²⁵. For example, femoral cortical bone would include 1000 mg/cm³ and 1750 mg/cm³ to bracket the densitometric range of interest^{26,27}.

In DCTC, segmentation is applied to identify the region or volume of interest within the individual density standards in the phantom. *Segmentation* is the creation of a digital region of interest. Each region of interest is summarized by a mean CT Number [HU] measurement to use in the derivation of a DCTC equation.

3.2.2 Existing databases of X-ray mass attenuation coefficient tables

Two databases have been created by the National Institute of Standards and Technology (NIST)^{28,29}: (1) the X-ray Attenuation and Absorption for Materials of Dosimetric Interest (XAAMDI), and (2) the XCOM: Photon Cross Sections Database. XAAMDI includes nominal densities of selected tissues and their accompanying relevant values, including

energy and mass attenuation coefficients²⁸. These tables build on prior work from the International Commission on Radiation Units (ICRU)³⁰. In them, each tissue density comes from a detailed review of the available literature, including averaged data-driven measurements when possible. Perhaps the best detailed in ICRU 44 is cortical bone, 1920 mg/cm³, which came from data for 24 adults (20—74 years old) and considered bones including the skull, vertebral column, pelvis, humerus, and femur³⁰. The XCOM database expands on the selected tissues from the XAAMD database by interpolating and combining X-ray mass attenuation values, based on photon cross section, for user-specified chemical compositions to provide data tables. K_2HPO_4 is an example of a material for which the XCOM database can provide energy-specific mass attenuation coefficients outside of the scope of the XAAMD database.

The versions of the tables from NIST examine the range of energies used in clinical practice at a finer resolution than the tables from ICRU. The mass attenuation coefficients can be plotted as a function of energy on a log—log plot. The NIST tables identify the discontinuities in these curves within the range of energies relevant to the clinic, such as cortical bone. One challenge with off-the-shelf densitometric reference phantoms is the proprietary nature of the materials used as density references, which limits the ability of XAAMD to provide tables for the densitometric reference phantoms.

3.2.3 Determining theoretical effective energy

In CT imaging, the X-ray source is a polychromatic beam consisting of multiple X-ray wavelengths. Accurate measurement of CT density is dependent on identifying the single X-ray wavelength, or monochromatic energy, equivalent to the combined X-ray wavelengths present in the polychromatic beam. The monochromatic energy, or *effective*

energy, is difficult to measure directly³¹. Some approaches to DCTC are dependent on determining the theoretical effective energy. *Slice effective energy* describes the effective energy specific to one axial CT slice. *Scan effective energy* describes the effective energy averaged over multiple slices or potentially the entire stack.

Several factors may affect effective energy by reducing X-ray intensity between the X-ray source and the X-ray detector, including (1) the examination object the X-ray passes through, (2) the X-ray current, and (3) the volumetric overlap and averaging of collected X-ray attenuation measurements. High-density anatomical features, such as cortical bone, may act as a high-pass filter, with higher mass absorption rates preventing lower-energy X-rays from passing through and increasing the effective energy. Variable current algorithms minimize patient exposure to unnecessary radiation in routine clinical practice. Introducing a different current for each CT image slice also creates slice-specific scan effective energies, and increases the difficulty of estimating effective energy. Section 3.1.2 mentioned measurement overshoot can be reduced by selecting the slice increment and slice thickness to capture overlapping CT measurements. This approach may also impact the estimation of scan effective energy. Two different methods for determining theoretical scan effective energy will be presented later in this chapter as the first step in the relevant approach to DCTC.

3.2.4 Phantom-based density approach to DCTC for equivalent BMD

In phantom-based DCTC, a CT scan of a densitometric reference is captured and a calibration relationship is derived. Mean CT Number [HU] measurements of regions of interest within the densitometric standards are captured from the CT calibration scan, to derive a CT Number-to-density relationship. Then a linear regression is performed

between the nominal density values of the individual densitometric standards (x-axis) and the mean CT Number [HU] measurements (y-axis). The resulting equation is then rearranged to express density [g/cm³] in terms of CT Number [HU].

$$CT\ Number\ [HU] = m * \rho + b \quad (3.12)$$

$$CT\ Number\ [HU] - b = m * \rho + b - b \quad (3.13)$$

$$\frac{CT\ Number\ [HU] - b}{m} = \frac{m * \rho}{m} \quad (3.14)$$

$$\rho = \frac{1}{m} * CT\ Number\ [HU] - \frac{b}{m} \quad (3.15)$$

In this method, the CT scan of the densitometric reference is assumed to have the same scan effective energy as the examination object. Ideally, the CT scan of the patient or examination object and the CT scan of the phantom are processed, reconstructed, and post-processed using the same hardware, acquisition parameters, correction algorithms, reconstruction kernels, and post-processing steps²⁵.

3.2.5 Example phantom-based density approach to DCTC

The ASTM standard E1935-97 provides guidance on the design of a phantom for use in DCTC. Consistent with this standard, a custom femoral density phantom (FDP) was created from four calcium hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$, abbreviated HA) densitometric standards (CIRS Inc, Norfolk, VA). These standards were selected to bracket the range of apparent densities relevant to human femoral trabecular (100–400 mg/cm³) and cortical (1000–1750 mg/cm³) bone^{32,33}. For this densitometric phantom the 100 and 400 mg/cm³ plugs were custom manufactured (parts RDH 357 Y-23 and RDH

362 Y-24) and the 1000 and 1750 mg/cm³ plugs were off-the-shelf (parts 06217 and 06221).

This phantom has been designed to be scanned *offline*, or separate from the patient. Therefore, a CT examination of the phantom submerged in deionized water was captured offline with CT scan acquisition and reconstruction parameters consistent with the UW—Madison Hospital clinical protocol for virtual colonoscopies. The densitometric standards were segmented in Mimics v.23 (Materialise, Leuven, Belgium). The mean CT Number [HU] of each segment was recorded in a table with the nominal density [mg/cm³] of the densitometric standard (Table 3-1).

Table 3-1: Example segmentation measurements from a CT examination of the femoral density phantom by densitometric standard.

Known HA Density [mg/cm ³]	Mean CT Number [HU]	Standard Deviation CT Number [HU]
100	120.17	21.38
400	491.57	33.81
1000	1212.19	81.26
1750	1959.26	188.51

A linear regression was performed for these data with mean CT Number [HU] on the y-axis and nominal density [mg/cm³] on the x-axis (Figure 3-3). The calibration curve is then derived following the process in equations (3.12)–(3.15). Depending on the material used in the densitometric phantom, this relationship calculates bone mineral density equivalent to the reference material, typically calcium hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$, abbreviated HA), or dipotassium phosphate (K_2HPO_4).

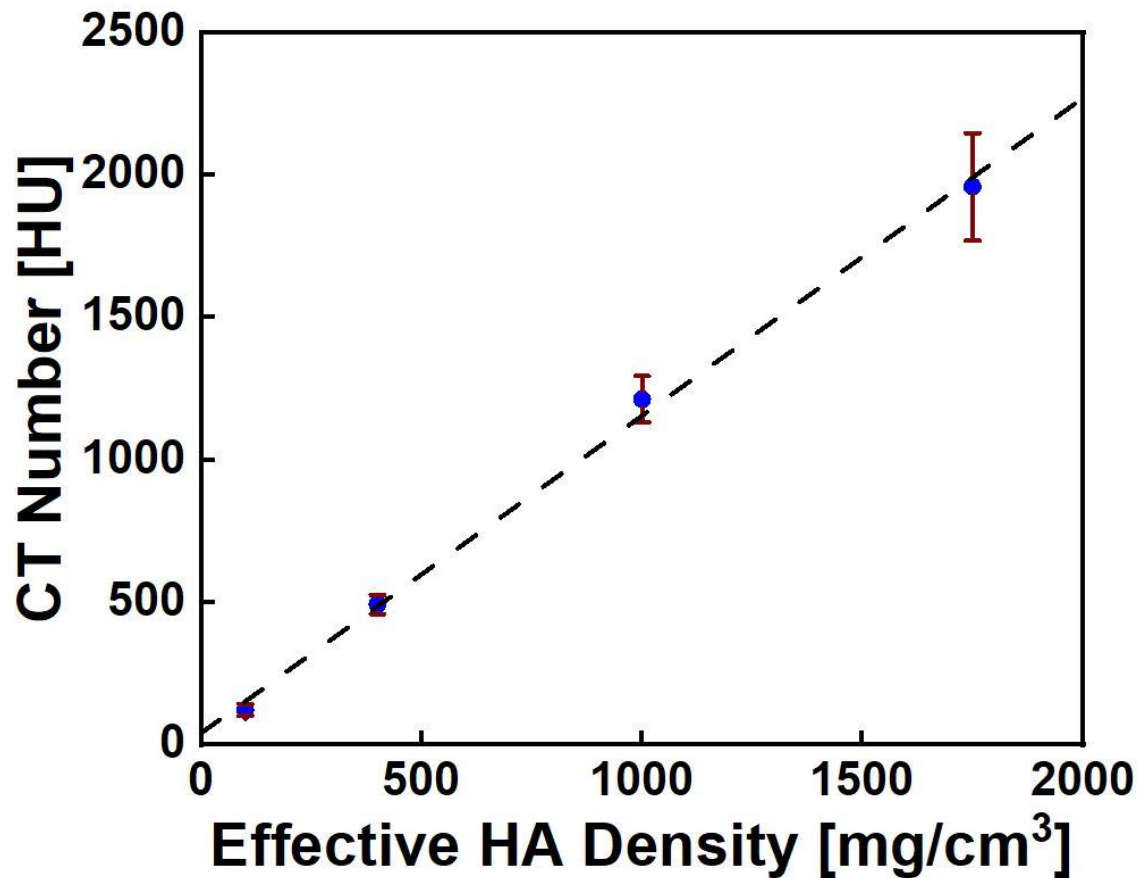


Figure 3-3: Example of a calibration curve for a calcium hydroxyapatite phantom. The resulting linear regression equation was $CT\ Number\ [HU] = 1.1155 * \rho \left[\frac{mg}{cm^3} \right] + 39.473$. The error bars represent the standard deviation for each densitometric standard segmentation mask CT Number [HU].

3.2.6 Phantomless density approach to DCTC for equivalent BMD

In the absence of a phantom as a densitometric reference, the nominal tissue densities from NIST's XAAMD database Table 2 may be assumed. Patient tissues can be identified near the anatomy of interest and segmented to determine mean CT Number [HU] for a region of interest. The same process can then be followed to determine the DCTC equation by constructing a linear regression between nominal density [mg/cm³] (x-axis)

and internal tissue segment mean CT Number [HU] (y-axis) and rearranging as demonstrated in equations (3.12)–(3.15). Depending on the CT scanner, the resulting intercept may need to be horizontally adjusted by the rescale intercept from the image metadata.

3.2.7 Attenuation approach to DCTC method for equivalent BMD

Attenuation-based DCTC has 2 steps: (1) simultaneously the effective energy and a linear attenuation to CT Number [HU] relationship are determined and (2) density is determined by multiplying the linear attenuation by theoretical mass attenuation coefficients. One approach to determining CT scan effective energy is described by ASTM E1935-97²⁵. In this approach, energy-specific semi-empirical linear attenuation values are found by multiplying the measured density by the energy-specific theoretical mass attenuation coefficients from the tables. A least-squares fit between the linear attenuation values and CT Number [HU] is then computed for at least three materials for each energy shown in equation (3.16), and the coefficient of determination for each fit is recorded.

$$\mu = m * CT\ Number\ [HU] + b \quad (3.16)$$

The effective energy is assumed to align with the maximum observed coefficient of determination³¹. Once the maximum coefficient of determination is identified and the scan effective energy is determined, the energy-specific linear relationship between CT Number [HU] and linear attenuation coefficients, μ , is also determined. Density can then be determined by using the relationship between linear attenuation, μ , and mass attenuation coefficients, $\frac{\mu}{\rho}$ repeated from equation (3.5).

$$\rho = \frac{\mu}{\left(\frac{\mu}{\rho}\right)} \quad (3.17)$$

3.2.8 Comparison of the density and attenuation DCTC methods

ASTM E1935-97 presents an illustrative example comparing two methods for density calibration, referred to here as the *density approach* and *attenuation approach*, respectively, to DCTC²⁵. This example illustrates that the two methods result in density calibrations that differ by a small amount (Table 3-2).

Table 3-2: Example CT density measurement comparison of the density and attenuation methods for polyamide and polycarbonate from ASTM E1935-97. Observe that the results for the attenuation method could be rounded to the same values resulting from the density method, indicating that the extra work of the attenuation method may not be worth the increase in accuracy. Additionally, the attenuation method is precise to within 1%, and the density method is precise to within 5% relative to the published density. Note that (1) polyamide and polycarbonate have densities that fall within the range of human cortical femoral bone which makes this example relevant to this dissertation, and (2) this example was derived on an industrial CT scanner, which may differ from clinical CT scanners, particularly in terms of resolution.

	CT Value	Density Method ρ [g/cm ³]	Attenuation Method ρ [g/cm ³]	Published ρ [g/cm ³]
Polyamide	1272	1.20	1.15	1.14
Polycarbonate	1273	1.20	1.23	1.21

In this illustration, the attenuation method is accurate to within 1%, and the density method is accurate to within 5%. Rounding the results from the attenuation method by one significant figure would yield the same result as the density method. This illustration may be relevant to DCTC for bone mineral density because the densities of polyamide and polycarbonate fall within the density range of femoral cortical bone (1 g/cm³ to 1.75 g/cm³). This illustration may not be relevant due to the difference in scale of the industrial CT scanner used in the example, which may be substantially different from clinical CT scanners. Following the format of this illustration from ASTM E1935-97, this dissertation includes illustrations more relevant to the clinic in subsection 3.2.10 and chapter 6.

3.2.9 Mass-fraction-model approach to DCTC for integral BMD

A *mass fraction* expresses each component of a mixture as a ratio between the mass of that component and the total mass of the mixture, equation (3.23). In a *mass-fraction-model* (MFM) the sum of mass fractions is set equal to 1, equation (3.25). Andrew Michalski described a mass fraction model approach to DCTC in his appendix³⁴ using the following steps: (1) determine both the effective energy and the energy-specific CT Number-to-mass attenuation relationship, (2) determine the CT Number-to-material density relationship, and (3) establishing a MFM that includes bone mineral density and bone marrow. This section describes these steps in detail.

A process for determining slice effective energy has been described previously^{31,34}. Inputs for the method include (1) mean CT Number [HU] measurements for regions of interest for a minimum of three materials and (2) tabulated energy specific mass attenuation coefficients associated with each material for scanner-relevant X-ray energies. To determine the effective energy, iterative linear regressions are constructed between the mean CT Number [HU] measurements (x-axis) and energy-specific mass attenuation coefficients for each energy (equation (3.18)). A vector of the coefficient of determination (R^2) for each regression is then constructed across energies.

$$\frac{\mu}{\rho} = m * CT\ Number\ [HU] + b \quad (3.18)$$

The determined effective energy is assumed to have the coefficient of determination closest to one in the vector. The resulting linear regression at the determined effective energy is also the CT Number-to-mass attenuation coefficient relationship.

In this semi-empirical method for DCTC, determining the CT Number-to-material density relationship requires several equations presented earlier in this. Recall the relationship between linear attenuation, mass attenuation coefficients, and density, which was introduced as equation (3.4):

$$\mu = \frac{\mu}{\rho} * \rho \quad (3.19)$$

and the definition of the Hounsfield Scale, previously shown in equation (3.10).

$$CT\ Number\ [HU] = \frac{\mu_m - \mu_{water}}{\mu_{water} - \mu_{air}} * 1000 \quad (3.20)$$

Substituting the relationship between linear attenuation, mass attenuation coefficients, and density into the Hounsfield Scale creates a relationship between CT Number [HU] and material density:

$$CT\ Number_m[HU] = \frac{\left(\frac{\mu}{\rho_m} * \rho_m - \frac{\mu}{\rho_{water}} * \rho_{water}\right)}{\frac{\mu}{\rho_{water}} * \rho_{water}} \quad (3.21)$$

Rearranging to solve for material density in terms of CT Number yields:

$$\rho_m = \frac{\frac{CT\ Number_m * \mu}{1000} * \rho_{water} + \frac{\mu}{\rho_{water}} * \rho_{water}}{\frac{\mu}{\rho_m}} \quad (3.22)$$

Here, $\frac{\mu}{\rho}$ are energy- and material-specific mass attenuation coefficients and ρ_{water} is the density of liquid water at standard temperature and pressure.

The MFM requires several inputs: the effective energy, the linear attenuation-to-CT Number relationship, and the material density-to-CT Number relationship. With these values, a MFM can be set up to differentiate between *integral BMD* and bone marrow, also known as triglyceride. Let a mass fraction be set up between the mass of the material and the mass of the voxel as shown in equation (3.23).

$$w_i = \frac{mass_{material}}{mass_{voxel}} \quad (3.23)$$

Then two equations can be set up to form a system, equations (3.24)–(3.25).

$$\frac{\mu}{\rho_{voxel}} = \sum_i w_i \frac{\mu}{\rho_i} \quad (3.24)$$

$$\sum_i w_i = 1 \quad (3.25)$$

Filling in these two equations yields:

$$\frac{\mu}{\rho_{voxel}} = w_{K_2HPO_4} \frac{\mu}{\rho_{K_2HPO_4}} + w_{triglyceride} \frac{\mu}{\rho_{triglyceride}} \quad (3.26)$$

$$w_{K_2HPO_4} + w_{triglyceride} = 1 \quad (3.27)$$

Substituting the mass fraction into these two equations yields:

$$\frac{\mu}{\rho_{voxel}} = \frac{mass_{K_2HPO_4}}{mass_{voxel}} \frac{\mu}{\rho_{K_2HPO_4}} + \frac{mass_{triglyceride}}{mass_{voxel}} \frac{\mu}{\rho_{triglyceride}} \quad (3.28)$$

$$\frac{mass_{K_2HPO_4}}{mass_{voxel}} + \frac{mass_{triglyceride}}{mass_{voxel}} = 1 \quad (3.29)$$

Rearranging equation (3.30) to isolate the mass ratio for triglyceride or bone marrow:

$$\frac{mass_{triglyceride}}{mass_{voxel}} = 1 - \frac{mass_{K_2HPO_4}}{mass_{voxel}} \quad (3.31)$$

And substituting into equation (3.32):

$$\frac{\mu}{\rho_{voxel}} = \frac{mass_{K_2HPO_4}}{mass_{voxel}} \frac{\mu}{\rho_{K_2HPO_4}} + \left(1 - \frac{mass_{K_2HPO_4}}{mass_{voxel}}\right) \frac{\mu}{\rho_{triglyceride}} \quad (3.33)$$

Then distribute:

$$\frac{\mu}{\rho_{voxel}} = \frac{mass_{K_2HPO_4}}{mass_{voxel}} \frac{\mu}{\rho_{K_2HPO_4}} + \frac{\mu}{\rho_{triglyceride}} - \frac{mass_{K_2HPO_4}}{mass_{voxel}} \frac{\mu}{\rho_{triglyceride}} \quad (3.34)$$

Rearranging to solve for $mass_{K_2HPO_4}$, or integral BMD:

$$\frac{\mu}{\rho_{voxel}} - \frac{\mu}{\rho_{triglyceride}} = \frac{mass_{K_2HPO_4}}{mass_{voxel}} \left(\frac{\mu}{\rho_{K_2HPO_4}} - \frac{\mu}{\rho_{triglyceride}} \right) \quad (3.35)$$

$$mass_{K_2HPO_4} = mass_{voxel} \frac{\left(\frac{\mu}{\rho_{voxel}} - \frac{\mu}{\rho_{triglyceride}} \right)}{\left(\frac{\mu}{\rho_{K_2HPO_4}} - \frac{\mu}{\rho_{triglyceride}} \right)} \quad (3.36)$$

Applying the definition of density:

$$\rho_{K_2HPO_4} = \frac{mass_{K_2HPO_4}}{volume_{voxel}} \quad (3.37)$$

It is shown that:

$$\rho_{K_2HPO_4} = \left(\frac{mass_{voxel}}{volume_{voxel}} \right) * \frac{\left(\frac{\mu}{\rho_{voxel}} - \frac{\mu}{\rho_{triglyceride}} \right)}{\left(\frac{\mu}{\rho_{K_2HPO_4}} - \frac{\mu}{\rho_{triglyceride}} \right)} \quad (3.38)$$

Where:

$$\frac{\mu}{\rho_{voxel}} = m * CT \text{ Number } [HU] + b \quad (3.39)$$

Since two linear equations are multiplied together, this DCTC method yields a calibration curve. In this example, integral BMD is derived in terms of dipotassium phosphate, $\rho_{K_2HPO_4}$. Alternately, DCTC could be derived in terms of calcium hydroxyapatite, ρ_{HA} . Also note that the CT Number-to-density [mg/cm³] relationship in this derivation is specific to the Hounsfield Scale, and therefore this relationship may need to be modified if the CT scanner being calibrated does not use the Hounsfield Scale.

3.2.10 Comparison of DCTC methods

The four methods for DCTC presented in this chapter have not yet been compared against each other in a clinical context. In this example, the femoral density phantom

(FDP) is the gold standard to compare against. This example uses the standards as the examination object and applies phantomless DCTC methods to calibrate the scan of the FDP. The standards have a known density. Once calibrated, the measurements can be compared to the density and the error can be quantified.

A phantom-based DCTC equation was derived based on mean CT Number [HU] measurements (y-axis) of a CT scan of a femoral density phantom (FDP). A linear regression was performed for known densities [g/cm^3] (x-axis) and mean segment CT Number [HU] measurements. The resulting equation was then rearranged as described previously, equations (3.12)–(3.15). Note that CT scanner, scan acquisition, and reconstruction parameters were kept constant between the scans of the patient and the FDP.

For illustrative purposes, phantomless mean CT Number [HU] measurements were captured for regions of interest within an axial slice of a patient CT scan, including air, adipose, aortic blood, skeletal muscle, and cortical bone. The slice effective energy was determined using both methods previously described. In the top row, semi-empirical linear attenuation coefficients of the tissue segmentations are iteratively correlated with the CT Numbers [HU], Figure 3-4. In the bottom row, tissue segmentations were iteratively correlated with tabulated energy-specific mass attenuation coefficients, Figure 3-4, and the coefficient of determination (R^2) was calculated. In both approaches, the slice effective energy was taken to be the maximum coefficient of determination. Using the slice effective energy, DCTC equations were found following three of the previously described approaches: the density method, the attenuation method, and the MFM method.

Table 3-3: Representative region-of-interest tissue segmentations from a CT scan of a 79-year-old patient at the proximal femur level.

Region of Interest/ Tissue Reference	CT Value	Assumed NIST Nominal Density
Air	−956.79	1.205
Adipose	−96.34	950
Aortic Blood	30.43	1060
Skeletal Muscle	19.82	1050
Cortical Bone	1215.01	1920

To assess accuracy, I assumed longitudinal X-ray tube stability and similar calibration equations between LightSpeed Model GE CT scanners. Then I applied phantomless calibration equations to calibrate a CT examination of densitometric standards in the FDP. Figure 3-5 shows the resulting density measurements for several calibration methods. None of the methods was a close match for all four femoral bone density standards. Accuracy was assessed across the femoral density range by averaging the normalised percent difference between the known density and measured density across plugs, as shown in equation (3.40).

$$Accuracy = \sum_1^n \left(\left(\frac{\rho_{known} - \rho_{measured}}{\rho_{known}} \right) * 100 \right) \quad (3.40)$$

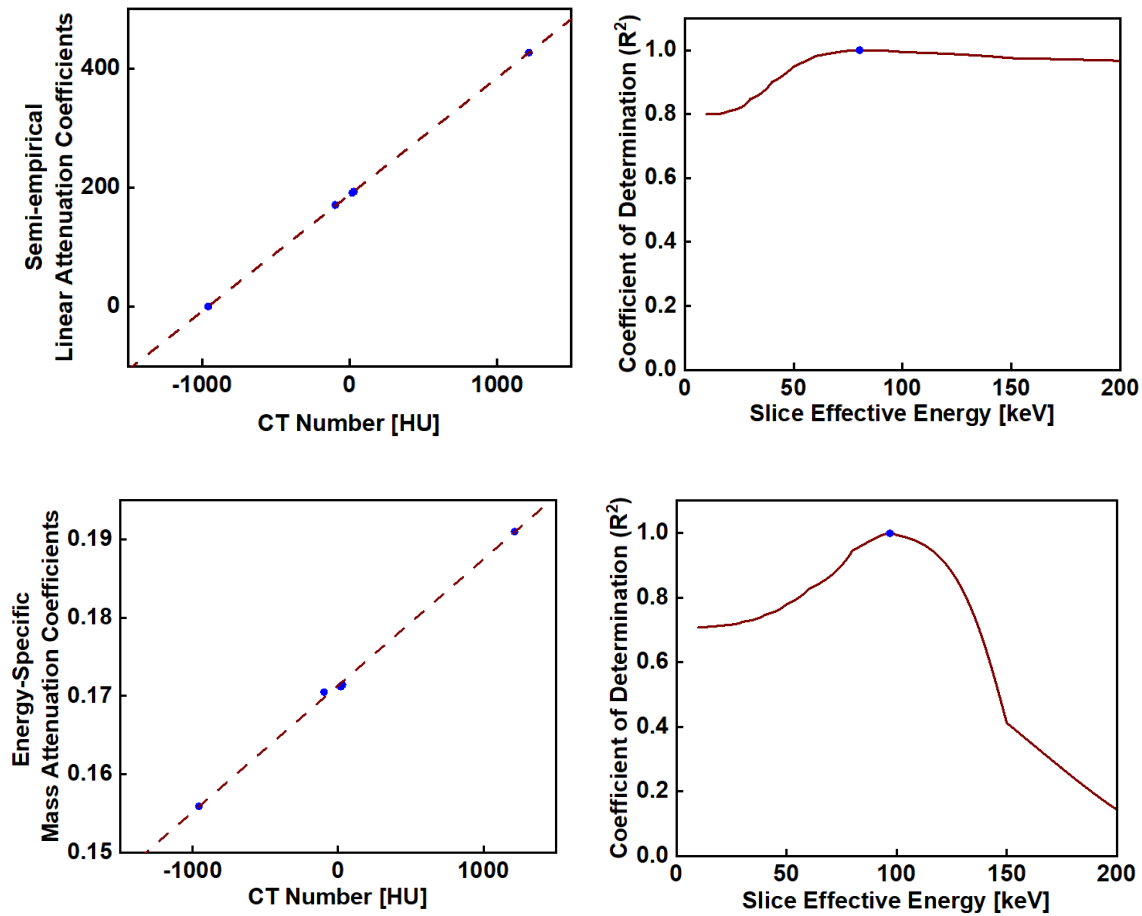


Figure 3-4: Representative slice effective energy calculations for tissue segmentations from the CT-scan of a 79-year-old patient. The top row presents the effective energy calculations for linear attenuation coefficients and the bottom row presents the effective energy calculations for the mass attenuation coefficients. In the left two plots, the blue dots represent different tissue segmentation measurements with mean CT Number [HU] on the x-axis and linear or mass attenuation coefficients respectively on the y-axis. In the right two plots, the blue dot identifies the maximum coefficient of determination across the slice effective energy [keV] range.

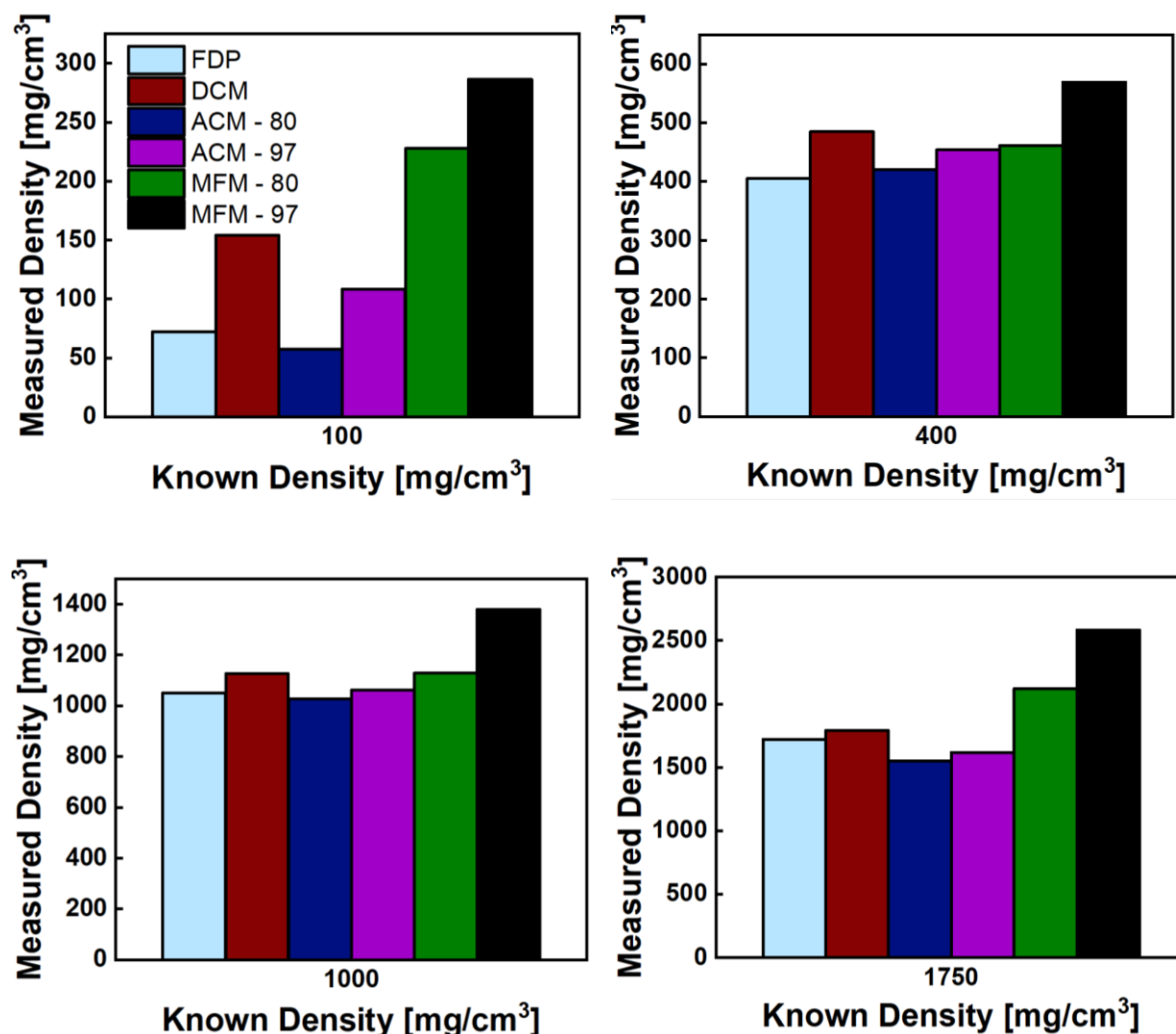


Figure 3-5: Representative density [mg/cm³] measurements for bracketing trabecular and cortical femoral bone standards for multiple calibration methods: femoral density phantom (FDP), phantomless density (ASTM DCM), phantomless attenuation (ASTM ACM–80, 97), and phantomless mass fraction model (MFM–80, 97). Where applicable, calculations were repeated for both effective energies respectively. Observe that the measured accuracy for the FDP results are within the 5% precision expected based on the example provided by ASTM E1935-97. Tissue-based calibration equations are based on the segmentation values for a 79-year-old patient from Table 3-3. This figure shows the most direct approach of performing a linear regression on nominal densities [mg/cm³] (x-axis) and mean CT Number [HU] measurements is the least likely to introduce error, although not the most precise approach available.

Note the FDP calibration had the best accuracy, 5.7%. In this analysis, the MFM calibration method had the highest overshoot in densitometric measurements. Using

linear or mass attenuation coefficients did noticeably impact the effective energy assumed. The MFM calibration method is more sensitive to the effective energy than the attenuation calibration method. Overall, the most direct approach of constructing a linear regression between nominal densities [mg/cm^3] (x-axis) and mean CT Number [HU] (y-axis). In the example presented in this section, this approach was also the least likely to introduce error. However, the attenuation methods may still be more precise.

3.3 Phantomless calibration facilitates QCT analyses

Phantomless DCTC shows promise for bringing quantitative CT-based measurements into clinical practice. Because they are capable of volumetric density measurements, CT scanners present an excellent foundation for physics-based modelling. DCTC remains an ongoing challenge limiting QCT analyses of the femoral bone. This chapter compared the “Standard Test Method for Calibrating and Measuring CT Density” from the ASTM to a MFM applied for DCTC. The most direct approach of performing a linear regression of nominal densities [mg/cm^3] (x-axis) against mean CT Number [HU] measurements is the least likely to introduce error but not the most precise approach available. The examples created for this section included segmented regions of interest limited to one CT slice for one patient, captured on only one CT scanner. However, the findings are likely relevant beyond the scope of the illustration. Later examples in this dissertation will consider additional patients. In any case, conducting similar benchmark evaluations across more patients and clinics may broaden the foundation of support for quantitative CT analyses.

Notably, ASTM’s “Standard Test Method for Calibrating and Measuring CT Density” has not been referenced in much of the literature in this area. The absence of this key reference may be due to the lack of awareness in the field, the cost associated with

accessing the standard, or a lack of clinical trust in this standard. Many of the guidelines recommended in the standard agree with the recommendations and observations from recent literature^{35,36}, including those in Troy et al.'s "Practical considerations for obtaining high quality quantitative computed tomography data of the skeletal system" and Brunnquell et al.'s "Sources of error in bone mineral density estimates from quantitative CT". By referencing this standard and providing a clinically relevant benchmark example, this dissertation aims to provide fundamental terminology to support further development of DCTC and CTPSFE.

4 CTPSFE analysis

Today, the best non-invasive methods for predicting risk of femoral fracture are derived from CTPSFE analyses. These methods are reasonably mature³⁷, and one existing implementation of these methods is currently the most accurate diagnostic for osteoporosis available in the clinic³⁸. Notably, CTPSFE analysis derived femoral strength predictions can consider all relevant variables to predicting fracture risk: femur geometry, bone mineral density, microarchitecture, the applied loads, and external interactions with the environment³⁹. Resolutions insufficient to capture microarchitecture are one limitation of patient CT scans captured during routine clinical practice. Despite this limitation, Verhulp et al. showed good agreement between CT scans of cadaveric femurs captured on both clinical- and micro-CT scanners⁴⁰. Further, Adams et al. showed that CTPSFE analysis conducted on routine clinical CT scans can yield fracture predictions at least as accurate as the current gold-standard diagnostic, DXA³⁸.

The simplest description of finite element analysis (FEA) includes 3 inputs: (1) geometry, (2) material properties, and (3) boundary and loading conditions. Clinical meaning is derived from the results through post-processing. This chapter discusses these three inputs and post-processing to survey CTPSFE analysis techniques specific to femoral fracture risk prediction. I also introduce terminology and details relevant to CTPSFE analyses and describe an example pipeline for retrospectively assessing deidentified patient cohorts (Figure 4-1). The steps in this pipeline are (1) capture patient CT scans, (2) extract patient geometry through image segmentation, (3) discretize the geometry for FEA, (4) spatially map material properties from CT scan data to the

discretized patient model, (5) define boundary and loading conditions to simulate the stance phase in walking or sideways fall loading, and (6) estimate the minimum force likely to result in fracture. Within each of these steps, differing software packages and research labs implement varying methods. As there is not yet a consensus across research labs, this chapter focuses on describing the details available in the literature relevant to two specific implementations^{7,17,41}.

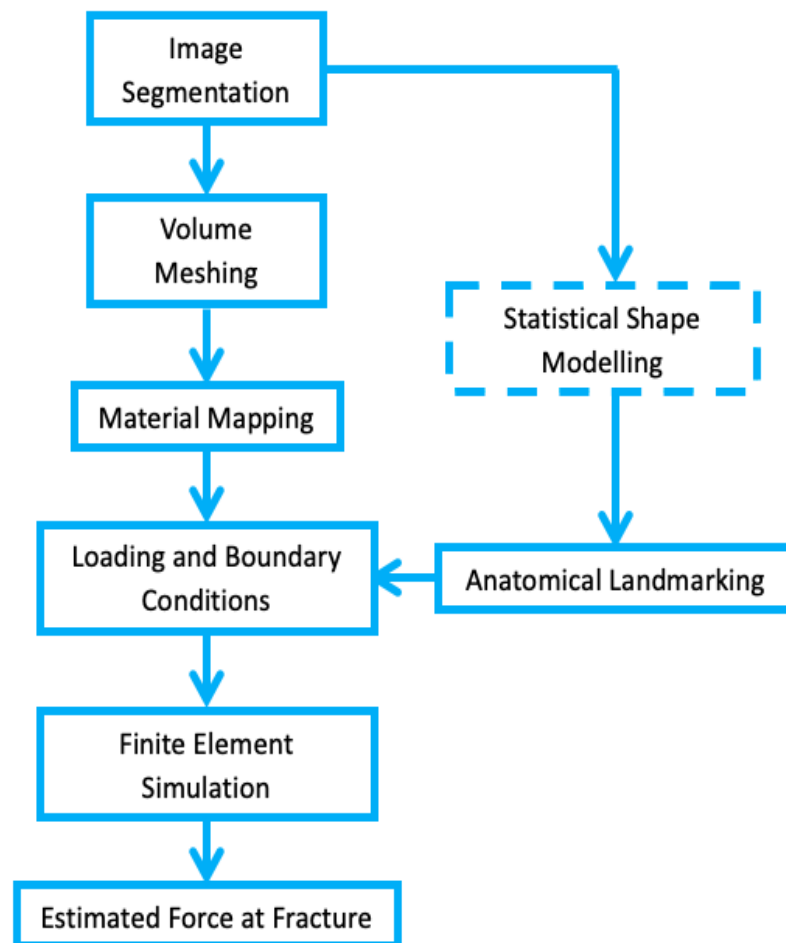


Figure 4-1: A visualization of computed tomography-based patient-specific finite element modelling pipeline steps. Statistical shape modelling is an optional step conducted for some data sets as indicated by the dashed outline.

4.1 Geometry

This section discusses three aspects of generating patient-specific femoral geometry. First, anatomical planes provide geometric reference terms for geospatially identifying locations specific to the patient (Figure 4-2). Second, image segmentation describes the process for manually generating a digital patient-specific geometry from CT data. Third, discretization and mesh generation describe the process of breaking up the patient-specific geometry into smaller sections to facilitate the simulations. The patient-specific geometry is the first step in building a CTPSFE model.

CTPSFE analyses are sensitive to geometry indicating the importance of accurately representing patient-specific anatomy such as the femur. Taddei et al. showed that errors in the geometric representation of the bone were always the dominant variable in resulting stress predictions⁴². Their study also showed that the variation in variable output from their method for building CTPSFE models of a femur from clinical CT data never exceeded 9%. Increasingly accurate representations of patient-specific anatomy are facilitated by increases in the fidelity of CT scanners and computational tools for creating geometry from CT data.

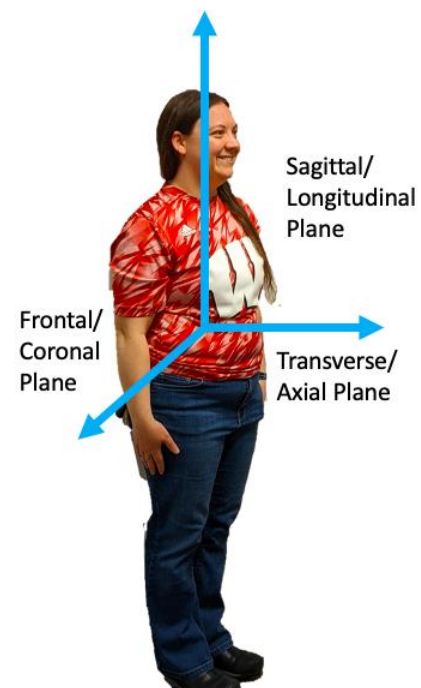


Figure 4-2: Representative example of the anatomical planes applied to a human.

4.1.1 Anatomical planes

When considering anatomy, there are three planes of the body: (1) the Coronal or Frontal plane; (2) the Sagittal or Longitudinal plane; and (3) the Transverse, axial, or Horizontal

plane (Figure 4-2). The Coronal plane divides the body into the anterior (front side) and posterior (back side) portions. The Sagittal plane divides the body into right and left portions. The transverse plane divides the body into upper and lower halves. Medial refers to the side of the anatomy closest to the central Sagittal plane, and lateral refers to the side of the anatomy further from the central Sagittal plane. Proximal implies close to the centre of the body. The proximal femur refers to the part of the femur near the hip joint. Conversely, distal refers to the direction away from the centre of mass. The distal femur being the part of the femur near the knee.

4.1.2 Image segmentation

Image segmentation is the process of identifying a region of interest on an image stack and creating a corresponding digital geometry. The segmented region of interest is typically called a mask (Figure 4-3). There are several different approaches to creating a digital patient-specific femoral geometry, depending on the availability of resources and expertise.

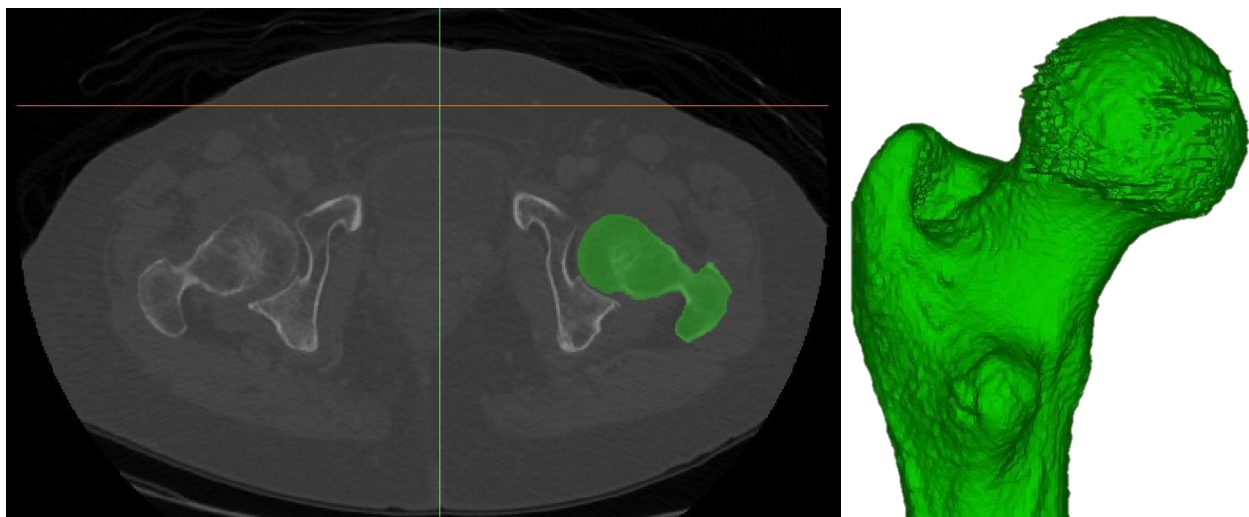


Figure 4-3: Representative axial CT slice and segmented proximal femur for a 71-year-old patient.

One semi-automatic segmentation approach, called thresholding, begins by isolating a specific range of CT Numbers [HU] for inclusion in a mask. The mask is refined by isolating the region of interest from the rest of the areas within the full density range. Finally, the mask edges are manually checked slice by slice in at least two planes until a closed volume is created. The closed volume can be checked automatically by “filling” in the centre, with failure resulting in the entire image becoming part of the mask and success resulting in only the femoral mask becoming filled in. One software package that supports this approach is Mimics (Materialise, Leuven, Belgium).

Region growing is an alternative semi-automatic segmentation approach, with the potential to reduce the time required to segment a femur. This approach begins by manually planting starter “seeds” within the patient femur. The user then iteratively steps through growing those seeds, with automated sets of user-specific iteration advances, until as much of the femur as possible is cleanly included in the mask. The mask is then updated to form a closed volume by manually editing the mask in each slice in at least two planes. One software package that uses this approach is ITK-Snap (ITK-Snap, University of Pennsylvania). In ITK-Snap, the mask is always displayed as hollow. When the full patient femur is not captured within the image stack, the bottom of the femoral mask is represented as open, and ITK-Snap will not allow it to be filled. To form a closed geometry required for later steps, the second to bottom slice can be filled in.

These two methods produce results of comparable accuracy. In general, the time required to segment a patient's femur is dependent on the disease state of the patient, the distance between the acetabulum and the femoral head, and the expertise of the operator. Mimics has created a built-in tool for automatically segmenting a patient's femur.

This tool may not yet be able to accurately segment the femur of a patient with osteoporosis, which limits its utility in the application of these methods to the cohort described in Chapter 5.

4.1.3 Mesh generation

Once the segmentation is exported from the image analysis software, the next step is to *discretize* the patient-specific femoral geometry into discrete finite elements connected by nodes. This process outputs a *mesh* or a collection of elements with nodes that represents the shape of the geometry. Automated mesh generators are widely available, both in commercial off the shelf software applications (ANSYS, Abaqus, HyperWorks, etc.) and open-source software applications (TetGen, etc.). There are several options for element geometry when creating a mesh, including tetrahedral elements (tets), hexahedral elements (hexes), wedges, and shells. Tetrahedral elements are frequently chosen because they can more accurately cover arbitrary geometries (Figure 4-4).

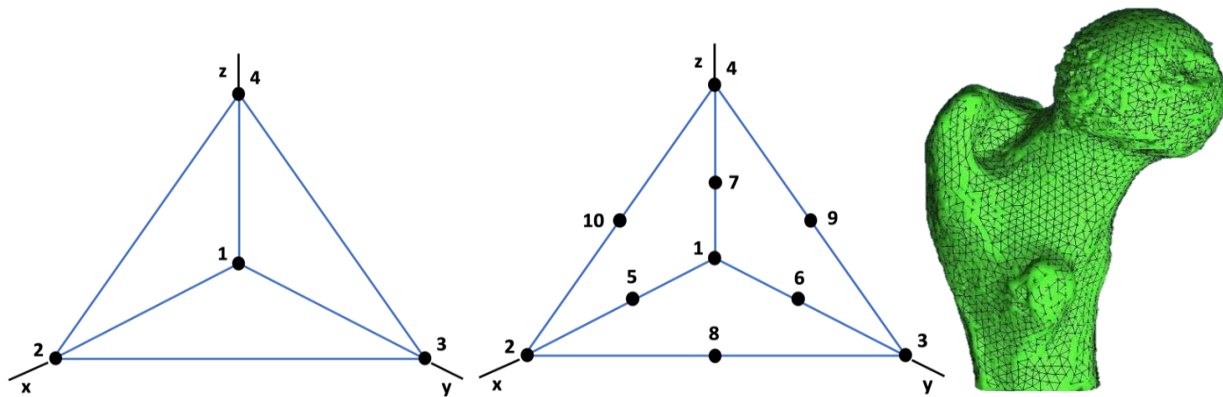


Figure 4-4: Examples of representative tetrahedral elements, tet-4 on the left, tet-10 in the middle and a patient femur with an applied tetrahedral mesh on the right.

Four- and ten-node tetrahedral elements are displayed in Figure 4-4. In FEA, one potential source of error when using tetrahedral elements is *shear locking*. Four node

tetrahedral elements have linear shape functions which cannot accurately model the curvature present when a material experiences the shear stress associated with bending. To avoid this source of error, ten-node tetrahedral elements are recommended due to their accompanying quadratic shape functions. After creating a tet-10 mesh for the CTPSFE model, the material properties can be derived and geospatially assigned throughout the mesh as described in the next section.

4.2 Material properties

The ability of CTPSFE analyses to simulate the results of cadaveric experiments may be sensitive to an accurate definition of the material properties of bone. Schileo et al. showed that accurately defining heterogeneous material properties reduced the measure element-by-element strain field error between specimen-specific finite element model results and cadaveric experimental results used for validation for a quasi-static side fall loading study of three specimen⁴³. Their method for accurately defining material properties relies on several empirically derived densitometric relationships. This section discusses the definitions of these densitometric relationships, and the methods used to find them.

4.2.1 Deriving phantom-based effective BMD

Two examples of deriving calibration equations are presented below. The first and simpler method was derived for use with plastic densitometric phantoms designed to mimic the radiodensity of calcium hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$, abbreviated HA). In this method, an examination of the phantom is captured using a CT scanner, and densitometric standards are segmented. The mean CT Number [HU] of each segment is

recorded in a table with the nominal density [mg/cm^3] of the densitometric standard (Table 4-1). A linear regression is constructed from these data, with mean CT Number [HU] on the y-axis and nominal density [mg/cm^3] on the x-axis (Figure 4-5). The resulting equation is called the *linear calibration curve*.

Table 4-1: Example segmentation measurements for each densitometric standard in a CT examination of the CIRS Model 004 CT Simulator for Bone Mineral Analyses, a vertebral density phantom.

Known HA Density [mg/cm^3]	Mean CT Number [HU]	Standard Deviation CT Number [HU]
50	37.48	49.93
100	87.74	35.34
150	147.49	31.09

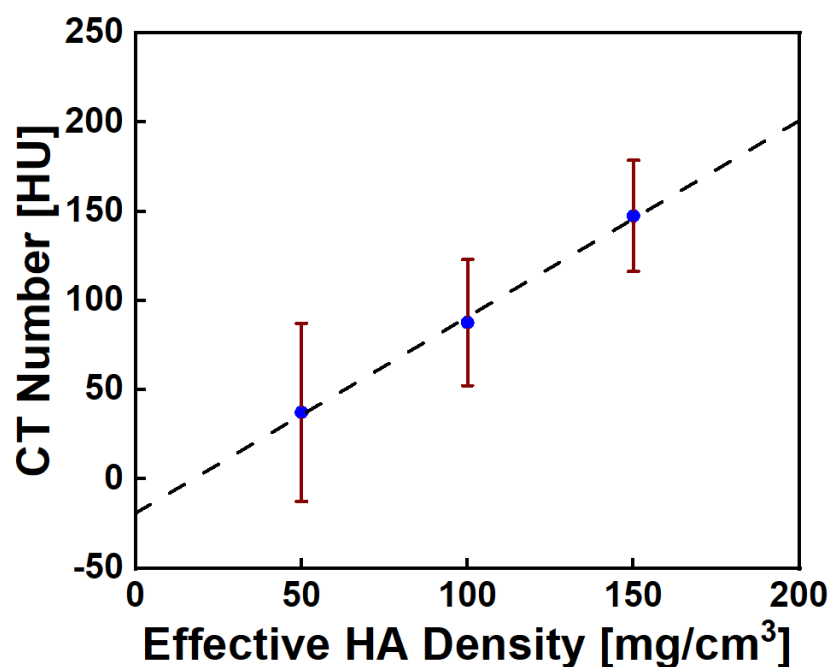


Figure 4-5: Example of a calibration curve linear regression for a calcium hydroxyapatite densitometric phantom. In this example, the resulting linear regression equation was $CT\ Number\ [HU] = 1.10 * \rho \left[\frac{mg}{cm^3} \right] - 19.11$. Note that error bars represent standard deviations of the CT Number [HU].

The second method was designed for use with a liquid phantom, before K_2HPO_4 equivalent plastic was available. For consistency, the same method is now used with plastic densitometric phantoms designed to mimic the radiodensity of K_2HPO_4 . First, an examination of the phantom is captured using a CT scanner, and densitometric standards are segmented. Second, the mean CT Number [HU] of each segment is recorded in a table with both water-equivalent densities and K_2HPO_4 -equivalent densities (Table 4-2). The water-equivalent density values are then subtracted from the CT Numbers [HU] to obtain the radiological density of the K_2HPO_4 (Table 4-2). The linear regression is then constructed with the radiological density of K_2HPO_4 on the y-axis and the nominal density of the K_2HPO_4 standards on the x-axis (Figure 4-6). The result of the linear regression still requires a correction to account for the physical consideration of the volume of water displaced by the addition of K_2HPO_4 . Provided in the QCTPro Software manual, these offset values for the Mindways Model 3 are a slope correction of -0.2174 and an intercept correction of $+999.6$. Once these offsets have been applied, the BMD calibration curve has been derived.

Table 4-2: Representative segmentation mean CT Number [HU] measurements for a 71-year-old patient with accompanying phantom-specific water and K_2HPO_4 density tabulated values. The far-right column presents the CT Number [HU] minus water as a prerequisite for plotting.

Water Density	K_2HPO_4 Density	CT Number [HU]	CT Number-Water
1012.2	-51.8	-53.3	-1065.5
1057	-53.4	-0.6	-1057.6
1103.6	58.9	205.2	-898.4
1119.5	157	350.9	-768.6
923.2	375.8	459.6	-463.6

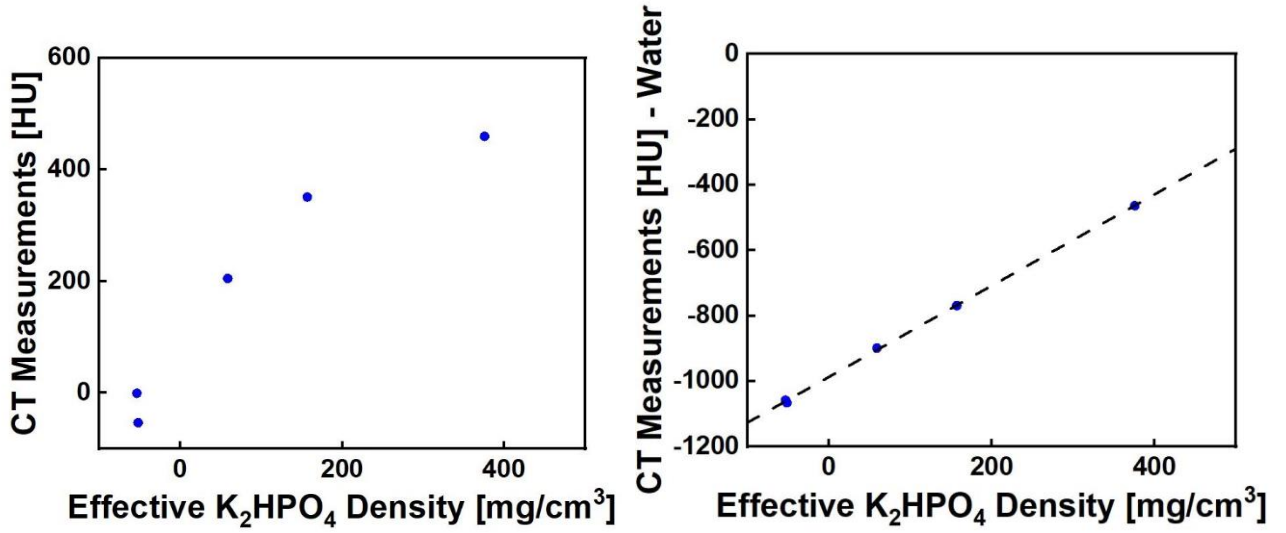


Figure 4-5: Example plots demonstrating why water CT Number [HU] measurements must be subtracted in order to regain linearity to find the linear calibration curve for the inline Mindways Model 3 BMD phantom.

For both methods, the calibration equation needs to be rearranged to solve for QCT density as a function of CT Number [HU], which will be used in the material mapping process. The correct algebra is shown in equations (4.1)–(4.4) below.

$$CT\ Number = m * \rho_{QCT} + b \quad (4.1)$$

$$CT\ Number - b = m * \rho_{QCT} + b - b \quad (4.2)$$

$$(CT\ Number - b)/m = \frac{m * \rho_{QCT}}{m} \quad (4.3)$$

$$\rho_{QCT} = \frac{1}{m} (CT\ Number - b) \quad (4.4)$$

4.2.2 Empirically derived density-elasticity relationships

To accurately apply nonhomogeneous material properties, a density-elasticity relationship or a series of density-elasticity relationships is used to map measured CT Number [HU] values to the patient-specific finite element mesh. Specific to the methods used in this

dissertation, a series of densitometric empirical relationships was derived. This subsection (1) explains the terms necessary to understand these relationships; (2) presents the densitometric relationships defined in the literature that are relevant to this CTPSFE method; and (3) discusses the experimental process for finding these densitometric relationships. Several of the densitometric relationships described here were proposed by Schileo et al.⁴³.

There are different ways to measure density depending on the application. *Radiological density*, ρ_{QCT} , refers to the quantitative CT measurement relative to a reference density from a phantom. For example, Schileo et al. scanned a European Spine phantom (ESP) on a GE Brightspeed CT scanner in helical mode at a peak voltage of 120 kVp, and a tube current of 160 mA with a slice spacing and slice reconstruction of 0.625 mm and pixel dimensions of 0.3125 mm * 0.3125 mm. Their measured radiological density relationship was $\rho_{QCT} \left[\frac{g}{cm^3} \right] = 0.007764 * CT \text{ Number } [HU] - 0.056148$.

According to Morgan et al., *apparent density*, ρ_{app} , is a measurement of *wet mass* divided by *bulk volume*⁴⁴. Following a previously published method⁴⁴, Schileo et al. calculated the *bulk volume* of a bone core, $V = \pi * \frac{D^2}{2} * L$, using averages of the values from six measurements of diameter and length⁴³. The diameter and length measurements were assumed to be captured using calipers. *Wet mass* was found following a multi-step process. First, the marrow was removed by washing the specimens in a 10% bleach ultrasound bath at 37°C for three periods of ten minutes, rinsed with water between baths. Second, specimens were repeatedly dried and weighed, taking measures after 60 seconds under air jets until consecutive readings differed by less than 0.5%.

According to Schileo et al., *ash density*, ρ_{ash} , is a measurement of the *ash weight* of a bone specimen divided by the same specimen's bulk volume. The same bulk volume measurement was used to calculate both apparent density and ash density. The *ash weight* of each bone specimen was found by burning each specimen at 650°C for 24 hours in a muffle furnace, letting the specimens cool in the furnace for 24 hours, and weighing the resulting ash for each specimen.

The *elastic modulus*, E , is a measure of the bone's ability to resist elastic deformation under load and is expressed in terms of density in this context. To find this relationship, Morgan et al. used experimental, computational, and analytical methods to identify a femoral-neck-specific density-elastic modulus. In their study, tissue specimens came from 61 donors with no medical history of either metabolic bone disease or cancer. Specimens were also examined radiologically to rule out evidence of damage or bone pathologies. Specialized protocols were used to obtain 8 mm diameter on-axis bone specimens parallel to the trabecular orientation from specific anatomic sites, including the proximal femur^{45–47}. Apparent elastic moduli were found using both uniaxial tension testing and compression testing of bone cores. For tension testing, the *apparent elastic modulus* was defined as the slope at zero strain of a quadratic curve fit to the stress—strain curve from 0 to 0.2% strain⁴⁴. For compression, an extensometer captured four apparent elastic modulus measurements at four different positions around the specimen circumference, approximately 90 degrees apart, and averaged to find the specimen's apparent elastic modulus⁴⁴. After mechanical testing, six specimens from the femoral neck were micro-CT scanned at a resolution of 22 μm . CT-based linear FEA with a tissue modulus equal to 1.0 GPa for all elements was conducted on a voxel mesh of each specimen⁴⁴. To find the

finite element computed apparent elastic modulus, the ratio of the experimental apparent elastic modulus to the apparent modulus was calculated from the FEA⁴⁴.

Two theoretical relationships already derived provided the starting point for the analytical portion of Morgan's investigation. The first, by Cowin et al.⁴⁸, incorporates specimen-specific architectural information and consists of three orthotropic stiffness matrix entries of the form:

$$C_{iiii} = \hat{c}_{iiii}(E_t, \phi, \lambda_i, II) \quad (4.5)$$

$$C_{iijj} = \hat{c}_{iijj}(E_t, \phi, \lambda_i, \lambda_j, II) \quad (4.6)$$

$$C_{ijij} = \hat{c}_{ijij}(E_t, \phi, \lambda_i, \lambda_j, II) \quad (4.7)$$

In these entries, i and j are the indices 1, 2, 3 with the limitation that i cannot equal j (note 1, 2, 3 are the principal axes of an orthogonal coordinate system), E_t is the tissue apparent elastic modulus, λ_i is the normalized mean intercept length eigenvalue associated with the i th direction (describes anisotropy), and $II = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3$. Morgan et al. noted that E_t is merely a scaling factor⁴⁴. These relationships were further developed and simplified by Kabel et al.⁴⁹, who determined dependence on volume fraction is a power law, $\phi^{1.6}$, and that λ_i and E_t are both contained in the leading coefficient.

The second theoretical relationship that Morgan et al. used⁵⁰ does not incorporate specimen-specific architectural information:

$$E = 1240E_t\phi^{1.8} \quad (4.8)$$

with E as the predicted, on-axis elastic modulus. Morgan et al. calculated E_t by setting the predicted theoretical elastic modulus equal to the experimentally measured elastic modulus and solving for E_t given either λ_i and ϕ for the first relationship or only ϕ for the second relationship. In their study of these relationships, Morgan et al. demonstrated

these density-elasticity relationships are specific to anatomic site, which is critical to keep in mind when applying these methods to differing anatomic sites. Schileo et al. applied Morgan et al.'s previously derived relationship between the elastic modulus and apparent density⁴⁴: $E \text{ [GPa]} = 6.850 * \rho_{app}^{1.49} \text{ [g/cm}^3\text{]}$.

Three additional relationships are required to define material properties: (1) a relationship between radiological density and ash density, (2) a relationship between ash density and apparent density, and (3) a Poisson's ratio for bone. Previously, some studies assumed radiological density is equal to ash density⁴³. However, several studies have reported the need for a linear correction to find ash density from radiological density⁴³. There may be several reasons this linear correction is needed. Schileo et al. suggested two potential reasons: (1) phantoms cannot perfectly mimic bone attenuation coefficients, and (2) phantom inserts are homogeneous, while bone is non-homogeneous at the resolution of clinical CT scanners⁴³. Not noted by Schileo et al. in their paper, another potential reason is the spatial inhomogeneity inherent in CT physics²⁵.

In their 2008 study, Schileo et al. found the following empirical relationship between radiological density and ash density, based on pooled trabecular and cortical femoral bone results from 60 cylindrical core specimens:

$$\rho_{QCT} \left[\frac{mg}{mm^3} \right] = 1.14 * \rho_{app} \left[\frac{mg}{mm^3} \right] - 0.09. \quad (4.9)$$

Additionally, Schileo et al. found a constant ratio, 0.598 ± 0.036 , between ash density and apparent density⁴³. Wirtz et al. review literature on known bone material properties and found the average value, 0.3, for the Poisson's ratio of cortical bone⁵¹. Schileo et al. used this value for Poisson's ratio in their study.

Rearranging the density-elasticity relationships to summarize the equations needed for the next section on material mapping, equations (4.10)-(4.11) show the example empirical relationships between radiological density and CT Number [HU] derived from the ESP; radiological density and ash density; ash density and apparent density; and apparent density and elastic modulus

$$\rho_{QCT} \left[\frac{\text{g}}{\text{cm}^3} \right] = 0.007764 \text{ CT Number [HU]} - 0.056148 \quad (4.12)$$

$$\rho_{ash} = 0.8772\rho_{QCT} + 0.07895 \quad (4.13)$$

$$\rho_{app} = \left(\frac{1}{0.6} \right) \rho_{ash} \quad (4.14)$$

$$E = 6850\rho_{app}^{1.49} \quad (4.15)$$

Recall that the application of these equations are limited based on the agreement of specific details including CT scanner, CT scan acquisition protocols, calcium hydroxyapatite reference phantoms, femoral neck material properties, pooled femoral trabecular and cortical bone material properties.

4.2.3 Assignment of apparent elastic moduli onto the FE mesh

The ability of CTPSFE models to predict stress and strain measurements, when compared to experimental results on the surface, is also dependent upon how the material properties are assigned throughout the model⁵². Taddei et al. developed and made available in the public domain a software application, Bonemat v.3.2. This software application is capable of geospatially mapping CT Number [HU] values onto an FE mesh, in terms of nonhomogeneous tissue apparent elastic moduli. Their method transforms the CT Numbers [HU] into a Young's modulus continuum field before performing the

numerical integration over each element's volume⁵². Note that Bonemat requires a tetrahedral or hexahedral finite element mesh.

4.3 Boundary and loading conditions

This section discusses how to formulate real-life scenarios likely to cause femoral fragility fracture, such as standing, walking, or a sideways fall. The process of walking, or *gait cycle*, is divided into two movements: the stance phase and the swing phase. The gait cycle includes heel contact, foot-flat, midstance, heel-off, toe-off, midswing, and heel contact. In the midstance position, the patient is supporting their entire body weight on one leg, placing increased stress on that femur. A sideways fall scenario, also routinely considered, includes the patient falling with their thigh against the ground and their bodyweight loading the femur in a direction that does not typically occur. Boundary and loading conditions relevant to the stance phase of walking and the sideways fall scenarios are shown in Figure 4-7. The following subsections describe literature relevant to specific boundary and loading conditions, the experiments that have been conducted to validate CTPSFE results, and the development of CTPSFE simulation methods.

4.3.1 Stance phase of walking loading

Instrumented femoral hip implants provide the best data available to understand the gait patterns and ground reaction forces within the femur during standing and walking. In the early 2000s, Bergmann et al. designed and surgically placed instrumented femoral hip implants in four patients (3 men, 1 woman)⁵³. The hip contact forces from routine activities identified by Bergmann et al. have been applied as loading conditions in CTPSFE analyses of stance loading^{17,24}.

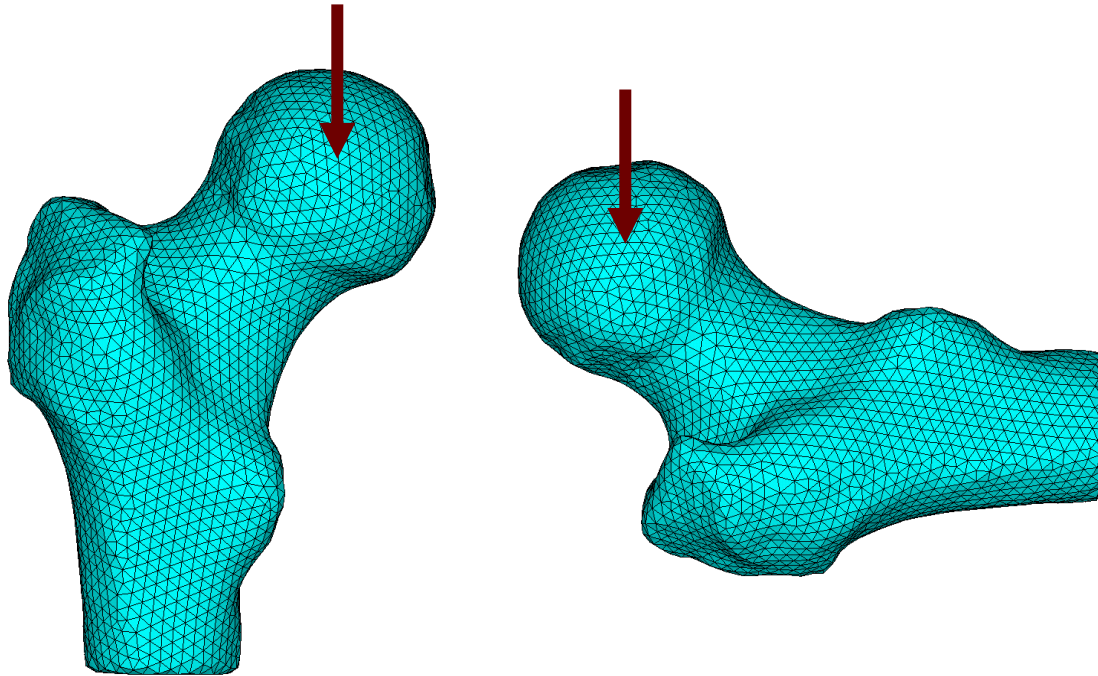


Figure 4-7: Representative stance (left) and sideways fall (right) loading conditions for a 71-year-old patient. Note that the load is applied to the node closest to the geometrical centre of the femoral head.

4.3.2 Sideways fall loading

The forces within the femur during a sideways fall have not been measured. In a sideways fall, the patient falls landing with their hip against the ground. The combined force of impact and bodyweight creates a combined loading on the femoral neck, including compression and bending. This can be simulated by applying loading and boundary conditions. A loading force is applied as a point force to the centre of the femoral head or as a distributed load to the surface of the femoral head, to simulate the force of bodyweight. The lateral side of the patient's femur can land anywhere on the ground. This can be simulated by imposing a contact constraint that does not allow the femur to move through the ground. A no-friction slider, that allows the femur to translate along the ground without experiencing friction and without passing through the ground, is one example of

how to apply this boundary condition. The knee joint is constrained against translation and rotation in all directions. This final boundary condition prevents rigid-body motion in the simulation and creates a statically determinate problem.

Qasim et al. showed that the sideways fall loading condition resulted in a higher fracture-control case stratification accuracy when using CTPSFE analysis towards femoral fracture risk prediction¹⁷. In a follow-up study, Altai et al. looked at differing boundary conditions with additional sideways fall loading angles beyond those applied by Qasim and found similar accuracy regardless of loading conditions, with a maximum improvement in stratification accuracy of 3% above the results found by Qasim et al.^{17,41}.

4.3.3 Experimental validation of sideways fall loading

An example cadaveric experiment is provided by Helgason et al. in 2016⁵⁴. Fresh frozen proximal femoral specimens were prepared by potting their shaft in an aluminium cylinder using polymethylmethacrylate (PMMA). The distal end of the specimen was placed in a hinge constraint such that the distal end of the specimen was free to rotate in the frontal plane. The trochanter was placed under the upper material testing machine platen and the femoral head was placed on top of the lower material testing machine platen. PMMA pads were formed to the femur and were placed between the material testing machine platens and the trochanter and femoral head respectively. The no-friction slider was created by placing ball bearing plates under the lower platen. Strain field measurements can be captured either by strain gages or by digital image correlation.

There are several limitations to cadaveric experiments. The specimens are typically fresh frozen to preserve the femoral specimens until the experiment, which can affect the

mechanical properties. For example, if stored improperly (not in saline-soaked gauze, not sealed, etc.) or left too long in the freezer, the mechanical properties of the femur can become compromised. As these cadaveric tests are destructive, each specimen can only be tested once.

4.3.4 Specimen-specific numerical simulation of a sideways fall

Verification and validation are critical to extracting clinically relevant data from finite element simulations⁵⁵. Conceptually, verification confirms numerical accuracy or that the applied method for solving the equations is correct. Verifying finite element software packages is typically accomplished by confirming results against a series of benchmark problems, such as a beam in bending. Validation checks the numerical prediction accuracy with respect to the physical phenomenon being replicated. Cadaveric experiments, such as those discussed in the previous section, can serve as validation for numerical modelling methods.

To compare strain field results between the cadaveric experiments and the numerical simulations, the numerical simulations need to replicate the experiments. CT scanning the specimens submerged in deionized water before the experiments allows for CT-based specimen-specific finite element models to be generated. Note that in this context the requirement to submerge the specimens comes from the prevalence of water within the living patients these methods seek to characterize. The CT scans provide the basis for extracting geometry, meshing, and assigning inhomogeneous material properties. Setting up the loading and boundary conditions is a multi-step process that begins with defining a coordinate system. Wu et al., as representatives of the standardization and terminology

committee of the International Society of Biomechanics, proposed a standard femur-specific coordinate system with the origin placed at the centre of the femoral head in 2002⁵⁶. In their coordinate system, the x-axis points forward from the femoral head perpendicular to the Frontal plane, the y-axis points up normal to the Axial plane, and the z axis points lateral (right femur) normal to the Longitudinal plane.

The coordinates and the node closest to the centre of the femoral head must be identified in order to both place the origin and apply loading at the centre of the femoral head. From the geometry, the coordinates of the centre of the femoral head can be found by performing an analytical spherical fit to the surface model of the femoral head. One software application that can facilitate an analytical spherical fit is 3-Matic (Materialise, Leuven, Belgium). The node at the centre of the femoral head can be identified by first isolating the number and coordinates of the nodes in the femoral head and then using the distance equation to identify which of those nodes is closest to the coordinates of the centre of the femoral head from the analytical spherical fit. Creating the list of nodes and their coordinates in the femoral head can be accomplished in programs such as Mechanical APDL (Ansys Inc, PA, USA). Using the distance formula to identify the node with the closest coordinates to the coordinates of the centre of the femoral head can be automated in programs such as MATLAB (v.2018b, The MathWorks, Inc., Natick, MA, US).

Once the coordinate system is defined, the loading and boundary conditions are applied, and the simulation is run. A no-friction slider may be simulated by applying a constraint to the most lateral node on the y-z plane. Typically, the digital femoral geometry is cut off just above where the cement for the hinge would end in the experimental setup.

Thus, the fixed boundary conditions at the distal end of the femur are simulated by identifying the nodes in the elements on the distal end of the femoral model and constraining them against translation and rotation in all directions. A load may be applied at the centre of the femoral head at the same angle the load was applied during the experiments used to validate the numerical simulations. Once the CTPSFE analyses methods are validated in comparison to the results from cadaveric experiments, these methods can then be adapted for use in subject- or patient-specific finite element modelling.

4.3.5 Patient-specific numerical simulations of sideways fall loads

One advantage to CTPSFE analyses is the ability to run more than one simulation. For example, several previous studies showed that including multiple loading conditions in the CTPSFE analyses improves the stratification accuracy over aBMD for a cohort of female femoral fracture and control patients^{17,24,41}. Typical ranges of sideways fall loading angles include 0°(lateral)-30°(medial), 0°(posterior)-30°(anterior)^{17,24}; or 0°(lateral)-30°(medial) and 30°(posterior)-30°(anterior)⁴¹. Other advantages include the ability to examine the effects of differing boundary conditions⁴¹ and methodological determinants¹⁷.

Each patient cohort collected has a differing level of detail available to be included in the analyses and a differing variety of missing information that may need to be filled in to conduct the analyses. For example, patient CT scans included in femoral fracture and control cohorts may or may not include the full femur. The cohort studied by Falcinelli et al. included CT scans of the full femur²⁴. The cohort studied by Qasim et al. included CT

scans of the proximal femur^{17,41}. Opportunistically gathered patient cohorts include scans collected for other reasons, such as virtual colonoscopies. Due to the retrospective nature of these cohorts, each patient CT scan includes a different amount of the femur. An example of this kind of cohort is that studied by Winsor et al. where many of the CT scans end near patient's trochanter⁵⁷. Having coordinates for the knee centre is methodologically important for two reasons: (1) Qasim et al. have shown that using the knee centre to derive the coordinate system increases the stratification accuracy of CTPSFE analyses derived femoral strength¹⁷, and (2) Altai et al. have shown that applying boundary conditions at the knee centre instead of at the cut-off distal end below the proximal femur increases stratification accuracy⁴¹.

If a CT scan of the full femur is not available, statistical shape modelling may be applied to estimate full femur geometry. A statistical shape modelling software application, publicly available as MAPClient, uses a three-step process to create full femur estimates^{58,59} (Figure 4-8). First an iterative closest point algorithm⁶⁰ aligns via a rigid-body transformation the centre of the patient femoral mesh with an averaged statistical shape model of a proximal femur from a cadaveric database. The statistical shape model of the proximal femur from the cadaver database is replaced with the full femur statistical shape model of the same femur from the cadaver database. Next, the full femur statistical shape model is deformed along the principal components to create a fitted whole femur that accurately represents the patient proximal femur geometry and has a realistic overall shape. The patient estimate full femur statistical shape model can then be exported and used to generate a knee centre coordinate and a coordinate system (Figure 4-9). While this method has been demonstrated to improve the accuracy of existing methods for

CTPSFE derived femoral fracture risk prediction, an examination of the accuracy found surface errors of over 1.5 mm in the proximal region and concluded the reconstruction is likely unsuitable for patient specific finite element modelling⁶¹.

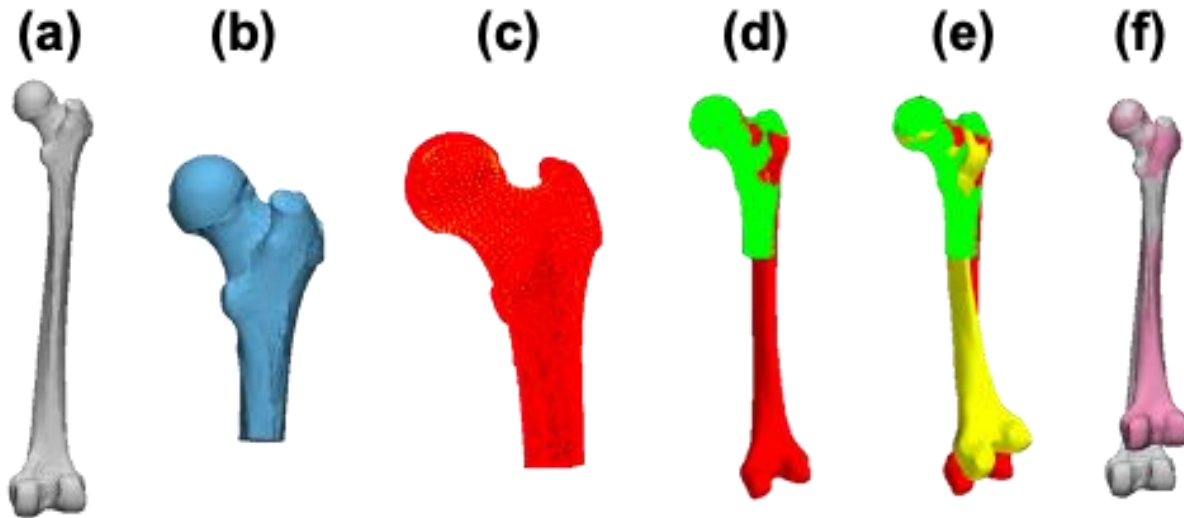


Figure 4-8: Representative images demonstrating the method for estimating the full femur using the statistical shape modelling software, MapClient. For demonstration purposes, (a) a cadaveric femur with data for the full femur (b) was cut to the top third of the femur, (c) the cadaveric specimen (yellow) aligned with the partial femur (red) from the MapClient cadaver database, (d) keeping the cadaveric specimen in place (green) the MapClient partial femur was replaced with the MapClient full femur (red), (e) principal component analysis was applied to estimate the full femur (yellow), and (f) the exported full femur estimate (pink) is compared to the original full cadaveric femur (grey).

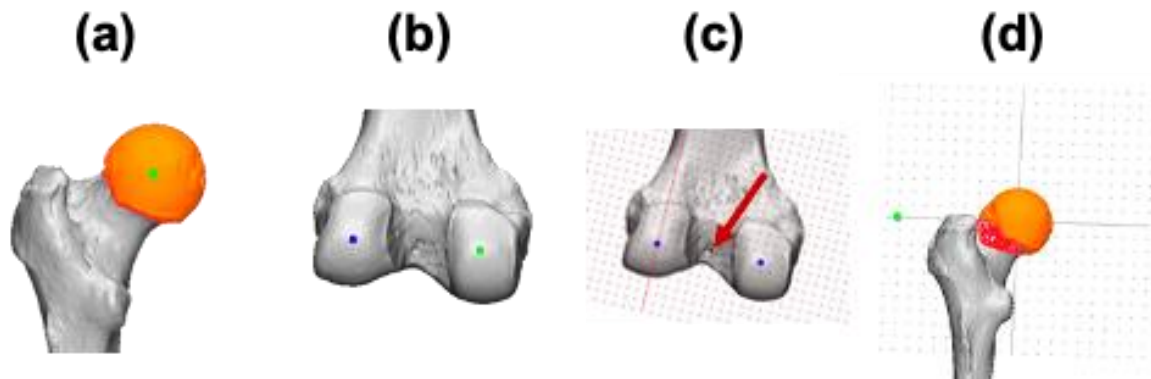


Figure 4-9: Representative image demonstrating the process of deriving a coordinate system using anatomical mapping. (a) The centre of the femoral head is identified using analytical spherical fitting. (b) The points where the condyles would touch the table if the femur were laid on the table are identified. (c) A plane is created using the centre of the femoral head and the two points on the condyles, a grid is overlaid, and the grid guides the identification of the midpoint between the two condyle points or the estimated knee centre. (d) A new plane is created using the centre of the femoral head, the estimated knee centre and one of the condyles. On this plane, a point in the lateral direction is identified to define the boundary conditions' coordinate system with the centre of the femoral head and the estimated knee centre.

4.4 Post-processing

This section describes CTPSFE simulation results post-processing. In FEA, post-processing follows a paradigm to transform highly detailed and complex outputs into a format both meaningful and easily understood by the user. In the context of femoral fracture prediction, the paradigm is related to what can be shown experimentally. This section describes one paradigm and the experiments it was based on.

4.4.1 Estimating the femoral fracture load and minimum fall strength

In addition to being a heterogeneous, anisotropic material, bone is also a load rate-dependent (viscoelastic) material⁶². These complexities complicate identifying an appropriate failure criterion for bone in general and specifically for the femoral bone. Continuing to place an emphasis on being able to validate the numerical model, Schileo

et al. investigated failure criteria to reproduce the elastic limit behaviour observed during their quasi-static, cadaveric experiments⁶³. Several aspects of their study supported the use of a strain-based failure criterion.

Schileo et al. proposed using a tensile strain elastic limit of 0.73% and a compressive strain elastic limit of 1.04%, based on experimental results from Bayraktar et al.^{63,64}. In their 2008 study, Schileo et al. compared three failure criteria: the maximum principal strain failure criterion, the von Mises stress failure criterion, and the maximum principal stress failure criterion. This study showed the maximum principal strain failure criterion defined a failure risk level more consistent with the experimental findings. Prior to Schileo et al. proposing the maximum principal strain failure criterion which models the elastic limit characteristics of bone⁶³, several studies contributed support towards using the maximum principal strain failure criterion. Chang et al. demonstrated that bone strength displays isotropy under mono-axial loading conditions⁶⁵. Keaveny et al. demonstrated femoral bone strength is invariant with respect to density^{45,47,66}. Two studies noted the tensile/compressive asymmetry of bone strength^{66,67}.

Schileo et al. implemented the maximum principal strain criterion using a three-step process⁶³. First each element is assigned a tensile or compressive predominance based on the absolute values of the first and third principal strain. Next, the corresponding tensile or compressive elastic limiting value is selected. Finally, the femoral failure risk factor (safety factor) is calculated as the strain limit over the maximum strain. If one static load is applied in the simulation, then there is one resulting femoral failure risk factor. In the case that multiple simulations are run for varying static loads^{17,24,41}, additional post-processing steps are required to determine the most critical load condition. For example,

Altai et al. found the minimum femoral failure risk factor across the loading conditions simulated and called this value the Minimum Fall Strength (MFS)⁴¹.

4.5 CTPSFE limitations

Understanding the methodological limitations of CTPSFE analyses is critical for their development, maintenance, and successful application. With such detailed and complex methodologies, limitations manifest in a variety of ways. Experimental limitations are critical to keep in mind due to the need for model validation. Examples of experimental limitations are: freezing and preparing cadaveric specimens for experiments; MTS machines are limited to applying loading in one direction while maintaining continuous contact throughout the test; and the inability to destructively test each specimen more than once. Experimental limitations will impact empirical relationships both those in the material mapping process and those in the previously proposed failure criterion. Morgan et al. demonstrated that the material mapping empirical relationships are specific to a variety of variables including anatomic site, and direction of loading⁴⁴. Similarly, the limitations on failure criterion related empirical relationships are one reason why there is not yet a universally accepted researcher consensus on a failure criterion for bone.

Outside of the need for model validation, there are also limitations arising from the clinical context. Patient data available for research are limited. Due to the radiation exposure, CT scans are not typically captured of healthy individuals. The individuals currently undergoing routine CT scans include patients at risk for cancer^{68–76}, subjects of clinical trials^{77–90}, and astronauts undergoing osteoporosis monitoring⁹¹. From an osteoporosis monitoring perspective, the lack of a large data set of patients, both healthy and affected, has so far prevented using FEA-based methods to diagnose osteoporosis

related fracture risk in the clinic in a parallel way to the current DXA diagnostic framework⁹². Additionally, the studies that have been done were conducted on limited cohorts with population specific details. The limited number of available patients makes it challenging both to meet power requirements for assessing particular research questions and to assess the ability of the methods to accurately diagnose a diverse group of patients.

When considering CT scans themselves, there are a variety of limitations. CT scans have multiple sources of potential error³⁶. Studies facilitated by the capture of CT scans frequently do not report relevant scanning acquisition and reconstruction parameters. This may be due to researchers not knowing certain scanning acquisition and reconstruction parameters were relevant at the time of their study. For example, the discussion of the relevance of reconstruction kernels only recently became a frequent topic in the literature^{12,13,93}. It is unclear from the literature both if the empirical material mapping relationships are specific to CT scan acquisition and reconstruction kernels or if the empirical material mapping relationships can be more generally applied to CT scan data regardless of acquisition and reconstruction parameters. Additionally, whether the CT scan data, were gathered prospectively or retrospectively impacts whether or not DCTC data specific to BMD may be available. Approaches to estimate CT calibration data specific to BMD will be discussed in greater detail later in this dissertation.

Software applications may temporarily impose limitations that can be overcome with time and continued development. Deep learning methods are continuing to develop and are increasingly applied to medical image segmentation. These methods were not yet at a point where they could be utilized to segment patient femurs for geriatric patients with

advanced stages of osteoporosis when the fracture cohort for this dissertation was being segmented.

4.6 CTPSFE shows promise for diagnosing patient femoral fracture risk

CTPSFE analyses continue to show promise for a variety of applications, including the assessment of femoral fragility fracture risk. Future developments to improve CTPSFE methods may come from a variety of imaging, experimental, or computational technical developments. CTPSFE methods are dependent on BMD specific DCTC data, which continues to be a challenge in this area of research. Prior to 2019, the discussion of phantomless or internal tissue-based DCTC in the literature was limited to proprietary methods. While contributing to developing phantomless or internal tissue-based DCTC methods, this dissertation initiates a discussion in the literature of design decisions relevant to the design of phantomless or internal-tissue based DCTC as a framework for the future development of site-specific methods.

5 Evaluation of patient tissue selection methods for deriving equivalent density calibration for femoral bone quantitative CT analyses

Please note Chapter 5 is already available as a published manuscript⁵⁷.

5.1 Introduction

Over 300,000 people experience an osteoporotic femoral fracture in the U.S. every year⁹⁴. Despite available treatments, osteoporosis remains underdiagnosed⁹⁵, inspiring research towards a better understanding of osteoporotic fracture. In addition, the stratification accuracy of the prognostic standard of care (bone densitometry) is too low to reliably diagnose osteopenic patients, and to decide when to adopt second-line treatments such as Denosumab or Teriparatide^{37,95}. This calls for more accurate prognostic methodologies. Various groups proposed quantitative computed tomography (QCT) based patient specific finite element analyses (FEAs) for improved osteoporotic hip fracture risk assessment^{7,17,96}. These FEAs have been shown to predict risk of hip fracture more accurately than areal bone mineral density (BMD)¹⁷. Retrospective reanalysis of patient computed tomography (CT) scans will further assist in the development of techniques to predict risk of osteoporotic fracture, potentially leading to improved prognostic accuracy. However, these models depend on the estimation of bone material properties, derived from CT X-ray attenuation. In phantom-based calibration, this is achieved by placing an inline calibration phantom under the patient or by scanning offline a calibration phantom immediately after the patient, using the same CT scan settings. Phantom-based calibration is the gold standard in the development of patient-specific

FEAs. However, scanning the patient with an inline phantom is not a standard clinical practice, and delayed offline retrospective calibration is not always possible due to clinics regularly purchasing new CT scanners. Phantomless CT scan calibration, derived from patient tissues, may be a feasible alternative.

Before considering literature on existing phantomless methods, several variables should be identified and defined. There are several points in the process of capturing a CT scan that affect density assessment including: underlying theory and definitions, the chemical composition of the object being scanned, the acquisition settings, and the reconstruction algorithms. Considering underlying theory, clinical CT images describe materials' X-ray attenuation in greyscale in terms of the Hounsfield Scale (in units HU),

$$CT\ Number = ((\mu_T - \mu_{water})/(\mu_{water} - \mu_{air})) * 1000\ [HU]. \quad (5.1)$$

Here μ , X-ray attenuation from the object, represents

$$\mu(E) = a_1 PE(E) + a_2 CS(E) = m_1 \mu_1(E) + m_2 \mu_2(E) \quad (5.2)$$

where E is the X-ray energy level, PE is the photoelectric basis function, CS is the Compton scattering effect basis function, and μ_1, μ_2 are any two independent materials². Compton scatter affects the definition of the Hounsfield scale such that X-ray attenuation measurements are roughly linearly proportional to density⁸. By definition this provides the basis for a linear estimate of the relationship between X-ray attenuation measurements and BMD⁸. CT Numbers are not numerically unique and thus a plastic-composite mimicking BMD results in a similar measurement to scanning actual bone. The variables can be simplified so that density can be calculated from X-ray attenuation measurements, by scanning a phantom of known chemical composition at a single energy. After initial X-ray attenuation measurements have been captured, reconstruction algorithms

generate an image of a specific density range with a particular anatomy of interest (i.e. soft tissue or bone). All of these variables impact the derivation of a conversion, between BMD and CT X-ray attenuation, that can be derived from CT X-ray attenuation measurements of a calibration phantom scanned in line with the patient^{97,98}. Recently, some studies have begun to discuss how specific details of clinical CT scan protocols affect density estimates by examining repeatability²⁶, patient positioning⁹⁹, and reconstruction kernel^{93,100,101}.

Different inline calibration phantoms have appeared in previous studies^{13,102–108}. These phantoms contain either calcium hydroxyapatite^{13,107,108} ($Ca_{10}(PO_4)_6(OH)_2$, abbreviated HA), or dipotassium phosphate^{102–106} (K_2HPO_4). When these phantoms are CT scanned, HA or K_2HPO_4 equivalent density is generally ρ_{QCT} for an inline phantom or ρ_{CT} for an offline phantom. The material specific abbreviations are ρ_{HA} or $\rho_{K_2HPO_4}$, respectively⁹⁸. Each phantom contains inserts with different known densities, such as 0, 50, 100 and 200 mg/cm³ of HA^{107,108}. After scanning the phantom and segmenting the density references, both a calibration factor and a calibration equation can be calculated. The calibration equation for a HA phantom can be calculated using a linear regression with CT Number [HU] on the y-axis and known density [mg/cm³] on the x-axis and then algebraically rearranging the equation to result in:

$$\rho_{HA} = (CT\ Number - b)/m \quad (5.3)$$

where m [HU/(mg/cm³)] and b [HU] are the slope and intercept, respectively, from the linear regression. When density-reference phantoms are used, the derivation of the calibration equation naturally characterizes and accounts for CT number variations due to factors including manufacturer, model and protocol¹⁰⁹. The use of stable, standardized

references in modern density phantoms can provide a comparison for analyses across clinics. However, in the case of an inline phantom that is externally located under the patient, the phantom will be subjected to patient-moderated spectra variable with patient composition, size, and geometric position¹¹⁰. While scanning an offline phantom removes this variation, this calibration method does not capture differences, such as those created by dosage-reducing variable current algorithms. Initially intended to create a standardized reference to characterize variations in CT number, differences in phantoms now introduce additional variances and limitations into the comparison of clinical assessment techniques. For example, Cann et al. used K_2HPO_4 in place of HA and argued K_2HPO_4 results in a slightly lower calibration slope than HA at equivalent densities^{6,15}, underestimating cortical bone density. They specifically pointed out that this difference is more pronounced at higher densities, visually demonstrated by Knowles et al.⁹⁸. Phantomless calibration, by definition, removes the variations created by scanning a phantom, retains the potential to create a scan-specific calibration equation, and increases accuracy over an inline phantom by using patient tissues as the density reference which are closer to the bone than a phantom could be placed.

To enable density assessment of patient scans where phantom-based calibration data were not captured, three approaches to phantomless calibration have been used in clinical research¹¹¹: (1) using CT Numbers [HU] directly^{69,112–114}; (2) using a calibration factor^{107,108,115,116}; and, (3) substituting tissues as a calibration reference^{15,16,32–38,18–23,26,31,32–38}. The first approach, using CT Numbers [HU] directly, is most accessible within the limitations of current clinical practice. Unfortunately, in order to be considered quantitative, the relevant BMD thresholds would have to be specific to each CT scanner

and protocol specific. Trying to derive relevant FEAs based thresholds in terms of CT Numbers poses challenges, such as requiring incalculable amounts of patient case studies. In the second approach, using a calibration factor, a general calibration factor (GCF) is calculated as the ratio of QCT-derived BMD divided by CT Numbers [HU] and then rearranged to extrapolate phantomless BMD through multiplying CT Number [HU] by GCF¹⁰⁷. While this approach is CT-scanner and -protocol specific, it is neither scan-specific nor precise enough for FEAs. The third approach, substituting tissues as calibration references is scan specific, and has been applied in FEAs of the femur^{13,90,92,103,118}. This method is limited by the assumption that internal patient tissues have the same density in every patient⁹². Previously, a variety of tissues served as the basis for deriving phantomless calibration: fat and muscle^{105,106,110,120–123}; air and blood^{90,103,118}; air and fat^{90,103,118}; air, fat, and muscle¹³; and air, fat, blood, muscle, and cortical bone^{13,92}. Many factors are known to influence the ability of CT Numbers [HU] to measure tissues: hydration levels¹⁰⁶, patient pathologies¹²⁴, heterogeneous distributions of muscle and fat¹⁰⁶, and IV contrast^{105,125}. Further, CT is unable to assess some pathologies known to affect CT Number, such as fatty atrophy of muscle¹²⁰.

While there is no standard method for determining which tissues to use as the basis for phantomless calibration, the literature provides some rationale for choosing specific tissues. Boden et al. showed that fat and muscle offer reliable internal reference standards for measuring vertebral bone density with QCT using tabulated reference densities from White^{110,126}. More recently, Michalski et al. used tabulated and standardised mass attenuation coefficients from the National Institute of Standards and Technology (NIST)^{28,92}. Some researchers have attempted to determine their own ground

truth values using a system-of-equations approach, finding: -69 mg/cm^3 for fat and 77 mg/cm^3 for muscle¹²¹; or -840 mg/cm^3 for air, -80 mg/cm^3 for fat, and 30 mg/cm^3 for muscle¹³. The limitation to deriving ground truth values, in lieu of using the standardized tables, is the unknown amount of pathological variation in the base cohort.

In the absence of phantom-based calibration data, computational researchers commonly estimate a linear relationship between a specific density and CT Number, based on available literature. Two such densities include ash density, ash mass divided by bulk sample volume, and apparent density, wet mass without marrow divided by bulk sample volume⁹⁸. Several studies are available where researchers empirically derived linear relationships between either ash density or apparent density measurements of bone and CT number^{27,43,127–132}. Ford et al. demonstrated a method for estimating a linear relationship between apparent density and CT Number for trabecular bone and cortical bone in mg/cm^3 ,

$$\rho_{app} = 1.106HU + 68.4, \quad (5.4)$$

before using the relationship in a computational study¹³³. Though not demonstrated in literature, another approach would be to estimate soft tissue density by estimating a theoretical calibration slope $CT_{theoretical} = 1.025 \frac{HU}{\frac{mg}{cm^3}}$ derived from theoretical air (1.205 mg/cm^3 , -1024 HU) and theoretical water (1000 mg/cm^3 , 0 HU). Both of these density estimation methods do not take into consideration CT scanner performance parameters or the anatomical area, as phantom-based or tissue-based phantomless calibration estimates do.

The method for deriving Young's modulus (E), a measurement of material stiffness, from CT data for use in patient-specific FEAs is sensitive to the relationship between a

specific density and CT Number due to a power-law relationship⁹⁸. CT data are converted to ρ_{ash} using equation (5.5) depending on the equivalent density, then to ρ_{app} using equation (5.6), and finally to Young's modulus using equation (5.7)⁴³.

$$\rho_{ash} = 0.8772 * \rho_{CT} + 0.07895 \quad (5.5)$$

$$\rho_{app} = 0.598 * \rho_{ash} \quad (5.6)$$

$$E = 6850 * \rho_{app}^{1.49} \quad (5.7)$$

In addition to being specific to the phantom's reference material, these relationships are also specific to anatomic site, in this case the femur⁴⁴. This suggests a need for a method flexible enough to consider anatomic site when selecting reference tissues for phantomless calibration.

The aim of this retrospective study was to demonstrate, verify, and validate a method for selecting patient tissues from which to derive density for use in femur strength prediction. Using the selected tissue combinations, we present a method for using phantomless calibration to estimate bone material properties for predictions of femoral fracture risk. Using a 2²-factorial design, we tested repeatability with and without theoretical data points and with and without including multiple scans for each patient. For verification, we compared patient-specific results against a custom offline CIRS BMD phantom and an inline Mindways Model 3 BMD calibration phantom. For validation, we compared patient-specific results against the inline Mindways Model 3 BMD phantom for the patients whose scans included the phantom.

5.2 Materials and methods

Patient scans were selected for a density-related sensitivity analysis from data gathered previously related to a cohort of 408 patients gathered at the University of Wisconsin—

Madison hospital. Scans from this cohort were previously identified to examine femoral fracture in an age-matched, case-control study. Full details of that study are available in Lee et al. 2017⁷⁴. Retrospective CT scan analysis was compliant with Health Insurance Portability and Accountability Act and approved by the UW-Madison Institutional Review Board (protocol number 2016-0168).

The pre-fracture cases analysed comprised 43 patients, with 26 female patients (ages 50–93 years) and 16 male patients (ages 56–95 years). The average time to fracture after CT scan was 1 year, with the minimum occurring the same year and the maximum occurring within 4 years. The control cases analysed consisted of 168 patients, with 108 female patients (ages 50–90 years) and 60 male patients (ages 50–91 years).

5.2.1 Method of selecting patient scans for analysis

Scans analysed were limited to those captured on a GE Lightspeed family CT scanner (Table 5-1). All scans analysed were captured at 120 kVp, and 1.25 mm slice thickness. The 258 scans analysed (Table 5-1) included images of 211 individual patients, both male and female (aged 50 to 95 years). Patients with surgical hardware were excluded from the study. Our goal in this selection was to cover a broad range of data so that the phantomless calibration to be broadly applicable, so we processed all data that met our inclusion criteria.

Table 5-1: The analysed cohort considered 211 distinct patients (bottom), with up to 258 scans (top). These scans were captured on nine different GE CT scanner models at the University of Wisconsin—Madison hospital.

GE CT scanner model	Number of scans	Number of patients with multiple scans	Pre-Fracture (Control)	Male (Female)
LightSpeed 16	45	14	4 (41)	13 (32)
LightSpeed Pro 16	22	5	9 (13)	7 (15)
LightSpeed Pro 32	1	0	1 (0)	1 (0)
LightSpeed Ultra	139	49	3 (136)	56 (83)
LightSpeed VCT	35	10	26 (9)	12 (23)
Discovery CT750 HD	8	4	5 (3)	3 (5)
Optima 580	1	1	1 (0)	1 (0)
Optima 660HD	6	4	2 (4)	1 (5)
Revolution GSI	1	1	0 (1)	1 (0)
Totals	258	88	52 (206)	95 (163)
LightSpeed 16	39	-	2 (37)	12 (27)
LightSpeed Pro 16	19	-	7 (12)	7 (12)
LightSpeed Pro 32	1	-	1 (0)	1 (0)
LightSpeed Ultra	108	-	3 (105)	41 (67)
LightSpeed VCT	30	-	23 (7)	12 (18)
Discovery CT750 HD	8	-	5 (3)	3 (5)
Optima 660HD	6	-	2 (4)	1 (5)
Totals	211	-	43 (168)	77 (134)

5.2.2 CT scanning protocol

Images were collected during routine abdominopelvic CT scans performed using 16- to 64-Multi-Detector CT scanners (LightSpeed Series, GE Healthcare). Hospital routine includes daily calibration scans on each machine to ensure the accuracy of the CT attenuation values. Standard scanning parameters for routine abdominopelvic CT scans are 120 kVp tube voltage, 1.25 mm slice thickness, 0.625 mm slice spacing, a medium or body type filter, a standard convolution kernel, and low doses of current, either static (50-100 mA) or modulated (noise index, 50; range 30-300 mA).

5.2.3 Inline quantitative equivalent density calibration using the Mindways Model

3 BMD calibration phantom

Eight out of the 408 patient scans included an inline effective K_2HPO_4 density calibration phantom (Model 3 phantom, Mindways Software, Inc., Austin, TX). Of those eight, three patients had existing surgical hardware and could not be analysed. Therefore, the analyses in this paper were limited to five patients. The calibration process for this phantom is described in detail by Mindways¹³⁴. Manual calculation of the calibration slopes for the five patients scanned with the inline calibration phantom was conducted (Table 5-2). A power analysis for a two-sample pooled t-test was conducted in MATLAB and the necessary sample size to meet 99% power ranged between 2 and 5 for the majority of the 40 phantomless slope combinations considered, with 3 outliers requiring a sample size of 8.

Table 5-2: BMD [mg/cm³] from equivalent density [mg/cm³] equations for the five patients scanned with the inline K₂HPO₄ Mindways Model 3 BMD calibration phantom.

Patient	CT Scanner	Calibration Slope [HU/(mg/cm ³)]	Calibration Y-Intercept [HU]	Conversion Slope [(mg/cm ³)/HU]	Conversion Y-Intercept [mg/cm ³]
1	LightSpeed VCT	1.03	5.59	0.97	-5.43
2	LightSpeed Pro 16	1.06	3.32	0.94	-3.12
3	LightSpeed Pro 16	1.05	13.54	0.95	-12.88
4	LightSpeed Pro 16	0.99	5.87	1.01	-5.93
5	Discovery CT750HD	1.06	-8.5	0.94	8.00
	Average values	1.04	3.96	0.96	-3.87

5.2.4 Offline equivalent density calibration using a custom BMD phantom

Retrospectively, we scanned offline a custom phantom with four HA density plugs at 100, 400, 1000 (part: 06217), and 1750 (part: 06221) mg/cm³ (CIRS Inc, Norfolk, VA) submerged in water. Scan settings were 120 kVp, 1.25 mm slice thickness, 0.625 mm slice spacing, 100 mA, and a standard reconstruction kernel on the Discovery 750HD. HA plug densities were selected to be representative of human femoral bone²⁶. Plugs were segmented by creating a virtual cylinder with a 10-pixel diameter across 10 slices in the centre of the plug using Mimics v. 21 (Materialise, Leuven, Belgium). Linear regressions were calculated for CT Number (HU) as a function of known density, ρ_{HA} . Resulting equations were:

$$CT\ Number\ [HU] = 1.100 * \rho_{HA} + 26.29. \quad (5.8)$$

$$(CT\ Number - 26.29)/1.100 = (1.100 * \rho_{HA} + 26.29 - 26.29)/1.100 \quad (5.9)$$

$$\rho_{CT} = 0.9091 * HU - 23.90 \quad (5.10)$$

5.2.5 Identify most consistent reference densities across patients

We analysed phantomless calibration on 258 scans and considered five patients' nominal density references, including adipose tissue, aortic blood, skeletal muscle, urine, and air. Tissue segmentations were captured as virtual cylinders, with a diameter of 10 pixels and a depth of 10 slices, using Mimics v21.0 (Materialise, Leuven, Belgium). Due to the small size of the femoral artery, the virtual cylinder captured was reduced to a diameter of 8 pixels. For consistency, all virtual cylinders were created such that the centre of the virtual cylinder was around the same axial slices as the centre of the femoral head. An example of the virtual cylinder placement is shown in Figure 5-1. Quality checks were conducted to ensure each virtual cylinder contained a volume of at least 100 voxels (ASTM E1935 2019). We were unable to segment urine in the patient's bladder for 167 out of the 258 scans due to empty bladders. Blood was also difficult to segment due to their small sizes, resulting in measured values outside of 40 ± 20 HU for 46/258 left patient arteries and 47/258 right patient arteries. Table 5-3 shows the nominal density values assumed for the linear regression of HU and tissue density²⁸. The 258 patients included in this study were segmented by a single operator. To assess the precision of results at the segmentation, BMD and FEAs levels, the five patients with inline phantoms were also segmented by three different operators. Inter- and intra-operator repeatability were calculated¹³⁵.

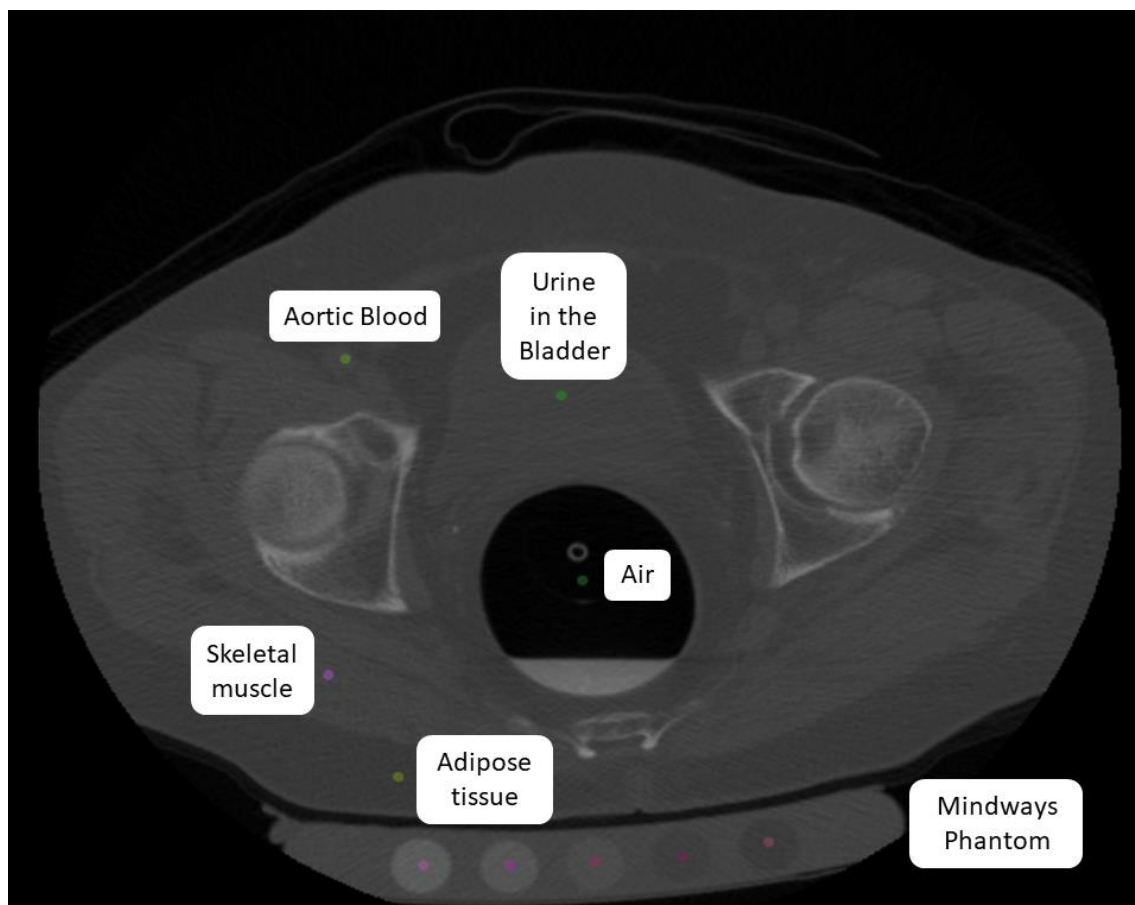


Figure 5-1: Representative axial slice of a CT scan of an 85-year-old patient at the proximal femur level. The patient is lying supine. A Mindways Model 3 BMD calibration phantom is visible underneath the patient.

Table 5-3: Nominal density values from NIST²⁸

Reference Material	Nominal Density [mg/cm ³]
Theoretical Air (not segmented)	1.205
Adipose Tissue	950
Aortic Blood	1060
Theoretical Water (not segmented)	1000
Skeletal Muscle	1050

Each patient had up to nine potential data points that could be used for line fitting: theoretical air, segmented air, adipose tissue (right and left), aortic blood (right and left), skeletal muscle (right and left), and theoretical water. Any combination of at least two and

up to nine data points could be used to derive a linear regression for the HU versus nominal density relationship, 502 possible combinations for each of the 258 scans. A custom MATLAB (v.2018b, The MathWorks, Inc., Natick, MA, US) script was developed to: (1) calculate all possible linear regressions, (2) discard all ill-conditioned calibration slope results, and (3) conduct a numerical analysis to sort density combination calibration slope results across patients. Ill-conditioned calibration slopes occurred when the algorithm fit a line with two values for the same tissue (i.e. right and left adipose). Sorting was accomplished by minimizing the sum of the squared error between the density calibration slope and a theoretical calibration slope, as in equation (5.11) below:

$$\sum_1^n (m - 1.025)^2. \quad (5.11)$$

Recall from the introduction that the theoretical calibration slope [1.025 HU/(mg/cm³)] is derived from theoretical air (1.205 mg/cm³, -1024 HU) and theoretical water (1000 mg/cm³, 0 HU). After discarding over-constrained combinations, the best 10 combinations and the worst combination were identified for further analysis.

5.2.6 Experimental design to test repeatability of tissue identification

Patient tissue segmentations were organized to form two groups: “Scans” included all scans eligible for processing for all patients, and “Patients” included only one scan for each patient. To form the Patients group, results from duplicate scans for patients were removed, such that the results for CT scanners with fewer patient scans were kept, except in the case of the Optima 580 and Revolution GSI, each of which only had one patient scan. A 2² factorial designed experiment was conducted by running the MATLAB script used to identify the most consistent reference densities across patient populations, with

two levels for each group including and excluding values for theoretical air and water in the combinatorial analysis.

5.2.7 Finite element model BMD and Femur Strength

Five finite element models were developed for each patient to investigate the impact of different calibration equations on BMD and femoral strength (FS) calculations model I: patient specific inline K_2HPO_4 calibration; model II: the average of the patient specific inline K_2HPO_4 calibrations; model III: the offline HA calibration; model IV: phantomless calibration derived from air, aortic blood, and skeletal muscle (AABSM); and, model V: phantomless calibration derived from air and adipose (AA). One femur was segmented for each patient: four were segmented in Mimics v19.0 or 21 (Materialise, Leuven, Belgium) and one was segmented in ITK-Snap (ITK-Snap 3.6.0, University of Pennsylvania). Each geometry was discretized into ten-node tetrahedral elements using ICEM CFD 16.2 (ICEM CFD 16.2, Ansys Inc., PA, USA) with a maximum edge length of 3 mm based on a previous mesh convergence study¹³⁶. Note that each patient had the same mesh for all models.

Elastic moduli were mapped onto the meshed bone using the equations described in the introduction and Bonemat (V3.2, Istituto Ortopedico Rizzoli, Bologna, Italy). BMD was calculated for each model as the summation across groups of the density in each material group, multiplied by the number of elements with that material group. Femur strength was calculated using a sideways fall loading scenario with fixed boundary constraints at the estimated knee centre and a simulated planar bearing at the lateral coordinate on the trochanter^{17,41}. A concentrated point load, 1000 N, was applied to the centre of the femoral head in thirty-three different force directions from -30° to 30°

(posteriorly to anteriorly directed) in the transverse plane and 0° to 30° (x-axis to medially directed) in the frontal plane⁴¹. FEAs strain results were post-analysed using a maximum principal strain failure criterion, with limiting values at 0.73% for tensile and 1.04% for compressive strains as previously defined by Bayraktar et al.⁶⁴. FS was defined as the minimum force (N) at failure across all 33 side-fall loading conditions. All FEAs were conducted in ANSYS 16.2 (Ansys Inc, PA, USA).

5.2.8 Statistical analysis

The mean and standard deviation were calculated for patient tissue segmentation measurement results in HU for both the “Scans” and “Patients” groups. Once patient specific density calibration slopes were calculated, statistical measurements were mean, standard deviation, and 95% confidence interval. Bland-Altman analyses were conducted for the five patients with inline Mindways Model 3 BMD phantoms included in their scans. The hypotheses that no statistically significant difference exists between calibration methods were tested using a students’ t-test ($\alpha = 0.01$) for the calibration slopes, the calibration intercepts, BMD, and FS. Normality was tested using the Shapiro-Wilk test in IBM SPSS Statistics for Windows, version 26 (IBM Corp., Armonk, N.Y., USA); however, all other statistics were calculated in MATLAB 2018b.

5.3 Results

Phantomless calibration was valid when compared against inline phantom calibration for FS, BMD, calibration equation (Figure 5-2, Figure 5-3, Figure 5-4). The algorithm produced calibration equation results consistent with those from inline phantom calibration (Figure 5-5). Intra- and inter-operator repeatability found the method highly

repeatable for FS, BMD, and calibration equations (Table 5-4). Adipose was the most repeatable tissue segmented, and the bladder was the least repeatable (Table 5-4).

The AABSM combination produced the best slope result for 3 of the 4 categories in the 2² factorial designed experiment. The 4th category, excluding multiple scans per patient and theoretical air and water, found the AA combination produced the best slope. The first category, including theoretical air and theoretical water for all scans (n = 258), found AABSM scan specific slope values [HU/(mg/cm³)] of mean $\pm$ std dev (lower – upper) = 1.021 ± 0.006 (1.008 – 1.034) and found measured air and theoretical water produced the worst combination, with slope values of 1.379 ± 6.185 (-10.99 – 13.75). The second category, including theoretical air and theoretical water for 1 scan per patient (n = 211), found AABSM scan specific slope values of mean $\pm$ std dev (lower – upper) = 1.021 ± 0.006 (1.009 – 1.034) for the best combination and found measured air and theoretical water produced the worst combination, with slope values of 1.468 ± 6.856 (-12.24 – 15.18). The third category, excluding theoretical air and theoretical water for all scans (n = 258), found AABSM scan specific slope values of mean $\pm$ std dev (lower – upper) = 1.017 ± 0.010 (0.998 – 1.037) for the best combination and found aortic blood and skeletal muscle produced the worst result, with values of 0.893 ± 2.151 (-3.458 – 5.195). The final category, excluding theoretical air and water for 1 scan per patient (n = 211), found AA scan specific slope values of mean $\pm$ std dev (lower – upper) = 0.975 ± 0.010 (0.956 – 0.994) and found aortic blood and skeletal muscle produced the worst result, with values of 0.839 ± 2.149 (-3.458 – 5.137).

For FS results, the AABSM calibration resulted in a 6.9% bias over scan specific inline calibration, a 7.3% bias over averaged inline calibration, and a 22% bias over offline

calibration; and the AA calibration resulted in a 9.9% bias over scan specific inline calibration, a 10% bias over averaged inline calibration, and a 25% bias over offline calibration (Figure 5-2). For BMD results, the AABSM calibration resulted in a 3.9% bias over scan specific inline calibration, a 3.7% bias over averaged inline calibration, and a 17% bias over offline calibration; and the AA calibration resulted in a 6.1% bias over scan specific inline calibration, a 6.0% bias over averaged inline calibration, and a 19% bias over offline calibration (Figure 5-3). When considering the calibration slopes directly, the AABSM and AA combinations resulted in biases of 2.6% and 6.3% over scan specific inline calibration, respectively (Figure 5-4). For the calibration intercepts, the AABSM and AA combinations resulted in biases of 110% and 110% over scan specific inline calibration, respectively (Figure 5-4). When comparing scan specific results for all 211 patient scans against the scan specific inline calibration, the ten best AABSM slope combinations all resulted in the majority of patients falling within the range demonstrated by the inline calibration (Figure 5-5). The three best AA slope combinations did not fall within the range demonstrated by the inline calibration; however, the inter-quartile range for the next seven best did fall within the range demonstrated by the inline calibration (Figure 5-5). All intercepts for the ten best combinations for both AABSM and AA fell within the range demonstrated by the inline calibration (Figure 5-5). Biases for the best ten tissue combination results for all four categories, compared with the inline calibration slope, were found to be less than or equal to $0.068 \pm 0.064 \text{ HU}/(\text{mg}/\text{cm}^3)$ for the five patients with inline calibration available. The resulting 40 calibration slopes and the scan specific inline calibration slopes were found to be normally distributed using a Shapiro-Wilk test.

Differences in FS between calibration methods were only statistically significant for AABSM versus the average of the inline calibrations ($p < 0.01$). Differences in BMD between calibration methods were not statistically significant for either phantomless calibration combination (AABSM and AA) and the inline phantom calibration ($p = 0.03, 0.10$). However, differences in BMD between calibration methods were statistically significant for both phantomless calibration combinations (AABSM and AA) versus the average of the inline calibrations ($p = 0.003, 0.002$) and the offline phantom ($p = 0.004, 0.003$). Differences in calibration equation followed the same trend. For the slopes, differences were not statistically significant between either phantomless calibration combination (AABSM and AA) and the inline phantom ($p = 0.04, 0.17$). Conversely, differences were statistically significant between both phantomless calibration combinations (AABSM and AA) versus the average of the inline calibrations ($p < 0.001, 0.001$) and the offline phantom ($p < 0.001, 0.001$). For the intercepts, differences were not statistically significant between either phantomless calibration combination (AABSM and AA) and the inline phantom ($p = 0.08, 0.26$). Continuing with the trend, differences were statistically significant between both phantomless calibration combinations (AABSM and AA) versus the average of the inline calibrations ($p < 0.001, 0.001$) and the offline phantom ($p < 0.001, 0.001$).

Both average intra-operator and inter-operator repeatability were better for AABSM than for AA when analysing FS, BMD, or calibration equation (Table 4). Segmentation CT Number [HU] results found similar means and standard deviations for tissues compared between the “all scans” and “one scan per patient” categories, respectively: adipose -

99.12 ± 9.44 and -98.98 ± 9.62 ; aortic blood 52.42 ± 17.28 and 51.83 ± 17.41 ; and muscle 43.58 ± 13.42 and 44.01 ± 14.05 .

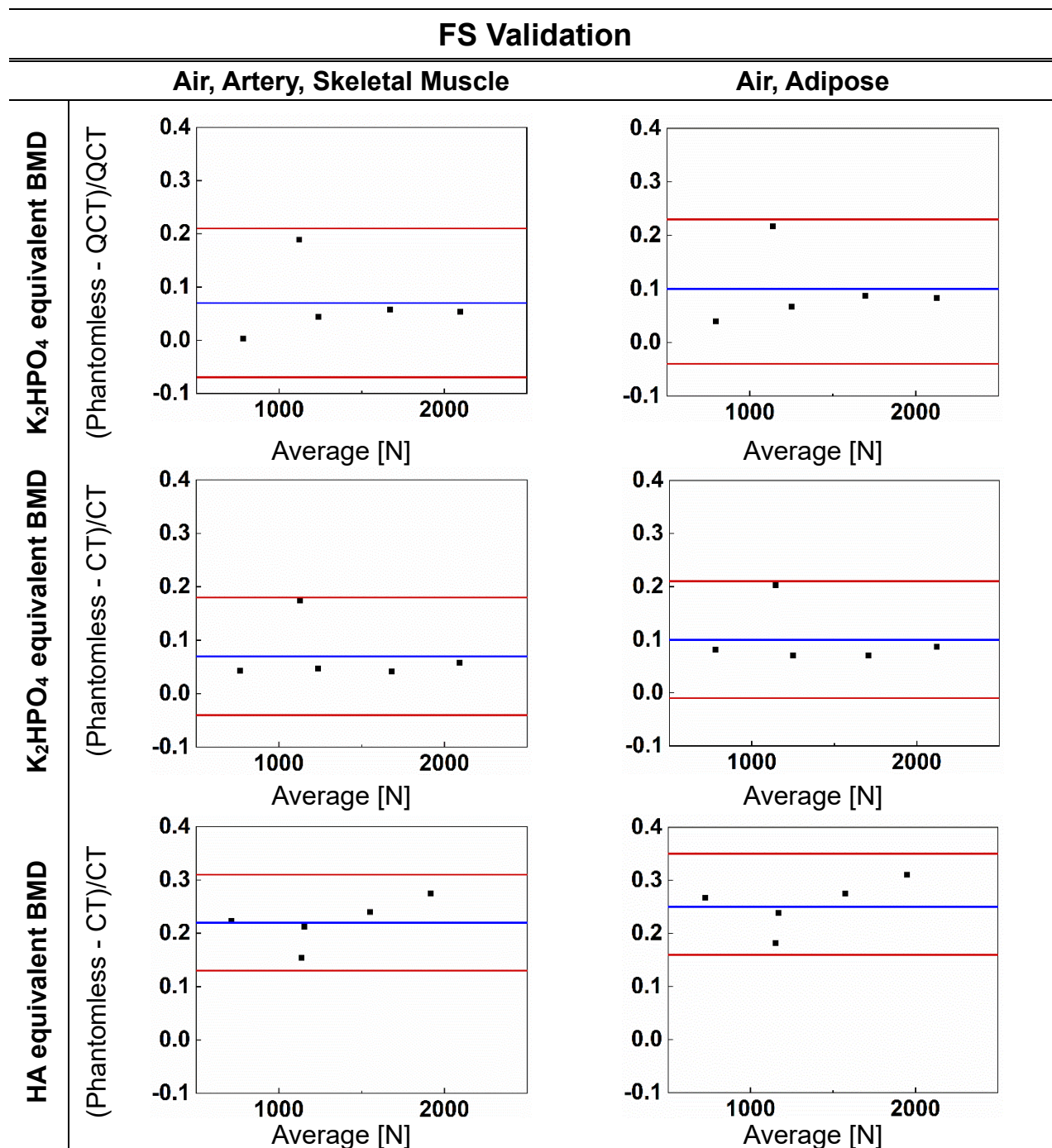


Figure 5-2: FS results derived using phantomless calibration displayed the least bias when compared against results derived using the average of the patient and scan-specific K₂HPO₄ calibration as shown by Bland-Altman analyses. Overall results using phantomless calibration were more consistent with results from the K₂HPO₄ phantom than the HA phantom. The blue lines are the means and the red lines are the 95% confidence interval.

BMD Validation

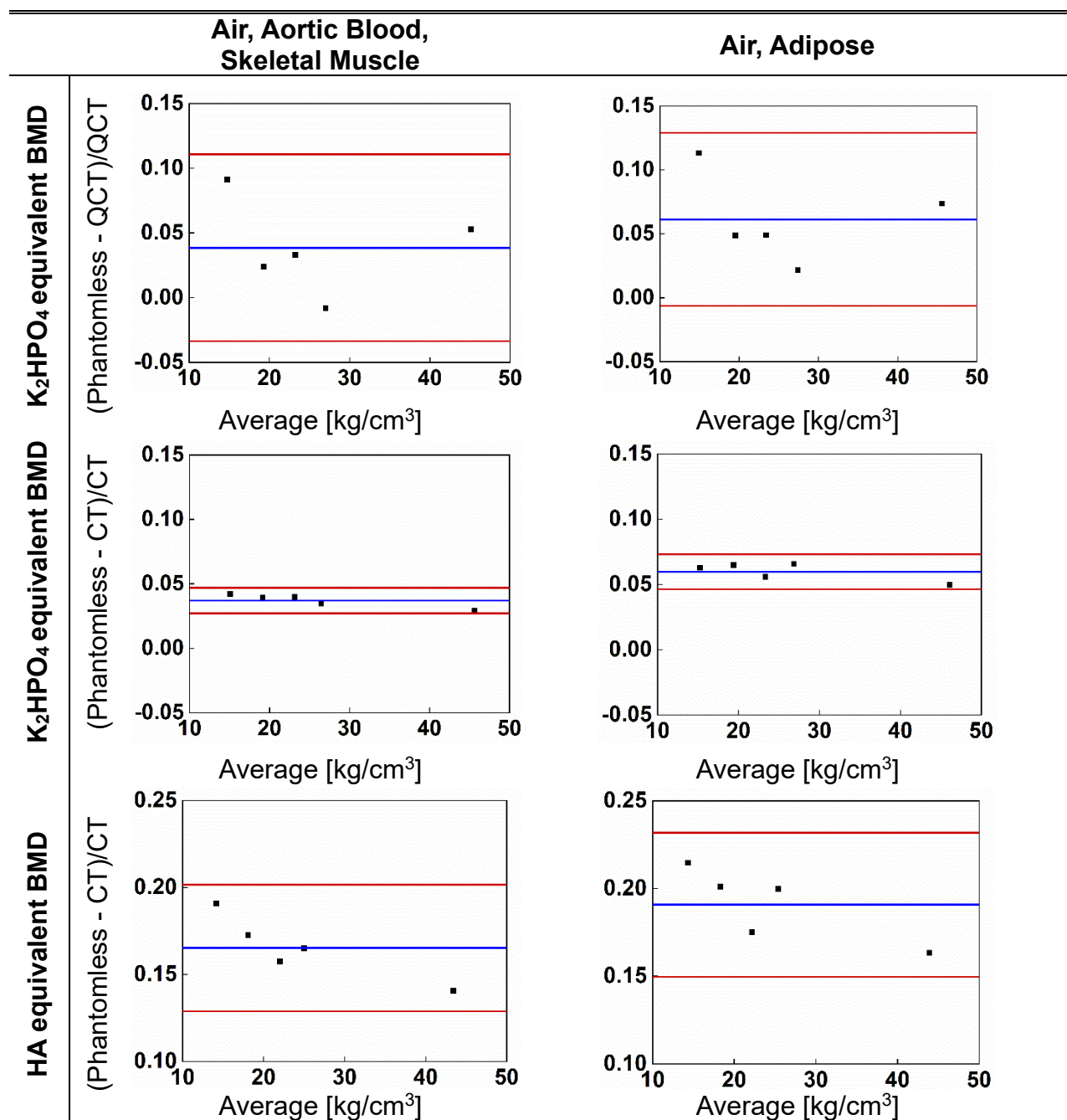


Figure 5-3: BMD results derived using phantomless calibration displayed the least bias when compared against results derived using the average of the patient and scan specific K₂HPO₄ calibration as shown by Bland-Altman analyses. Overall results using phantomless calibration were more consistent with results from the K₂HPO₄ phantom than the HA phantom. The blue lines are the means and the red lines are the 95% confidence interval.

Calibration Equation Validation

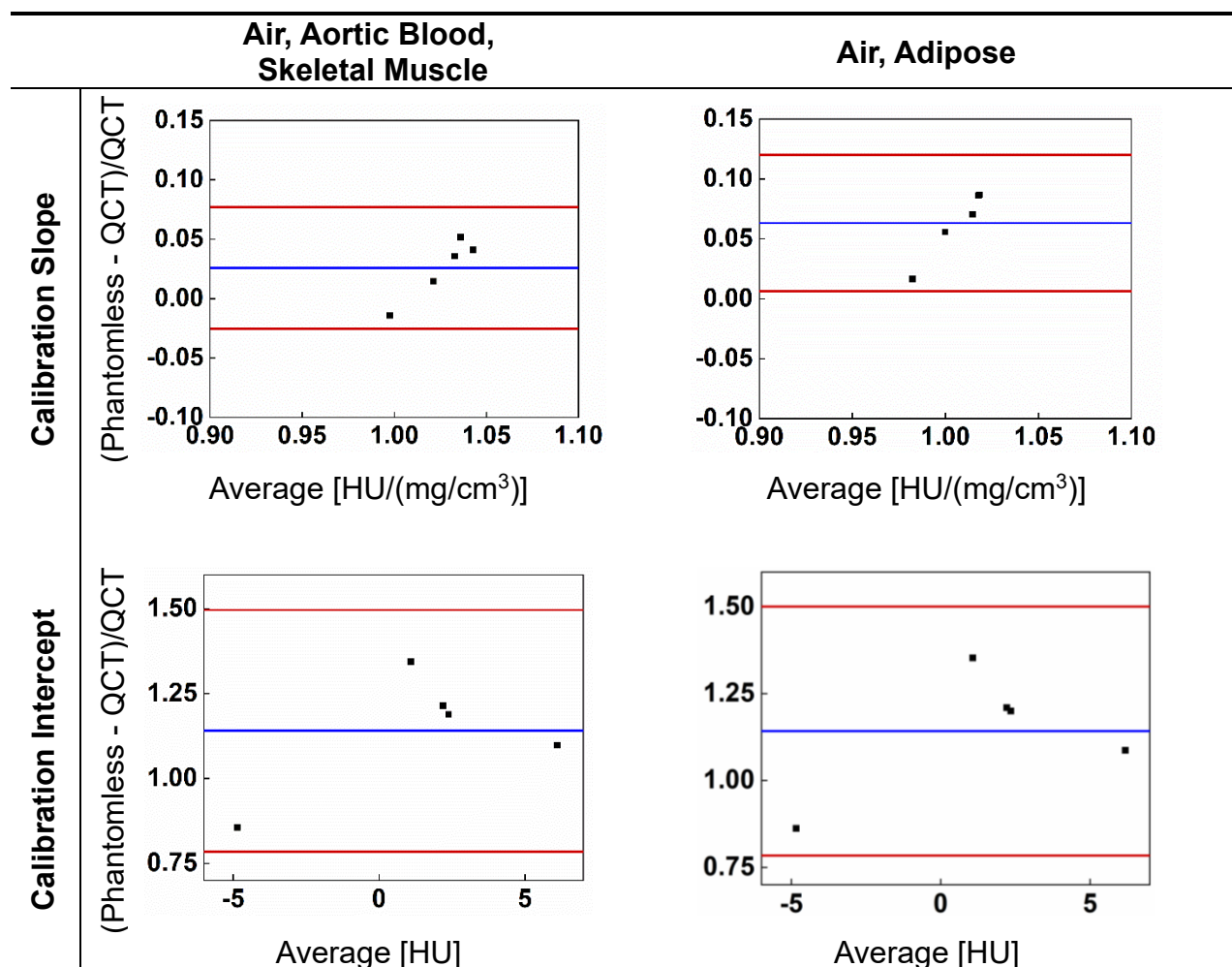


Figure 5-4: Phantomless calibration slopes derived from air, aortic blood and skeletal muscle segmentations displayed less bias than those derived from air and adipose when compared with patient and scan specific K_2HPO_4 calibration as shown by Bland-Altman analyses. While both sets of phantomless calibration intercepts displayed similar and large bias, all averages were within the performance expectations for a GE CT scanner (0 ± 7 HU)¹⁰⁹. The blue lines are the mean and the red lines are the 95% confidence interval.

Calibration Equation Verification

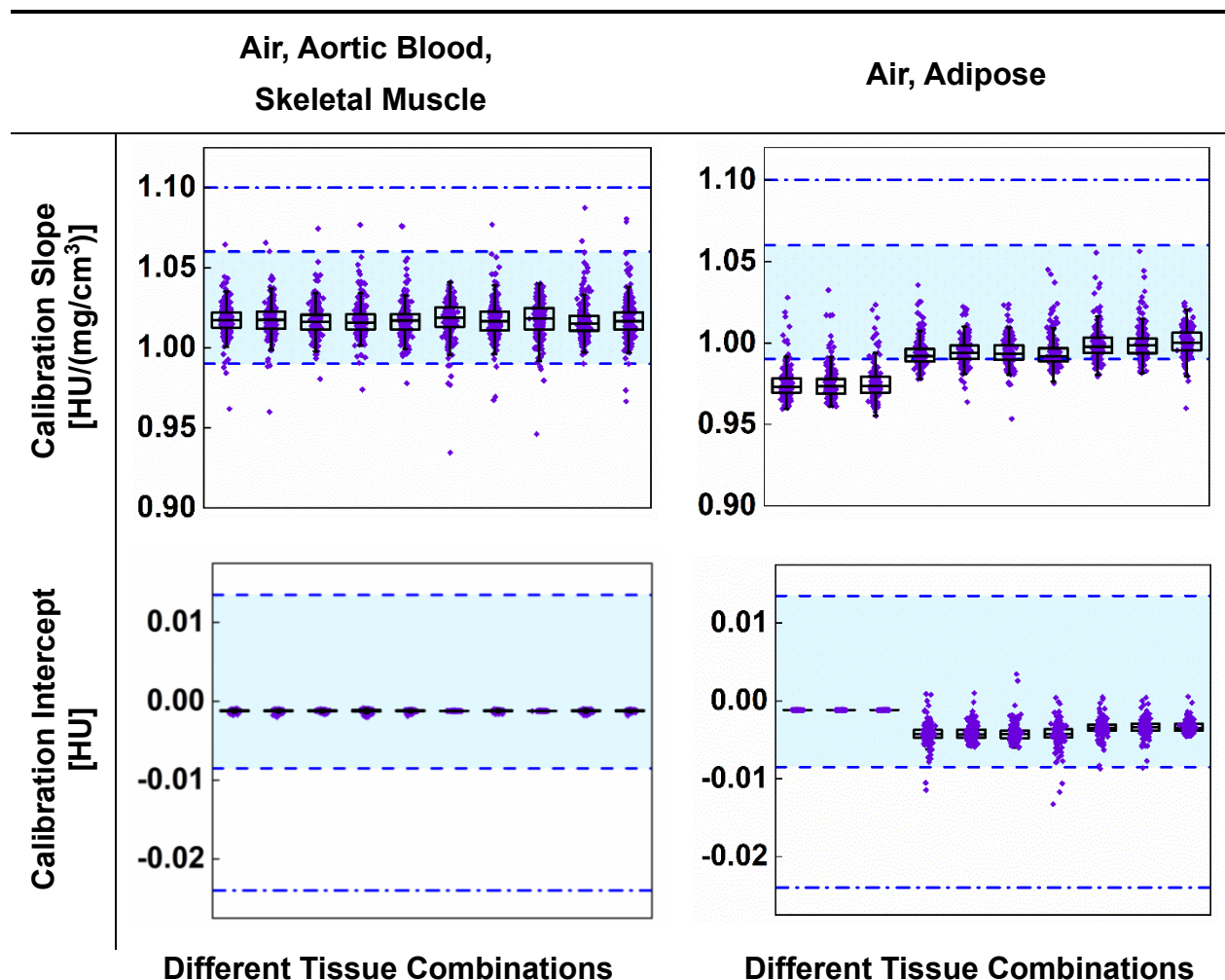


Figure 5-5: These plots compare the ten best combinations of tissues in terms of calibration slopes and intercepts. Boxplots are overlaid on scatter plots of the patient specific calibration slopes and intercepts (purple points). For the slopes plots, the three blue lines include: the dashed lines for the minimum (0.99 HU/(mg/cm³)) and maximum (1.06 HU/(mg/cm³)) slopes across patients from the K₂HPO₄ calibration phantom, and the dash-dot line is for the calibration slope for the custom phantom scanned offline in water (1.10 HU/(mg/cm³)). All slopes are in HU/(mg/cm³). For the intercept plots, the three blue lines include: the dashed lines for the minimum (-0.0085 HU) and maximum (0.0135 HU) patient specific results for the K₂HPO₄ phantom, and the dash-dot line is calibration intercept for the custom phantom scanned in water (-0.0239 HU).

Table 5-4: Intra- and Inter-operator reanalysis precision error (root-mean-square) for FS, BMD, calibration equation, and tissue segmentations at the femur for $n = 5$. Coefficients of variation (CV_{RMS} , in %) and standard deviations (SD_{RMS} , in absolute units) are presented.

Measurement	Intra-operator		Inter-operator	
	CV_{RMS} (%)	SD_{RMS}	CV_{RMS} (%)	SD_{RMS}
Femoral Strength [N]				
AABSM	0.28	9.67	0.42	13.27
AA	1.27	63.85	4.14	224.44
Inline	0.37	11.68		
Bone Mineral Density [kg/cm³]				
AABSM	0.22	0.13	0.39	0.20
AA	0.65	0.48	1.52	1.28
Inline	0.57	0.57		
Slope [HU/(mg/cm³)]				
AABSM	0.22	0.00	0.41	0.01
AA	0.41	0.01	1.01	0.02
Inline	0.60	0.01		
Intercept [HU]				
AABSM	2.70	0.08	4.56	0.14
AA	6.20	7.45	17.03	20.12
Inline	9.31	2.22		
Segment [HU]				
Adipose	1.68	3.77	1.78	3.98
Air	0.35	7.44	0.93	20.10
Aortic Blood	10.09	10.67	13.93	15.99
Skeletal Muscle	10.25	7.19	11.24	8.12
Bladder	50.82	3.67	68.11	12.21

5.4 Discussion

The main aim of this study was to demonstrate, verify, and validate a method for selecting basis patient tissues for deriving an equivalent density equation in femoral bone QCT analyses. As an example, this method identified AABSM as the best combination of tissues for phantomless calibration. This method was shown to be valid for FS, BMD, and calibration equation results. The validity of phantomless calibration for FEAs of the femur is consistent with other studies^{92,103,118}. To verify this method, results for 258 scans were

shown to be within the range of those from the inline calibration of five scans. This method shows promise for use in the retrospective analysis of patient cohorts without available calibration data and can be applied opportunistically to any CT scan.

This study differs from previous studies in several ways, including different CT scanners, CT scan protocols, tissues used as the basis for phantomless calibration, assumed tissue densities, methods of segmentation, and FEA pipelines. Focusing in on which tissues are used as the basis for phantomless calibration, this study's selection of the AABSM combination of tissues is different from prior combinations in literature for FEAs of the femur, including: fat and muscle¹²⁰; air and fat^{90,103,118}; air, fat, and muscle¹³; and air, fat, blood, muscle, and cortical bone^{13,92}. The variety of different combinations shows the need for a universally accessible objective method, such as that presented in this study, for identifying the best tissues for use as the basis for phantomless calibration within the existing constraints of CT scanners and CT scan protocols for the application specific anatomic site. Algorithms for decision making, such as that presented in the current study, can be more robust than correlation approaches, such as those presented by Eggermont et al.¹³.

Despite the differences in FEA pipelines, the bias introduced by phantomless calibration is comparable across studies, with all other variables held constant within the respective studies. This study's calculated FS mean absolute difference, 90 N (6.9%), was similar to recent studies, such as Lee et al. 30 N (0.8%)¹⁰³, and Michalski et al. -40 N (17%)³⁹. The calculated BMD biases 0.92 kg/cm³ (0.04%) were larger than a recent study on a more developed method presented by Lee et al., 2 mg/cm³ (0.9%)¹⁰³. Note that differences observed in FS measurements were expected to be greater than

differences observed in BMD measurements for two reasons. First, differences that appear small when examining preliminary results (i.e. segmentation, calibration equation, BMD) are amplified by the power-law component of the density-elastic modulus relationship (equation (5.7)), making FEAs sensitive to changes in the calibration equation. Second, the sideways fall load case is more sensitive to changes in mechanical properties of materials due to the stress gradient from bending in the combined-loading. Both the results of this study and the results from literature show greater differences in FS biases than BMD biases. From a clinical perspective, this drives the reasonable assumption that variables known to affect CT Number [HU] or BMD measurement would have an amplified effect on FS.

Recent studies have proposed the use of QCT derived FEAs for improved osteoporotic hip fracture risk prediction^{7,17,96} and the use of phantomless calibration in this context^{92,103}. Limited studies have been conducted to identify and quantify the impact of relevant factors. Michalski et al., who conducted part of their analysis on ten full body cadavers, iteratively correlated ROI specific CT Numbers across energy levels, setting the example of taking these factors into account during the development of their phantomless calibration method⁹². Several authors have noted the improvements in phantomless calibration results, due to the decreased distance between the patient and the reference^{103,106,110}. The current study controlled for some factors known to create variations in CT Number [HU] by limiting data analysed to scans captured on GE LightSpeed CT Scanners with 120 kVp, variable mA, slice thickness of 1.25 mm, slice increments of 0.625 mm, and a standard reconstruction kernel. Lee et al. used similar inclusion criteria, identifying 120 kVp and a standard reconstruction kernel as the most

important imaging technique factors and their decision to analyse a single protocol as a limitation¹⁰³. Although attempting to work with a standardized protocol, Eggermont et al. found that a small number of their patients were scanned with a different reconstruction kernel, allowing them to make relevant observations (1) changing reconstruction kernel had no significant effect on phantom-based or air-fat-muscle calibration, and (2) changing reconstruction kernel resulted in significantly higher failure loads when using their non-patient specific calibration¹³. Michalski et al. observed that by using consistent imaging acquisition and a single imaging protocol, there were fewer confounding variables when measuring methodological precision⁹². Beyond the limitation of only considering one clinical protocol, this study was also limited to pre-fracture cases that went on to experience femoral fragility fracture.

The current study's segmentation method may be less repeatable than the segmentation methods presented in other studies. Where this study conducted manual segmentation using the mean CT Number [HU] over the digital volume, other studies used higher fidelity segmentation methods. Examples relevant to multiple studies include: Lee et al., who have automated their segmentation using gradient-profile algorithms independent of absolute attenuation^{90,103,118}, or the popular histogram and peak fitting approach^{13,92,106,110,122}. Boden et al. designed the histogram and peak fitting approach specifically to overcome the challenge of reliably locating a conventional ROI to calculate the mean CT Number [HU] of the digital volume¹¹⁰. This implies that methods using this approach would naturally account for the heterogeneity included in patient tissues and improving the precision of phantomless calibration. The differences in segmentation methods are a major reason why this method was less repeatable than those presented

previously in literature (Table 5-5). This comparison shows using a higher fidelity segmentation method may improve the repeatability of the current study's phantomless calibration method.

Table 5-5: Comparison of precision errors between the current study and literature.

Measurement	Intra-operator		Inter-operator	
	CV _{RMS} (%)	SD _{RMS}	CV _{RMS} (%)	SD _{RMS}
Femoral Strength				
Lee ¹⁰³ [N]			0.4	20
Michalski ⁹² [N]			6.0	84
Current Study [N]	0.28	9.67	0.42	13
Bone Mineral Density				
Lee ¹⁰³ [mg/cm ³]			< 0.3	≤ 1
Michalski ⁹² [mg/cm ³]	4.3	12	5.3	11
Current Study [kg/cm ³]	0.22	0.13	0.39	0.20

This study showed phantomless calibration results were close to results derived from the Mindways Model 3 BMD inline phantom, which relies on K₂HPO₄ as a reference material. Further, the phantomless calibration derived results were not significantly different from the inline calibration derived results and were significantly different from both the averaged inline calibration and the offline calibration. Both the inline phantom, which ranges from -53.4 to 375.8 of equivalent K₂HPO₄, and this phantomless calibration technique, require extrapolation in order to define *in vivo* BMD¹³⁴. The potential for extrapolation errors has been raised as a concern in several studies^{13,106,122}. In their phantomless study, Lee et al. demonstrated their method to calibrate CT scans was equivalent to traditional phantom-based calibration¹⁰³. If assumptions are made about the density of bone and included when deriving phantomless calibration, the results become less accurate, as shown by the correlation analysis in the pilot study written by Eggermont et al.¹³.

There were several limitations to this study. CT scans of the proximal femur region include a limited choice of tissues to segment: adipose tissue, skeletal muscle, aortic

blood, and in some cases the bladder is empty. In addition to population variance across patients, tissues also depend on a variety of patient specific variables such as: hydration level¹⁰⁶, patient pathologies¹²⁴, heterogenous distributions of muscle and fat¹⁰⁶, i.v. contrast^{105,125}, exercise habits, and body mass index. The cohort studied here did not include patient details about exercise habits, body mass index or comorbidities. Future studies should consider a more detailed examination of factors known to cause variance across patients and a larger sample size to further develop the phantomless calibration methodology. In this study, GE LightSpeed family CT scanners were used to demonstrate the calibration process. CT scanners from other manufacturers were not analysed due to lack of available data. Future work should consider a multi-centre study comparing the same model of CT scanner across different hospitals and consider CT scanners from other manufacturers. Also of note was the small sample size of available calibration curves for comparison.

This study did not examine several potential confounding variables. When reassigning pre-fracture/control pairings, researchers were not blind to CTXA, a method for measuring areal BMD from CT data, mathematically equivalent to dual-energy X-ray absorptiometry, density measurements. Stratification accuracy between pre-fracture and control cases when using phantomless calibration was not examined. Additional confounding factors may have been present such as: other diseases, routine exercise habits, differences in body-mass index/height/weight, comorbidities, or different pathologies. These were not considered due to lack of readily available cohort information. Several of these variables could be considered in a prospective study or in a reanalysis of retrospective data prospectively gathered. A more systemic method of

randomly assigning controls to pre-fracture cases could be developed and implemented to mitigate the potential alignment of CTXA density measurements between pre-fracture and control cases. Future studies could be designed to fully test stratification accuracy between pre-fracture and control cases when using phantomless calibration.

Overall, results derived from the phantomless calibration slopes were a valid substitute for those derived from the inline calibration. When considering FS, the phantomless calibration resulted in a small 7% increase over inline calibration. For BMD, the phantomless calibration resulted in a small 4% increase over inline calibration. The phantomless calibration slopes were consistently comparable with the range demonstrated by the patient specific Mindways Model 3 BMD phantom calibration slopes, with our best method displaying a small bias of $0.028 \pm 0.054 \text{ HU}/(\text{mg}/\text{cm}^3)$. The study shows the proposed method for phantomless calibration is valid for FEA studies of retrospective cohorts lacking calibration data. This method can be applied opportunistically to CT scans captured for analyses other than hip fracture. Further examination of the error introduced when the proposed method for phantomless calibration is applied in patient specific FEA derived FS should be conducted.

5.5 Acknowledgements

The authors would like to thank the Whitaker Foundation for financial support of this project, the Institute for International Education for overseeing the Whitaker International Program Fellowship, and the funding from the European Commission H2020 CompBioMed2 project (Grant agreement ID: 823712). Additional thanks go to: Leticia Campello, Ashley Pernsteiner, and Emily Cameron for assistance with repeatability assessments for the phantomless methodology; Carrie Bartels for assistance capturing

CT calibration scans; and Keenan Brown for helpful discussions regarding phantom-based calibration and his experience with GE LightSpeed family CT scanners.

5.6 Conflict of interest

This study was supported by the Whitaker Foundation. The authors declare that they do not have any financial or personal relationships with other people or organizations that could have inappropriately influenced this study.

6 Comparison of internal tissue-based phantomless densitometric CT calibrations and error assessment

As shown in Chapter 5, the proposed approach to tissue-based phantomless DCTC has the potential to enable both opportunistic evaluation of osteoporosis and retrospective analysis of patient CT scans lacking phantom-based densitometric calibration data. As published, this method is available in the public domain and remains recommended for use. The previous chapter thoroughly discussed the limitations of tissue-based phantomless calibration in general and of that study in particular. This chapter will broaden the discussion by examining the proposed method in the context presented in Subsection 3.2.10.

As noted previously, ASTM E1935-97 recommends selecting densitometric standards that bracket the range of densities being measured. In the case of tissue-based phantomless calibration, including femoral bone as a reference tissue will likely increase the error introduced for several reasons. ICRU 44 and NIST^{28,30} assume a cortical bone nominal density of 1920 mg/cm³. The data this is based on comes from 24 adults (20–74 years old) and may not be a reasonable assumption for either a geriatric population or a predominately female population³⁰. In any case, assuming a density for cortical bone increases the risk of analysing the assumption instead of the actual patient bone.

Since the previous illustration from Chapter 3 was created including cortical bone tissue in all calculations, a follow-up illustration is included here to repeat all the same calculations, excluding cortical bone as an internal reference tissue. First, slice effective energy is calculated following both approaches described in Chapter 3: iteratively linearly

regressing CT Number [HU] and linear attenuation coefficients as recommended by ASTM E1935-97²⁵, and iteratively linearly regressing CT Number [HU] and mass attenuation coefficients as recommended by Michalski et al.³⁴. Note that the iterative linear regressions include segmentations for air, adipose, aortic blood, and skeletal muscle. The resulting plots are included in Figure 6-1.

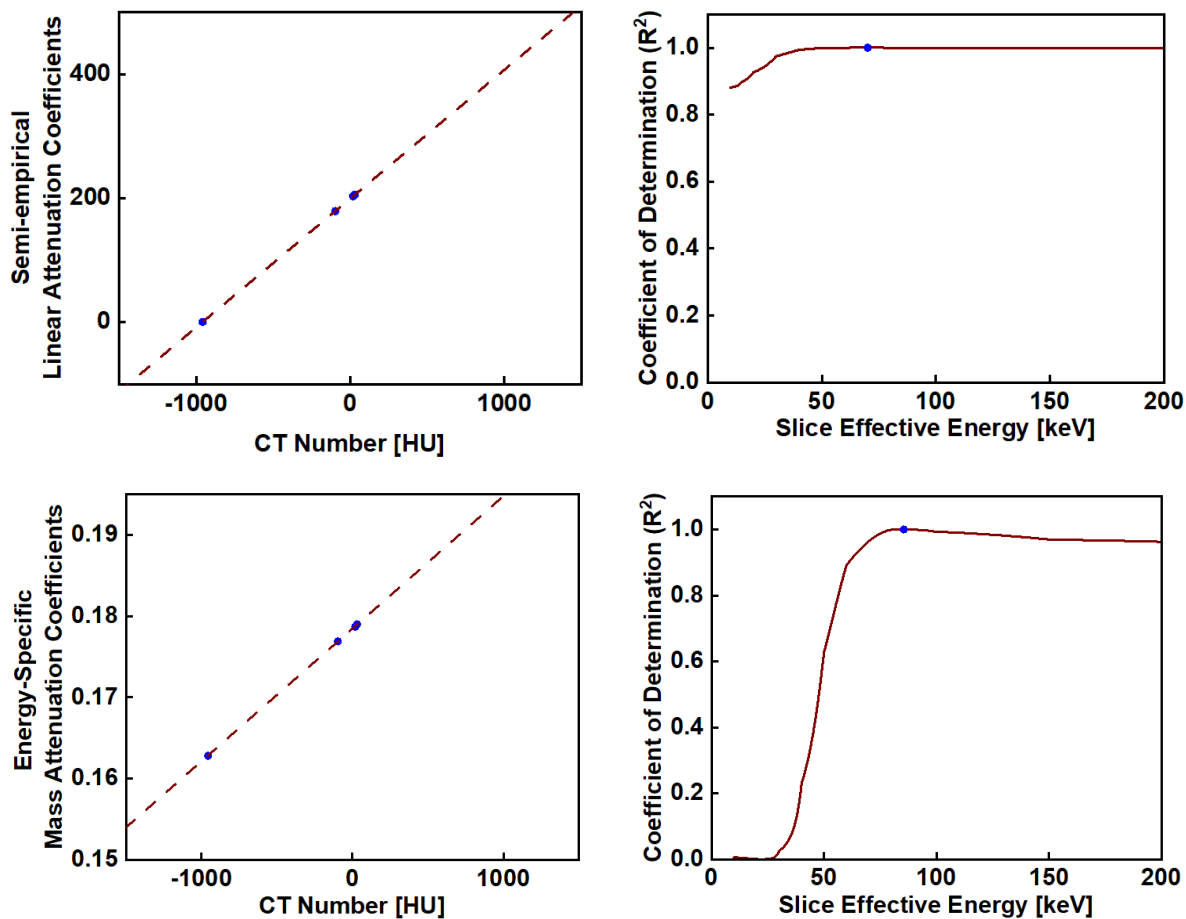


Figure 6-1: Slice effective energy calculations for tissue segmentations from the CT-scan of a 79-year-old patient.

After calculation of slice effective energy, nine densitometric calibration equations were derived for three different combinations of internal patient tissues: (1) air, adipose, aortic blood, and skeletal muscle; (2) air, aortic blood, and skeletal muscle; and (3) air,

and adipose. For each of these combinations, three calibration equations were derived following the density and attenuation approaches for each slice effective energy described previously. Then to compare densitometric results, each DCTC equation was applied to estimate the bracketing densities of femoral trabecular and cortical bone (Figure 6-2). Consistent with the results in the previous chapter, the air, aortic blood, and skeletal muscle combination introduced the least amount of error across the range of densities relevant to human femoral bone. Similar to the example from Chapter 3, this analysis shows that the corrections based on slice-specific energy improves accuracy (Figure 6-3). Whether this difference is clinically significant depends on the application.

The prior illustrations have been limited to regions of interest on one CT slice of one patient. Another useful illustration is conducting the same analyses for the segmentation data from Chapter 5 (Figure 6-4). In this illustration, measured BMD from both tissue-based phantomless DCTCs result in higher measured BMD than the measurement from the FDP for three out of four densities. Overall, the results in this chapter continue to support the use of tissue-based phantomless DCTC, and the proposed method continues to show promise for facilitating QCT analyses.

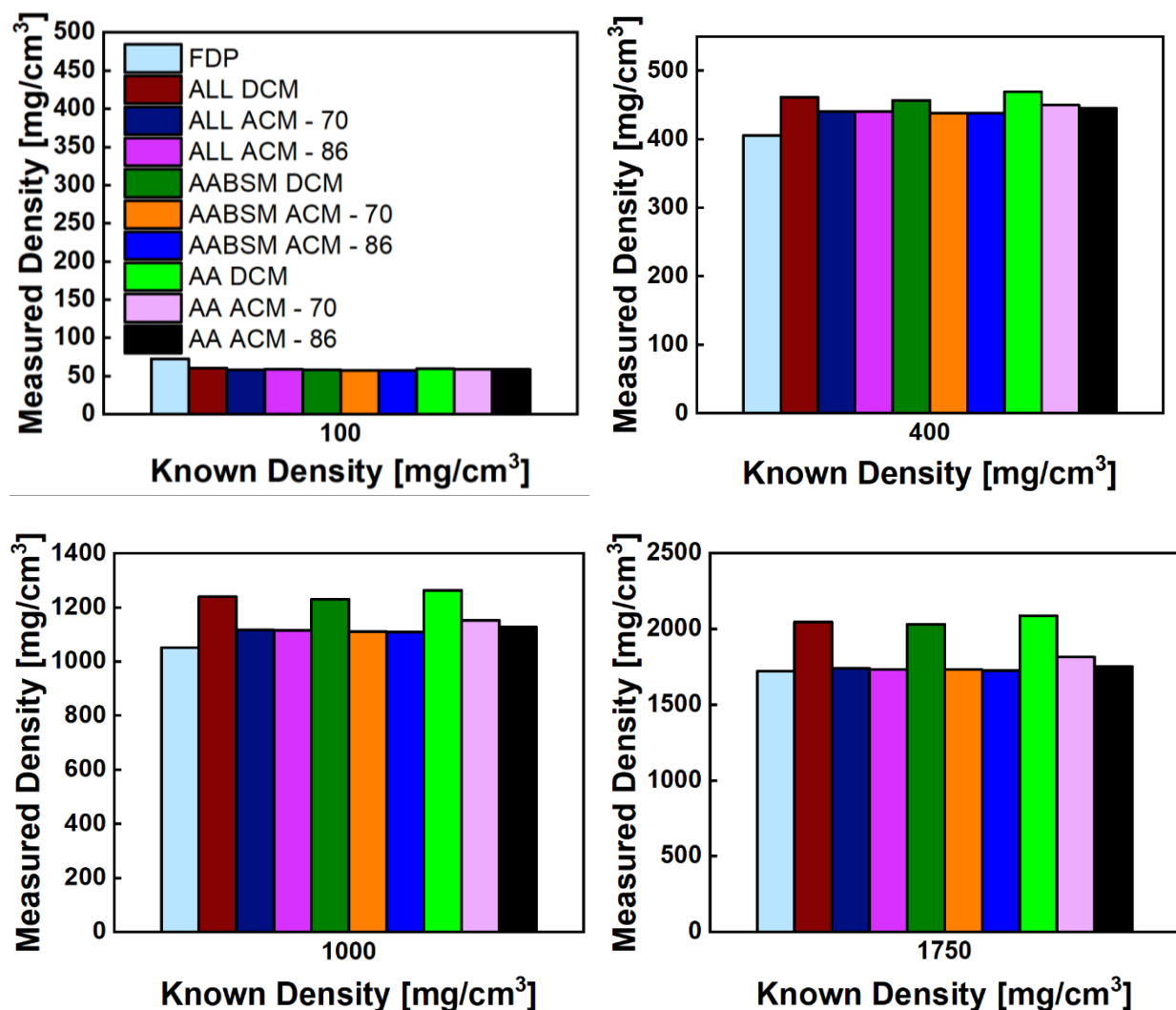


Figure 6-2: Representative density [mg/cm^3] measurements for bracketing trabecular and cortical femoral bone standards for multiple calibration references: femoral density phantom (FDP); density and attenuation calibration methods for internal tissue-based phantomless calibration derived from air, adipose, aortic blood, and skeletal muscle; density and attenuation calibration methods for internal tissue-based phantomless calibration derived from air, aortic blood, and skeletal muscle; and density and attenuation calibration methods for internal tissue-based phantomless calibration derived from air, and adipose. When relevant, attenuation approaches were calculated two for differing slice effective energies.

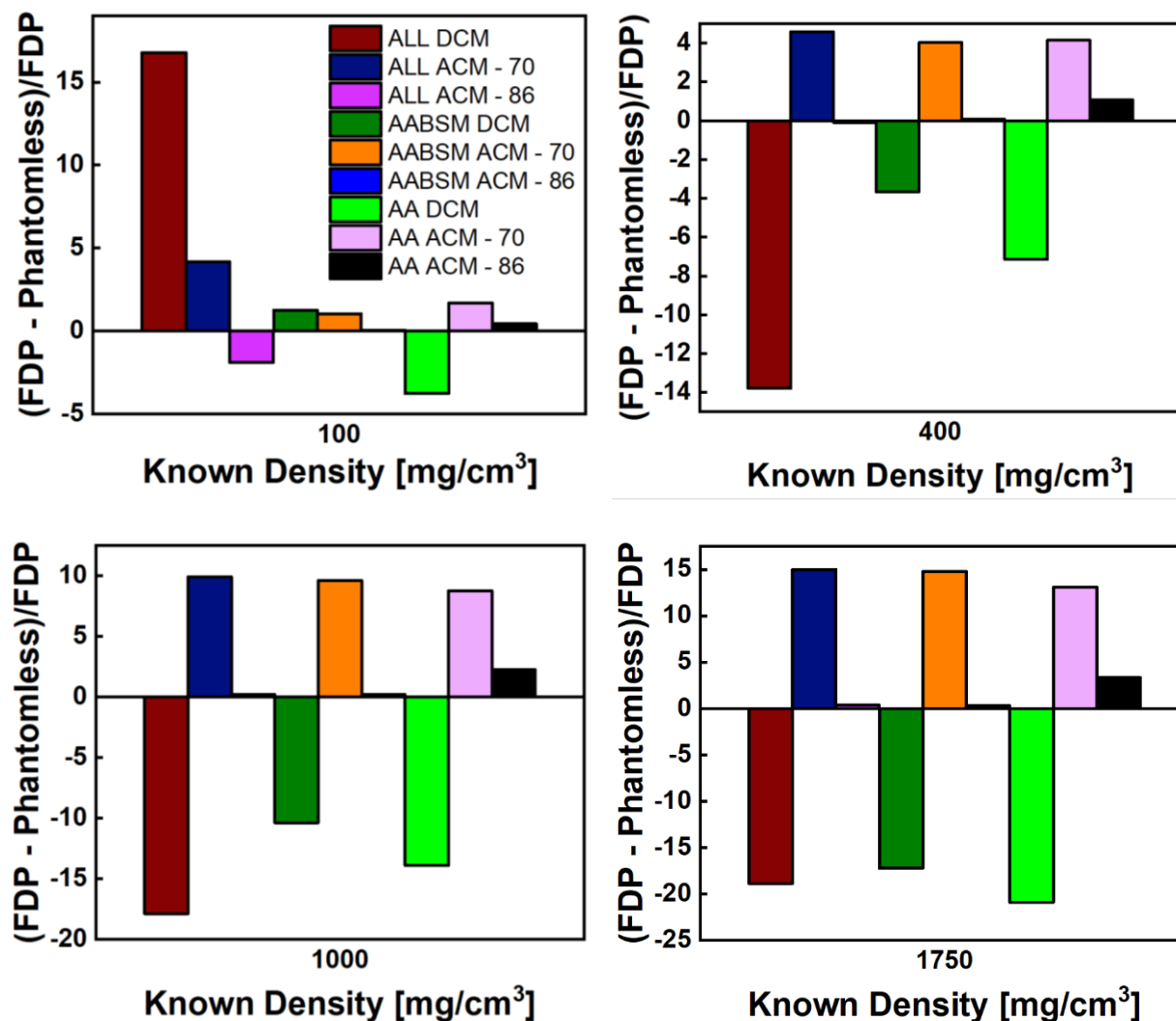


Figure 6-3: Percent difference calculations between the femoral density phantom (FDP) and phantomless density measurements for bracketing trabecular and cortical femoral densities. Phantomless measurements include density and attenuation calibration methods derived from air, adipose, aortic blood, and skeletal muscle; density and attenuation calibration methods derived from air, aortic blood, and skeletal muscle; and density and attenuation calibration methods derived from air, and adipose. When relevant, attenuation approaches were calculated two for differing slice effective energies. Note that differences range from -21% to 17%.

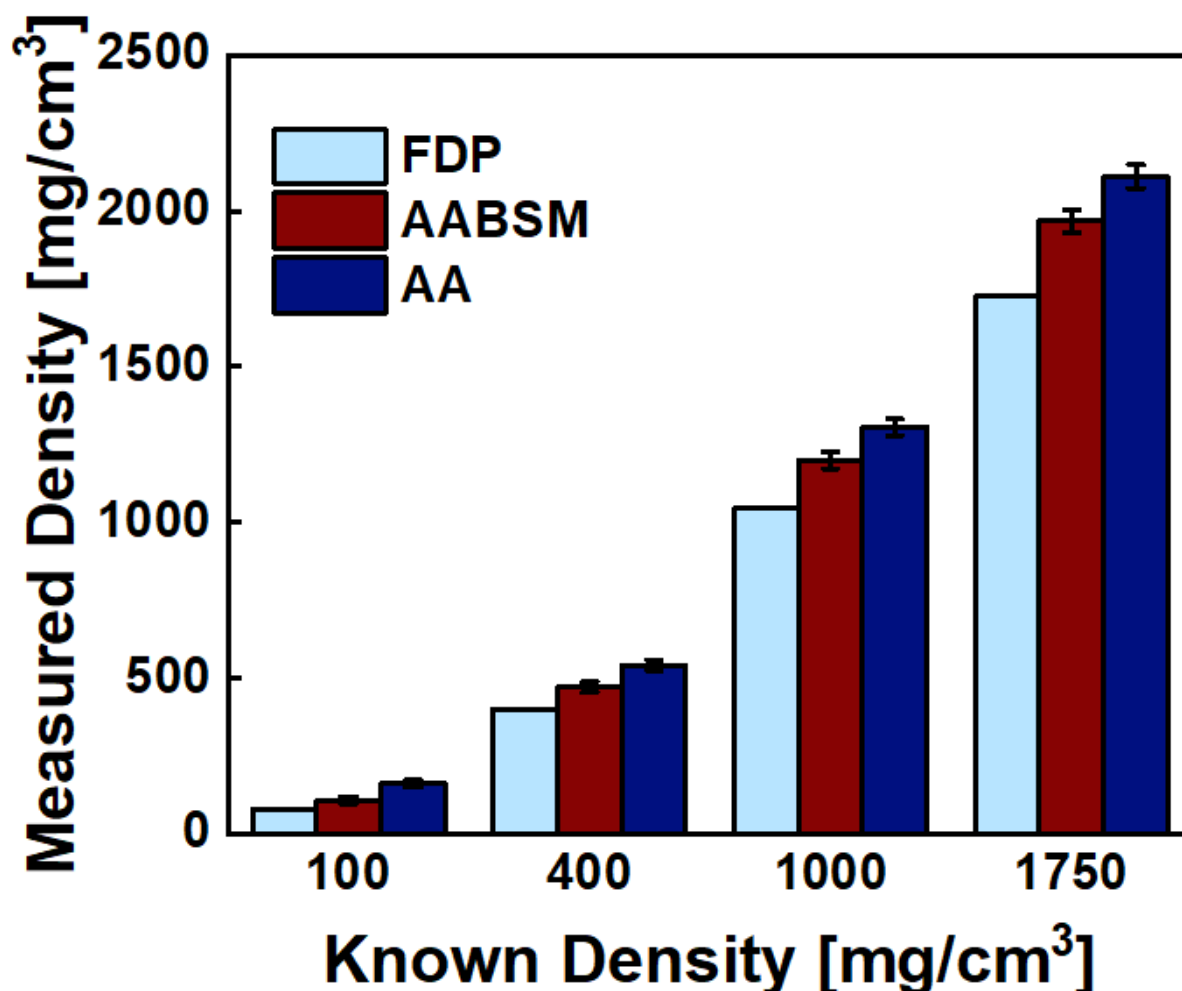


Figure 6-4: Representative densitometric CT measurements for trabecular and cortical human femoral bone bracketing densities. The light blue bars represent measurements derived from the femoral density phantom. The red and dark blue bars represent internal tissue-based phantomless DCTC derived measurements, for 258 CT scans of 211 patients, from the air, aortic blood, and skeletal muscle (AABSM); and air, and adipose (AA) combinations respectively. Note that the phantom-based measurements slightly underpredict the nominal density [mg/cm^3] and that the AABSM measurements are closer to them than the AA measurements.

7 Discussion

Two CTPSFE pipelines to predict femoral fragility fracture are available as clinical services. The first, VirtuOst software (O.N. Diagnostics, Berkeley, CA), is FDA approved and is clinically available in the USA. The second, “CT to Strength” Service non-invasive bone strength estimation from INSIGEO Institute for *in silico* Medicine and the University of Sheffield, can return an accurate assessment within 48 hours. To overcome the need for phantom-based DCTC data, Lee et al. validated an approach to phantomless DCTC in their 2017 study¹⁰³. This method is proprietary and therefore the clinical accessibility of this method is limited to clients of VirtuOst. A public-domain method allows for widespread collaboration across clinics, and potentially greater accessibility to patients. The objective of this dissertation was to demonstrate, verify, and validate a public domain method for selecting tissues to use as the basis for tissue-based phantomless DCTC for use in femoral QCT analyses.

A thorough understanding of the relevant definitions, limitations, and assumptions of both X-ray-based clinical imaging and CTPSFE was required. Chapter 3 presented the definitions, limitations, and assumptions of X-ray-based clinical imaging. X-rays were shown to be sensitive to energy [kVp], current [mA], and thickness [cm]. DXA is known to be sensitive to thickness of tissue present around the bone. Clinical CT scanners were shown to measure density relative to water using the Hounsfield Scale. The impacts of slice thickness, slice overlap, and differing reconstruction kernels were discussed.

Several approaches to approximating DCTC were described and illustrated in the examples in Chapter 3. Using the same structure as the examples in ASTM E1935-97

“Standard Test Method for Calibrating and Measuring CT Density”, an illustration in this dissertation compared and contrasted densitometric results from a variety of calibration approaches: a femoral density phantom; a density approach with phantomless DCTC derived from air, adipose, aortic blood, skeletal muscle and cortical bone; and an attenuation approach with internal tissue-based phantomless DCTC derived from air, adipose, aortic blood, skeletal muscle and cortical bone. The illustration showed that two different approaches to estimating the slice effective energy found 80 keV and 97 keV respectively, almost 20% difference in results. As expected, the femoral density phantom yielded the most accurate results across the range of densities relevant to human trabecular and cortical femoral bone. The density approach was more accurate than the MFM approach but less accurate than the attenuation approaches, regardless of which effective energy was assumed. Overall, the benefits from assuming an effective energy were uncertain and did not meaningfully reduce the potential error introduced by these approaches.

The final example in Chapter 6 showed a strong correlation between left and right adipose and a weaker correlation between left and right skeletal muscle tissue segmentation measurements when paired within the patient. The strong correlation in adipose makes sense and agrees with the Chapter 5 finding that adipose is the most consistent tissue across patients. The weaker correlation for skeletal muscle also makes sense given that most people experience a muscle imbalance. The corresponding Bland-Altman analysis showed zero bias between right and left adipose and skeletal muscle tissue segmentation measurements across patients.

Chapter 4 presented the relevant definitions, limitations, and assumptions for CTPSFE analyses. Medical image software applications show medical image data in three planes—the coronal, sagittal, and axial planes. A representative segmented patient femur was shown with the description of how various software applications support segmentation of patient femurs and what approaches are sufficiently accurate to yield trustworthy results. The empirical relationships that form the foundation of the material property mapping were explored, starting with scanning a densitometric reference phantom and deriving the CT Number-to-radiological-density relationship. The experiments conducted to derive the relationships between radiological density and ash density, ash density and apparent density, and apparent density and Young's modulus were explored in detail. Many of the methodological details in these experiments may vary from lab to lab; however, the ones discussed in this dissertation not only serve as the foundation for the VirtuOst software (O.N. Diagnostics, Berkely, CA) but have also been agreed upon by two other leading research groups in the field^{43,137}. Importantly, the material mapping strategy applied in the study was shown to bring the results of CT-based specimen-specific finite element analyses closer to the experimental results⁴³.

The stance phase during walking and sideways fall loading are two loading conditions identified as relevant to assessing femoral fracture risk. The most commonly applied loading conditions for the stance phase during walking come from a study by Bergmann et al.⁵³. Although empirical data are not available for sideways fall loading, 33 loading angles worth considering have been identified⁴¹. An advantage of applying FEA methods to study femoral fracture risk is the ability to simulate more than one loading condition, considering experimental specimens can only be broken once. 3D-printed

patient-specific femur geometries could be printed and leveraged in a variety of ways, including testing the experimental setup or assessing precision of results through the repetition of experiments. Whether stance phase during walking or sideways fall loading is being simulated, experimental validation is important and has been conducted on human femoral specimens in several studies. Some examples were discussed. A limitation of this dissertation is the lack of researcher experience with conducting validation experiments with cadaveric specimens; however, early on in graduate training some experimentation with 3D printed versions of the cadaveric specimens was conducted and observed.

CTPSFE analyses have been shown to be more accurate if boundary conditions are applied at the knee centre rather than making a cut and applying them close to the bottom of the trochanter. In prospectively captured data, the entire patient femur can be included in the CT scan by design and protocol. However, in retrospective case studies, patient CT scans frequently contained limited amounts of the femur. To overcome this gap in information, Ju Zhang created MapClient, which employs statistical shape modelling and customizable pipelines to extrapolate the missing portion of the femur. Using the estimated full femur from statistical shape modelling based on the available part of the patient femur and a database of 200 cadavers, MapClient can provide an estimated full femur capable of facilitating anatomical mapping to establish a coordinate system. While using this coordinate system has been shown to result in higher femoral fracture-control case stratification accuracy, after further studies the authors have reported that the accuracy of the statistical shape mapping tool is not sufficient to be used for CTPSFE analyses⁶¹.

The ability to derive clinically meaningful outcomes from the models is as important as being able to set up and run the models. Chapter 4 summarized the research behind applying a maximum principal strain failure criterion and the different approaches towards identifying a strong fracture status classifier. The best classifier to identify patients at risk for femoral fragility fracture prior to fracture is still an open question. Minimum Fall Strength (MFS) shows promise, especially when applying the recent stochastic modelling approaches^{138–140}.

The CTPSFE analyses methods described in Chapter 4 lacked a method for internal tissue-based phantomless DCTC, limiting the ability to retrospectively analyse patient cohorts without available phantom-based DCTC data. Designed for use in industry rather than clinical use, ASTM E1935-97 “Standard Test Method for Calibrating and Measuring CT Density” provides a good foundation for estimating DCTC. The National Institute of Standards and Technology provides reasonable assumptions for nominal human tissue densities. The outstanding question in this puzzle was what combination of internal tissues should be used as the basis for estimating DCTC.

The demonstrated method for segmenting internal tissues, quality-checking tissue segment mean CT Number [HU] measurements and optimizing with the Hounsfield Scale as the gold standard, resulted in a reasonable estimate for a DCTC equation. When the sparsely available cohort-specific phantom-based DCTC measurements were treated as representative boundary measurements, the tissue-based phantomless DCTC equations were not only verified but also were shown to produce smaller measurement variance across patients and CT scanners than the representative boundary measurements. This makes sense because both the patient and the inline phantom are in the path of the X-ray

when capturing DCTC data from an inline phantom. The resulting phantom-based measurement has higher density measurements, due to the distance from CT scanner isocentre in the field of view and due to the increase in thickness the X-rays travel through. The bias measured between the demonstrated tissue-based phantomless DCTC and the inline phantom-based DCTC-derived clinical outcomes were comparable with other studies in the field.

Chapter 6 presents an updated version of the example from Chapter 3; however, in Chapter 6, internal tissue-based phantomless DCTC-based measurements were derived without including cortical tissue segmentation measurements. The error in measured density across the range relevant to human trabecular and cortical femoral bone only came down slightly when excluding the cortical tissue segmentation measurements (Figure 7-1). The finding that the attenuation method did not result in an increase in accuracy large enough to be clinically meaningful remained consistent.

While the error introduced by tissue-based phantomless DCTC-based density measurements is larger than measurements from the FDP-based DCTC, there are several advantages to tissue-based phantomless DCTC. First, the reduction in variation of tissue-based phantomless DCTC equations result in a more consistent measurement and potential data for patient evaluation. Second, patient tissues will always be present in patient CT scans. Tissue-based phantomless DCTC adds data to the CTPSFE analyses and may improve the stratification accuracy of clinical assessments. Additionally, the quality checks presented in Chapter 5 facilitate the public-domain use of this method.

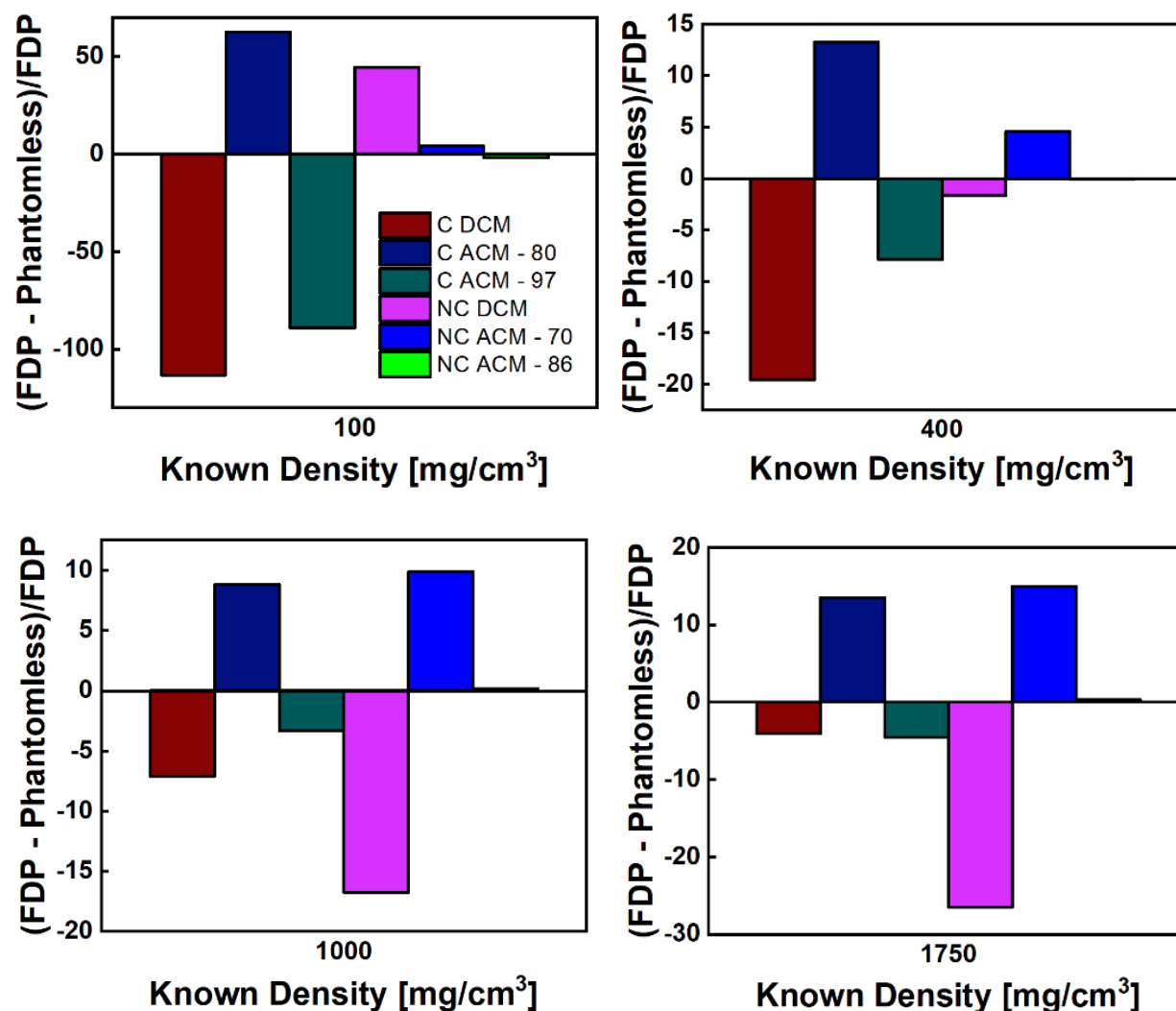


Figure 7-1: Percent difference calculations for phantomless density measurements with femoral density phantom (FDP) based density measurements as the gold standard. Cortical (C) and No Cortical (NC) are compared for the Density and Attenuation Calibration Method percent difference phantomless density measurements. Note that the No Cortical Attenuation Method with the Mass-Absorption Coefficient assumed effective energy has the least percent difference across the trabecular and cortical femoral bracketing densities.

Limitations to tissue-based phantomless DCTC are the dependency of the quality checks on the NIST tissue nominal density tables and the available information about the CT scanners. Internal tissues are dependent on a variety of patient-specific factors such as health status, hydration, and the presence of a contrast agent. The validations

conducted in Chapter 5 were limited to LightSpeed family CT scanners from GE, a single clinical protocol of CT scan acquisition and reconstruction parameters, and a small number of patients. All patients included in the study in Chapter 5 were undergoing routine abdominopelvic CT examinations due to cancer risk.

Tissue-based phantomless DCTC shows promise for application in CTPSFE and other quantitative CT-based analyses. The method demonstrated, verified, and validated in Chapter 5 is the first public-domain method. This method enables the retrospective evaluation of patient CT data and the opportunistic assessment of osteoporosis.

8 Scientific knowledge generated

Several new pieces of scientific knowledge were generated. At a high level, the first public-domain method for tissue-based phantomless DCTC was demonstrated, verified, and validated. Tissue segmentation measurements of air, adipose, aortic blood, and skeletal muscle were captured. Through statistical analyses, these measurements provide a basis for quality-checking future patient-specific tissue segmentation measurements for the broadest representation possible with the patient cohort. Through combinatorics and optimization, combinations of patient tissues were identified to serve as a basis for tissue-based phantomless DCTC. While this work recommended two specific combinations, the verification space showed that a wide variety of combinations may be valid. Making the segmentation measurements publicly available upon request allows for future researchers to test a specific combination with these data and then apply that combination to another patient cohort⁵⁷. The breadth of these results can further enable the development and application of tissue-based phantomless DCTC in clinically relevant quantitative CT analyses.

In addition to the patient tissue segmentation measurements, the work in this dissertation may be the only comparison of tissue-based phantomless DCTC to a FDP-based DCTC. By reversing the assumption that offline phantom-based DCTC data is relevant to patient CT scans, this work could also apply the patient-specific, tissue-based phantomless DCTC data to the measured mean CT Numbers [HU] for the four densitometric standards in the FDP. Through this approach, measured density could be

compared between the phantom-based and phantomless results, and accuracy could be quantified.

9 Future work

There are several directions of future research relevant to tissue-based phantomless DCTC. The data available for the development of this tissue-based phantomless DCTC were collected using machines from only one manufacturer. Since the method proposed is general to CT systems with the expectation that any manufacturer-specific CT rescale intercept may need to be corrected for and the tissues are chosen to align with the Hounsfield Scale, the method is presumed to be applicable to any clinical CT scanner operating on the Hounsfield Scale. As a public-domain method, global researchers can work together to test this presumption by applying the proposed method in their own clinical contexts, gradually widening the validation space and building clinical confidence in the methods.

Patients with surgical implants have not yet been considered in the studies on tissue-based phantomless DCTC methods. One reason for this is the X-ray scatter created by the implanted hardware. One advantage to the proposed method is that segmentation is performed manually, and specific tissue CT Number [HU] segmentation thresholds are provided. Through careful segmentation, it may be possible to apply this tissue-based phantomless method to characterize and analyse CT scans of patients with surgical implants (Figure 9-1). Initial research would need to assess the tissue-based phantomless method result in comparison to a phantom, ideally the femoral density phantom, which was designed to support CTPSFE models of femurs to aid in analyses surrounding surgical implants. Follow-on research could include CTPSFE analyses of the contralateral

femur and may be comparable to the Sheffield cohort which looks at the contralateral femur^{17,39,140,141}.

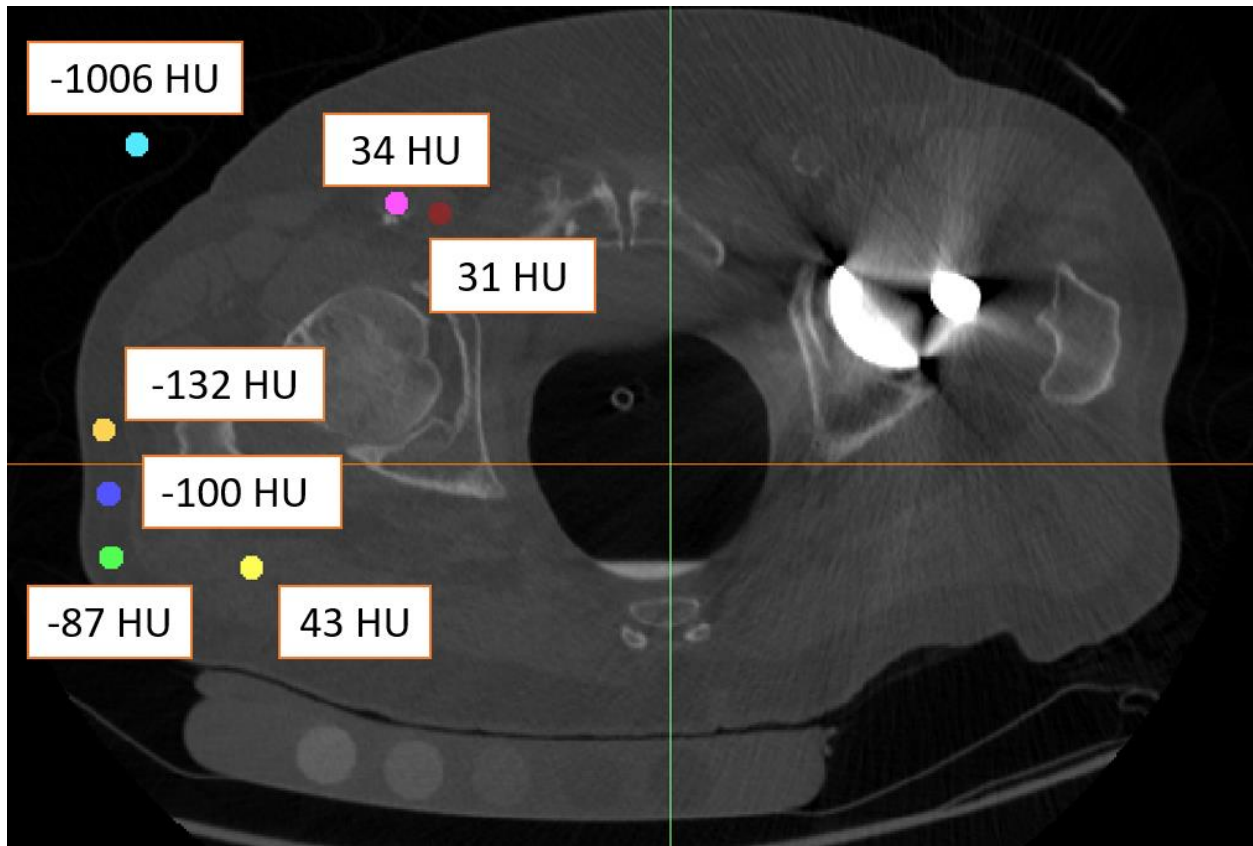


Figure 9-1: Representative CT slice of internal tissue segmentations for a 93-year-old patient with a surgical implant. Note that three adipose segmentations were necessary to get a result consistent with the thresholds established in Chapter 5, and that a result within that framework was possible.

There are still several research questions that may be answered with further analyses on the femoral fracture cohort for which the method was developed. A age- and sex-matched pre-fracture cohort was selected for a FEA study from the full fracture cohort of patient data retrospectively identified at UW-Madison. Of the 43 pre-fracture patients, 40 patient CT scans met the following inclusion criteria. First the scan contained the femur. Second the scan was captured at the prescribed clinical CT scan acquisition

protocol and reconstructed with a standard kernel. An additional sub-cohort of 40 patients were identified as age and sex matched controls.

Table 9-1: Breakdown of prefracture and control patients segmented from the Madison fracture cohort for later CTPSFE analyses.

	Pre-fracture	Control	Total
Female	26	26	52
Male	14	14	28
Total	40	40	80

The Madison cohort differs from the Sheffield cohort in at least two ways: the Sheffield cohort was captured on a GE Lightspeed CT scanner with a BonePlus reconstruction kernel, and the CTPSFE analyses were conducted on the contralateral femur^{16,17}. The Madison cohort includes femoral CT scans of patients up to 5 years prior to fracture, although most are within a year of fracture. This offers a unique opportunity to assess the stratification accuracy of CTPSFE derived MFS developed with the Sheffield cohort on the femur known to go on to fracture.

Phantomless DCTC continues to show promise for femoral QCT analyses. Pursuing future studies on additional patient cohorts has the potential to broaden the applicability of the methods. Additional studies assessing pre-fracture cohorts may yield insights relevant to our understanding of femoral fragility fracture. Phantomless DCTC has the potential to facilitate QCT analyses.

References

1. Howell JD. Early Clinical Use of the X-ray. *Trans Am Clin Climatol Assoc.* 2016;127:341-349.
2. Szczykutowicz TP. *Dual-Energy and Spectral Imaging*. Elsevier B.V.; 2014. doi:10.1016/b978-0-444-53632-7.00207-0
3. Cameron JR, Sorenson J. Measurement of bone mineral in vivo: An improved method. *Science (1979)*. 1963;142(3589):230-232. doi:10.1126/science.142.3589.230
4. Blake GM, Fogelman I. The role of DXA bone density scans in the diagnosis and treatment of osteoporosis. *Postgrad Med J.* 2007;83(982):509-517. doi:10.1136/pgmj.2007.057505
5. Yu EW, Thomas BJ, Brown JK, Finkelstein JS. Simulated increases in body fat and errors in bone mineral density measurements by DXA and QCT. *Journal of Bone and Mineral Research.* 2012;27(1):119-124. doi:10.1002/jbmr.506
6. Cann CE, Adams JE, Brown JK, Brett AD. CTXA hip - An extension of classical DXA measurements using Quantitative CT. *PLoS One.* 2014;9(3). doi:10.1371/journal.pone.0091904
7. Keaveny TM, Clarke BL, Cosman F, et al. Biomechanical Computed Tomography analysis (BCT) for clinical assessment of osteoporosis. *Osteoporosis International*. Published online 2020:1025-1048. doi:10.1007/s00198-020-05384-2
8. Kamalian S, Lev MH, Gupta R. *Computed Tomography Imaging and Angiography – Principles*. Vol 135. 1st ed. Elsevier B.V.; 2016. doi:10.1016/B978-0-444-53485-9.00001-5
9. Cormack AM. Representation of a function by its line integrals, with some radiological applications. *J Appl Phys.* 1963;34(9):2722-2727. doi:10.1063/1.1713127
10. Cormack AM. Representation of a function by its line integrals, with some radiological applications. II. *J Appl Phys.* 1964;35(10):2908-2913. doi:10.1063/1.1713127

11. Cormack AM. Reconstruction of densities from their projections, with applications in radiological physics. *Phys Med Biol.* 1973;18(2):195-207. doi:10.1088/0031-9155/18/2/003
12. Michalski AS, Edwards BW, Boyd SK. QCT Reconstruction Kernel has Important Quantitative Effects on Finite Element Estimated Bone Strength. *CMBES Proceedings.* 2016;39(1).
13. Eggermont F, Verdonchot N, van der Linden Y, Tanck E. Calibration with or without phantom for fracture risk prediction in cancer patients with femoral bone metastases using CT-based finite element models. *PLoS One.* 2019;14(7):1-13. doi:10.1371/journal.pone.0220564
14. Giambini H, Dragomir-Daescu D, Huddleston PM, Camp JJ, An KN, Nassr A. The Effect of Quantitative Computed Tomography Acquisition Protocols on Bone Mineral Density Estimation. *J Biomech Eng.* 2015;137(11):1-6. doi:10.1115/1.4031572
15. Goodsitt MM. Conversion relations for quantitative CT bone mineral densities measured with solid and liquid calibration standards. *Bone Miner.* 1992;19(2):145-158. doi:10.1016/0169-6009(92)90922-Z
16. Yang L, Udall WJM, McCloskey E V., Eastell R. Distribution of bone density and cortical thickness in the proximal femur and their association with hip fracture in postmenopausal women: A quantitative computed tomography study. *Osteoporosis International.* 2014;25(1):251-263. doi:10.1007/s00198-013-2401-y
17. Qasim M, Farinella G, Zhang J, et al. Patient-specific finite element estimated femur strength as a predictor of the risk of hip fracture: the effect of methodological determinants. *Osteoporosis International.* 2016;27(9):2815-2822. doi:10.1007/s00198-016-3597-4
18. Enns-Bray WS, Bahaloo H, Fleps I, et al. Biofidelic finite element models for accurately classifying hip fracture in a retrospective clinical study of elderly women from the AGES Reykjavik cohort. *Bone.* 2019;120:25-37. doi:10.1016/j.bone.2018.09.014

19. Fleps I, Pálsson H, Baker A, et al. Finite element derived femoral strength is a better predictor of hip fracture risk than aBMD in the AGES Reykjavik study cohort. *Bone*. 2022;154. doi:10.1016/j.bone.2021.116219
20. Nishiyama KK, Ito M, Harada A, Boyd SK. Classification of women with and without hip fracture based on quantitative computed tomography and finite element analysis. *Osteoporosis International*. 2014;25(2):619-626. doi:10.1007/s00198-013-2459-6
21. Orwoll ES, Marshall LM, Nielson CM, et al. Finite element analysis of the proximal femur and hip fracture risk in older men. *Journal of Bone and Mineral Research*. 2009;24(3):475-483. doi:10.1359/jbmr.081201
22. Amin S, Kopperdhal DL, Melton LJ, et al. Association of hip strength estimates by finite-element analysis with fractures in women and men. *Journal of Bone and Mineral Research*. 2011;26(7):1593-1600. doi:10.1002/jbmr.347
23. Kopperdahl DL, Aspelund T, Hoffmann PF, et al. Assessment of incident spine and hip fractures in women and men using finite element analysis of CT scans. *Journal of Bone and Mineral Research*. 2014;29(3):570-580. doi:10.1002/jbmr.2069
24. Falcinelli C, Schileo E, Balistreri L, et al. Multiple loading conditions analysis can improve the association between finite element bone strength estimates and proximal femur fractures: A preliminary study in elderly women. *Bone*. 2014;67:71-80. doi:10.1016/j.bone.2014.06.038
25. Committee WTOTB to T. Standard Test Method for Calibrating and Measuring CT Density 1. *American Society for Testing of Materials*. 2003;03(May):1-5. doi:10.1520/E1935-97R19.Copyright
26. Hubbell JH, Seltzer SM. Tables of X-Ray mass attenuation coefficients and mass energy-absorption coefficients (version 1.4). <http://physics.nist.gov/xaamdi>. Published online 2004.
27. M.J. Berger, J.H. Hubbel, S.M. Seltzer, J. Chang, J.S. Coursey, R. Sukumar, D.S. Zucker and KO. XCOM: Photon Cross Sections Database. <https://www.nist.gov/pml/xcom-photon-cross-sections-database>. doi:<https://dx.doi.org/10.18434/T48G6X>

28. White, D.R.; Booz, J. ; Griffith, R. V.; Spokas, J. J.; Wilson IJ. Tissue substitutes in radiation dosimetry and measurement. *International Communications Radiation Units Measurement*. 1989;23(1):1-184.
doi:<https://doi.org/10.1093/jicru/os23.1.Report44>
29. Michael R. Millner, William H. Payne, Robert G. Waggener, William D. McDavid, Michael J. Dennis and VJS. Determination of effective energies in CT calibration. *Med Phys*. 1978;5(6):543-545. doi:<https://doi.org/10.1118/1.594488>
30. Zioupos P, Cook RB, Hutchinson JR. Some basic relationships between density values in cancellous and cortical bone. *J Biomech*. 2008;41(9):1961-1968.
doi:10.1016/j.jbiomech.2008.03.025
31. Carter DR, Hayes WC. The compressive behavior of bone as a two-phase porous structure. *J Bone Joint Surg*. 1976;59-A(7):954-962. doi:10.1007/978-1-4471-5451-8_116
32. Michalski AS. A quantitative computed tomography approach towards opportunistic osteoporosis screening. Published online 2020.
<http://hdl.handle.net/1880/111749>
33. Crookshank M, Ploeg HL, Ellis R, MacIntyre NJ. Repeatable calibration of Hounsfield units to mineral density and effect of scanning medium. *Advances in biomechanics and applications*. 2013;1(1):15-22. doi:10.12989/aba.2013.1.1.015
34. Troy KL, Edwards WB. Practical considerations for obtaining high quality quantitative computed tomography data of the skeletal system. *Bone*. 2018;110:58-65. doi:10.1016/j.bone.2018.01.013
35. Brunnquell CL, Winsor C, Aaltonen HL, Telfer S. Sources of error in bone mineral density estimates from quantitative CT. *Eur J Radiol*. 2021;144.
doi:10.1016/j.ejrad.2021.110001
36. Viceconti M, Qasim M, Bhattacharya P, Li X. Are CT-Based Finite Element Model Predictions of Femoral Bone Strengthening Clinically Useful ? *Curr Osteoporos Rep*. Published online 2018:1-8. doi:10.1007/s11914-018-0438-8
37. Adams AL, Fischer H, Kopperdahl DL, et al. Osteoporosis and Hip Fracture Risk From Routine Computed Tomography Scans: The Fracture, Osteoporosis, and CT

- Utilization Study (FOCUS). *Journal of Bone and Mineral Research*. 2018;33(7):1291-1301. doi:10.1002/jbmr.3423
38. Taylor M, Viceconti M, Bhattacharya P, Li X. Finite element analysis informed variable selection for femoral fracture risk prediction. *J Mech Behav Biomed Mater*. 2021;118(February):104434. doi:10.1016/j.jmbbm.2021.104434
 39. Verhulp E, van Rietbergen B, Huiskes R. Comparison of micro-level and continuum-level voxel models of the proximal femur. *J Biomech*. 2006;39(16):2951-2957. doi:10.1016/j.jbiomech.2005.10.027
 40. Altai Z, Qasim M, Li X, Viceconti M. The effect of boundary and loading conditions on patient classification using finite element predicted risk of fracture. *Clinical Biomechanics*. 2019;68(June):137-143. doi:10.1016/j.clinbiomech.2019.06.004
 41. Taddei F, Martelli S, Reggiani B, Cristofolini L, Viceconti M. Finite-element modeling of bones from CT data: Sensitivity to geometry and material uncertainties. *IEEE Trans Biomed Eng*. 2006;53(11):2194-2200. doi:10.1109/TBME.2006.879473
 42. Schileo E, Dall'Ara E, Taddei F, et al. An accurate estimation of bone density improves the accuracy of subject-specific finite element models. *J Biomech*. 2008;41(11):2483-2491. doi:10.1016/j.jbiomech.2008.05.017
 43. Morgan EF, Bayraktar HH, Keaveny TM. Trabecular bone modulus-density relationships depend on anatomic site. *J Biomech*. 2003;36(7):897-904. doi:10.1016/S0021-9290(03)00071-X
 44. Kopperdahl DL, Keaveny TM. Yield strain behavior of trabecular bone. *J Biomech*. 1998;31(7):601-608. doi:10.1016/S0021-9290(98)00057-8
 45. DL K, EF M, TM K. Quantitative computed tomography estimates of the mechanical properties of human vertebral trabecular bone. *J Orthop Res*. 2002;20:801-805. doi:10.1016/S0736-0266(01)00185-1
 46. Morgan EF, Keaveny TM. Dependence of yield strain of human trabecular bone on anatomic site. *J Biomech*. 2001;34(5):569-577. doi:10.1016/S0021-9290(01)00011-2
 47. Cowin SC. The Relationship Between the Elasticity Tensor and the Fabric Tensor. *Mechanics of Materials*. 1985;4(1):137-147.

48. Kabel J, Van Rietbergen B, Odgaard A, Huiskes R. Constitutive relationships of fabric, density, and elastic properties in cancellous bone architecture. *Bone*. 1999;25(4):481-486. doi:10.1016/S8756-3282(99)00190-8
49. Yang G, Kabel J, Van Rietbergen B, Odgaard A, Huiskes R, Cowin SC. The Anisotropic Hooke's Law for Cancellous Bone and Wood. *J Elast*. 1998;53(2):125-146. doi:10.1023/A:1007575322693
50. Wirtz DC, Schiffers N, Forst R, Pandorf T, Weichert D, Radermacher K. Critical evaluation of known bone material properties to realize anisotropic FE-simulation of the proximal femur. *J Biomech*. 2000;33(10):1325-1330. doi:10.1016/S0021-9290(00)00069-5
51. Taddei F, Schileo E, Helgason B, Cristofolini L, Viceconti M. The material mapping strategy influences the accuracy of CT-based finite element models of bones: An evaluation against experimental measurements. *Med Eng Phys*. 2007;29(9):973-979. doi:10.1016/j.medengphy.2006.10.014
52. Bergmann G, Bergmann G, Deuretzbacher G, et al. Hip forces and gait patterns from routine activities. *J Biomech*. 2001;34:859-871. doi:http://dx.doi.org/10.1016/S0021-9290(01)00040-9
53. Helgason B, Gilchrist S, Ariza O, et al. The influence of the modulus-density relationship and the material mapping method on the simulated mechanical response of the proximal femur in side-ways fall loading configuration. *Med Eng Phys*. 2016;38(7):679-689. doi:10.1016/j.medengphy.2016.03.006
54. Viceconti M, Olsen S, Nolte LP, Burton K. Extracting clinically relevant data from finite element simulations. *Clinical Biomechanics*. 2005;20(5):451-454. doi:10.1016/j.clinbiomech.2005.01.010
55. Wu G, Siegler S, Allard P, et al. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion-part I: ankle, hip, and spine. *J Biomech*. 2002;35:543-548. doi:10.1016/s0301-2115(02)00202-6
56. Winsor C, Li X, Qasim M, et al. Evaluation of patient tissue selection methods for deriving equivalent density calibration for femoral bone quantitative CT analyses. *Bone*. 2021;143(July 2020). doi:10.1016/j.bone.2020.115759

57. Zhang J, Malcolm D, Hislop-jambrich J, Thomas CDL, Nielsen P. Automatic Meshing of Femur Cortical Surfaces from Clinical CT Images. In: *Lecture Notes in Computer Science*. Vol 7599. Springer, Berlin, Heidelberg; 2012:40-48. doi:10.1007/978-3-642-33463-4
58. Zhang J, Sorby H, Clement J, et al. The MAP Client: User-Friendly Musculoskeletal Modelling Workflows. In: *Lecture Notes in Computer Science, Vol 8789*. Springer, Cham; 2014:182-192. doi:10.1007/978-3-319-12057-7_21
59. Besl PJ, McKay ND. Method for registration of 3-D shapes. *Sensor Fusion IV: Control Paradigms and Data Structures*. 1992;1611(May):586-606. doi:10.1117/12.57955
60. Zhang J, Besier TF. Accuracy of femur reconstruction from sparse geometric data using a statistical shape model. *Comput Methods Biomech Biomed Engin*. 2017;20(5):566-576. doi:10.1080/10255842.2016.1263301
61. Courtney AC, Wachtel EF, Myers ER, Hayes WC. Effects of loading rate on strength of the proximal femur. *Calcif Tissue Int*. 1994;55(1):53-58. doi:10.1007/BF00310169
62. Schileo E, Taddei F, Cristofolini L, Viceconti M. Subject-specific finite element models implementing a maximum principal strain criterion are able to estimate failure risk and fracture location on human femurs tested in vitro. *J Biomech*. 2008;41(2):356-367. doi:10.1016/j.jbiomech.2007.09.009
63. Bayraktar HH, Morgan EF, Niebur GL, Morris GE, Wong EK, Keaveny TM. Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. *J Biomech*. 2004;37(1):27-35. doi:10.1016/S0021-9290(03)00257-4
64. Chang WCW, Christensen TM, Pinilla TP, Keaveny TM. Uniaxial Yield Strains for Bovine Trabecular Bone are Isotropic and Asymmetric. *Journal of Orthopaedic Research*. 1999;17:582-585.
65. Keaveny TM, Morgan EF, Niebur GL, Yeh OC. Biomechanics of Trabecular Bone. *Annual Reviews of Biomedical Engineering*. Published online 2001:307-333.
66. Niebur GL, Feldstein MJ, Yuen JC, Chen TJ, Keaveny TM. High-resolution finite element models with tissue strength asymmetry accurately predict failure of

- trabecular bone. *J Biomech.* 2000;33(12):1575-1583. doi:10.1016/S0021-9290(00)00149-4
67. Michalski AS, Besler BA, Burt LA, Boyd SK. Opportunistic CT screening predicts individuals at risk of major osteoporotic fracture. *Osteoporosis International*. Published online 2021. doi:10.1007/s00198-021-05863-0
 68. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N. Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. *Ann Intern Med.* 2013;158(8):588-595. doi:10.7326/0003-4819-158-8-201304160-00003
 69. Ziemlewicz TJ, Binkley N, Pickhardt PJ. Opportunistic Osteoporosis Screening: Addition of Quantitative CT Bone Mineral Density Evaluation to CT Colonography. *Journal of the American College of Radiology.* 2015;12(10):1036-1041. doi:10.1016/j.jacr.2015.04.018
 70. Brett AD, Brown JK. Quantitative computed tomography and opportunistic bone density screening by dual use of computed tomography scans. *J Orthop Translat.* 2015;3(4):178-184. doi:10.1016/j.jot.2015.08.006
 71. Lenchik L, Weaver AA, Ward RJ, Boone JM, Boutin RD. Opportunistic Screening for Osteoporosis Using Computed Tomography: State of the Art and Argument for Paradigm Shift. *Curr Rheumatol Rep.* 2018;20(12). doi:10.1007/s11926-018-0784-7
 72. Pickhardt PJ, Lee SJ, Liu J, et al. Population-based opportunistic osteoporosis screening: Validation of a fully automated CT tool for assessing longitudinal BMD changes. *British Journal of Radiology.* 2019;92(1094). doi:10.1259/bjr.20180726
 73. Lee SJ, Anderson PA, Pickhardt PJ. Predicting future hip fractures on routine abdominal CT using opportunistic osteoporosis screening measures: A matched case-control study. *American Journal of Roentgenology.* 2017;209(2):395-402. doi:10.2214/AJR.17.17820
 74. Aggarwal V, Maslen C, Abel RL, et al. Opportunistic diagnosis of osteoporosis, fragile bone strength and vertebral fractures from routine CT scans; a review of approved technology systems and pathways to implementation. *Ther Adv Musculoskelet Dis.* 2021;13:1-19. doi:10.1177/1759720X211024029

75. Ziemlewicz TJ, Maciejewski A, Binkley N, Brett AD, Brown JK, Pickhardt PJ. Opportunistic Quantitative CT Bone Mineral Density Measurement at the Proximal Femur Using Routine Contrast-Enhanced Scans: Direct Comparison With DXA in 355 Adults. *Journal of Bone and Mineral Research*. 2016;31(10):1835-1840. doi:10.1002/jbmr.2856
76. Black DM, Cummings SR, Karpf DB, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. *Lancet*. 1996;348(9041):1535-1541. doi:10.1016/S0140-6736(96)07088-2
77. Reginster JY, Minne HW, Sorensen OH, et al. Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. *Osteoporosis International*. 2000;11(1):83-91. doi:10.1007/s001980050010
78. Langdahl BL, Libanati C, Crittenden DB, et al. Romosozumab (sclerostin monoclonal antibody) versus teriparatide in postmenopausal women with osteoporosis transitioning from oral bisphosphonate therapy: a randomised, open-label, phase 3 trial. *The Lancet*. 2017;390(10102):1585-1594. doi:10.1016/S0140-6736(17)31613-6
79. Harris ST, Watts NB, Genant HK, et al. Effects of Risedronate Treatment on Vertebral and Nonvertebral Fractures in Women With Postmenopausal Osteoporosis<SUBTITLE>A Randomized Controlled Trial</SUBTITLE>. *JAMA*. 2003;282(14):1344. doi:10.1001/jama.282.14.1344
80. Bone HG, Wagman RB, Brandi ML, et al. 10 years of denosumab treatment in postmenopausal women with osteoporosis: results from the phase 3 randomised FREEDOM trial and open-label extension. *Lancet Diabetes Endocrinol*. 2017;5(7):513-523. doi:10.1016/S2213-8587(17)30138-9
81. McClung MR, Brown JP, Diez-Perez A, et al. Effects of 24 Months of Treatment With Romosozumab Followed by 12 Months of Denosumab or Placebo in Postmenopausal Women With Low Bone Mineral Density: A Randomized, Double-Blind, Phase 2, Parallel Group Study. *Journal of Bone and Mineral Research*. 2018;33(8):1397-1406. doi:10.1002/jbmr.3452

82. Keaveny TM, Crittenden DB, Bolognese MA, et al. Greater Gains in Spine and Hip Strength for Romosozumab Compared With Teriparatide in Postmenopausal Women With Low Bone Mass. *Journal of Bone and Mineral Research*. 2017;32(9):1956-1962. doi:10.1002/jbmr.3176
83. Keaveny TM, McClung MR, Wan X, Kopperdahl DL, Mitlak BH, Krohn K. Femoral strength in osteoporotic women treated with teriparatide or alendronate. *Bone*. 2012;50(1):165-170. doi:10.1016/j.bone.2011.10.002
84. Whitmarsh T, Treece GM, Gee AH, Poole KES. Mapping Bone Changes at the Proximal Femoral Cortex of Postmenopausal Women in Response to Alendronate and Teriparatide Alone, Combined or Sequentially. *Journal of Bone and Mineral Research*. 2015;30(7):1309-1318. doi:10.1002/jbmr.2454
85. Seeman E, Delmas PD, Hanley DA, et al. Microarchitectural deterioration of cortical and trabecular bone: differing effects of denosumab and alendronate. *Journal of Bone and Mineral Research*. 2010;25(8):1886-1894. doi:10.1002/jbmr.81
86. Keaveny TM, Hoffmann PF, Singh M, et al. Femoral bone strength and its relation to cortical and trabecular changes after treatment with PTH, alendronate, and their combination as assessed by finite element analysis of quantitative CT scans. *Journal of Bone and Mineral Research*. 2008;23(12):1974-1982. doi:10.1359/jbmr.080805
87. Cummings SR, Black DM, Thompson DE, et al. Effect of Alendronate on Risk of Fracture in Women With Low Bone Density but Without Vertebral Fractures<SUBTITLE>Results From the Fracture Intervention Trial</SUBTITLE>. *JAMA*. 2003;280(24):2077. doi:10.1001/jama.280.24.2077
88. Imai K, Ohnishi I, Matsumoto T, Yamamoto S, Nakamura K. Assessment of vertebral fracture risk and therapeutic effects of alendronate in postmenopausal women using a quantitative computed tomography-based nonlinear finite element method. *Osteoporosis International*. 2009;20(5):801-810. doi:10.1007/s00198-008-0750-8

89. Schwaiger BJ, Facchetti L, Gersing AS, et al. Vertebral and femoral bone mineral density and bone strength in prostate cancer patients assessed in phantomless PET/CT examinations. *Bone*. 2017;101:62-69. doi:10.1016/j.bone.2017.04.008
90. Michalski AS, Amin S, Cheung AM, et al. Hip load capacity cut-points for Astronaut Skeletal Health NASA Finite Element Strength Task Group Recommendations. *NPJ Microgravity*. 2019;5(1):1-10. doi:10.1038/s41526-019-0066-3
91. Michalski AS, Besler BA, Michalak GJ, Boyd SK. CT-based internal density calibration for opportunistic skeletal assessment using abdominal CT scans. *Med Eng Phys*. 2020;78:55-63. doi:10.1016/j.medengphy.2020.01.009
92. Giambini H, Dragomir-Daescu D, Nassr A, Yaszemski MJ, Zhao C. Quantitative Computed Tomography Protocols Affect Material Mapping and Quantitative Computed Tomography-Based Finite-Element Analysis Predicted Stiffness. *J Biomech Eng*. 2016;138(9):1-7. doi:10.1115/1.4034172
93. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. *Journal of Bone and Mineral Research*. 2007;22(3):465-475. doi:10.1359/jbmr.061113
94. Siris ES, Chen YT, Abbott TA, et al. Bone mineral density thresholds for pharmacological intervention to prevent fractures. *Arch Intern Med*. 2004;164(10):1108-1112. doi:10.1001/archinte.164.10.1108
95. Katz Y, Dahan G, Sosna J, Shelef I, Cherniavsky E, Yosibash Z. *Scanner Influence on the Mechanical Response of QCT-Based Finite Element Analysis of Long Bones*. Vol 86.; 2019. doi:10.1016/j.jbiomech.2019.01.049
96. Engelke K, Libanati C, Fuerst T, Zysset P, Genant HK. Advanced CT based in vivo methods for the assessment of bone density, structure, and strength. *Curr Osteoporos Rep*. 2013;11(3):246-255. doi:10.1007/s11914-013-0147-2
97. Knowles NK, Reeves JM, Ferreira LM. Quantitative Computed Tomography (QCT) derived Bone Mineral Density (BMD) in finite element studies: a review of the literature. *J Exp Orthop*. 2016;3(1):36. doi:10.1186/s40634-016-0072-2

98. Szczekutowicz TP, DuPlissis A, Pickhardt PJ. Variation in CT number and image noise uniformity according to patient positioning in MDCT. *American Journal of Roentgenology*. 2017;208(5):1064-1072. doi:10.2214/AJR.16.17215
99. Michalski AS, Edwards WB, Boyd SK. The Influence of Reconstruction Kernel on Bone Mineral and Strength Estimates Using Quantitative Computed Tomography and Finite Element Analysis. *Journal of Clinical Densitometry*. 2019;22(2):219-228. doi:10.1016/j.jocd.2017.09.001
100. Giambini H, Dragomir-Daescu D, Huddleston PM, Camp JJ, An KN, Nassr A. The Effect of Quantitative Computed Tomography Acquisition Protocols on Bone Mineral Density Estimation. *J Biomech Eng*. 2015;137(11):114502. doi:10.1115/1.4031572
101. Lee YH, Kim JJ, Jang IG. Patient-Specific Phantomless Estimation of Bone Mineral Density and Its Effects on Finite Element Analysis Results: A Feasibility Study. *Comput Math Methods Med*. 2019;2019:1-10. doi:10.1155/2019/4102410
102. Lee DC, Hoffmann PF, Kopperdahl DL, Keaveny TM. Phantomless calibration of CT scans for measurement of BMD and bone strength—Inter-operator reanalysis precision. *Bone*. 2017;103:325-333. doi:10.1016/j.bone.2017.07.029
103. Lee S, Gong H hyeon, Hyun J young, et al. Estimation of stature from femur length measured using computed tomography after the analysis of three-dimensional characteristics of femur bone in Korean cadavers. *Int J Legal Med*. 2017;131(5):1355-1362. doi:10.1007/s00414-017-1556-z
104. Kaesmacher J, Liebl H, Baum T, Kirschke JS. Bone mineral density estimations from routine multidetector computed tomography: A comparative study of contrast and calibration effects. *J Comput Assist Tomogr*. 2017;41(2):217-223. doi:10.1097/RCT.0000000000000518
105. Mueller DK, Kutscherenko A, Bartel H, Vlassenbroek A, Ourednicek P, Erckenbrecht J. Phantom-less QCT BMD system as screening tool for osteoporosis without additional radiation. *Eur J Radiol*. 2011;79(3):375-381. doi:10.1016/j.ejrad.2010.02.008

106. Budoff MJ, Malpeso JM, Zeb I, et al. Measurement of phantomless thoracic bone mineral density on coronary artery calcium CT scans acquired with various CT scanner models. *Radiology*. 2013;267(3):830-836. doi:10.1148/radiol.13111987
107. Mao SS, Luo Y, Fischer H, Budoff MJ, Li D. Routine coronary calcium scan can precisely measure vertebral bone density without a quantitative calibration phantom. *J Comput Assist Tomogr*. 2016;40(1):126-130. doi:10.1097/RCT.0000000000000330
108. Cropp RJ, Seslija P, Tso D, Thakur Y. Scanner and kVp dependence of measured CT numbers in the ACR CT phantom. *J Appl Clin Med Phys*. 2013;14(6):338-349. doi:10.1120/jacmp.v14i6.4417
109. Boden SD, Goodenough DJ, Stockham CD, Jacobs E, Dina T, Allman RM. Precise measurement of vertebral bone density using computed tomography without the use of an external reference phantom. *J Digit Imaging*. 1989;2(1):31-38. doi:10.1007/BF03168013
110. Mao SS, Li D, Luo Y, Syed YS, Budoff MJ. Application of quantitative computed tomography for assessment of trabecular bone mineral density, microarchitecture and mechanical property. *Clin Imaging*. 2016;40(2):330-338. doi:10.1016/j.clinimag.2015.09.016
111. Pickhardt PJ, Lee LJ, Muñoz Del Rio A, et al. Simultaneous screening for osteoporosis at CT colonography: Bone mineral density assessment using MDCT attenuation techniques compared with the DXA reference standard. *Journal of Bone and Mineral Research*. 2011;26(9):2194-2203. doi:10.1002/jbmr.428
112. Christensen DL, Nappo KE, Wolfe JA, et al. Ten-year fracture risk predicted by proximal femur Hounsfield units. *Osteoporosis International*. 2020;31(11):2123-2130. doi:10.1007/s00198-020-05477-y
113. Christensen DL, Nappo KE, Wolfe JA, et al. Proximal femur hounsfield units on CT colonoscopy correlate with dual-energy X-ray absorptiometry. *Clin Orthop Relat Res*. 2019;477(4):850-860. doi:10.1097/CORR.0000000000000480
114. Habashy AH, Yan X, Brown JK, Xiong X, Kaste SC. Estimation of bone mineral density in children from diagnostic CT images: A comparison of methods with and

- without an internal calibration standard. *Bone*. 2011;48(5):1087-1094. doi:10.1016/j.bone.2010.12.012
115. Summers RM, Baecher N, Yao J, et al. Feasibility of simultaneous CT colonography and fully-automated bone mineral densitometry in a single examination. *J Comput Assist Tomogr*. 2011;35(2):212-216. doi:10.1097/RCT.0b013e3182032537
 116. Therkildsen J, Thygesen J, Winther S, et al. Vertebral Bone Mineral Density Measured by Quantitative Computed Tomography With and Without a Calibration Phantom: A Comparison Between 2 Different Software Solutions. *Journal of Clinical Densitometry*. 2018;21(3):367-374. doi:10.1016/j.jocd.2017.12.003
 117. Fidler JL, Murthy NS, Khosla S, et al. Comprehensive Assessment of Osteoporosis and Bone Fragility with CT Colonography. *Radiology*. 2016;278(1):172-180. doi:10.1148/radiol.2015141984
 118. Saffarzadeh M, Hightower RC, Talton JW, Miller AN, Stitzel JD, Weaver AA. Multicenter analysis of CIREN occupant lumbar bone mineral density and correlation with age and fracture incidence. *Traffic Inj Prev*. 2016;17:34-41. doi:10.1080/15389588.2016.1203429
 119. Boomsma MF, Slouwerhof I, van Dalen JA, et al. Use of internal references for assessing CT density measurements of the pelvis as replacement for use of an external phantom. *Skeletal Radiol*. 2015;44(11):1597-1602. doi:10.1007/s00256-015-2206-5
 120. Weaver AA, Beavers KM, Hightower RC, Lynch SK, Miller AN, Stitzel JD. Lumbar Bone Mineral Density Phantom-less Computed Tomography Measurements and Correlation with Age and Fracture Incidence. *Traffic Inj Prev*. 2015;16(0 2):153-160. doi:10.1016/j.physbeh.2017.03.040
 121. Gudmundsdottir H, Jonsdottir B, Kristinsson S, Johannesson A, Goodenough D, Sigurdsson G. Vertebral bone density in Icelandic women using quantitative computed tomography without an external reference phantom. *Osteoporos Int*. 1993;3(2):84-89. doi:10.1007/BF01623378
 122. Therkildsen J, Thygesen J, Winther S, et al. Vertebral Bone Mineral Density Measured by Quantitative Computed Tomography With and Without a Calibration

- Phantom: A Comparison Between 2 Different Software Solutions. *Journal of Clinical Densitometry*. 2018;21(3):367-374. doi:10.1016/j.jocd.2017.12.003
123. Aubrey J, Esfandiari N, Baracos VE, et al. Measurement of skeletal muscle radiation attenuation and basis of its biological variation. *Acta Physiologica*. 2014;210(3):489-497. doi:10.1111/apha.12224
 124. Wilmink JT, Roukema JG, Van den Burg W. Effects of IV contrast administration on intraspinal and paraspinal tissues: A CT study. 2. Visual assessment. *American Journal of Neuroradiology*. 1988;9(1):191-193.
 125. White DR. Tissue substitutes in experimental radiation physics. *Med Phys*. 1978;5(6).
 126. Rho JY, Hobatho M c., Ashman RB. Relations of mechanical properties to density and CT numbers in human bone. *Med Eng Phys*. 1995;17(5):347-355. doi:10.1007/978-94-010-0367-4_13
 127. Ciarelli MJ, Goldstein SA, Kuhn JL, Cody DD, Brown MB. Evaluation of orthogonal mechanical properties and density of human trabecular bone from the major metaphyseal regions with materials testing and computed tomography. *Journal of Orthopaedic Research*. 1991;9(5):674-682. doi:10.1002/jor.1100090507
 128. McBroom RJ, Hayes WC, Edwards WT, Goldberg RP, White AA. Prediction of vertebral body compressive fracture using quantitative computed-tomography. *J Bone Joint Surg*. 1985;67-A(8):1206-1214.
 129. Esses SI, Lotz JC, Hayes WC. Biomechanical properties of the proximal femur determined in vitro by single-energy quantitative computed tomography. *Journal of Bone and Mineral Research*. 1989;4(5):715-722. doi:10.1002/jbmr.5650040510
 130. Lotz, Jeffrey; Gerhart, Tobin; Hayes W. Mechanical Properties of Trabecular Bone from the Proximal Femur: A Quantitative CT Study. *J Comput Assist Tomogr*. 1990;14(1):107-114.
 131. Kaneko TS, Bell JS, Pejicic MR, Tehranzadeh J, Keyak JH. Mechanical properties, density and quantitative CT scan data of trabecular bone with and without metastases. *J Biomech*. 2004;37(4):523-530. doi:10.1016/j.jbiomech.2003.08.010

132. Vivanco JF, Burgers TA, Garcia-Rodriguez S, et al. Estimating the density of femoral head trabecular bone from hip fracture patients using computed tomography scan data. *Institution of Mechanical Engineers*. 2014;228(6):616-626. doi:10.1177/0954411914540285
133. Ford CM, Keaveny TM, Hayes WC. The effect of impact direction on the structural capacity of the proximal femur during falls. *Journal of Bone and Mineral Research*. 1996;11(3):377-383. doi:10.1002/jbmr.5650110311
134. Mindways Software Inc. Qct pro TM. Published online 2013.
135. Glüer CC, Blake G, Lu Y, Blunt¹ BA, Jergas¹ M, Genant¹ HK. Accurate assessment of precision errors: How to measure the reproducibility of bone densitometry techniques. *Osteoporosis International*. 1995;5(4):262-270. doi:10.1007/BF01774016
136. Helgason B, Perilli E, Schileo E, Taddei F, Brynjólfsson S, Viceconti M. Mathematical relationships between bone density and mechanical properties: A literature review. *Clinical Biomechanics*. 2008;23(2):135-146. doi:10.1016/j.clinbiomech.2007.08.024
137. Keyak JH, Kaneko TS, Tehranzadeh J, Skinner HB. Predicting proximal femoral strength using structural engineering models. *Clin Orthop Relat Res*. 2005;(437):219-228. doi:10.1097/01.blo.0000164400.37905.22
138. Bhattacharya P, Altai Z, Qasim M, Viceconti M. A multiscale model to predict current absolute risk of femoral fracture in a postmenopausal population. *Biomech Model Mechanobiol*. 2019;18(2):301-318. doi:10.1007/s10237-018-1081-0
139. Aldieri A, Bhattacharya P, Paggiosi M, et al. Improving the Hip Fracture Risk Prediction with a Statistical Shape-and-Intensity Model of the Proximal Femur. *Ann Biomed Eng*. 2022;50(2):211-221. doi:10.1007/s10439-022-02918-z
140. Aldieri A, Terzini M, Audenino AL, et al. Personalised 3D Assessment of Trochanteric Soft Tissues Improves HIP Fracture Classification Accuracy. *Ann Biomed Eng*. 2022;50(3):303-313. doi:10.1007/s10439-022-02924-1
141. Cristofolini L, Baleani M, Schileo E, et al. Differences between contralateral bones of the human lower limbs: A multiscale investigation. *J Mech Med Biol*. 2014;14(3). doi:10.1142/S0219519414500328

Appendix – Clinical handout for phantomless calibration

This document is intended to highlight specific details from “Evaluation of patient tissue selection methods for deriving equivalent density calibration for femoral bone quantitative CT analyses” (Winsor et al. 2021). Please note some reference in here that are relevant may also come from Lee/Keaveny et al. 2017 or Michalski/Boyd et al. 2020.

When forming a cohort, results will be more precise if patients are grouped by specific CT scan acquisition and reconstruction parameters (Bligh 2009). Further to this, I also recommend stratifying according to variables known to impact the calibration equation so that

Variables known to impact the calibration equation in general include scan voltage, slice thickness, and reconstruction kernel. Note that my cohort only had sufficient validation data for 120 kVp, 1.25 mm slice thickness with 50% overlap, and a standard reconstruction kernel. (Other lurking currently unknown impactful variables may appear.)

Once you’ve identified your desired study cohort, the next step is segmentation. Segmentation can be challenging, for this method, I recommend finding the approximate centre of the femoral head and trying to keep the area of interest in the same places for 5 slices in each direction. For air, urine, skeletal muscle, and adipose tissue, I recommend a diameter of about 10 pixels. Aortic Blood is a bit harder, and I found a diameter of 8 pixels led to a better measurement.

To create scan-specific calibration measurements, I linearly regressed the tissue-specific segmentation measurements [HU] (dependent variable) against the nominal density from the National Institute of Standards and Technology Tables of X-Ray Mass Attenuation Coefficients and Mass Energy Absorption Coefficients Table 2. This density calibration approach is discussed in ASTM E1935-97: Standard Test Method for Calibrating and Measuring CT Density.

Theoretical air (not segmented)	1.205
Adipose tissue	950
Aortic blood	1060
Theoretical water (not segmented)	1000
Skeletal muscle	1050

I ran this for one patient with air, aortic blood and skeletal muscle to get an example equation:

$$Y = 0.974 * x - 984.859$$

At this point, I recommend modifying the intercept by adding the values of the Rescale Intercept from the Dicom header. For GE LightSpeed scanners, this value is typically 1024. So my updated equation would be:

$$Y = 0.974 * x + 39.15.$$

For the ability to process many patients at a time, I have set up the accompanying matlab script. In the event that the rescale slope changes, the script may need to be tweaked.

Appendix – Matlab function for calibration equation derivation

```
function [slope intercept r2] = mycalibrationequation(x, y);
```

```
    [curve, G] = fit(x,y,'poly1');
```

```
    slope = curve.p1;
```

```
    intercept = curve.p2+1024;
```

```
    r2 = G.rsquare;
```

```
end
```