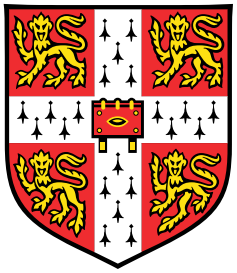
Automated Modelling of Cortical Bone from Clinical CT



Rose Alicia Pearson

Supervisor: Dr. Graham Treece

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Queens' College November 2017

I would like to dedicate this thesis to all who have taught me. Thank-you for your passion and patience. You shaped me more than you know.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole nor in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Rose Alicia Pearson
November 2017

Acknowledgements

Although a thesis has only one author, it has not been a solitary affair. I would like to thank Graham Treece, my supervisor, for providing a steady hand to guide me. Andrew Gee and Ken Poole for your interest in my work, and for sharing your respective expertise in shape modelling and medicine. Tom Turmezei and Tristan Whitmarsh for the ideas and company shared in our office. Enrico Schileo for the spine validation dataset. Patrick Gosling and Anna Langley for keeping my computer running smoothly, and Rachel Fogg, Louise Segar and Diane Hazell for prompt and consistent admin support throughout my PhD.

I would also like to thank my friends and family both here in Cambridge, and at home in New Zealand. You are the reason my life is so full. A special thanks to Andy Howell, Philip Withnall, Matt Kingston, Hannah Thompson, Harry Fogg, Cota Valdivieso, Hannah Fogg, Evan Miles and Becky Howard for making Cambridge a home away from home, and to Mathew Falloon for all the supplies - I wouldn't have made it without them. Mum and Dad thank-you for the constant support, and never judging my mistakes. And finally to my brother, Nelson we've been through a lot. I'm just so glad there is more to come.

Finally, I would like to acknowledge and thank the Commonwealth Scholarship Commission for funding my studies. This thesis would not have been possible without your support.

Publications

Journal publications arising from this dissertation

Material from Chapters 2 and 3 has been accepted for publication in: Pearson, R. A. and Treece, G. M. (2017). Measurement of the bone endocortical region using clinical CT. *Medical Image Analysis*, Accepted

Abstract

Osteoporosis is an age-related skeletal disease characterised by an increased incidence of fragility fractures, which poses a significant and increasing healthcare burden (NIH 2001). The extent of the endocortical region and cortical bone mineral density (cBMD) are of interest as both have been linked to fracture risk. Non-invasive in-vivo clinical CT-based techniques capable of measuring the cortical bone attributes of thickness, density and mass over a bone surface have already been proposed. Several studies have robustly shown these methods to be capable of producing cortical thickness measurements to a sub-millimetre accuracy, but existing methods are unable to provide high quality cBMD estimates, and are not designed to measure any attributes over the endocortical region of cortical bone.

In this thesis I develop the endocortical smoothly spatially constrained CBM (SSC-CBM) method, which is a cortical bone mapping (CBM) based technique capable of providing a smooth, spatially varying estimate of cBMD and measuring the extent of the endocortical region. Its performance is assessed using simulated QCT data from three simulated phantoms with differing bone architecture, a paired dataset of ex-vivo QCT and HR-pQCT scans across 72 proximal femurs, and a second paired dataset of ex-vivo scans across 20 vertebral bodies. The simulated data allows for inaccuracies in CBM measurements resulting from beam hardening effects to be considered for the first time: I show that beam hardening leads to an underestimation in cortical thickness and an overestimation in trabecular BMD (tBMD) and that these inaccuracies can be reduced through adjustments to the CBM optimisation process.

A new technique for analysing HR-pQCT scans was developed for this thesis: high resolution tissue classification (HRTC). It was used to validate the CBM measurements in the paired QCT and HR-pQCT datasets. HRTC acts to extract the macro-architectural cortical, endocortical, and sub-surface trabecular bone features from the micro-architectural detail included in the HR-pQCT scan. It was used in place of the well known full-width half-maximum (FWHM) and other established HR-pQCT techniques for its ability to provide localised measurements of the endocortical region, which it defines as extending from the first endocortical pore to the first trabecular pore in agreement with other HR-pQCT techniques. 72 HR-pQCT scans of proximal femurs were examined to compare the HRTC and FWHM measurements. This indicates that the HRTC method is less susceptible to cortical thickness and tBMD errors, due to high levels of beam hardening, than the FWHM method. The HRTC further distinguishes itself from the FWHM method for its ability to measure the mean cBMD, which has a greater clinical relevance than the peak cBMD measured by the FWHM method as it includes the impact of porosity.

Over QCT scans of the proximal femur, I demonstrate that the extent of the endocortical region can be measured to an accuracy of (-0.15 ± 0.71) mm, and that local cBMD measure-

ments are possible down to densities of $300\,\mathrm{mg/cm^3}$, while maintaining comparable cortical thickness, mass surface density and tBMD measurements to established techniques. I also validate CBM methods over the vertebrae for the first time and demonstrate that, in this case, cortical thickness can be measured with an accuracy of (0.10 ± 0.30) mm and endocortical thickness (-0.20 ± 0.53) mm.

Two clinical trials involving teriparatide are used to examine the ability of the endocortical SSC-CBM method to detect changes in bone architecture on typical clinical CT data over the proximal femur and lumbar spine. The first trial demonstrates, for the first time, the ability of the endocortical SSC-CBM method to detect changes in the dense cortical and endocortical bone over the proximal femur. This could be attributed to localised increases in intracortical remodelling and endosteal apposition. The second trial shows that, despite the challenging architecture of the lumbar vertebrae, the endocortical SSC-CBM method can detect regionally significant changes in the cortical and endocortical thickness. A third cross-sectional fracture discrimination trial is used to investigate regions of significant difference in the more accurate cBMD and new endocortical thickness measurement between fracture cases and healthy controls, which confirms that fracture incidence is associated with regions of significant decrease in both measures.

In conclusion, the endocortical SSC-CBM method developed in this thesis is capable of measuring the extent of the endocortical region over the proximal femur and lumbar spine and improving the quality of cBMD measurements over the proximal femur, while still measuring cortical thickness, mass surface density and tBMD of the cortical and subsurface trabecular region accurately. These measurements can be used to assess treatment effectiveness and to examine localised regions of differing bone architecture between fracture and control populations. Changes in these measurements may be attributed to changes in endosteal apposition over the proximal femur and lumbar spine, and intracortical remodelling over the proximal femur.

Table of contents

List of figures

List of tables

Nomenclature

Greek Symbols

$\mathcal{O}$	Landau big O notation	

Superscripts

$\bar{x}$	The mean of x	

Subscripts

cb	Cortical bone subscript	
db	'Dense' cortical bone subscript	
E	Endosteal edge subscript	
ec	Endocortical region subscript	
Edb	Dense endosteal edge subscript	
Etb	Trabecular endosteal edge subscript	
P	Periosteal edge subscript	
st	Soft tissue subscript	
tb	Trabecular bone subscript	

Other Symbols

$erf()$	Gauss error function	
$H()$	Unit step function	
$max(\boldsymbol{x})$	The maximum of vector $\boldsymbol{x}$	
$med(\boldsymbol{x})$	The median of vector $\boldsymbol{x}$	
$R()$	Ramp function	
x	Position along a density profile	mm

y	Density at position along density profile	mg/cm^3

Acronyms / Abbreviations

$aBMD$	Areal bone mineral density	mg/cm^2
BMD	Bone mineral density	mg/cm^3
BV/TV	Trabecular bone volume of total volume	
CBM	Cortical bone mapping	
$cBMD$	Cortical bone mineral density	mg/cm^3
CI	Confidence interval	
CSA	Cross-sectional area	mm^2
CT	X-ray computed tomography	
cTh	Cortical thickness	mm
$CTXA$	Computed tomography X-Ray absorptiometry	
$dBMD$	'Dense' cBMD	mg/cm^3
DMB	Degree of mineralisation of bone	
dTh	'Dense' cortical thickness	mm
DXA	Dual-energy X-ray absorptiometry	
$ecTh$	Endocortical thickness	mm
FBP	filtered back-projection	
$FWHM$	Full-width half-maximum	
GLM	General linear model	
GNM	Gauss-Newton method	
$HR-pQCT$	High resolution quantitative computed tomography	
$HRTC$	High resolution tissue classifier	
HU	Hounsfield unit	
LAD	Locally affine deformation	
LMM	Levenberg-Marquardt method	

MSD	Mass surface density	mg/cm^2
PSF	Point spread function	
QCT	Quantitative computed tomography	
ROI	Region of interest	
SPM	Statistical parametric mapping	
$SSC-CBM$	Smoothly spatially constrained CBM	
STD	Standard deviation	
$tBMD$	Trabecular bone mineral density	mg/cm^3
th	Thickness	mm
TZS	Transitional zone segmentation	

Chapter 1

Background

1.1 Introduction

Osteoporosis is an age-related skeletal disease characterised by a reduction in bone mineral density (BMD) and a deterioration in bone quality [27, 89]. Individuals with osteoporosis are predisposed to suffering fragility fractures [27, 70], which are caused by low impact events, such as standing, walking or falling from standing [129]. These most commonly occur at the hip, spine, distal forearm and proximal humerus [61], which are shown in Figure 1.1. While recovery from forearm and humerus fractures is generally complete, hip fractures are associated with slow and often incomplete recoveries, and increased rates of mortality and morbidity [61, 64]. Vertebral fractures can result in acute pain and loss of function with subsequent fractures producing increasing disability [61].

According to various prospective studies of northern European populations, the life-time risk of suffering osteoporotic fractures at age 50 ranges from 20–25 % for men and 45–55 % for women [1, 64, 71]. The high rates of fracture incidence combined with the associated morbidity and mortality [61] make osteoporosis a significant health care burden [23, 52, 61, 135], which is most significant for hip [23, 52, 135] and vertebral fractures [23, 52]. This burden is set to increase with the ageing world population, and so fracture risk prediction, treatment assessment, and the monitoring of osteoporosis disease progression are all of particular research interest.

Bone strength is determined by its material composition and architecture [39, 111], which are compromised in individuals with osteoporosis [61, 64]. In this thesis, I develop and validate an improved method of measuring localised bone architecture from non-invasive quantitative computed tomography (QCT) scans that can be used in a clinical setting. Validation was performed over the proximal femur and lumbar spine. The technique was then used to investigate the architectural changes associated with teriparatide therapy [28] over the lumbar spine and proximal femur, and the architectural traits associated with increased fracture risk over the proximal femur.

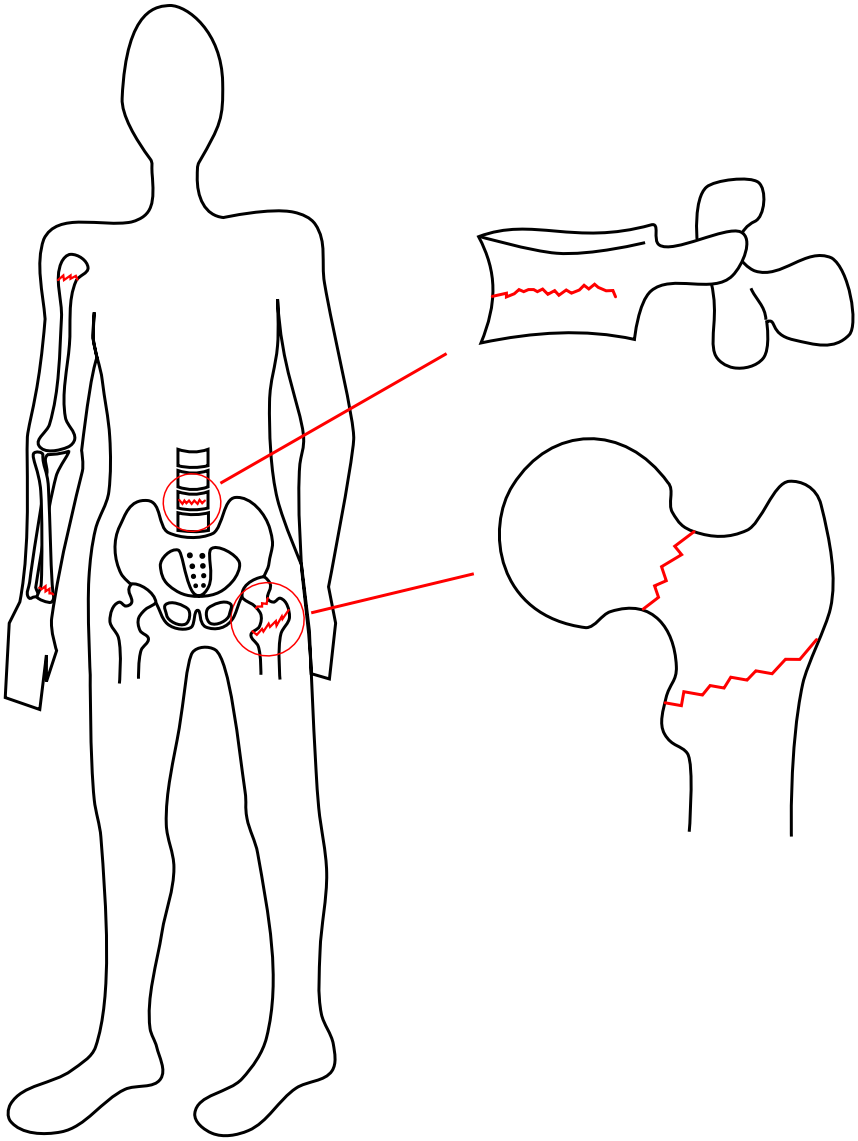


Fig. 1.1 Sites of common fragility fractures with the vertebrae and proximal femur shown in detail.

1.2 Bone architecture and composition

The skeleton is a framework of interconnected bones through which movement is achieved. Ideally, the individual bones should be as light as possible, while still being strong enough to withstand the tensile and compressive loads required to achieve movement [14, 111]. Bone tissue can be categorised by its material composition and architecture. Materially it is composed of a type I collagen scaffold in which calcium hydroxyapatite crystals are deposited [14, 111]. Its architecture is determined by the arrangement and connectivity of the collagen scaffold at the microscopic and macroscopic level [111]. The density of calcium deposited on the scaffold is termed the degree of mineralisation of bone (DMB). This is

typically around 1100 mg/cm^3 but it varies with the rate of remodelling and the function of the bone [4, 10, 39]. Different imaging techniques can be used to image different aspects of bone. In all techniques, there is a trade-off between resolution and invasiveness, and a limit of what can be imaged. Clinical QCT scans are currently the best placed non-invasive technology with which the bone architecture and mineral content of the proximal femur and lumbar spine can be measured in a clinical setting.

1.2.1 Cortical and trabecular bone

Bone is classified as either cortical or trabecular based upon its morphology [76, 109], which differs as shown in Figure 1.2. Cortical bone forms the outer envelope of bone, and provides most of its overall stiffness [14, 95]. It has a dense laminar structure that is interspersed by Haversian and Volkman canals [116, 138], which form pathways for nutrient transport and sites for remodelling. In contrast, trabecular bone is a porous scaffold of interconnected rods and plates that provides the inner structure and compressive strength of bone [14, 111, 117]. The individual pores contain hematopoietic marrow, while the overall scaffold structure surrounds the stromal marrow contained in the medullary cavity [76].

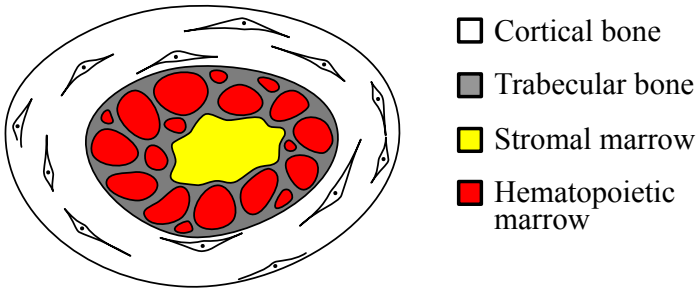


Fig. 1.2 Bone composition and architecture.

Cortical bone makes up about 80 % of the skeleton by mass, although this percentage varies with function. Its thickness also varies over the skeleton with thicknesses of 0.15 – 0.85 mm for vertebrae [35, 37, 106], and 0.15–6 mm for the proximal femur [123–125]. Porosity is the proportion of non-ossified tissue within the total bone volume. In cortical bone, it is typically around 15 % [109]. This is affected by the quantity of Haversian and Volkman canals, and the size and quantity of any remodelling cavities [138]. The remaining 20 % of bone is trabecular, which can be categorised by the thickness and connectivity of individual trabeculae, and the density and porosity of the overall scaffold. The thickness of individual trabecular rods typically varies between 100–400 μm [45]. The spongy structure of trabecular bone leads to porosities of 70 % [109], and a corresponding decrease in the overall density. The connectivity of trabecular bone is typically defined by the number and anisotropy of the connections, which can only be measured by imaging technologies capable of detecting individual trabeculae.

1.2.2 Bone remodelling

Bone remodelling is the process by which bone repairs and renews itself through primary remodelling and secondary remodelling [76]. In primary remodelling damaged bone is removed and replaced by osteoclasts, osteoblasts and osteocytes, which are specialised cells. This occurs in four stages as shown in Figure 1.3: activation, resorption, reversal and formation. Secondary remodelling is the gradual increase in mineralisation that occurs between the conclusion of the formation stage and the beginning of the next activation stage [14, 111].

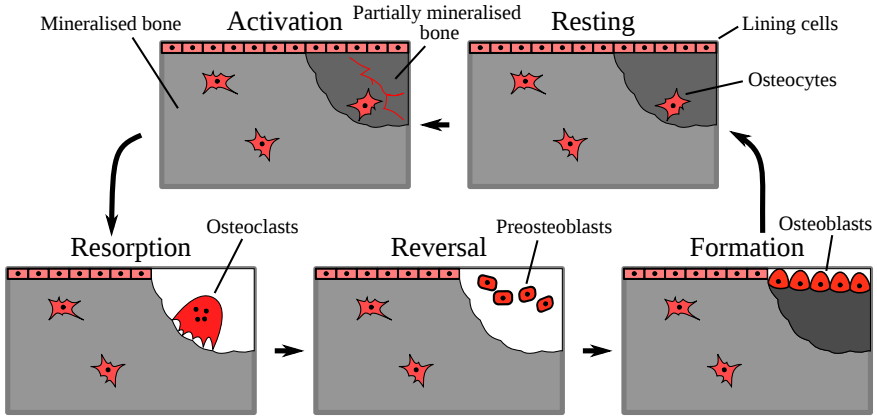
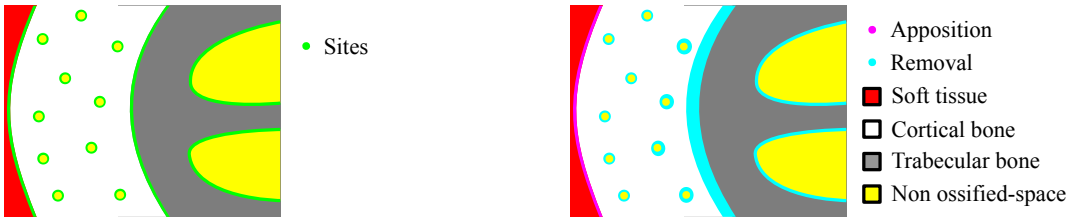


Fig. 1.3 Remodelling of bone.

Activation is the first stage of the primary bone remodelling process, and it is precipitated by osteocytes that have detected micro-damage or changing loading patterns. In response pre-osteoclasts begin fusing to form multi-nucleated osteoclasts. In the resorption phase, osteoclasts orchestrate the resorption of mineralised bone by adhering to the bone surface then producing an acidic solution and enzymes. This is followed by the reversal phase where unspecialised cells involved in the bone remodelling process remove the debris produced during the resorption phase. The final stage of the remodelling process is formation, where osteoblasts are recruited to the reabsorbed area and produce an un-calcified bone matrix then promote its mineralisation [14, 107, 134].

The periosteal and endosteal surfaces are thin membranes of connective tissue that control the synthesis of collagen and mineralised bone over the outer and inner surfaces of cortical bone [76, 134]. The periosteal membrane controls the level of remodelling along the outer cortical surface and the endosteal membrane controls the level along the inner cortical surface [76]. Throughout this thesis, the endocortical region is defined as the region of cortical and trabecular bone where remodelling is primarily controlled by the endosteal membrane. Remodelling also occurs along the surface of individual trabeculae in trabecular bone and the Haversian and Volkman canals within cortical bone as shown in Figure 1.4a. The ratio of formation to resorption in each remodelling cycle determines the net change in bone along these surfaces [14]. This varies with age, but tends to become negative in later life [92, 111] especially among post-menopausal women [36, 41, 122]. A prolonged negative balance slowly leads to an erosion of individual trabeculae, thinning of cortical bone, and



(a) Sites of bone remodelling (b) The effects of bone remodelling later in life

Fig. 1.4 The sites and effects of bone remodelling in later life.

the gradual formation of remodelling cavities within cortical bone [92, 111, 139]. It has been suggested that remodelling is greatest on the endocortical surface in later life, which would lead to increased porosity concentrated within the endocortical region [111, 122]. A slight positive balance along the periosteal surface has also been observed in some studies in later life [122]. This would lead to gradual widening of the bone through a process called periosteal apposition in an adaptation thought to provide improved bending strength in response to decreased cortical thickness [111]. The cumulative effect of these theorised adaptations is illustrated in Figure 1.4b.

1.3 Imaging techniques

Various techniques exist for measuring bone structure and composition including: dual-energy X-ray absorptiometry (DXA), QCT, high resolution peripheral QCT (HR-pQCT), μCT, μMRI and histological methods. Of these, DXA and QCT can be used in the clinical study of osteoporosis over the proximal femur and lumbar spine, while HR-pQCT can be used to provide cadaveric validation of the QCT techniques over these regions.

1.3.1 DXA imaging

DXA is an imaging technique that produces a two-dimensional planar projection of the three-dimensional BMD. It is used to measure the areal BMD (aBMD), which is the mean two-dimensional projection of BMD over a section of bone [90, 135]. This has consistently been identified as the single greatest indicator of bone strength [2, 48, 60, 63, 70], and so it has become the standard diagnostic tool for osteoporosis with the World Health Organisation (WHO) publishing guidelines defining osteoporosis and osteopenia in terms of the aBMD T-score within the femoral neck [48, 135].

Unfortunately, aBMD measurements predict fracture risk with a low sensitivity [62] with fewer than half those suffering osteoporotic fractures exhibiting aBMD defined osteoporosis [58, 110]. Studies show aBMD only accounts for 60–70 % of the variations in bone strength [2], as the inherent two-dimensional nature of DXA renders it unable to measure three-dimensional features of bone architecture [48, 90] or focal variations in bone structure.

1.3.2 QCT imaging

QCT is a medical imaging modality that provides three-dimensional BMD measurements [24, 32, 44, 45] from a standard X-ray computed tomography scanner by using a density calibration standard to convert the CT image values into BMD values [44]. The calibration standard is typically generated by scanning a calibration phantom with known BMD values [8, 32, 69]. X-ray CT imaging produces a regular volumetric voxel grid of X-ray attenuations that are mathematically reconstructed from a series of X-ray images taken over a range of orientations [53]. Typically the resulting voxel resolution is no better than 0.3 mm in-plane and 1.0 mm out-of-plane [45].

QCT can be used to measure many three dimensional aspects of architecture [14, 44, 70] including: trabecular density [123]; cortical thickness [75, 90, 123–125]; cortical density [123]; and cortical mass surface density [123, 125]. Unfortunately, QCT resolution is still too low to directly resolve most cortical bone and trabecular bone features, which become indistinct after blurring. This makes the clinical measurement of these features challenging.

1.3.3 HR-pQCT imaging

HR-pQCT uses the same basic approach as QCT to produce three-dimensional BMD measurements at a higher resolution and with an acceptable radiation dosage over the distal radius and tibia in-vivo [12, 26]. The higher resolution is in part possible as HR-pQCT scans are generated from an increased number of X-ray projections. The still acceptable radiation dosage is achieved by limiting the field of view (FoV) of the scanner. HR-pQCT scans have a nominal isotropic resolution of approximately 80–90 µm [17, 32, 48, 68]. The restrictive FoV of HR-pQCT means it cannot be used in-vivo over the spine and proximal femur, but it can be used ex-vivo to produce paired HR-pQCT to QCT scans that can be used to validate the quality of the QCT based CBM measurement techniques.

HR-pQCT scans have been used to generate various macro- and micro-structural measurements of trabecular, cortical and endocortical features [14, 17–19, 32, 48, 66, 68, 72, 73, 86, 123–125, 138, 140]. These measurements need to be localised if they are to be compared with the surface based measurements produced with the CBM method. In previous validation studies of the CBM method, HR-pQCT measurements were achieved with the full-width half-maximum (FWHM) technique, which is capable of measuring the cBMD, tBMD and cortical thickness. A new technique is developed in this thesis to provide localised measurements of additional features.

1.4 QCT architecture measurements

Two common approaches exist for measuring architectural features of bone from QCT scans. The first uses morphological operations to compartmentalise the scan into cortical and trabecular regions, which can be used to derive mean architectural measures over the scan. The commercially available computed tomography X-ray absorptiometry (CTXA)

hip exam analysis protocol by Mindways[1] [21, 65] is arguably the most widely adopted example of this approach. The second uses a parametrised model of cortical bone to produce localised measurements of cortical and subsurface trabecular bone features over the periosteal surface [103, 118, 123]. It is this model-based approach that is the focus of this thesis as its ability to produce localised measurements allows for a more detailed study of the changes in bone architecture in individuals with osteoporosis.

1.4.1 Cortical bone mapping (CBM)

Cortical bone mapping (CBM) is a model-based technique in which many local measurements are made over the surface of the bone and combined to provide surface maps of the measured features. This allows focal regional variations between populations to be detected. Model based methods use a transect of cortical and sub-surface trabecular bone sampled from a QCT scan, a parametrised estimate of the underlying density, and an estimation of the imaging blur, to model the underlying cortical and subsurface trabecular features using deconvolution [123]. This is shown in the middle section of Figure 1.5. These methods do not attempt to resolve individual trabeculae and instead aim to measure the overall localised subsurface trabecular density.

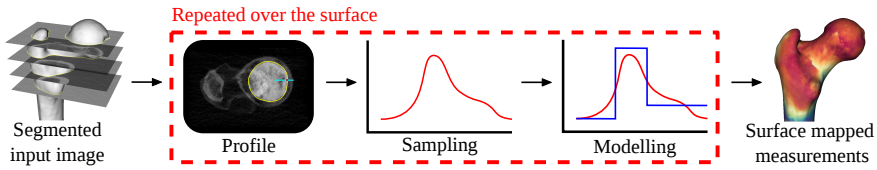


Fig. 1.5 The basic cortical bone mapping (CBM) pipeline.

The three tier model in Figure 1.5 has previously been used to represent the density in profile across the cortex, at each surface location [123–125]. This model measures the cortical thickness, the cortical bone mineral density (cBMD) and the trabecular bone mineral density (tBMD), where these measurements are single aggregate values across the entire cortical or trabecular region included in the profile. These measurements can be used to calculate the mass surface density (MSD), which is more resistant to errors stemming from image blur than other cortical measurements [123, 125].

The CBMv2 method is the current best CBM implementation. It can be used with the three tier rectangle model shown in the modelling section of Figure 1.5 to measure cortical thickness to sub-millimetre accuracy and tBMD to within $50\,\mathrm{mg/cm^3}$ over most of the proximal femur, but has a limited ability to adapt to local variations in cBMD [123]. Despite the additional information measured by this model, these techniques are only marginally more effective than aBMD-based methods at fracture risk prediction [99, 126, 137]. This indicates more features contributing to bone strength need to be considered.

[1] CTXA-Hip™ by Mindways Software Inc. http://qct.com/home/products/cxta-hip/

1.4.2 Measuring the endocortical region

The endocortical region is an area of clinical interest, as it has been suggested that it is the primary site of remodelling in later life when most fragility fractures occur [111, 122, 138]. Recently remodelled bone has a reduced DMB [9, 39], and remodelling in later life is associated with a negative balance which leads to increased porosity [36, 41, 92, 122]. Together these effects would lead to a decrease in BMD over the endocortical region. This has been observed with the gradual trabecularisation of the cortical bone, where cortical bone takes on the appearance if not the function of trabecular bone [138, 139]. Figure 1.6 shows one possible manifestation of this process with progressively decreasing BMD across an endocortical region that is centred about the endosteal surface. I propose an adaptation of the rectangle model to measure the transition in density across the endosteal surface separately to the periosteal surface as a gradual and approximately linear density change, which is also shown in Figure 1.6.

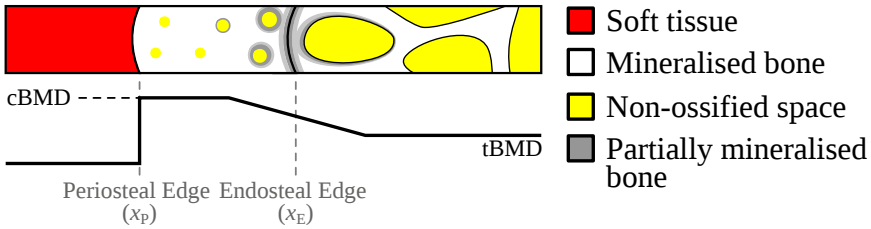


Fig. 1.6 Bone macro-structure and the endocortical model.

1.4.3 Measuring cBMD

cBMD is a significant indicator of bone strength and fracture risk [3, 39, 82]. It is affected by both the DMB and porosity of cortical bone [3, 9, 39, 82]. cBMD increases with DMB but decreases with porosity. Fracture risk generally decreases with DMB and increases with porosity, so a decrease in cBMD due to either mechanism increases fracture risk. The relative contributions of each to cBMD cannot be distinguished at the comparatively low resolution of QCT; but, it is possible to measure the average cBMD, which is still a significant indicator of bone strength and fracture risk [3, 39, 82]. Improving the quality of this measurement is another focus of this thesis.

1.5 Research overview

The aim of this work is to improve the CBM technique to better reflect the cortical and subsurface trabecular features of bone by more accurately adapting to variations in the cortical density, and by adapting the model to include a measure of the endocortical region as shown in Figure 1.6.

1. **Clinical QCT methods**
   The CBM method is extended to consider both the endocortical region and regional

variations in cBMD. A simple simulator is extended, so that it can simulate QCT scans of several synthetic datasets that represent a wide range of bone architectures. These simulated scans are used to evaluate the performance of the extended CBM method.

2. **HR-pQCT cadaveric validation of CBM**
   Techniques for evaluating the cortical and subsurface trabecular macro-structural features from HR-pQCT are developed and evaluated over a cadaveric HR-pQCT dataset. The performance of the extended CBM method is then evaluated using paired cadaveric QCT and HR-pQCT datasets of the spine and proximal femur.

3. **Clinical applications**
   The clinical relevance of the new CBM method is established by investigating its ability to measure changes in bone architecture over the proximal femur and lumbar spine in response to pharmaceutical intervention, and by investigating its ability to detect differences in the bone architecture of individuals who have suffered a fragility fracture.

4. **Conclusions**
   The performance and clinical applicability of the new technique are discussed before avenues for further research are highlighted.

Chapter 2

QCT and CBM assessment

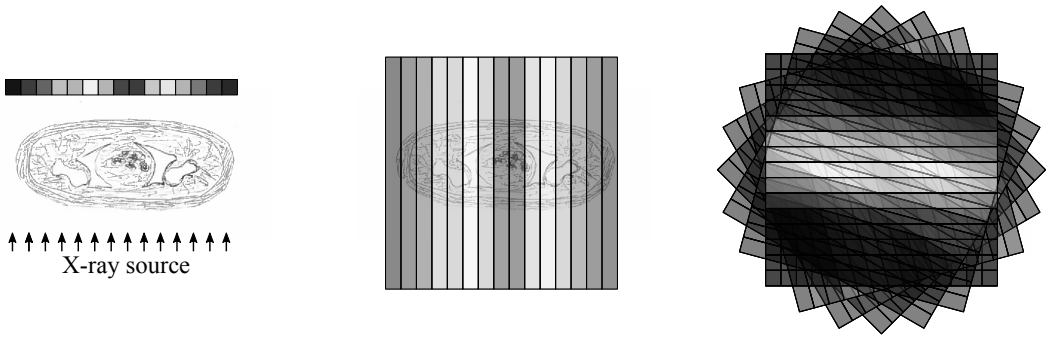
2.1 Introduction

Trabecular bone, cortical bone and the endocortical region are all important contributors to bone strength [60, 117, 122, 130]. The strength of trabecular bone can be characterised by its density and interconnectivity [2, 60], while the strength of cortical bone and the endocortical region can be characterised by their density, thickness and porosity [2, 6, 33, 75, 88, 124]. The aim of this chapter is to improve on existing techniques for measuring these features from clinical QCT, and to assess the performance of the resulting method by applying it over simulated QCT scans of synthetic datasets with known bone architectures.

2.1.1 CT imaging

CT imaging is the underlying process used to generate QCT and HR-pQCT images. It produces three-dimensional voxel-based measurements. This process can be considered as occurring in two stages with image acquisition followed by image reconstruction. During image acquisition, a series of X-ray images are taken at a variety of orientations. In an X-ray image, X-rays are uniformly applied perpendicular to the imaging plane. The proportion of X-rays to penetrate the subject at each position in the image plane is recorded to produce a two-dimensional image as shown in Figure 2.1a. Different materials absorb different amounts of X-rays and produce different intensity shadows that depend on the internal structure of the image subject as shown in Figure2.1a–2.1b. Image reconstruction is the process by which the shadows from X-rays taken at different orientations are combined to define the underlying structures internal to the image subject. Reconstruction is generally achieved with some form of filtered back projection (FBP) which is illustrated at its most simple in Figure 2.1c, where each voxel is the normalised sum of all the projections that pass through it.

CT reconstruction produces voxel measurements, which represent the linear attenuation (μ) of the material within the voxel. A material's linear attenuation is the amount an X-ray beam is attenuated per unit distance as it travels through that material. The X-ray beams used in CT imaging contain a spectrum of different energies, which experience differing levels of linear attenuation. Broadly the linear attenuation caused by a material decreases with



(a) X-ray penetration (b) X-ray projections (c) Simple back projection

Fig. 2.1 A two-dimensional representation of X-ray based Computed Tomography.

increasing X-ray energy [57]. This causes the mean energy of an X-ray beam to increase as it penetrates through a material, which leads to a decrease in the net linear attenuation experienced by the X-ray beam. This phenomenon is called beam hardening, and it causes the linear attenuation measured by CT reconstruction to decrease with penetration depth. This is discussed in more detail alongside the CT imaging process in Section 2.3.

As the linear attenuation of a material varies with the energy of the X-ray, CT scans of the same material acquired with X-ray beams that have different energy profiles will measure different linear attenuations [57]. Ideally, the voxel measurements would be independent of such acquisition details. This is achieved in QCT and HR-pQCT by using a density calibration standard to convert linear attenuation into BMD [32]. This is attractive as BMD is an X-ray independent material property that has also been linked to fracture risk and bone health [2, 13, 60, 61]. The density calibration standard is generally experimentally derived by scanning a phantom with a known BMD profile and relating the experimentally measured linear attenuations to the known BMD values of the phantom [8, 32, 69]. Commonly used QCT density phantoms are discussed in more detail in Section 2.3.1.

2.1.2 The partial volume effect (PVE) and edge blur

Limitations in the X-ray beam, the X-ray detectors and image reconstruction algorithms introduce blur into QCT scans [33, 49, 51], which produces systematic errors in the measurement of thin structures due to partial volume effects (PVE) [60], and edge blur [114, 115]. A blurred image can be modelled as a convolution between the imaged subject and the point spread function (PSF) of the imaging system as shown in Figure 2.2 [33, 102]. Generally, the image blur can be approximated as smooth with finite support, which means it occurs over a finite range. The blur is generally isotropic in the x–y plane, but anisotropic between the x–y plane and z-axis due to differences between the projection angles and the slice spacing.

PVE occurs when small high contrast objects are ‘washed out’ by the surrounding values [105]. This occurs when the support of the PSF is greater than the width of the object, so it is always being averaged with some of its surroundings. This causes the density of the thin structure to be underestimated [33, 88], as illustrated by Figure 2.3c. Edge blurring

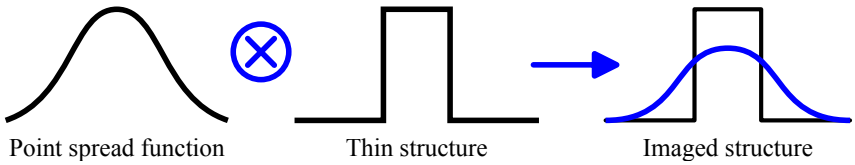


Fig. 2.2 The imaging process: convolution of a thin structure and the imaging system PSF.

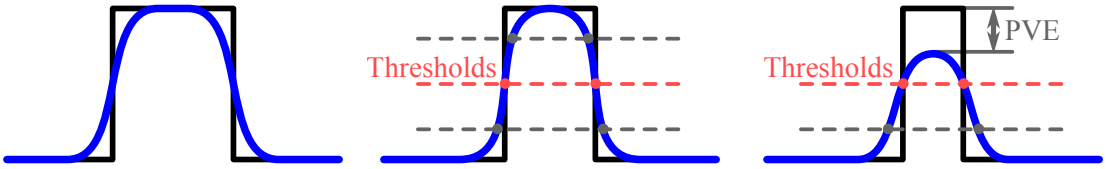


Fig. 2.3 PVE and edge blur resulting from the interplay between structure width and the support of the PSF in an image system. Correct thresholds shown in red.

occurs at sharp contrast boundaries and produces a smooth transition from low to high values also shown in Figure 2.3. This leaves the actual edge location ambiguous [115], which can lead to under- or overestimation of the thickness depending upon the threshold value used to define the edge location as shown in Figure 2.3b–2.3c.

Previous studies [6, 78, 112] have shown the 'correct' edge locations can be detected with a threshold mid-way between the edge values as exactly half the PSF is convolved with the high edge and half with the low edge at the edge boundary, but only if the PSF is symmetric. The value at the edge location can remain unchanged in the presence of PVE as shown in Figure 2.3c. This occurs when the thin structure is greater than half the width of the PSF support, so the PSF does not extend beyond the other side of the thin structure when it is centred on the edge of the thin structure. It follows that the edge value will be reduced for thinner structures as it will also be impacted by the values beyond the far edge of the thin structure. As such, thresholding can be used to identify the width of a thin blurred structure, but the threshold value and its relationship to the upper and lower values of the structure will vary when the structure's width is less than that of the support of the PSF.

2.1.3 Thresholding and morphological operations

Early methods of cortical bone thickness and density calculations relied extensively on thresholding [16, 49, 50, 75, 97, 105] and morphological operations [16]. Morphological operations use finite pixel filters to alter the connectivity or shape of objects in an image. Thresholding at its most simple is the use of a single globally fixed value to label pixels: all image pixels with a lower value are labelled 'not bone', while all others are labelled 'bone' [51]. Density can then be estimated as either the peak or mean value of the resulting 'bone' pixels. Thickness calculations prove more problematic due to the often fragmented boundary of the 'bone' region. Two early studies only produced a single mean measure of thickness by either estimating the cortex as a ring using two concentric circles and calculating

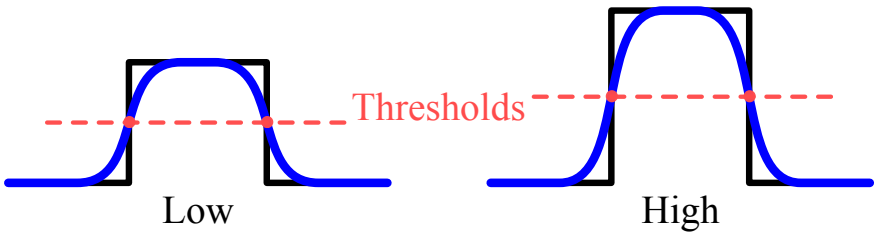


Fig. 2.4 Variations in the edge values of structures with two different densities.

the thickness as the difference in their radii [50], or calculating thickness using a ratio of the trabecular and cortical area [75].

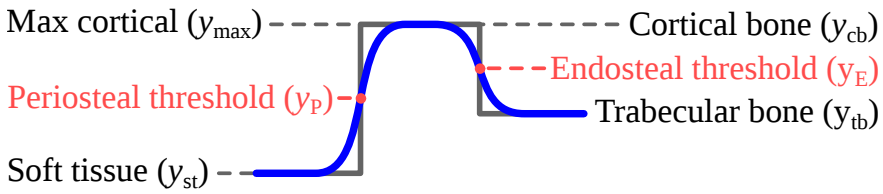
Two later studies attempted to address the 'fragmented boundary' issue. One used an additional fixed threshold in combination with morphological operations [16], while the other used an adaptive classifier [97]. These methods aimed to promote cortical connectivity while reducing trabecular connectivity. Both reported improvements, but still struggled with robustness. As a result, the adaptive classifier method required a manual correction stage [97]. The performance of these fixed thresholding methods was assessed in several clinical studies [16, 33, 49, 50, 75, 97, 105]. Two studies compared thresholding with manual techniques [16, 75], and showed the thickness measurements correlated well between techniques, but the relationship was weaker for thin structures. Hangartner et al. [50] performed a more rigorous assessment using a variable thickness fixed density phantom. This showed that the fixed thresholding method produced accurate measurements over thick regions, but systematically underestimated both the density and thickness measurements in thin regions.

2.1.4 Full-width half-max (FWHM) method

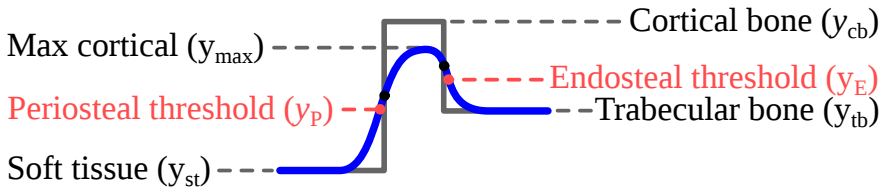
The 'optimal' threshold for a density boundary depends on the density of each side as illustrated by Figure 2.4. The inability of fixed thresholding techniques to adapt to variations in the optimal threshold over an image is a major limitation [33], which is addressed by the full-width half-max (FWHM) method shown in Figure 2.5. In the context of cortical bone, the FWHM method is an adaptive technique for calculating the periosteal and endosteal thresholds using Equation 2.1, where y_{st} is the density of soft tissue, y_{cb} is the cBMD, y_{tb} is the tBMD, y_P is the FWHM periosteal edge density threshold and y_E is the FWHM endosteal edge density threshold. The thickness is then the distance between the threshold locations. The maximum density is typically used to define y_{cb} as the name of the technique suggests.

$$y_P = \frac{1}{2}\left(\overline{y_{st}} + \max\left(y_{cb}\right)\right) \tag{2.1a}$$

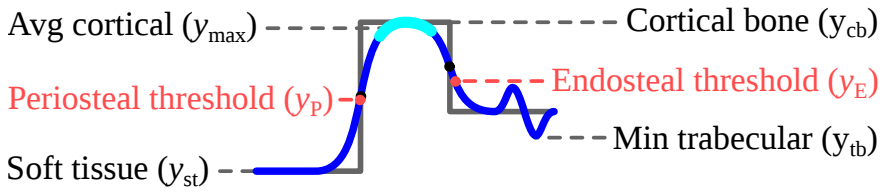
$$y_E = \frac{1}{2}\left(\max\left(y_{cb}\right) + \overline{y_{tb}}\right) \tag{2.1b}$$



(a) FWHM with no PVE



(b) FWHM with PVE



(c) FWHM with cBMD averaging and variable tBMD

Fig. 2.5 The FWHM method of applying adaptive thresholding to thin structures. Actual edge locations in black, and FWHM estimates in red.

FWHM has been extensively used to measure cortical thickness and density [33, 50, 88, 102, 124, 115]. Although generally similar, implementation details between studies vary notably when estimating the soft tissue, cBMD and tBMD. Early studies often used the peak cortical bone value, and the minimum soft tissue and trabecular bone values as shown in Figures 2.5a–2.5b [88, 102, 124]. This can result in the selection of an outlier value due to noise and natural variations in the density of trabecular bone. This will affect the calculated thickness and density as illustrated by Figure 2.5c [115]. The robustness of the FWHM method was improved in several implementations through some form of averaging [102, 115] as shown in Figure 2.5c, but this can lead to an underestimation of the cortical bone density due to PVE [102].

FWHM works well over a range of densities, but its accuracy remains limited to thick structures [33, 50, 88, 102, 115] as shown in Figure 2.5b. Experimental investigations show increasing thickness and density errors for structures with a thickness less than 1.5–3 $\times$ the FWHM of the image system PSF [33, 102]. When the cortical thickness is less than this, as is often the case in the proximal femur and vertebrae [124, 125], the PVE will reduce the maximum cortical density. This causes the FWHM method to underestimate its thresholds leading it to overestimate cortical thickness, and underestimate cBMD. Although the FWHM method is clearly limited when using QCT scans to measure cortical bone, it has been used as an unbiased gold standard for cortical measurements from HR-pQCT scans where the image resolution is much higher [123–125]. In several studies, the FWHM

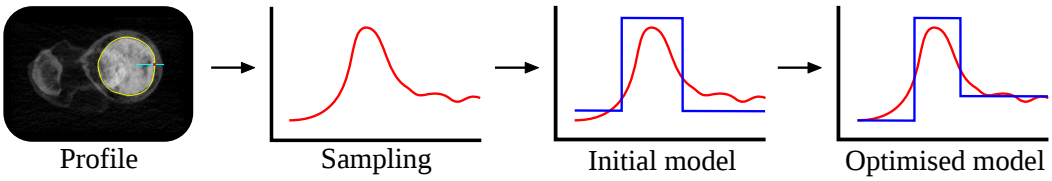


Fig. 2.6 Model-based characterisation of cortical features.

method was implemented using an optimiser to estimate the soft tissue density and tBMD values [123–125].

Several of the studies investigating the performance and limitations of the FWHM method noted coupling between the thickness and density measurements [33, 51, 81, 88], where incorrect thickness and density estimates still produce the correct product. The relationship results from the preservation of area inherent to convolution [33], but only when soft tissue and trabecular bone share the same density [103, 125]. This relationship was exploited in a study into the effect of ageing on cortical bone [81]. The peak cortical density within each scan was assumed to be the actual cBMD. This was used to calculate a 'corrected' thickness estimate [81] from the initial cortical thickness and cBMD measurements. A later study used variable thickness and density phantoms to derive imaged thickness versus actual thickness calibration curves, which were used to iterate towards the correct density and thickness measurements [51]. Its results show improved accuracy and precision over the FWHM method. However the method is limited by its assumption of a constant PSF throughout the image, and its reliance upon the thickness calibration curves, which may vary between scanners [91].

2.1.5 Model-based techniques

The limitations of the FWHM method can be addressed using model-based methods that directly model the imaging of cortical bone as the convolution between a PSF and the density profile of cortical bone. These methods show significant improvements in accuracy and precision over FWHM based methods [91, 118, 123]. The cortical bone is modelled along sampled transects perpendicular to its surface as a thin structure with soft tissue on one side and trabecular bone on the other side. The PSF and the model of the cortical bone profile are parametrised, so the convolution process can be mathematically represented and equated to the sampled transect as shown in Figure 2.6.

Numerical optimisation methods can then be used to find a locally optimal set of parameters that give the closest match between the convolution of PSF with the model, and the sampled transect. The model parameters represent anatomic features of the cortical and subsurface trabecular bone and properties of the system blur [91, 118, 123]. Although model-based methods are more accurate, they are also more complex, computationally expensive and have a less constrained solution space. Consequently, the performance of model-based methods can vary widely based on implementation-specific details, which include: the numerical optimiser used; the setup of the numerical optimiser; the model used; any parameter constraints used; and the initial parameter values used to begin the optimisation process.

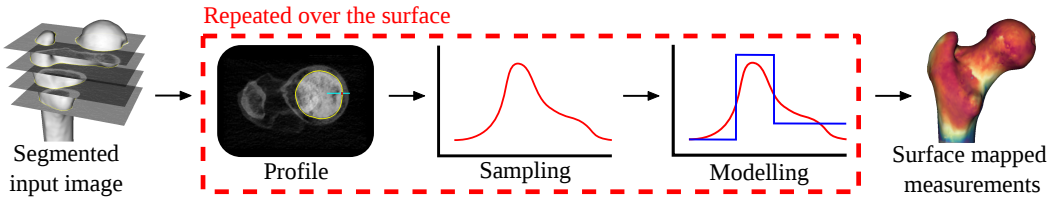


Fig. 2.7 The cortical bone mapping pipeline.

Streekstra et al. (2007) first used a model-based method to measure the thickness of cartilage over the wrist [118]. They estimated the image PSF from a CT scan of a phantom with known properties, so that an optimiser was only needed to solve for the optimal cortical bone model parameters [118]. The method was able to accurately measure cartilage thicknesses well below the FWHM of the image PSF [118]. This study was followed by an unconstrained implementation that produced lower bias results using an entirely unconstrained model that was validated through comparison with measurements of HR-pQCT scans [91]. Both implementations were applied to individual locations with no mechanism of applying these methods to a set of points defining the surface of a bone.

2.1.6 Cortical bone mapping (CBM)

Cortical bone mapping (CBM) is an automated technique for applying model-based methods at many locations over a bone's surface as shown in Figure 2.7. It produces many cortical and subsurface trabecular measurements, from which the focal variations in these features can be examined. The pipeline has been utilised in several studies to investigate the effect of constraint choice on method performance [123–125], which found that constraining the cBMD parameter led to improved model stability and accuracy over techniques with no constraints or a constrained PSF blur [124]. The assessment was performed using both simulated data and data from paired datasets of HR-pQCT and QCT scans [124].

cBMD varies locally with the porosity and DMB of cortical bone [3, 9, 39, 82], and so the global cBMD constraint will be incorrect in regions where the local cBMD differs from the global constraint [123, 125]. This can lead to errors in the cortical thickness measurements [125]. As such, methods of relaxing the cBMD constraint to better reflect the actual cBMD are desirable. This has been addressed in two studies, with additional corrective steps aimed at improving the quality of the cBMD constraint. In the first, the global cBMD was adapted, after comparing the initial model-based solution with the FWHM solution [125]; while in the second, the cBMD constraint was adapted based on the difference between the local blur estimate from the initial CBM measurement and a global estimate of the PSF blur [123]. The remainder of this chapter investigates possible improvements to these methods.

2.2 CBM Method

The CBM method uses a parametrised model to define the underlying physical attributes of the cortical bone, subsurface trabecular bone and surrounding soft tissue along each sampled transect.

2.2.1 Constrained three-tier rectangle model

Previous work has used the three-tier rectangle model shown in Figure 2.8a termed the 'rectangle model' throughout this thesis. The model has instantaneous transitions between the different tissues and a constant density over the entire cortical cross section as defined in Equation 2.2a, where y_{st}, y_{cb}, y_{tb}, x_P, and x_E respectively correspond to the density of soft tissue, cBMD, tBMD, the location of the periosteal edge, and the location of the endosteal edge. $x_E - x_P$ defines the cortical thickness (cTh).

$$y(x) = y_{st} + (y_{cb} - y_{st}) \cdot H(x - x_P) + (y_{tb} - y_{cb}) \cdot H(x - x_E) \tag{2.2a}$$

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2\sigma^2}} \tag{2.2b}$$

The blurred rectangle model, y_{blur}, is the convolution of the density model with the image point spread function (PSF) as shown in Equation 2.3. The image blur can be approximated with a Gaussian as shown in Equation 2.2b [103, 118, 123–125], where σ corresponds to the regionally variable standard deviation (STD) of the Gaussian PSF. The rectangle models are constrained to share the same globally fixed cBMD value, while the remaining model features and the PSF blur are varied until the resulting blurred rectangle model best matches the sampled density profile. This is achieved with the Levenberg-Marquardt optimisation method as discussed in section 2.2.3. The models can then be used to generate cortical thickness, cBMD, tBMD and MSD surface maps.

$$y_{blur}(x) = y(x) * g(x) \tag{2.3}$$

Convolution is an $\mathcal{O}(N^2)$ operation, where N is the number of density samples. This can be reduced to an $\mathcal{O}(N)$ formulation by making use of the fundamental theorem of calculus followed by the 'Sifting' property of the subsequent Dirac Delta functions as shown in Equation 2.4c [125]. The fundamental theorem of calculus is used to differentiate the rectangle function to produce Equation 2.4a. The 'Sifting' property of Dirac Delta functions is used to sample the Gaussian functions to produce Equation 2.4b, which can be simplified

into Equation 2.4c. $H(x)$ represents the unit step function, and erf() is the error function that results from integrating a Gaussian.

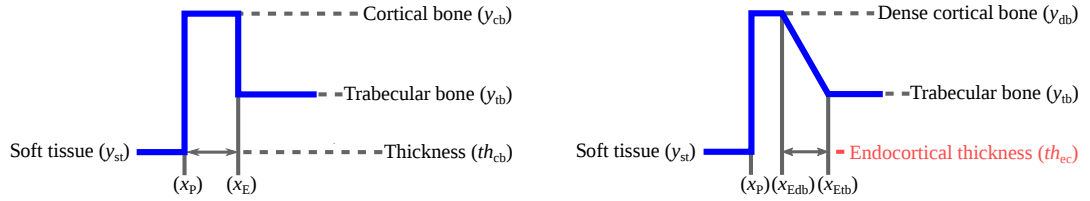
$$\begin{aligned} y_{\text{blur}}(x) &= y(x) * g(x) \\ &= \int_{-\infty}^{x} y'(\xi) * g(\xi) d\xi \qquad (2.4a) \\ &= \int_{-\infty}^{x} \left[(y_{\text{cb}} - y_{\text{st}}) \cdot \delta(\xi - x_{\text{P}}) + (y_{\text{tb}} - y_{\text{cb}}) \cdot \delta(\xi - x_{\text{E}}) \right] * g(\xi) \, d\xi \\ &= \int_{-\infty}^{x} \left[(y_{\text{cb}} - y_{\text{st}}) \cdot g(\xi - x_{\text{P}}) + (y_{\text{tb}} - y_{\text{cb}}) \cdot g(\xi - x_{\text{E}}) \right] d\xi \qquad (2.4b) \\ &= \frac{y_{\text{st}} + y_{\text{tb}}}{2} + \frac{y_{\text{cb}} - y_{\text{st}}}{2} \cdot \text{erf}\left(\frac{x - x_{\text{P}}}{\sigma\sqrt{2}} \right) + \frac{y_{\text{tb}} - y_{\text{cb}}}{2} \cdot \text{erf}\left(\frac{x - x_{\text{E}}}{\sigma\sqrt{2}} \right) \qquad (2.4c) \end{aligned}$$

The model fitting problem is ill-posed due to the large number of model parameters relative to the information available in each blurred density profile. This is the reason for constraining the cortical density to reduce the number of unconstrained parameters to five. The recently proposed CBMv2 method [123] relaxes the rigidity of the cBMD constraint by performing two CBM measurement steps: the first with the global cBMD constraint, and the second with a locally adjusted cBMD constraint based upon the difference between the local blur estimate from the initial CBM measurement and a global estimate of the PSF blur [123]. This allows for some variability in the cBMD, while maintaining the number of unconstrained parameters at five.

2.2.2 Endocortical model

The endocortical region is the interface between the cortical and trabecular bone. As discussed in Section 1.4.2, it is of particular interest as it is the primary site of remodelling in later life, when most fragility fractures occur [111, 138]. The increased endocortical remodelling has been linked to both a decrease in the DMB and an increase in porosity [9, 39, 138], both of which result in decreased BMD over the endocortical region. In response to this, the rectangle model was extended to include a linear finite width transition from the 'dense' cortical bone to trabecular bone as shown in Figure 2.8b. This extended model is termed the 'endocortical model' throughout the remainder of this thesis. The linear transition aims to reflect the increasing porosity that has been reported in the endocortical region [139].

This is achieved with a single additional parameter by splitting the endosteal edge in two as shown in Equation 2.5a. The meaning of y_{st}, y_{tb}, x_{P}, and σ remain unchanged compared to Equation 2.2a. y_{db} replaces the y_{cb} of the rectangle model and corresponds to the density of the dense region of cortical bone, which excludes the endocortical region. x_{Edb} and



(a) The rectangle model (b) The endocortical model

Fig. 2.8 The parametrised density profile models.

x_{Etb} correspond to the locations of the beginning and end of the endocortical region, and $x_{\mathrm{Edb}} - x_{\mathrm{Etb}}$ defines the width of the endocortical region (ecTh). $R(x)$ is the Ramp function.

$$\begin{aligned} y(x) &= y_{\mathrm{st}} + (y_{\mathrm{db}} - y_{\mathrm{st}}) \cdot H(x - x_{\mathrm{P}}) + \frac{y_{\mathrm{tb}} - y_{\mathrm{db}}}{x_{\mathrm{Etb}} - x_{\mathrm{Edb}}} R(x - x_{\mathrm{Edb}}) \\ &- \frac{y_{\mathrm{tb}} - y_{\mathrm{db}}}{x_{\mathrm{Etb}} - x_{\mathrm{Edb}}} \cdot R(x - x_{\mathrm{Etb}}) \end{aligned} \tag{2.5a}$$

This model can be used to measure the 'dense' cBMD (dBMD), the dense cortical bone thickness (dTh), the endocortical thickness and tBMD. As the endocortical model does not include cortical thickness or cBMD parameters, these must be derived from the measured values. Equation 2.6 defines the cortical thickness, th_{cb}, and cBMD, y_{cb} as the thickness of the periosteal edge to half-way along the endocortical region and the mean density over this region.

$$th_{\mathrm{cb}} = \frac{1}{2}(x_{\mathrm{Edb}} + x_{\mathrm{Etb}}) - x_{\mathrm{P}} \tag{2.6a}$$

$$y_{\mathrm{cb}} = y_{\mathrm{db}} \frac{x_{\mathrm{Edb}} - x_{\mathrm{P}}}{th_{\mathrm{cb}}} + \left(\frac{3}{4} y_{\mathrm{db}} + \frac{1}{4} y_{\mathrm{tb}}\right) \frac{\frac{1}{2}(x_{\mathrm{Etb}} - x_{\mathrm{Edb}})}{th_{\mathrm{cb}}} \tag{2.6b}$$

Convolution of the endocortical density model and the image PSF gives the blurred endocortical model. Once again, it is desirable to avoid performing an $\mathscr{O}\left(N^2\right)$ convolution at each step in the optimisation process. As with the rectangle density profile estimate, this can be reformulated into an $\mathscr{O}(N)$ equation. As shown in Equation 2.7a, the presence of the Ramp function in the model requires an additional integration-differentiation step before the 'Sifting' property can be applied to the resulting Dirac Delta functions in Equation 2.7b. This Produces the more complicated equation shown in Equation 2.7c. The required integral of

the error function is not supported in standard libraries, however it can be precomputed and accessed at runtime using a look-up table.

$$
\begin{aligned}
y_{\text{blur}}(x) &= y(x) * g(x) \\
&= \int_{-\infty}^{x} (y_{\text{db}} - y_{\text{st}}) \cdot \delta(\xi - x_{\text{P}}) * g(\xi)\, d\xi && \text{(2.7a)} \\
&+ \int_{-\infty}^{x} \int_{-\infty}^{\xi} \frac{y_{\text{tb}} - y_{\text{db}}}{x_{\text{Etb}} - x_{\text{Edb}}} (\delta(\psi - x_{\text{Edb}}) - \delta(\psi - x_{\text{Etb}})) * g(\psi)\, d\psi d\xi \\
&= \int_{-\infty}^{x} (y_{\text{db}} - y_{\text{st}}) \cdot g(\xi - x_{\text{P}})\, d\xi && \text{(2.7b)} \\
&+ \int_{-\infty}^{x} \int_{-\infty}^{\xi} \frac{y_{\text{tb}} - y_{\text{db}}}{x_{\text{Etb}} - x_{\text{Edb}}} [g(\psi - x_{\text{Edb}}) - g(\psi - x_{\text{Etb}})]\, d\psi d\xi \\
&= \frac{y_{\text{st}} + y_{\text{tb}}}{2} + \frac{y_{\text{db}} - y_{\text{st}}}{2} \cdot \text{erf}\left(\frac{x - x_{\text{P}}}{\sigma\sqrt{2}}\right) && \text{(2.7c)} \\
&+ \frac{(y_{\text{tb}} - y_{\text{db}})}{(x_{\text{Etb}} - x_{\text{Edb}})} \int_{-\infty}^{x} \left[\text{erf}\left(\frac{\xi - x_{\text{Edb}}}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{\xi - x_{\text{Etb}}}{\sigma\sqrt{2}}\right)\right] d\xi
\end{aligned}
$$

The Levenberg-Marquardt optimiser is also used to fit the endocortical model. As with the constrained rectangle model, the dBMD value of the model can be globally constrained before performing CBM to increase the stability of the optimiser. However the inclusion of the additional endocortical edge parameter further destabilises the already ill-posed problem. This can be counteracted to some extent, through the careful management of the optimisation process as discussed in Section 2.2.3.

2.2.3 CBM Optimisation

In the CBM method an optimiser is used to select the model parameters. These parameters should give the best possible alignment between the sampled image values and the expected image values, which are defined by the selected model parameters. Achieving the best alignment is a non-linear least squares problem in which the squared error between the sampled image data and the blurred parametrised model is minimised. In the implementation of the CBM method developed for this thesis the Levenberg-Marquardt method (LMM) is used as in previously published implementations of the CBM method [123–125]. The LMM begins with an initial guess as to the optimal model parameters that is iteratively improved to provide an ever more optimal set of model parameters [85]. When given appropriate initial conditions, the LMM will converge to a locally minimum set of parameters, but there is no guarantee that it will arrive at the global minimum. The quality of the local minimum reached is influenced by the optimiser settings and how the solution space is constrained. The LMM optimiser is just one of many optimisation methods. Preliminary work showed the LMM optimiser to perform favourably compared to the Powell [101] and the one-plus-one evolutionary [119, 120] optimisers, but a thorough comparison would be needed before the relative performance of the LMM method compared to other optimisers could be considered.

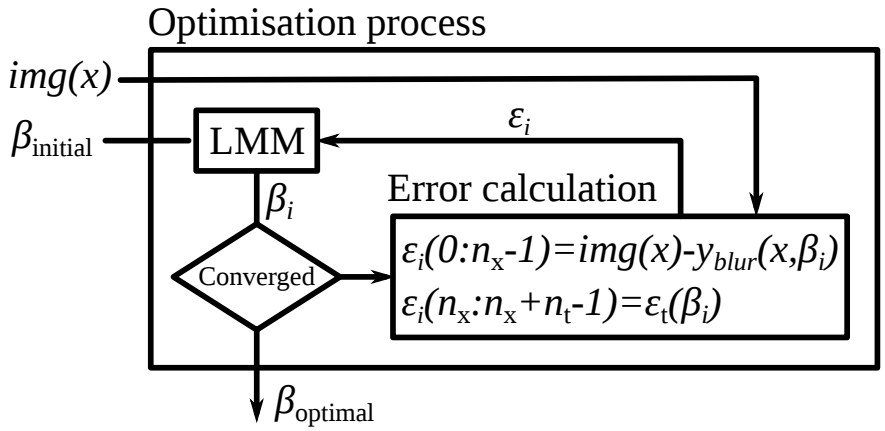


Fig. 2.9 The parameter optimisation process.

Theory

The LMM is a combination of the gradient descent method and the Gauss-Newton method (GNM) [85]. In each iteration, the updated parameters recommended by the gradient descent and GNM are calculated, and then combined as a weighted average. The LMM controls the weighting, so that the gradient descent method is dominant far from the optimal parameter values, and the GNM is dominant close to the optimal parameter values. This results in improved convergence over the gradient descent method and improved robustness over the GNM [79].

In the LMM, an error vector is used to internally update its estimate of the gradient and the relative weighting it applies to the gradient descent and GNM parameter values before calculating the next set of new parameter values. The error vector is calculated external to the LMM with a problem-specific cost function. This pipeline is shown in Figure 2.9, where i is the iteration; β are the model parameters; $img\,(x)$ are the sampled image values; ε is the error vector produced by the cost function; n_{x} is the number of sampled image values; and, n_{t} is the number of ways a model can be physically invalid.

Implementation

The Insight Segmentation and Registration Toolkit[1] (ITK) LMM implementation was used. This implementation requires specification of: the relative scaling to be applied to each different parameter; the convergence tolerance; the maximum number of iterations; and an initial estimate of the parameters. In each iteration, the LMM provides a new set of parameter values from which an error vector must be produced. The relative scaling determines how much each parameter is adjusted in each iteration. These scaling values were set to reflect the differing dynamic ranges of the position and density based model parameters.

[1] An open-source medical image analysis toolkit. https://itk.org

Cost function

As shown in Figure 2.9, the error vector is generated outside the LMM in each iteration. It contains information both about the alignment difference between the sampled densities and the blurred model, and the physical validity of the model. When a parameter is constrained, as y_{cb} and σ have been in the past, the parameter is removed from the iteratively updated set of parameters, and the fixed value is used instead for all error vector calculations. The first n_{x} error vector values correspond to the alignment difference errors, which are calculated as the signed difference between the blurred model and the sampled density transect. These error values are scaled by a weighting function, which acts to improve stability by targeting the fit quality around the periosteal edge as discussed in the following section.

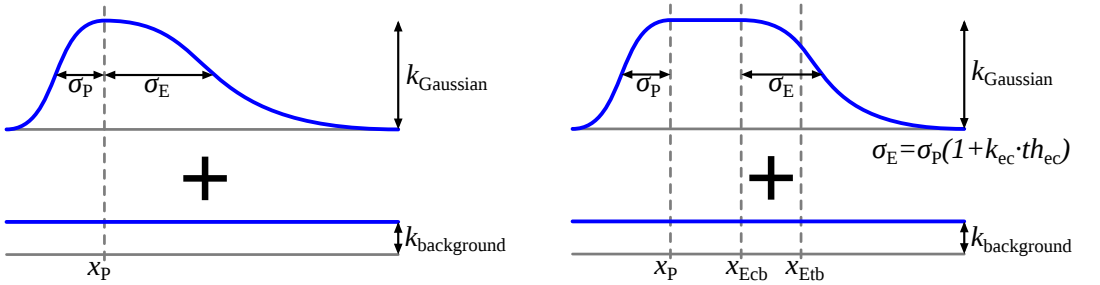
A density model is physically unrealistic if it has: a negative cortical thickness; a negative endocortical thickness; a cBMD less than the soft tissue density; a cBMD less than the tBMD; or, a negative or very large σ of the Gaussian blur. The endocortical thickness only applies to the optimisation of the endocortical model. Each of the parameter combinations capable of producing a physically invalid model are given an element in the error vector. Equation 2.8 is used to define the error associated with each possible physically unrealistic trait. Where, ε_t is the error reported in the error vector; k is the scaling factor associated with that trait; t is the model trait; and t_{min} is the model trait value below which the model is physically invalid. This approaches 0 for valid traits, and exponentially increases with the magnitude of an invalid trait.

$$\varepsilon_t = e^{k\cdot(t-t_{\text{min}})} \tag{2.8}$$

The relative weighting applied by the cost function to the model-to-image misalignment versus invalid model traits determines the relative importance given to each source of error. Decreasing the relative importance given to the model-to-image alignment errors reduces the quality of the final alignment of the models, while reducing the relative weighting given to invalid model traits increases the number of invalid models. The performance of several stages of the CBM method was assessed for a range of relative weighting values using synthetic QCT scans with a range of noise levels as discussed in Section 2.5.2. These results were used to select appropriate relative weighting values at each stage of the CBM pipeline.

Weighting function

Previous implementations of the CBM method use a single weighting function to improve the quality and stability of the rectangular model [123–125]. The weighting function controls the relative importance assigned to discrepancies between the PSF blurred model and the sampled image densities along the profile. The maximum value of the weighting function is aligned with the periosteal edge used to initialise the optimiser, so that discrepancies between the PSF blurred model and sampled data receive the greatest penalisation at the location of the initial periosteal edge estimation. As a result, the weighting function encourages the optimiser to converge on a model that is closely aligned with the underlying periosteal surface [123–125].



(a) Periosteal centred weighting function (b) Thickness adjusted weighting function

Fig. 2.10 The parametrised components used to construct weighting functions.

As the shape of the weighting function affects the fit of the final model with a tighter fit expected over portions of the profile assigned a higher relative weight, differently shaped weighting functions can be used to improve the quality of specific model traits. Various weighting functions with the same basic shape as the weighting function used in the initial CBMv2 implementation were compared in an effort to further improve the quality and overall stability of the CBM methods. In order to systematically investigate the impact that different aspects of the weighting function shape have on performance, a series of weighting functions were constructed from independently weighted background and Gaussian components as shown in Figure 2.10a. The $k_{\text{background}}$ and k_{Gaussian} values in the figure define the respective weighting applied to each of the weighting function components, while σ_{P} and σ_{E} are the STDs of the Gaussian component in the periosteal and endosteal directions.

The suitability of using a uniform weighting over the cortical bone region to further improve the quality of the cortical bone measurements was also investigated using the weighting function shown in Figure 2.10b. A range of periosteal centred and thickness adjusted weighting functions were generated by varying the weighting function parameters. As with the selection of the relative cost function weighting, the quality of the CBM measurements generated with different weighting functions were compared using simulated QCT scans produced with a range of noise levels as discussed in Section 2.5.2. These were used to select the weighting functions shown in Figure 2.11. The unaligned weighting function is periosteal centred and used for any generically initialised or unconstrained LMM optimisation, while the aligned weighting functions are thickness adjusted and used for LMM optimisations initialised with the locally optimal parameters from a previous LMM optimisation. These are used in the optimisation of the endocortical model by the Smoothly spatially constrained CBM (SSC-CBM) method described in Section 2.2.4.

2.2.4 Smoothly spatially constrained CBM (SSC-CBM)

cBMD is a regionally varying property of bone [3, 9, 39, 82, 123, 125] that has been linked to bone strength and fracture risk [3, 39, 82]. Unfortunately, the limited resolution of QCT means that cBMD cannot be directly measured over much of the proximal femur and vertebrae due to PVE, while the ill-posed nature of the optimisation problem means that the CBM method struggles to accurately measure both the cortical thickness and cBMD [123].

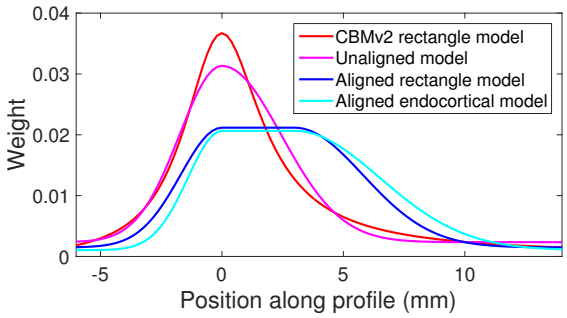


Fig. 2.11 The different weighting functions used by the CBM methods. These weighting functions are for an initial periosteal edge location at 0mm and an initial cortical thickness of 3mm.

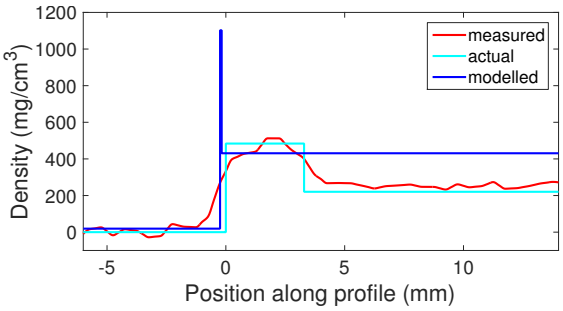


Fig. 2.12 Typical failure mode for the rectangular CBMv2 method.

The rectangular CBMv2 method introduced in Section 2.2.1 currently provides the best compromise between accurately measuring cortical thickness and measuring the cBMD, but its cBMD measurements are still of a much lower quality [123]. The CBMv2 method also has a low cBMD density failure mode where it grossly overestimates the cBMD and tBMD, and underestimates the cortical thickness as shown in Figure 2.12.

For these reasons, a new pipeline for adapting the dBMD constraint used in the endocortical CBM method to better reflect the underlying data was developed. Its design was based on the surface mapped dBMD values observed in the proximal femur HR-pQCT scans examined in Chapter 3, which appear to show the dBMD varying smoothly over the cortical surface. The pipeline aims to estimate the dBMD as a smoothly varying value, which is used to constrain the CBM method as shown in Figure 2.13.

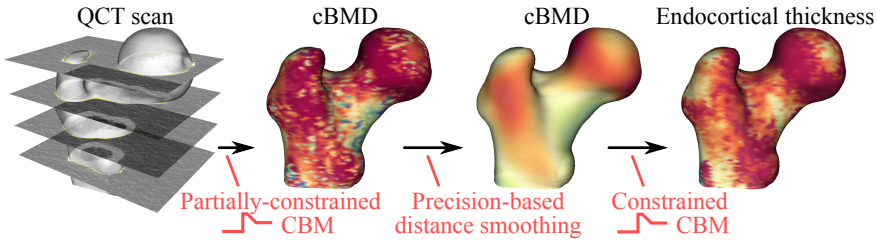


Fig. 2.13 The smooth spatially constrained CBM pipeline.

The pipeline contains two CBM measurement steps. The first is a partially constrained CBM measurement that is used to produce dBMD estimates with three LMM optimisations at each measurement location. The first optimisation uses a generically initialised and globally dBMD-constrained rectangle model. Its measurements are used to initialise the second optimisation, which is still a globally dBMD constrained rectangular model. The resulting measurements initialise the final LMM optimisation, which is left entirely unconstrained. The partially constrained configuration was used in place of the entirely unconstrained CBM configuration used in previous studies [91, 123] due to its greater stability. The resulting dBMD surface map is noisy, so it is spatially smoothed using Gaussian distance-weighted precision smoothing. The σ of the Gaussian distance filter is specified globally, while the precision value is calculated locally from estimates of the variance of the LMM model parameters. These smoothed locally-varying dBMD values are used as constraints in the second CBM measurement step, which involves two LMM optimisations at each measurement location. The first optimisation is of a generically initialised and locally dBMD-constrained rectangle model. Its resulting measurements are used alongside a generic endocortical thickness to initialise the endocortical model used in the second optimisation, which uses the same local dBMD-constraint.

The rectangular model is used in the inital LMM optimisations of each CBM measurement step to improve the overall stability of the method. A rectangular or endocortical model can be used to generate the dBMD constraints in the final unconstrained LMM optimisation of the first CBM measurement step. When used, the endocortical model is initialised with a generic endocortical thickness. The model selection will affect the dBMD constraint map and the resulting endocortical SSC-CBM measurements. As with the selection of the optimiser cost function values, QCT simulations produced with a range of noise levels were used to compare the performance of the measurements produced with the rectangular and endocortical dBMD constraints as discussed in Section 2.5.3.

2.3 QCT simulation

QCT simulations can be used to assess the performance of QCT methods [51, 102, 103, 118, 124, 123]. Most studies to have assessed QCT methods with simulations have only modelled the high level effects of the QCT imaging process as the convolution of a PSF and a density profile [118, 124, 123] with the possible addition of white noise [103, 102]. The exception is Hangartner et al. (2007), who simulated the imaging of a synthetic phantom composed of fat, water and bone with a second generation translate-rotate laboratory CT scanner assuming an ideal X-ray beam with a single energy of roughly 30 keV [51]. While no implementation details were provided, the use of an idealised single energy X-ray beam means that it could not be used to model beam hardening effects. In this section, the general approach described in Hangartner et al. (2007) was extended to simulate QCT imaging with beam hardening effects. This was used to assess the performance of the endocortical SSC-CBM method during its development and in comparison with the rectangular CBMv2 method over a wide range of bone architectures, resolutions, noise levels and levels of beam hardening.

2.3.1 Theory

As discussed in Section 2.1.1, QCT imaging involves the acquisition of, and reconstruction from, a series of X-ray projections taken over a range of orientations. Simulating the QCT imaging process requires a thorough understanding of X-rays, the linear attenuation of materials, and the interactions between X-rays and materials.

X-ray attenuation

X-rays are a form of electromagnetic radiation with wavelengths (λ) ranging from 0.1 nm to 10 nm and energies ranging from 100 eV to 100 keV [7, 121]. An X-ray beam's energy is attenuated exponentially while it travels through a material. The level of attenuation is related to the material composition, the amount of material, and the energy of the X-ray beam as defined by Equation 2.9 and illustrated in Figure 2.14. Where E is the energy spectrum of the X-ray beam, $I_{initial}$ is the initial intensity of the X-ray beam over the energy spectrum, I is the final intensity of the X-ray beam over the energy spectrum, $x_1 - x_0$ is the distance the X-ray beam has travelled through the attenuating material, and μ is the linear attenuation of the material at position x and energy E [7, 15, 74, 121].

$$I(E) = I_{\text{initial}}(E)\, e^{-\int_{x_0}^{x_1} \mu(E,x)dx} \tag{2.9}$$

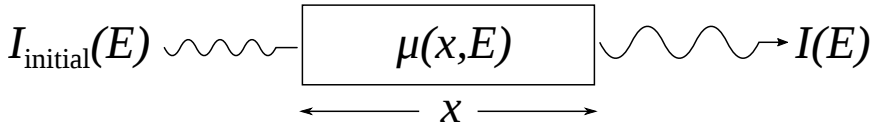
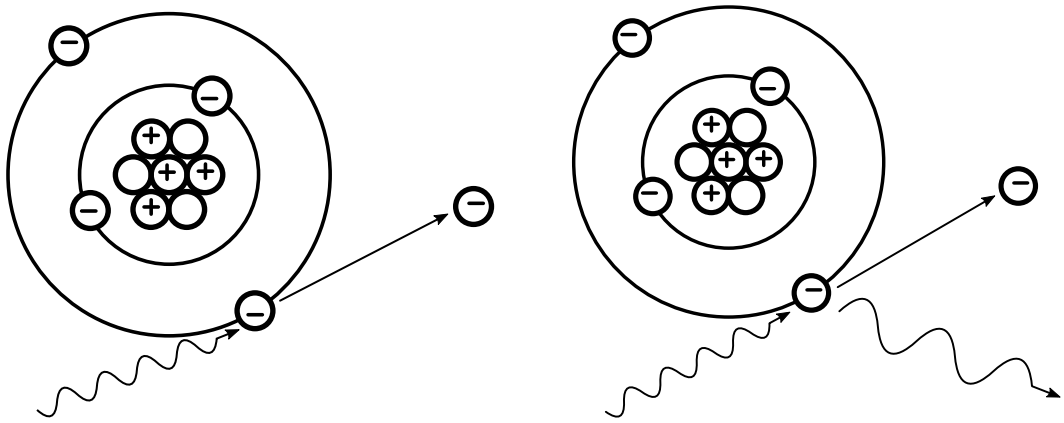


Fig. 2.14 X-ray beam attenuation.

X-ray attenuation is caused by several mechanisms in which some or all of the X-ray energy is transmitted to the material. The total attenuation experienced by the X-ray beam is the sum of the attenuations caused by each mechanism. The photoelectric effect and Compton scatter are the two most common mechanisms of X-ray attenuation over the range of X-ray energies utilised by QCT and HR-pQCT. Both mechanisms involve the excitation of an outer shell electron by an X-ray photon as shown in Figure 2.15. In the photoelectric effect all energy is transmitted to the electron, while in Compton scatter the X-ray photon continues at a reduced energy [7, 15, 74].

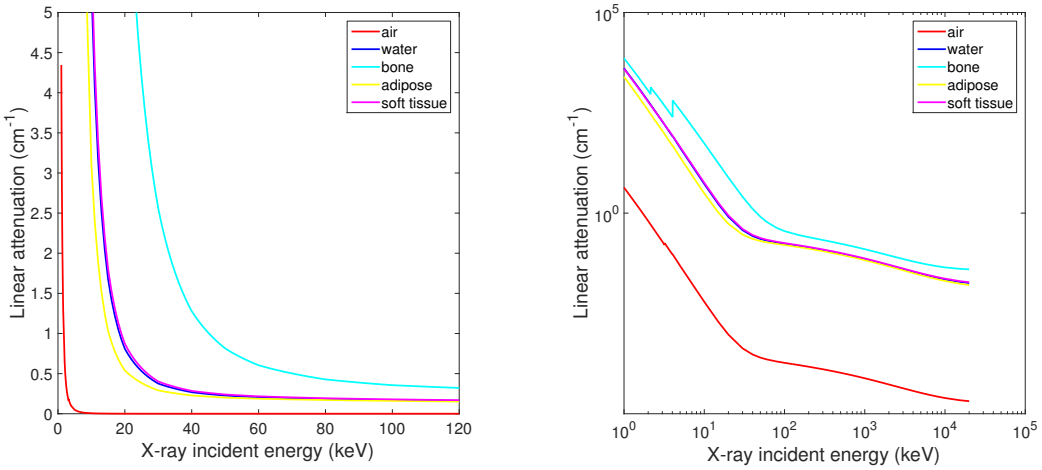
Linear attenuation varies with material composition and X-ray energy as the interactions causing attenuation occur at different energies for different elements. These values have been measured experimentally by Hubbell et al. (1996) for various materials over a range of X-ray energies [57]. Figure 2.16a shows the measured linear attenuations of air, water, mineralised bone, adipose tissue and soft tissue over a wide range of X-ray energies. Attenuation due to the photoelectric effect is related to X-ray energy by $\mu \propto \frac{\rho}{E^3}$, while attenuation due to Compton scatter is related to X-ray energy by $\mu \propto \frac{\rho}{E}$ [7, 74]. As such, the attenuation from the photoelectric effect dominates at lower X-ray energies, and the attenuation from Compton scatter dominates for higher energies. This is illustrated by Figure 2.16b where the transition



(a) Photoelectric effect (b) Compton scatter

Fig. 2.15 Common X-ray interactions.

in dominance of the photoelectric effect to dominance of Compton scatter is marked by the prominent decrease in slope that occurs between 20 keV and 100 keV. The discontinuities in the linear attenuation of some materials that occur for energies less than 10 keV are caused by the impact of the k-edges of that material on the photoelectric effect [7].



(a) Linear scale (b) Log scale

Fig. 2.16 Linear attenuation values for several materials against energy produced using the tables of X-ray mass attenuation and densities provided on the NIST Standard Reference Database [57].

X-ray sources

X-rays are typically produced using an X-ray vacuum tube in which electrons are released from a cathode and accelerated toward an anode where X-rays are released through either

bremsstrahlung or characteristic X-ray emission. The former produces a continuous spectrum of X-rays, while the latter produces a discrete emission spectrum of X-ray energies [7, 15, 74]. Together these energy spectrums form the characteristic curve of an X-ray beam. Approximate QCT and HR-pQCT characteristic curves are shown in Figure 2.17 for a QCT scan with a peak energy of 120 keV, and a HR-pQCT scan with a peak energy of 70 keV. These curves were based on those produced by SpecCalc Lite[2]. HR-pQCT typically use X-rays with an energy spectrum of 30–70 keV, while QCT typically use a wider spectrum of X-rays in the range of 30–120 keV. This means that QCT X-ray beams experience a lower net linear attenuation than the HR-pQCT X-ray beams due to the lower linear attenuation of high energy X-rays.

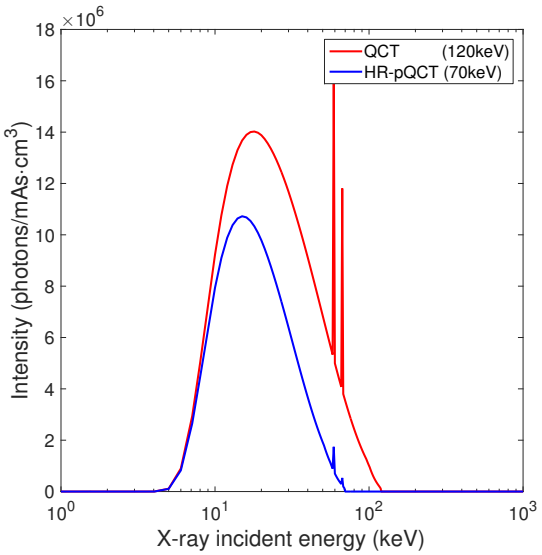


Fig. 2.17 An approximation of X-ray source characteristic curves for peak energies typical to QCT and HR-pQCT scan settings, which are generated with a piecewise function based on X-ray emission spectra from tungsten anodes with 1 mm of Aluminium shielding.

Hounsfield units

As mentioned in Section 2.1.1, the attenuation experienced by X-rays varies with their energy. This makes it difficult to compare material attenuation measurements from images produced by different X-ray beams. In response to this, CT measurements are normalised into Hounsfield Units (HU) using Equation 2.10. This reduces the variations in the CT measurements produced by different X-ray beams by ensuring water will always be 0 HU and air will always be −1000 HU. Unfortunately, the HU of other materials remain dependent on the energy of the incident X-ray beam as shown by the HU of bone, soft tissue and adipose tissue in Figure 2.18, and so calibration to convert HU to BMD is still desirable.

[2] A freely available software tool for calculating the X-ray emission spectra for tungsten anodes available from: http://spekcalc.weebly.com/

$$HU = 1000 \times \frac{\mu(E) - \mu_{\text{water}}(E)}{\mu_{\text{water}}(E) - \mu_{\text{air}}(E)} \tag{2.10}$$

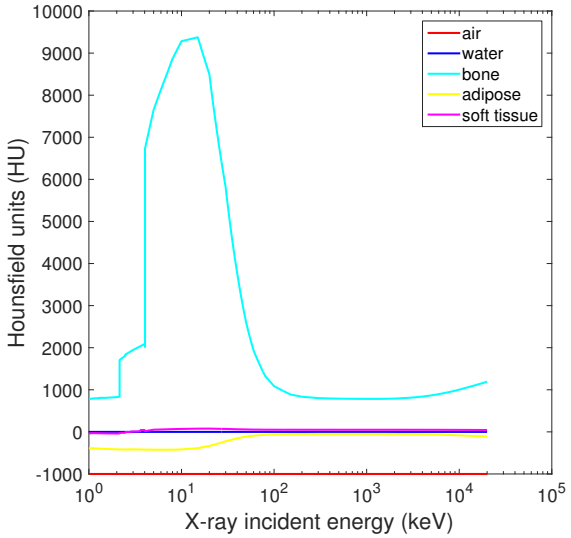


Fig. 2.18 Hounsfield unit values for several materials against energy calculated using the linear attenuations derived from the NIST Standard Reference Database [57].

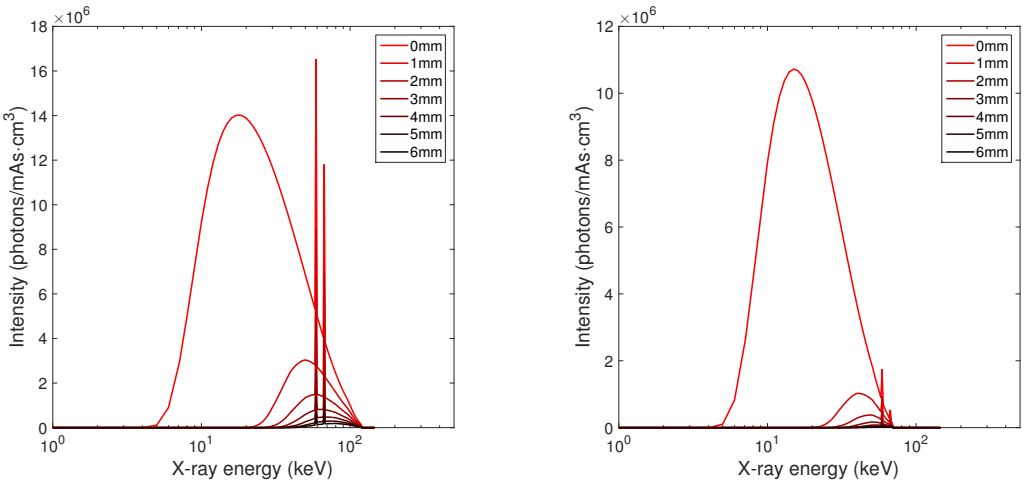
Beam hardening

As discussed in Section 2.1.1, beam hardening is an increase in the mean energy of the X-ray beams with increasing penetration depth, which occurs even as the net intensity of the X-ray beams decrease. Figure 2.19 shows this effect on the generic QCT and HR-pQCT source X-ray beams from Figure 2.17. The level of attenuation experienced by the different X-ray beam energies was calculated using the linear attenuations of mineralised bone derived from the NIST Standard Reference Database across the range of X-ray energies [57].

The decrease in linear attenuation of mineralised bone with X-ray energy shown in Figure 2.16, means that the linear attenuation measured in a voxel decreases with penetration depth. This also occurs for other materials, although the rate at which beam hardening occurs varies depending on the variation in linear attenuation of that material with X-ray energy. The differing linear attenuations of different materials means that a voxel will experience different levels of beam hardening depending upon quantity and composition of the material that surrounds it. This provides an added challenge when trying to relate the measured linear attenuation to clinically relevant material properties.

Beam hardening correction

Beam hardening correction can be used to partially remove the effects of beam hardening from the measured X-ray projections during CT reconstruction. This is achieved by using



(a) QCT (b) HR-pQCT

Fig. 2.19 The beam hardening of X-ray beams caused by penetration of mineralised bone calculated for the X-ray source beams from Figure 2.17 and linear attenuations derived from the NIST Standard Reference Database [57].

thickness based calibration to convert the measured integral of the linear attenuation along a projection line into the expected integral of the linear attenuation along that projection line in the absence of any beam hardening. This requires a mapping function as shown in Equation 2.11a to convert the measured integral of the linear attenuation along a projection line into the thickness of the material. This also requires a second linear scaling function to transform the distance measurement back into the integral of the linear attenuation of that material in the absence of any beam hardening as shown in Equation 2.11b. In these equations $\int \mu(x,\phi)\,dx$ is the total attenuation measured along the projection line at angle ϕ, f is the mapping function, $t_{\text{material}}(\phi)$ is the material thickness for the projection line at angle ϕ, and k is selected such that the measured and corrected linear attenuation integrals are equal where the effects of beam hardening are negligible. Unfortunately, this can only be applied to correct the beam hardening caused by a single material as different materials will have different mapping functions. Medical QCT scans typically perform beam hardening correction for water as the human body is predominantly water.

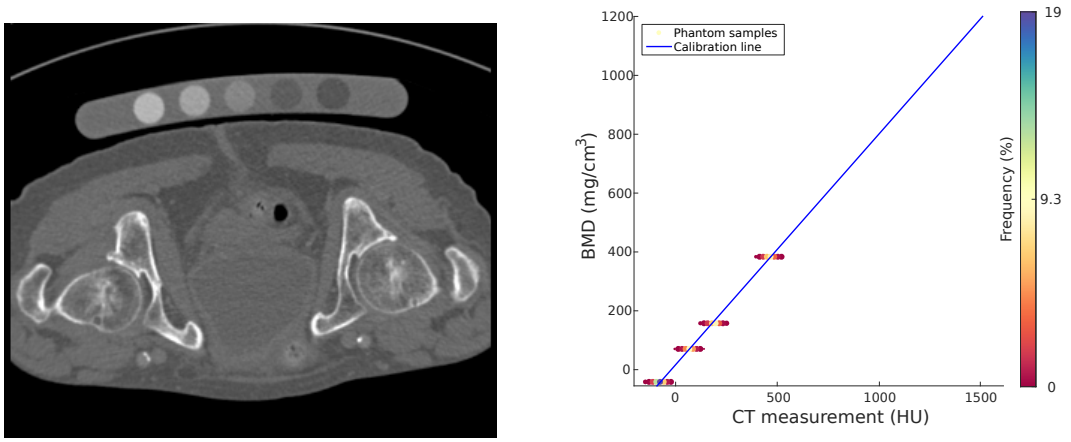
$$t_{\text{material}}(\phi) = f\left(\int \mu(x,\phi)\,dx\right) \tag{2.11a}$$

$$\int \mu_{\text{corrected}}(x,\phi)\,dx = k{\cdot}t_{\text{material}}(\phi) \tag{2.11b}$$

Density calibration

QCT and HR-pQCT scans are calibrated to relate the CT measurements into the clinically relevant BMD, which is preferred to the measured HU and linear attenuation as a material

property that is both independent of the X-ray energy spectrum and related to bone strength [2, 13, 60, 61]. BMD can only be extracted from HU measurements by directly accounting for the X-ray incident energy used in the scan. This is generally achieved by scanning an object of known BMD properties [32] and using it to define a calibration curve between BMD and HU as shown in Figure 2.20. Figure 2.20a is a scan from the proximal femur fracture discrimination study that is analysed in Section 4.4; it contains a typical example of the five-core Mindways solid phantom[3]. The cores have known BMD values and can be sampled to produce HU and BMD value pairs, which can be related with a linear regression fit as shown in Figure 2.20b.



(a) A QCT scan with a Mindways solid phantom

(b) The Mindways calibration curve

Fig. 2.20 BMD calibration of a QCT scan of the proximal femur with a Mindways solid phantom.

The five-core solid Mindways phantom is one of several phantoms that are used to calibrate the QCT scans analysed. The other phantoms used in this thesis are the European spine phantom (QRM-ESP)[4] and the bone density calibration phantom (QRM-BDC)[5]. Table 2.1 shows the number and range of densities included in each phantom. Of these, the QRM-ESP is designed to be scanned on its own but in the same scanner and with the same settings as the scans it is to calibrate, while the other scanners are designed to be scanned with the patient.

Calibration phantom limitations

Beam hardening affects the linear attenuation measurements within both the patient and calibration phantom. The effect will vary throughout the scan depending on the surrounding material, which reduces the accuracy of the calibration process. This is illustrated by considering the idealised calibration of a synthetic calibration phantom composed of twelve cores with radii of 10 mm and density values ranging between 0 mg/cm^3 to 1100 mg/cm^3 at

[3]Mindways Software Inc. http://qct.com/home/products/
[4]QRM Gmbh Germany http://www.qrm.de/content/products/bonedensity/esp.htm
[5]QRM Gmbh Germany http://www.qrm.de/content/products/bonedensity/bdc.htm

Phantom name	**# of Sites**	**min density** ($\mathbf{mg/cm^3}$)	**max density** ($\mathbf{mg/cm^3}$)
Mindways solid	5	-50	375
QRM-BDC	3	0	200
QRM-ESP	6	0	800

Table 2.1 The density ranges of calibration phantoms used in this thesis. The Mindways density values are derived and should not be interpreted as physical quantities [84].

intervals of $100\,\mathrm{mg/cm^3}$. Let the material composition of each core be modelled as a mixture of mineralised bone and water equivalent plastic that gives the appropriate BMD. Figure 2.21 shows the HU value at the centre of each core, which was calculated with Equations 2.9 and 2.10 given the material composition of each core, the radius of each core, and the energy distribution of the source QCT X-ray beam from Figure 2.17. The HU in the centre of each core was calculated separately assuming each core was scanned in a vacuum so that the relationship between beam hardening and BMD density could be considered without having to consider any secondary beam hardening effects from the arrangement of the cores within the phantom.

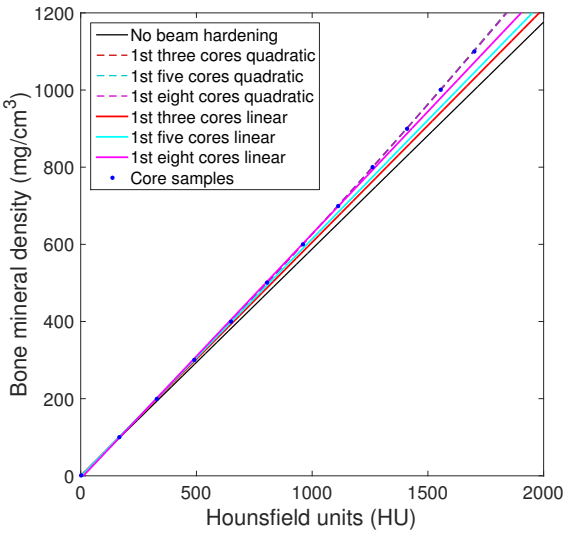


Fig. 2.21 The effect of beam hardening caused by the penetration of calibration cores on the relationship between HU and BMD calculated for the QCT X-ray source beam from Figure 2.17 and linear attenuations derived from the NIST Standard Reference Database [57].

The blue points in Figure 2.21 are the HU values calculated at the centre of each BMD core, while the black line shows the linear relationship between BMD and HU for a mono-energy X-ray beam that experiences no beam hardening. A comparison of these values shows that the beam hardening experienced by the core samples causes a convex relationship between core BMD and HU. The HU values from the centre of each BMD core were used

to generate linear and quadratic calibration curves over density ranges corresponding to the QRM-BDC, the Mindways solid, and the QRM-ESP phantoms. The solid red, magenta and cyan lines show the linear calibration curves, while the dashed red, magenta and cyan lines show the quadratic calibration curves. The linear calibration curves show the lower the density range of the phantoms, the greater the linear calibration curve underestimates larger densities. In contrast, all three quadratic calibration curves show a much greater correspondence over the full range of sampled core densities.

Beam hardening also occurs in the scan subject. The beam hardening caused by the outer muscle and soft tissue will largely be corrected as discussed in Section 2.3.1, but this will have little impact on the much greater levels of beam hardening caused by bone. Cortical bone thickness is typically less than 6 mm [123–125]. Figure 2.22 shows the effects of beam hardening after penetrating 2, 4 and 6 mm strips of cortical bone with variable BMD. This indicates that the amount of beam hardening varies with bone thickness, which means the amount of beam hardening within a CT scan varies with the macro-architecture of the bone. Hence although the relationship between BMD and HU can be assumed to be slightly convex, the variations caused by beam hardening within the patient will be independent from those observed in the calibration phantom.

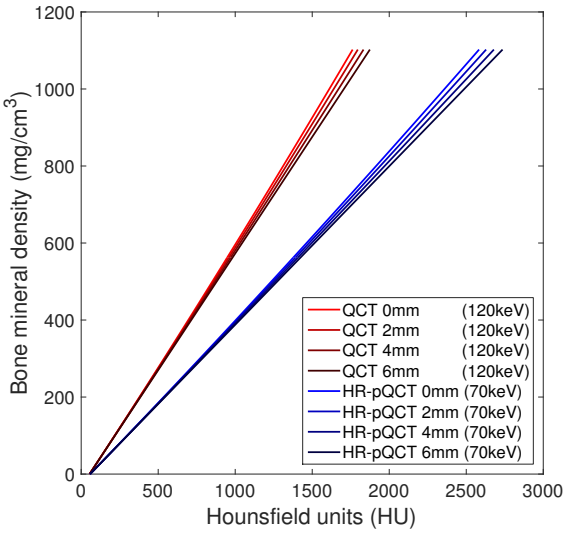


Fig. 2.22 The effect of beam hardening caused by the penetration of mineralised bone on relationship between HU and BMD calculated for the X-ray source beams from Figure 2.17 and linear attenuations derived from the NIST Standard Reference Database [57].

Material composition is another potential source of variations in the linear attenuation measurements throughout a scan. Bone is composed of a medullary canal surrounded by trabecular bone encased in a layer of cortical bone. The medullary canal is primarily composed of stromal marrow, which is similar to adipose tissue [22]. Trabecular bone is composed of mineralised bone and hematopoietic marrow, which produces blood cells. The mineralised bone forms a porous scaffold with the hematopoietic marrow contained within its cavities. In contrast, cortical bone is almost exclusively composed of mineralised bone

with no marrow. In adults, the stromal filled medullary canal exists in the shaft of long bones surrounded by cortical bone. The hematopoietic-containing trabecular bone exists at the proximal ends of long bones; and in flat bones such as the pelvis, sternum, vertebrae, skull, ribs and shoulder blades [22].

This means the proximal femur is composed of mineralised bone, hematopoietic marrow and stromal marrow, while the vertebral body is primarily composed of mineralised bone and hematopoietic marrow. Figure 2.23 shows two models that aim to consider the possible extent of non-linearities introduced into the relationship between linear attenuation and BMD by localised variation in material composition. Adipose tissue represents stromal marrow and soft tissue represents hematopoietic marrow. The left model contains the tissues expected in the proximal femur. The composition is modelled as entirely stromal marrow transitioning to trabecular bone composed of hematopoietic marrow with increasing quantities of mineralised bone transitioning to cortical bone. The right model contains the tissues expected in a vertebral body, which does not contain stromal marrow. It is modelled as trabecular bone composed of hematopoietic marrow with increasing quantities of mineralised bone content as it transitions to cortical bone.

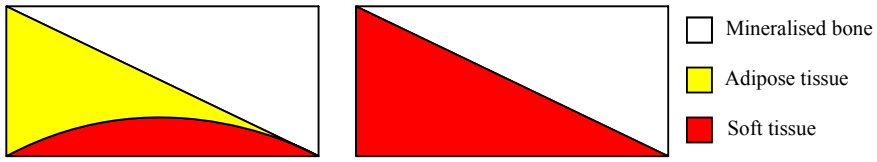
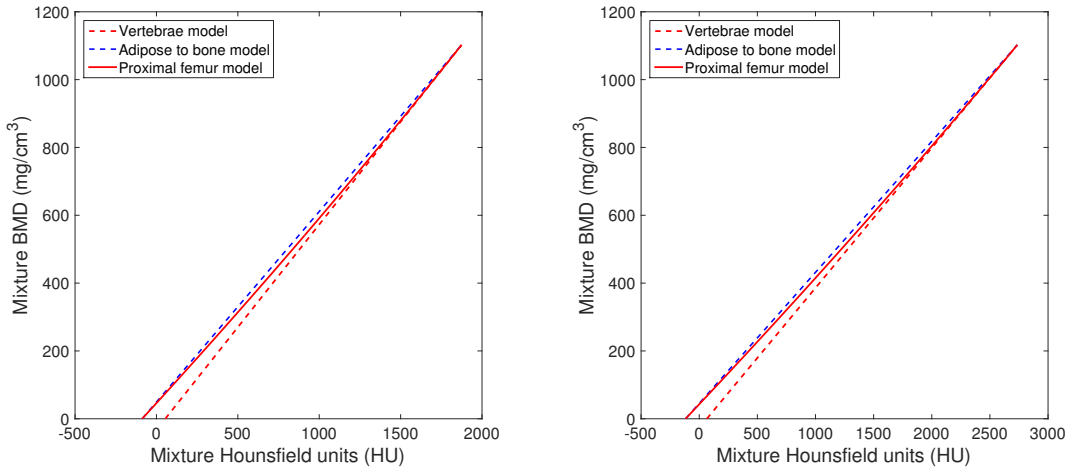


Fig. 2.23 A simple cross section model showing a plausible tissue composition of the proximal femur (left) and the vertebral body (right).

Figure 2.24 shows the simulated relationship between BMD and HU for typical QCT and HR-pQCT energy sources produced by the two varying tissue composition models. The QCT HU and HR-pQCT HU values were calculated for an infinitely narrow X-ray beam travelling vertically through the model using Equations 2.9 and 2.10, the tables of X-ray mass attenuation and densities provided on the NIST Standard Reference Database [57], and the X-ray sources introduced in Section 2.3.1. The solid red line shows the convex relationship produced by the model of the proximal femur including stromal and hematopoietic marrow, while the red dashed line shows the linear model produced by the model of the vertebral body containing only hematopoietic marrow. The dashed blue line shows the linear relationship produced by bone containing only stromal marrow. This is included to emphasise the non-linearity in the femur model curve. As these models of the macro-structural composition of bone are gross simplifications, Figures 2.24a and 2.24b should only serve to indicate that tissue composition can be another source of non-linearities in the relationship between linear attenuation and BMD.

The accuracy of phantom-based calibration is limited by variations in the relationship between linear attenuation and BMD, which can be caused by beam hardening affects and variations in tissue composition as shown in Figures 2.21–2.22 and 2.24. The accuracy is further compromised as these variations can be localised, and so the relationship between linear attenuation and BMD experienced by the calibration phantom may not be representative



(a) QCT (b) HR-pQCT

Fig. 2.24 The affect of non-linear marrow composition model on BMD calibration.

of that experienced by the patient. Finally, most calibration phantoms contain a limited range of low density values, as they were designed for the calibration of trabecular bone and not cortical bone. This limits their ability to accurately model the relationship between HU and BMD at high BMD values.

2.3.2 Method

The QCT imaging of a three-dimensional synthetic dataset was simulated slice-by-slice. The first stage of the CT imaging process is the acquisition of X-ray projections, which was simulated in each slice. This was followed by image reconstruction with a fairly standard implementation of the filtered back-projection algorithm. This method is an extension of a teaching tool developed in the Cambridge University engineering department[6] to support synthetic datasets with a mixture of materials at a single location.

Image acquisition

The acquisition of X-ray projections over a wide range of equally spaced angles was modelled in MATLAB[7] based on a first generation CT scanner with a flat detector array that rotates in-plane about synthetic data that is surrounded by air. The finite width of the X-ray beam and X-ray detectors means that each X-ray projection measures the attenuation over a finite width sliver. This leads to blurring, which was modelled in reverse by blurring each slice of the synthetic dataset with a Gaussian PSF, before measuring the attenuation along infinitely thin projection lines. The material properties along the projection lines were sampled at the Nyquist sampling rate dictated by the Gaussian PSF with linear interpolation between the

[6]Developed by Graham Treece for GG2: CT reconstruction and visualisation. E-mail: gmt11@cam.ac.uk.
[7]https://mathworks.com, version 2015a

nearest pixels. The synthetic datasets can contain mixtures of materials, which are encoded by recording the proportion of each material contained at each pixel location. The proportions of these materials are then blurred separately, before the total is normalised at each location.

The total amount of each material along each X-ray projection is calculated as the summed proportion of that material at each sample location. This was used to calculate the attenuated energy profile to reach the detector using: Equation 2.9; the linear attenuation values calculated from the the NIST Standard Reference Database [57]; and the source X-ray energy profile shown in Figure 2.17. The attenuated X-ray energy profile was converted into the number of photons to reach the detector based on the $\mathrm{mA} \cdot \mathrm{s}$ applied across the anode-cathode pair during the generation of the X-ray beam. This gives the number of photons at each energy, which can be integrated to give the total number of photons to reach the detector. This gives the idealised attenuation caused by the materials along the projection path, but includes none of the noise caused by the Poisson nature of the individual X-ray interactions and X-ray detections.

Throughout the modelled transformation, the X-ray interactions are not modelled at the atomic level. Instead, the Poisson nature of the X-ray interactions was incorporated into the image acquisition model based upon the total number of X-ray photons in each beam when calculating the number of photons at each energy level after attenuation. This also indirectly models the noise produced by scatter caused by these interactions. The Poisson nature of the detection process was introduced to the sum of all photons across all energies detected. In both cases, the number of X-ray photons was directly related to the α of the Poisson distribution. This means that an increase in number of X-ray photons caused by an increase in the $\mathrm{mA} \cdot \mathrm{s}$ of the simulation process will lead to a QCT image with reduced noise. The photon counts at each adjacent detector for a single projection angle define a single line of a sinogram. The complete sinogram is defined by the photon counts at each projection angle in a slice.

CT reconstruction

CT reconstruction to produce HU estimates was performed in three stages. First, beam hardening correction for water was applied to the sinogram produced by the simulation of the image acquisition process as is typical for medical QCT scans. This was performed using the approach described in Section 2.3.1. The attenuation to thickness mapping was generated using the X-ray mass attenuation and densities of water provided on the NIST Standard Reference Database [57]. Next, the CT reconstruction was performed using filtered back projection with a Ram-Lak filter to estimate the three-dimensional linear attenuation of the synthetic dataset from the corrected sinogram. The raised-cosine variation of the Ram-lak filter defined by Equation 2.12 was used with an alpha of 0.0001. Finally, the linear attenuation estimates were converted into Hounsfield units (HU) using Equation 2.10, and precomputed water and air linear attenuation profiles that were also generated from the NIST Standard Reference Database [57]. The HU values could then be converted into BMD values using a calibration standard. The approach used in this thesis is described in Section 2.4.6.

$$f(\omega) = \frac{|\omega|}{2\pi} \cos\left(\frac{\omega}{\omega_{\max}} \frac{\pi}{2}\right)^{\alpha} \tag{2.12}$$

2.4 Synthetic datasets

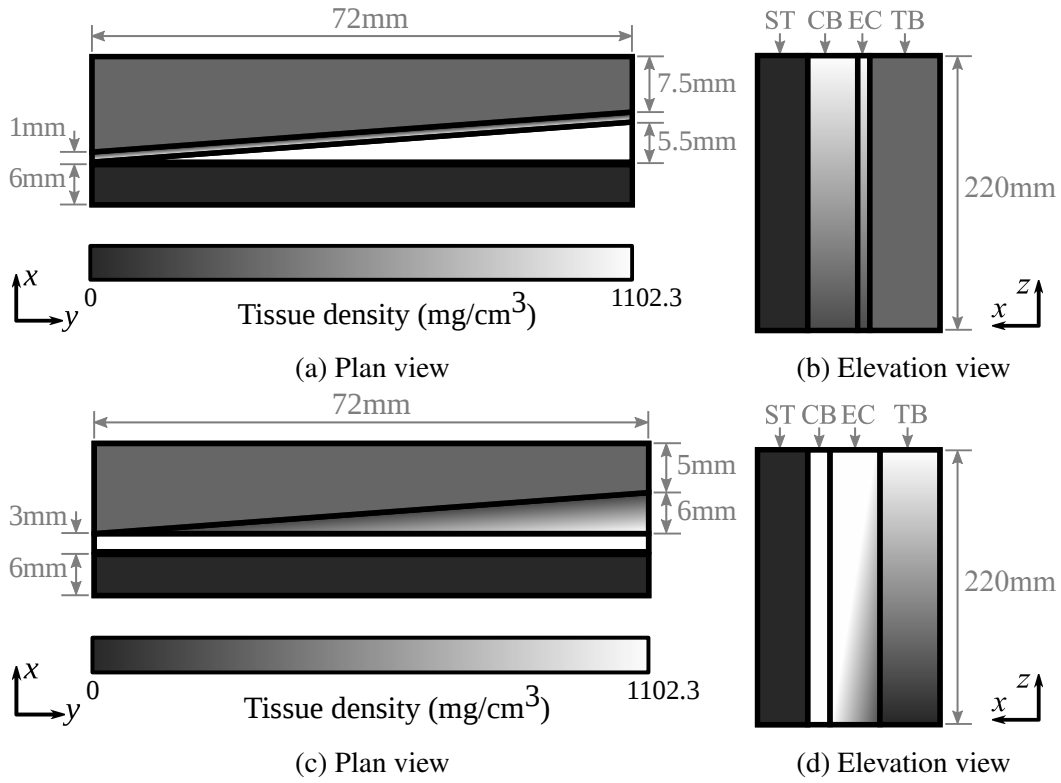
Simulated QCT and HR-pQCT scans of three synthetic datasets were used to assess performance during the development and assessment of the endocortical SSC-CBM method. Each dataset was composed of blended mixtures of bone and soft tissue combined to imitate cortical, endocortical and trabecular bone tissue, which were arranged to resemble a wide spectrum of bone architectures with precisely known dimensions and compositions. Each dataset allows different architectural components to vary: the cortical dataset has variable cortical thickness and cBMD values, the endocortical dataset has variable endocortical width and tBMD values, and the beam hardening dataset has variable levels of total BMD.

The dimensions of each synthetic dataset were selected to minimise size, while still limiting the architectural change between pixels to less than $5\,\mathrm{mg/cm^3}$ for density and 0.05 mm for thickness for a scan resolution of 0.5 mm. The imitation tissues were ordered to reflect the structure of human bone along the measurement profiles: soft tissue, cortical bone, endocortical bone, and trabecular bone. The cortical surface occurs at the transition from soft tissue to cortical bone. This was defined by a measurement mesh with nodes distributed in a regular grid pattern with a 0.5 mm resolution. The quoted cortical thicknesses and cBMD for each synthetic dataset are derived from the dense cortical and endocortical features using Equation 2.6. The properties of the imitation tissues were defined using the tables of X-ray mass attenuation and densities provided on the NIST Standard Reference Database [57].

The synthetic datasets were imaged in air using the method described in Section 2.3. Each dataset was extended by an additional 4 mm using a nearest neighbour filter to ensure the edges of the sampled region of the dataset were minimally affected by blurring between the imitation tissue values and the surrounding air during the simulation of the image acquisition process. In addition, each simulated QCT scan included a synthetic calibration phantom designed to mimic calibration phantoms used in clinical QCT scans. The synthetic calibration phantom was positioned next to each synthetic dataset as discussed in Sections 2.4.4 and 2.4.5 and used to calibrate the scan.

2.4.1 Cortical dataset

The synthetic cortical dataset was designed to examine the effects of variations in the cortical thickness and cBMD on performance over the conceivable range of cortical architectures. The dataset is rectangular with an outer dimension of $20\,\mathrm{mm} \times 72\,\mathrm{mm} \times 220\,\mathrm{mm}$ without the nearest neighbour border and architectural variations occurring linearly along the y and z axes as shown in Figure 2.25a–2.25b. Profile measurements are made along the x axis. The dBMD varies along the z axis from 200–$1102.3\,\mathrm{mg/cm^3}$, where $200\,\mathrm{mg/cm^3}$ is the density of the tBMD in this dataset and $1102.3\,\mathrm{mg/cm^3}$ is the density of fully mineralised bone as calculated from the tables of X-ray mass attenuation and densities of bone provided on the



(a) Plan view (b) Elevation view

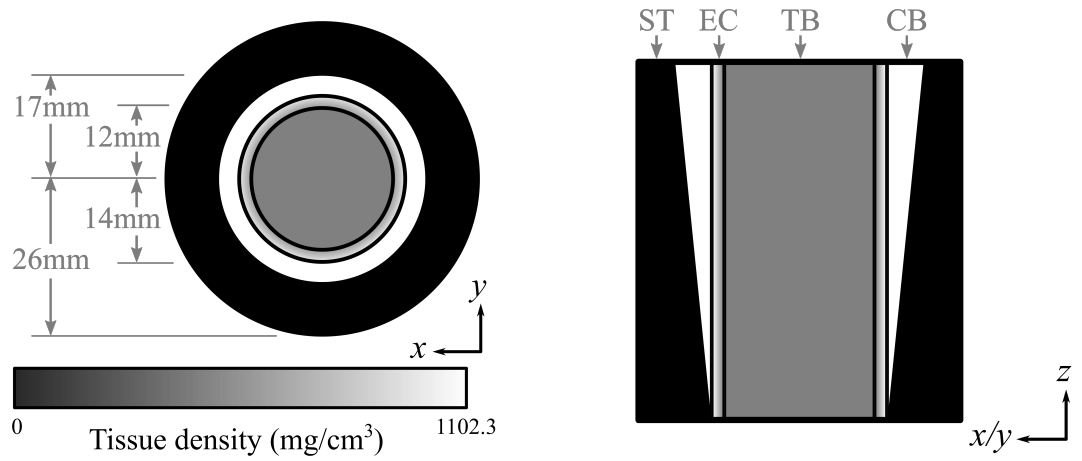
(c) Plan view (d) Elevation view

Fig. 2.25 The cortical (top), and endocortical (bottom) synthetic datasets without the nearest neighbour border.

NIST Standard Reference Database [57]. The dense cortical thickness varies along the y axis from 0.0–5.5 mm, which is combined with a constant endocortical width of 1 mm to give a mean cortical thickness that varies between 0.5–6.0 mm. This configuration was also used to produce a second synthetic dataset without an endocortical region and a cortical bone thickness that varies between 0–6 mm, which was used to consider the performance of the FWHM method at HR-pQCT resolutions in Chapter 3.

2.4.2 Endocortical dataset

The synthetic endocortical dataset was designed to examine how well the endocortical region can be measured and to what extent the tBMD affects this measurement. As with the synthetic cortical dataset, it is rectangular with outer dimensions of 20 mm × 72 mm × 220 mm without the nearest neighbour border and architectural variations occurring linearly along the y and z axes as shown in Figure 2.25c–2.25d. Profile measurements are made along the x axis. The tBMD varies along the z axis from 0–1102.3 $\mathrm{mg/cm^3}$, where 0 $\mathrm{mg/cm^3}$ is the density of the soft tissue and 1102.3 $\mathrm{mg/cm^3}$ is the density of the dBMD in the dataset. The endocortical width varies along the y axis from 0.0–6.0 mm, which is combined with a constant cortical thickness of 3 mm to give a mean cortical thickness varying between 3.0–6.0 mm.



(a) Plan view (b) Cross section view

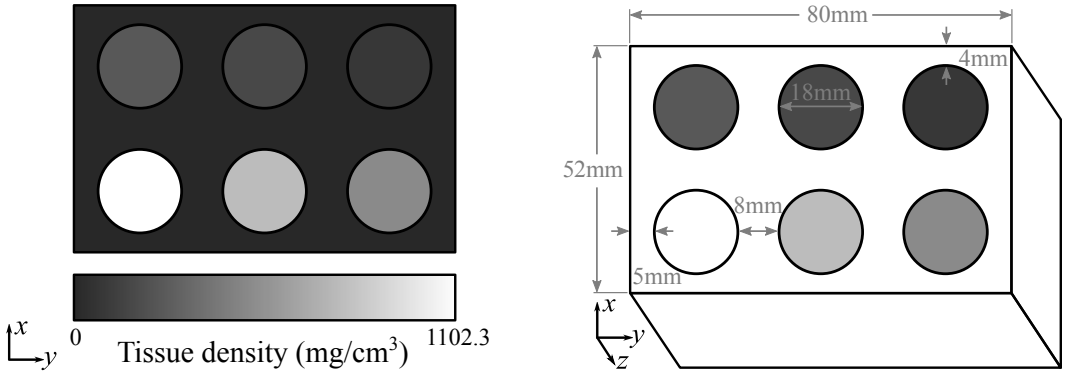
Fig. 2.26 The beam hardening synthetic dataset without the nearest neighbour border.

2.4.3 Beam hardening dataset

As discussed in Section 2.3.1, beam hardening leads to underestimation of densities inside objects and this underestimation increases with the penetration depth. The impact of this progressive density underestimation on the architectural measurements produced by the CBM methods has not been studied before. As such, the final synthetic dataset was designed to allow the impact of beam hardening on the performance of the different CBM methods to be examined. It was designed to contain variable levels of beam hardening with minimal architectural variations to facilitate identifying the impact of beam hardening on measurement accuracy. This was achieved using a cylinder with a variable diameter as shown in Figure 2.26. The only architectural variation in the dataset occurs along the z axis with an increase in dense cortical thickness from 0–6 mm and a corresponding decrease in the amount of surrounding soft tissue from 12–6 mm. Profile measurements are made radially inwards in the x-y plane, so that all measurements in a slice are of the same architecture. The dense cortical region has a dBMD of 1102.3 mg/cm^3, the trabecular region has a radius of 12 mm and a tBMD of 200 mg/cm^3, and the endocortical region has a width of 2 mm.

2.4.4 Synthetic calibration phantom

A synthetic calibration phantom was used to generate calibration curves and to study the impact of variations in tissue composition and architecture on the calibration process. The synthetic calibration phantom is shown in Figure 2.27. It was modelled after the QRM-BDC phantom but with core densities spanning the full range of BMD values occurring in bone. It contains six fixed density cores that each have a length of 220 mm and a radius of 9 mm. The cores are separated by 8 mm and arranged in a 2×3 grid. The cores are surrounded by soft tissue with a 4 mm border beyond the extent of the grid. The cores are composed of a mixture of water and bone where the ratios vary from 0 % bone to 96 % bone with densities of: 0 mg/cm^3, 100 mg/cm^3, 200 mg/cm^3, 500 mg/cm^3, 800 mg/cm^3, 1100 mg/cm^3. These



(a) Densities (b) Dimensions

Fig. 2.27 The synthetic calibration phantom with the nearest neighbour border.

were sampled and used to generate linear calibration curves, which allowed for the CBM density measurements to be compared with the actual BMD values. The calibration curve generation is discussed in Section 2.4.6.

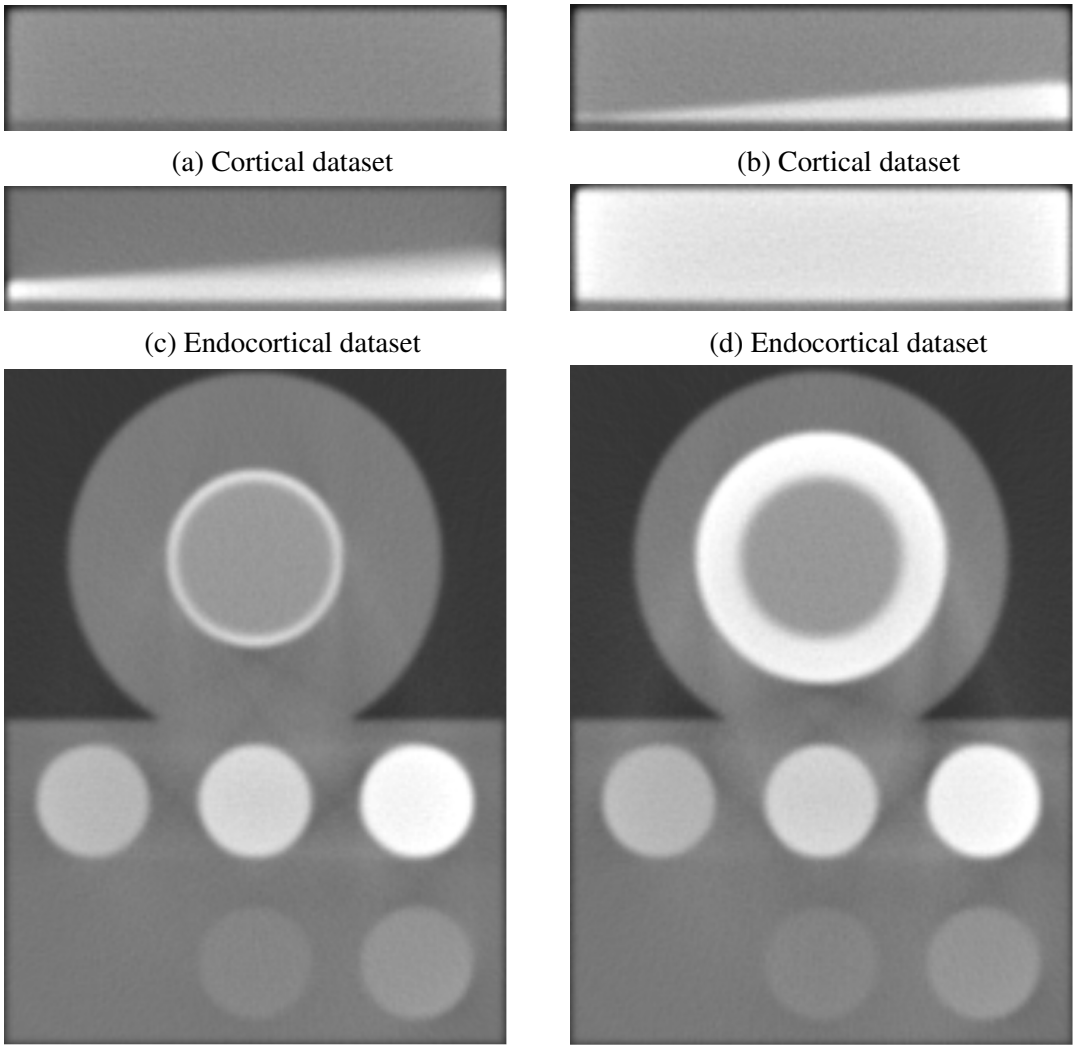
2.4.5 Simulated QCT scans

The synthetic datasets were transformed into simulated QCT scans by modelling the image acquisition process as discussed in Section 2.3 using the QCT X-ray source characteristic curve shown in Figure 2.17. During the simulated image acquisition, the datasets were sampled at a resolution of 0.05 mm then blurred and sub-sampled down to the desired imaging resolution using a Gaussian filter with a STD equal to the desired imaging resolution.

The top and bottom slices of the simulated 200 mA · s QCT scans of the cortical, the endocortical, and the beam hardening datasets are shown Figure 2.28. All synthetic datasets were scanned with a synthetic calibration phantom although these were cropped out of the cortical and endocortical slices in Figure 2.28. The detector spacing and number of X-ray photons were both varied between simulations, so that the performance of the CBM methods for blur values of 0.25 mm, 0.5 mm and 0.75 mm, and for the noise levels associated with 70 mA · s, 100 mA · s and 200 mA · s could be considered.

2.4.6 Density calibration

Although the simulated synthetic datasets could have been automatically calibrated using values sampled from within each simulated scan as the BMD values are known exactly, a synthetic calibration phantom was used to calibrate each simulated scan. The synthetic calibration phantom was included beside the synthetic datasets during the simulation process as shown in Figure 2.29. The cores in the simulated scan were sampled and used to calibrate the scans. This allowed for the complex interplay between the beam hardening observed in a calibration phantom and a scan subject to be examined.

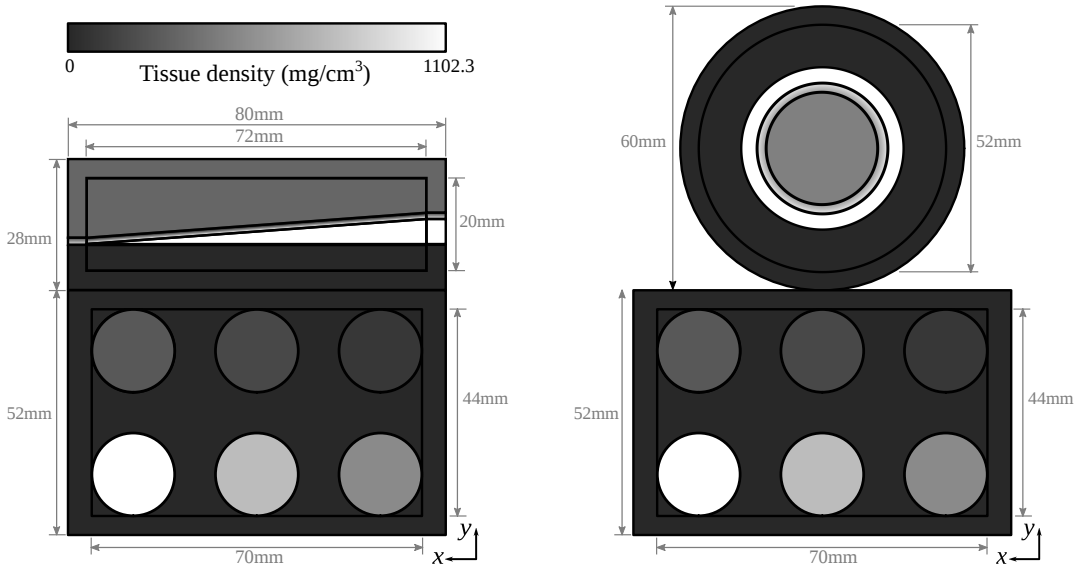


(a) Cortical dataset (b) Cortical dataset

(c) Endocortical dataset (d) Endocortical dataset

(e) Beam hardening dataset (f) Beam hardening dataset

Fig. 2.28 The top (right) and bottom (left) slices of the cortical, endocortical and beam hardening synthetic datasets simulated at 0.5 mm and 200 mA · s.

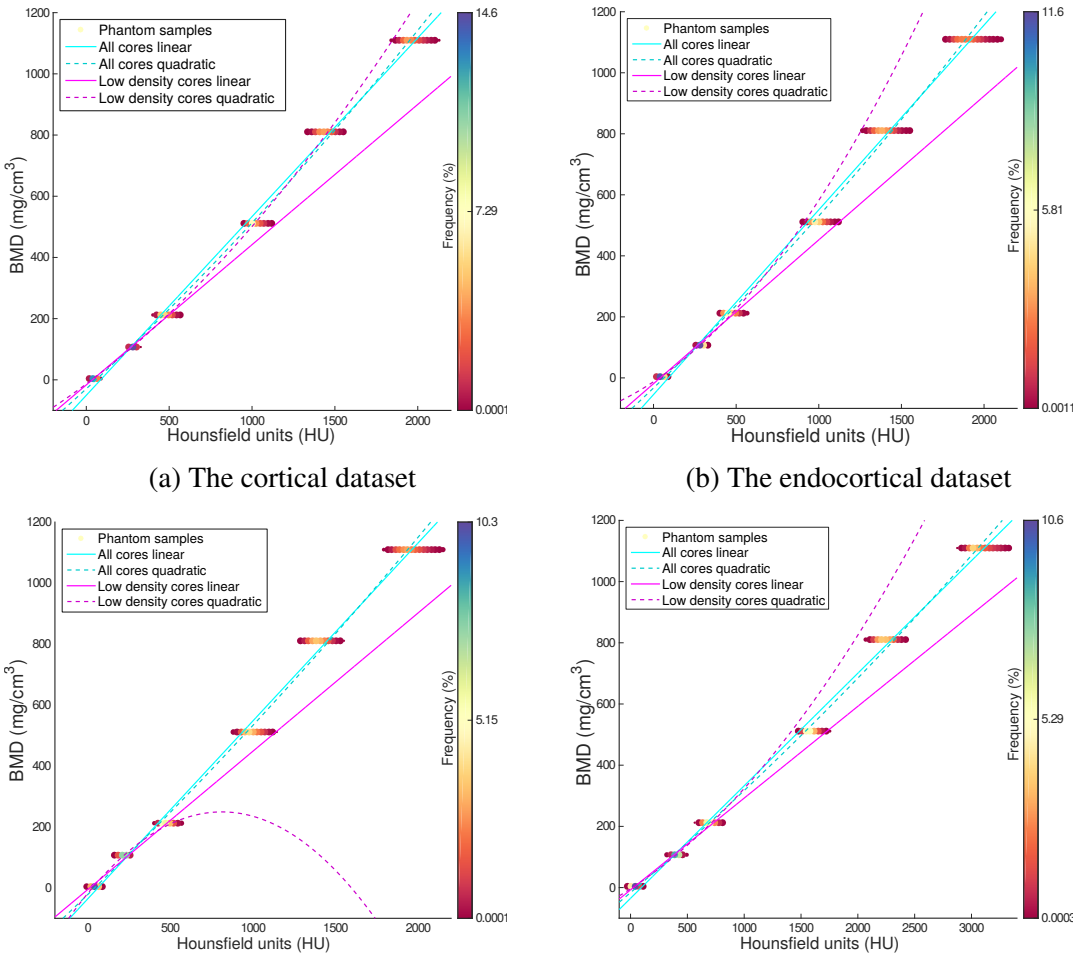


(a) Performance phantoms (b) Beam hardening phantom

Fig. 2.29 Plan views of the synthetic cortical, endocortical and beam hardening datasets with the calibration phantom dataset added during the simulation process.

Calibration curves were generated for each simulated scan using the HU values sampled from the central 50 % of each calibration core paired with its associated BMD value. The HU values were sampled in every slice at the pixel resolution, and show a convex relationship with the associated BMD values. Figures 2.30a-2.30c show the calibration values and curves for the 200 mA·s QCT scans of each phantom, while Figure 2.30d shows the calibration values and curve from a single slice of the 200 mA·s HR-pQCT cortical bone phantom without the endocortical region. The 200 mA·s QCT scan calibration values and curves are representative of the 100 mA·s and 70 mA·s QCT scan calibration curves.

Several calibration curves were generated from the sampled HU and BMD value pairs using linear and quadratic regression. These were used to assess the suitability of linear and quadratic calibration techniques, and the limitations of low density calibration phantoms. The regression curves generated from all six calibration cores are shown in blue and red, and the regression curves generated from the low density cores with densities in the range of 0–200 $\mathrm{mg/cm^3}$ are shown in dashed aqua and dashed fuchsia. Both curves generated from all cores represent the full range of densities fairly well. The quadratic curve is better able to represent the beam hardening experienced by the calibration phantom, but there is no guarantee this will be similar to the amount of beam hardening experienced by the subject's bone. The low density calibration curves poorly estimate high density values. The linear curve underestimates high density values, while the quadratic curve is poorly constrained for high densities and can produce wildly inaccurate high density estimates as shown in Figure 2.30c. Linear calibration curves were used in preference to quadratic curves over all datasets analysed in this thesis as the quadratic curves are susceptible to over-fitting high



(a) The cortical dataset

(b) The endocortical dataset

(c) The beam hardening dataset

(d) The cortical dataset with no endocortex

Fig. 2.30 Sampled calibration core values, and the linear and quadratic calibration curves for the simulated synthetic datasets.

density values when the calibration phantom only has low densities as with most datasets in this thesis.

Figure 2.30 shows that calibration phantoms should aim to span the same density range and experience similar beam hardening to the anatomical region of interest if the phantom is to provide high quality calibration data. The QRM-ESP achieves this by mimicking the density range and architecture of the lumbar spine. The phantom is scanned in isolation and used to calibrate other scans from the same scanner. Alternatively, the Mindways solid and QRM-BDC phantoms are shaped to allow flush placement alongside the scan subject, so that the phantom experiences similar levels of beam hardening to the subject. Unfortunately, these phantoms do not contain any high density BMD values as they were designed to calibrate trabecular bone and not cortical bone. This makes them less suitable for the calibration of cortical bone.

2.5 CBM assessment

The performance of the endocortical SSC-CBM and rectangular CBMv2 methods are investigated using the simulated QCT scans of the cortical, endocortical and beam hardening synthetic data. These were used: to make systematic adjustments to the endocortical SSC-CBM cost function as discussed in Section 2.5.2; to consider the impact of model choice on dBMD constraint generation as discussed in Section 2.5.3; to consider the performance of both methods in low and moderate noise environments as described in Sections 2.5.4 and 2.5.5; to consider the effect of scan resolution on both methods as described in Section 2.5.6; and to consider the performance of both methods in the presence of beam hardening as described in Section 2.5.7.

2.5.1 Experimental pipeline

Figure 2.31 shows the pipeline used to investigate the performance of the CBM methods. Each method was applied over simulated 200 mA · s, 100 mA · s and 70 mA · s QCT scans with resolutions of 0.25 mm, 0.5 mm and 0.75 mm of the three synthetic datasets. The resulting measurements were then compared against the architecture of each dataset to give the accuracies of each measurement. The measurements from both methods were smoothed with two iterations of mesh-connectivity-based precision smoothing. The precisions were extracted from the estimated model parameter variances of the final LMM optimisation at each measurement site. The smoothing was performed to remove outlier measurements resulting from very low precision optimisations. Unless stated otherwise, all simulations are for a resolution of 0.5 mm

The rectangular CBMv2 method requires a global estimate of the image system PSF blur, which was estimated as the mean PSF blur measured by the rectangular FWHM method in each scan. The endocortical SSC-CBM requires the radius of the distance filter, which should be selected based upon the expected distance over which the bone geometry varies. A value of 6 mm was used over each of the synthetic datasets.

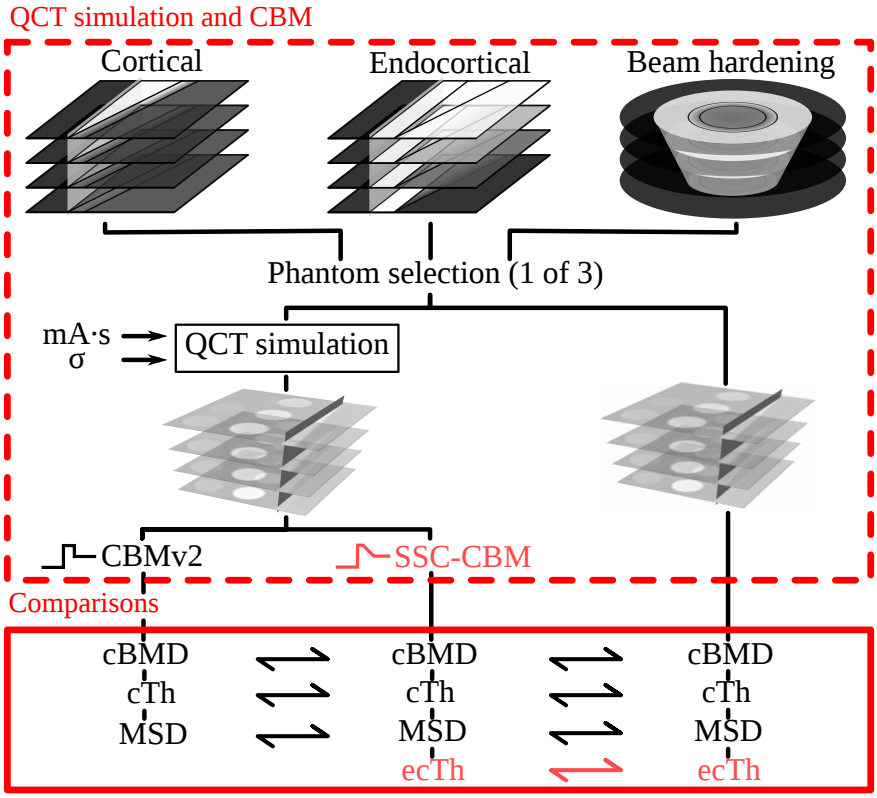


Fig. 2.31 The pipeline used to compare the different CBM methods using simulated QCT scans of synthetic datasets.

	$k_{\text{background}}$	k_{Gaussian}	σ_{P}	σ_{E}
Periosteal centred	0.075	0.925	1.75	2.45
Cortical centred rectangular	0.120	0.880	1.40	2.80
Cortical centred endocortical	0.050	0.950	1.70	3.50

Table 2.2 Weighting function parameters.

2.5.2 Cost function weighting

The experiment pipeline was used to investigate the effect of the relative weighting applied by the cost function on the performance of the endocortical SSC-CBM method. First the elements of the periosteal centred and thickness adjusted weighting functions were varied in isolation over a range of values. In each case the mean cortical thickness, cortical density, trabecular density, and endocortical thickness errors were plotted against the parameter values. This allowed for the iterative selection of the weighting function parameters shown in Table 2.2.

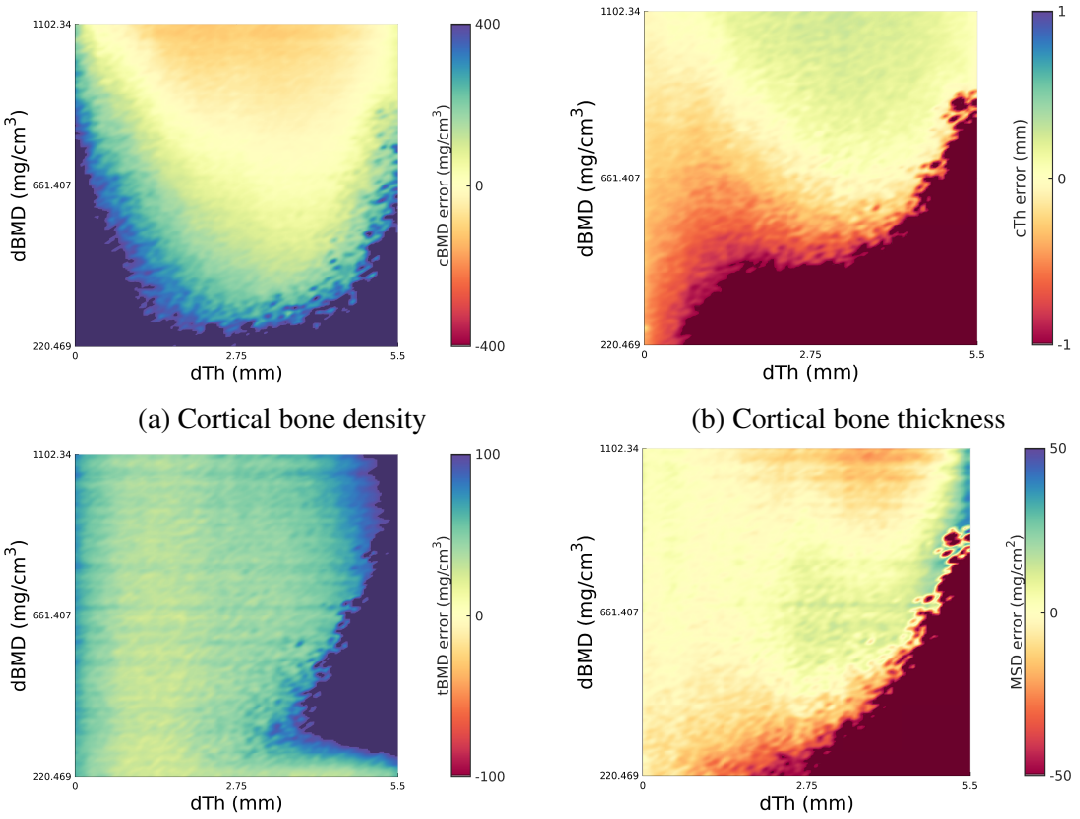
Next, the relative weight applied over the model-to-image misalignment errors, and the model validity errors was considered. This was varied in isolation at each stage of the CBM process starting with the first execution of the LMM. This showed that applying a lower model-to-image misalignment weighting to the first generically initialised LMM method, and a larger model-to-image misalignment weighting to all later uses of the LMM increased stability and reduced error. Exact values are not quoted as these are implementation specific.

2.5.3 dBMD constraint generation

The impact of using the rectangular or endocortical model to generate the dBMD constraint map on the overall performance of the SSC-CBM method was also considered using the experimental pipeline. Endocortical SSC-CBM measurements generated with both constraint maps were compared across the 70 mA·s, 100 mA·s and 200 mA·s QCT simulations of each synthetic dataset so that both accuracy and resilience to noise could be considered. The measurements generated with the endocortical dBMD constraints were generally more accurate over both datasets, and exhibited better precisions and resilience to noise over the cortical dataset. As a result, endocortical dBMD constraints were used in the endocortical SSC-CBM method throughout the rest of this thesis aside from over the two vertebral datasets considered in Chapters 3 and 4. In these datasets the additional stability afforded by the rectangular model warranted its use.

2.5.4 Low noise performance

The experimental pipeline was used to investigate the performance of the rectangular CBMv2 and the endocortical SSC-CBM methods in a moderately low noise environment using the 200 mA·s QCT simulations of the synthetic cortical and endocortical datasets. Figures 2.32–2.33 show rectangular CBMv2 and endocortical SSC-CBM measurement errors over the



(a) Cortical bone density

(b) Cortical bone thickness

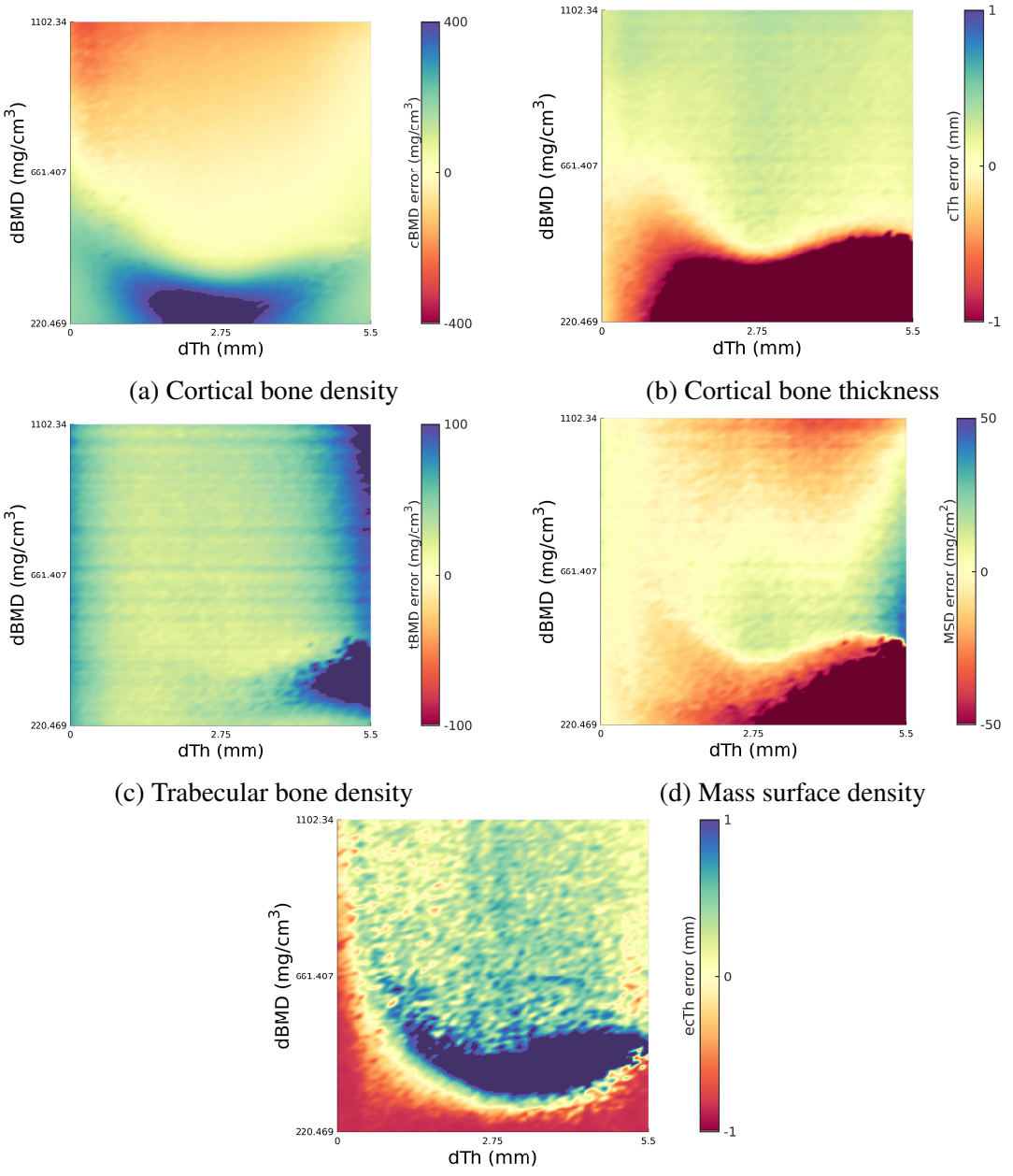
(c) Trabecular bone density

(d) Mass surface density

Fig. 2.32 Rectangular CBMv2 measurement errors over the low noise QCT simulation of the cortical dataset with saturated colour-maps where the method is unable to provide measurements that are at all accurate.

cortical dataset, and Figures 2.34–2.35 show the same errors over the endocortical dataset. The errors were calculated as the signed difference between the dataset architecture and measured values. The dBMD and dense cortical thickness errors of the endocortical SSC-CBM are very similar to the cBMD and cortical thickness errors so in the interest of space these are omitted. The colour-maps show the variations in error with bone architecture. Regions with saturated colour-maps indicate the bone architectures over which the respective methods are unable to provide measurements that are at all accurate.

Table 2.3 shows the bias and precision values of the errors generated from an architectural subset of each dataset. As in previous publications [123–125], the cortical dataset errors are calculated over three different cortical thickness ranges: 0.3–1 mm, 1–3 mm and 3–6 mm. As there is no variation in the cortical thickness over the endocortical dataset, these values are reported over endocortical thickness ranges instead. The architectural subsets of each synthetic dataset were chosen to better reflect the range of dBMD and tBMD values that commonly occur over the proximal femur and lumbar spine. dBMD values ranging from 551.2–1102.3 $\mathrm{mg/cm^3}$ in the cortical dataset and tBMD values ranging from 0–551.2 $\mathrm{mg/cm^3}$ in the endocortical dataset were used. The table also records the overall



(a) Cortical bone density

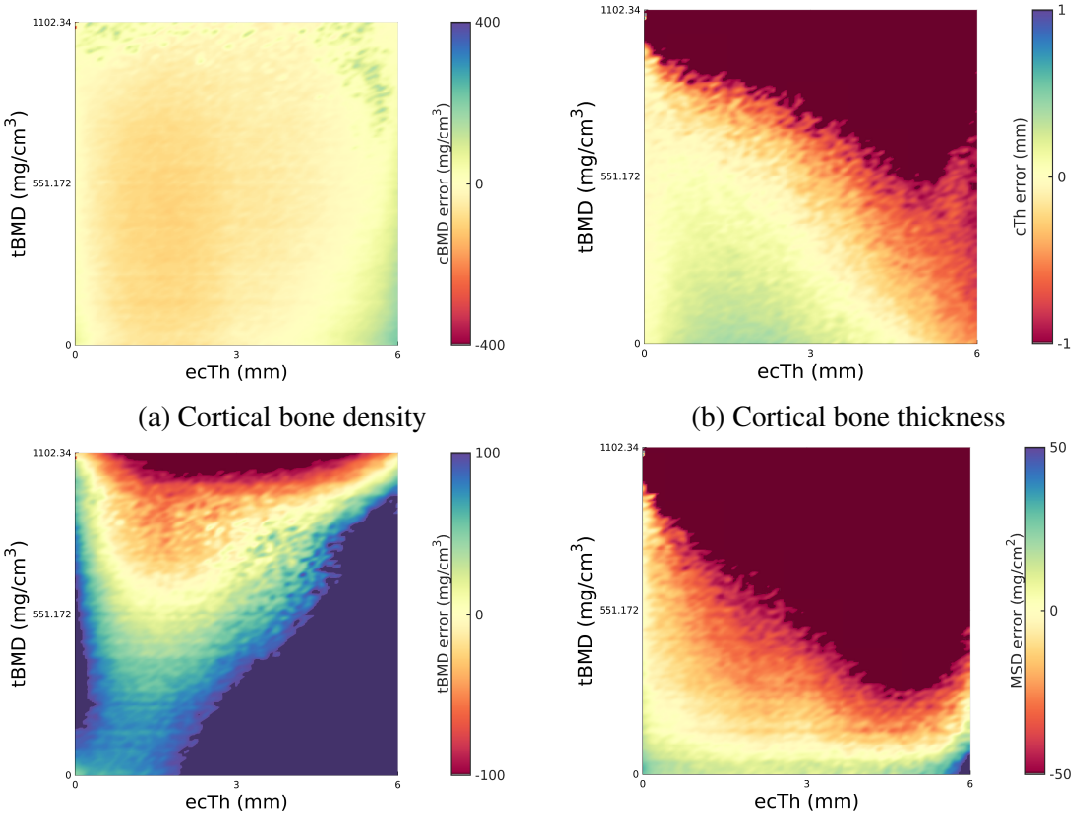
(b) Cortical bone thickness

(c) Trabecular bone density

(d) Mass surface density

(e) Endocortical thickness

Fig. 2.33 Endocortical SSC-CBM measurement errors over the low noise QCT simulation of the cortical dataset with saturated colour-maps where the method is unable to provide measurements that are at all accurate.



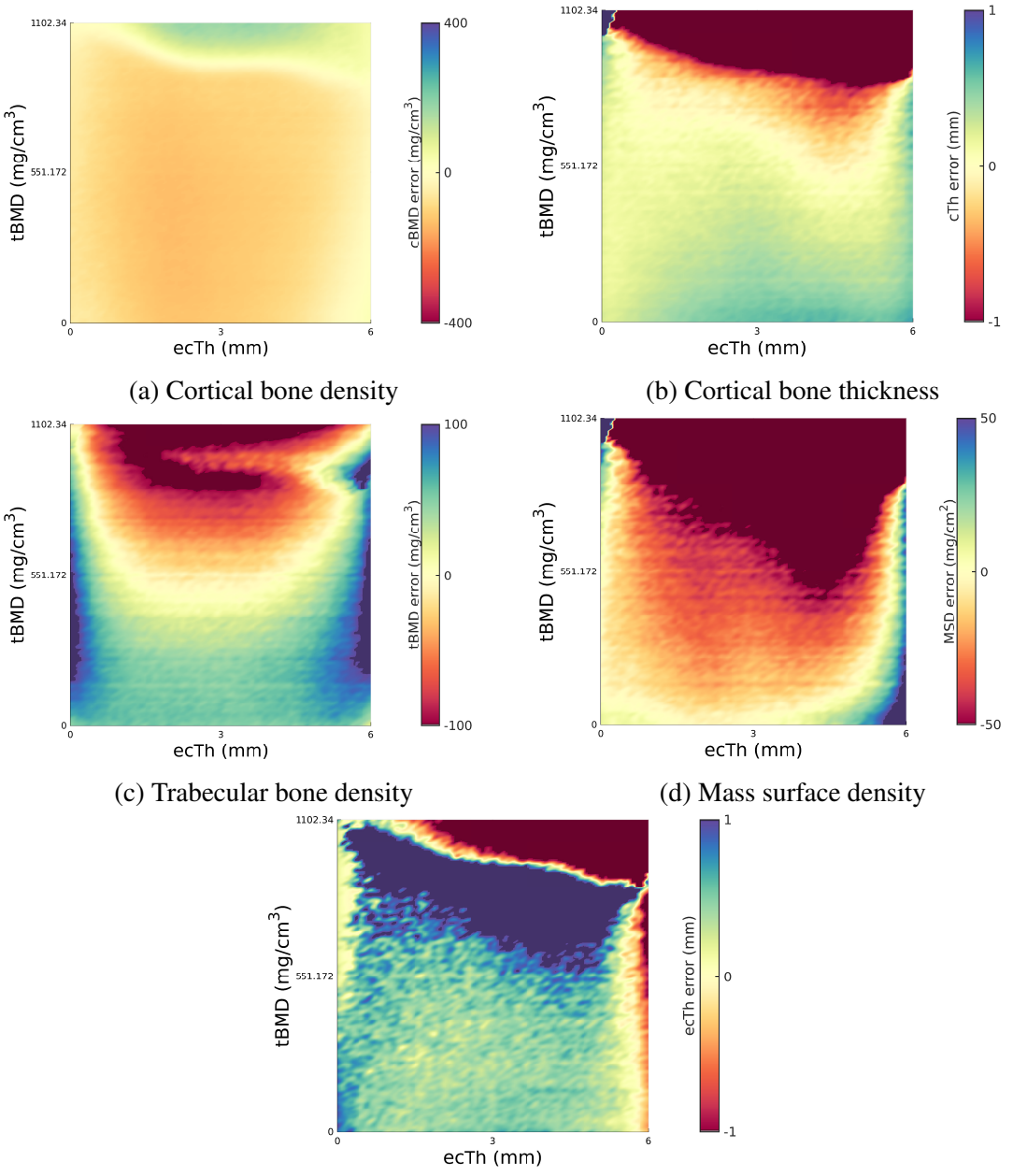
(a) Cortical bone density

(b) Cortical bone thickness

(c) Trabecular bone density

(d) Mass surface density

Fig. 2.34 Rectangular CBMv2 measurement errors over the low noise QCT simulation of the endocortical dataset with saturated colour-maps where the method is unable to provide measurements that are at all accurate.



(a) Cortical bone density

(b) Cortical bone thickness

(c) Trabecular bone density

(d) Mass surface density

(e) Endocortical thickness

Fig. 2.35 Endocortical SSC-CBM measurement errors over the low noise simulation of the endocortical dataset with saturated colour-maps where the method is unable to provide measurements that are at all accurate.

stability of each method, where stability is defined as the number of physically valid models before precision smoothing out of the total number of measurement sites within the considered architectural subset.

Figures 2.36–2.38 show the confidence interval (CI) plots of the CBM measurements against the actual values, which show the median, and 40th, 68th, 80th and 95th percentiles of the measurements against the actual values over the same architectural subsets that were used to calculate the biases and precisions. The CI plots were only generated for architectural values that varied in each synthetic dataset as shown in Figure 2.36 for the rectangular CBMv2 measurements, and in Figures 2.37–2.38 for the endocortical SSC-CBM measurements.

2.5.5 Moderate noise performance

The resilience of the CBM methods to noise was also considered with $70\,\mathrm{mA \cdot s}$ QCT simulations of the cortical and endocortical synthetic data. A selection of the resulting measurement errors across both datasets is shown in Figure 2.39 for the rectangular CBMv2 method and Figure 2.40 for the endocortical SSC-CBM method.

2.5.6 Resolution and performance

The effect of scan resolution on the performance of the CBM methods was also considered through a comparison of the measurement errors over 0.25 mm, 0.5 mm and 0.75 mm resolution low noise QCT simulations of the cortical and endocortical synthetic datasets. A selection of the measurement errors across the 0.25 mm and 0.75 mm resolution QCT simulations is shown in Figures 2.41 and 2.43 for the rectangular CBMv2 method and Figures 2.42 and 2.44 for the endocortical SSC-CBM method.

2.5.7 Beam hardening

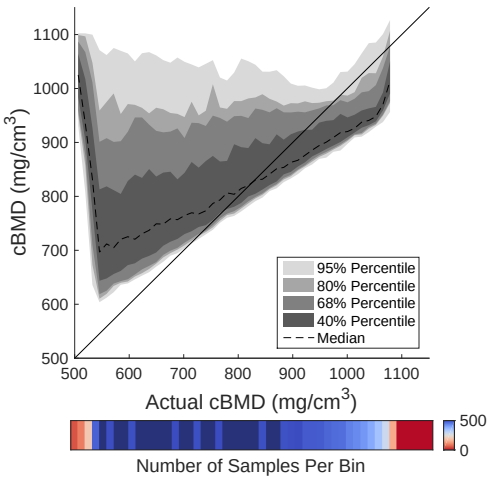
The $70\,\mathrm{mA \cdot s}$, $100\,\mathrm{mA \cdot s}$ and $200\,\mathrm{mA \cdot s}$ simulations of the synthetic beam hardening dataset were used to examine the effect of beam hardening on the measurement errors of the two CBM methods. Figure 2.45 shows the progressive reduction in density caused by beam hardening along a typical density profile sampled from the beam hardening phantom. Figure 2.46 shows the mean error in the cBMD, cortical thickness, tBMD and endocortical measurements across each slice of the simulated beam hardening scans.

2.6 Discussion

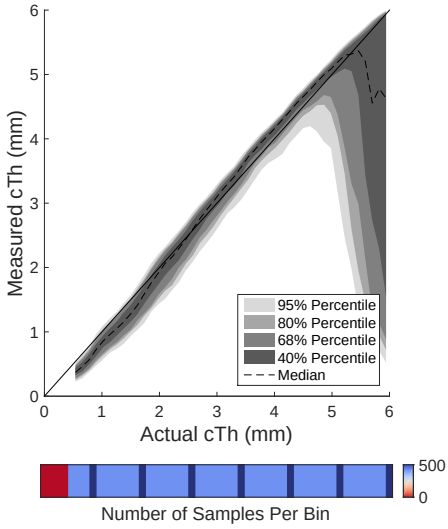
The CBM measurement errors in Figures 2.32–2.35 show that the endocortical SSC-CBM method provides low errors over a wider range of bone architectures than the rectangular CBMv2 method, and that the rectangular CBMv2 method is more sensitive to low dBMD and high tBMD values. This is reflected in the precision values reported in Table 2.3, which are generally lower for the endocortical SSC-CBM method. As discussed in Section 2.2.2, the endocortical model is less stable than the rectangular model as it has an additional parameter

Table 2.3 The biases, precisions and overall stability of the CBM measurements over the synthetic cortical and endocortical datasets. A dash is used for features not measured by the rectangular model.

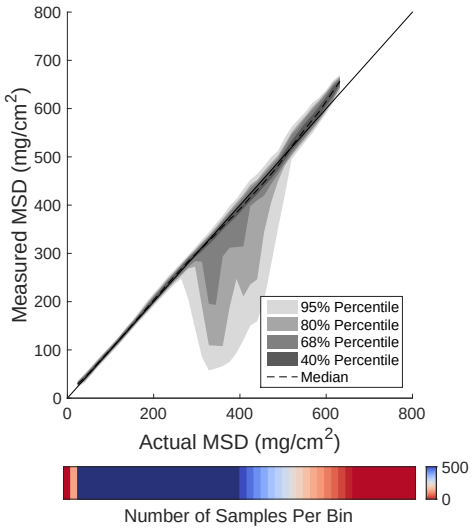
Quantity	Thickness Ranges (mm)	Cortical dataset		Endocortical dataset	
		Rectangle CBMv2	Endocortical SSC-CBM	Rectangle CBMv2	Endocortical SSC-CBM
Stability (%)	$0.0 \leq t < 6.0$	99.4	100	99.7	100
cBMD (mg/cm^3)	$0.3 \leq th < 1.0$	286±148	−84±112	−57±19	−89±12
	$1.0 \leq th < 3.0$	65±116	−89±65	−80±15	−127±8
	$3.0 \leq th < 6.0$	13±88	−51±48	2±55	−81±36
cTh (mm)	$0.3 \leq th < 1.0$	−0.18±0.12	0.13±0.13	0.15±0.08	0.19±0.06
	$1.0 \leq th < 3.0$	−0.12±0.22	0.19±0.17	0.15±0.14	0.31±0.11
	$3.0 \leq th < 6.0$	−0.30±1.11	0.24±0.07	−0.38±0.32	0.32±0.16
tBMD (mg/cm^3)	$0.3 \leq th < 1.0$	59±8	57±9	71±14	60±15
	$1.0 \leq th < 3.0$	39±5	31±5	73±29	35±21
	$3.0 \leq th < 6.0$	87±81	51±22	179±62	54±27
MSD (mg/cm^2)	$0.3 \leq th < 1.0$	2.0±2.0	2.1±2.2	−4.3±10.8	−10.5±8.2
	$1.0 \leq th < 3.0$	0.1±3.4	−3.5±5.4	−16.9±16.3	−22.5±10.7
	$3.0 \leq th < 6.0$	−16.1±61.1	−2.4±14.1	−36.1±34.0	−11.7±26.7
dBMD (mg/cm^3)	$0.3 \leq th < 1.0$	-	−138±131	-	−81±15
	$1.0 \leq th < 3.0$	-	−92±76	-	−132±12
	$3.0 \leq th < 6.0$	-	−50±49	-	−95±32
dTh (mm)	$0.3 \leq th < 1.0$	-	0.17±0.14	-	−0.08±0.09
	$1.0 \leq th < 3.0$	-	0.04±0.19	-	0.09±0.13
	$3.0 \leq th < 6.0$	-	0.07±0.11	-	0.13±0.24
ecTh (mm)	$0.3 \leq th < 1.0$	-	−0.09±0.33	-	0.54±0.11
	$1.0 \leq th < 3.0$	-	0.30±0.21	-	0.43±0.09
	$3.0 \leq th < 6.0$	-	0.34±0.19	-	0.37±0.27



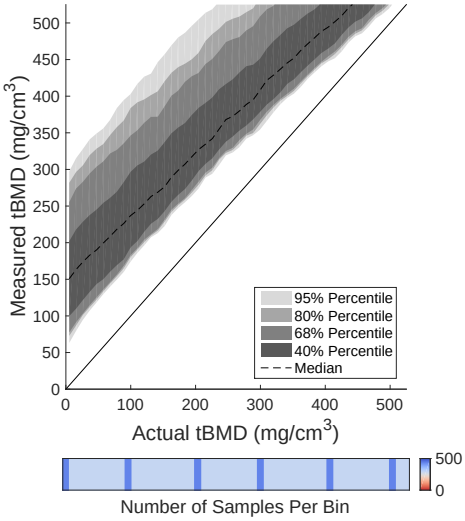
(a) Cortical bone density



(b) Cortical bone thickness

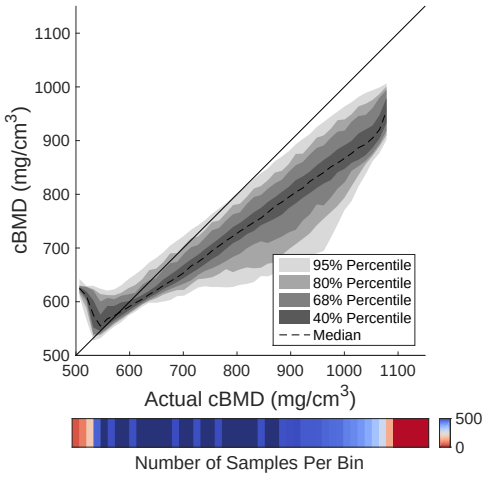


(c) Mass surface density

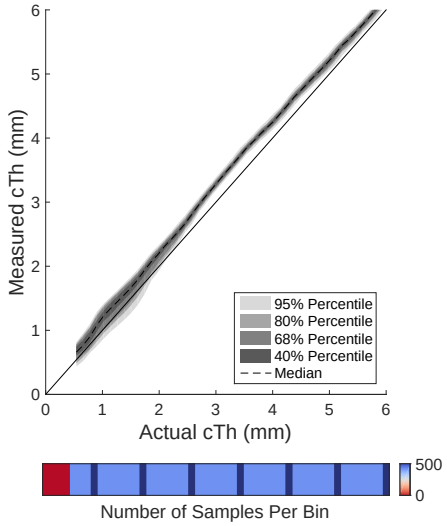


(d) Trabecular bone density

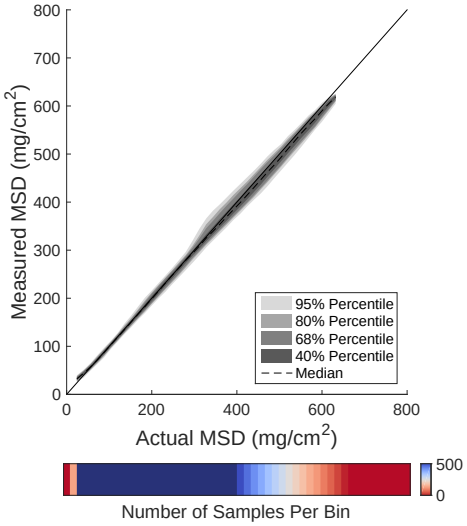
Fig. 2.36 CI plots of the rectangular CBMv2 measurements over the low noise cortical (a-c) and endocortical (d) datasets.



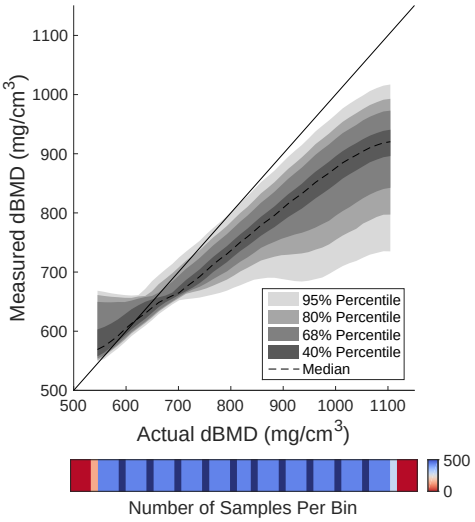
(a) Cortical bone density



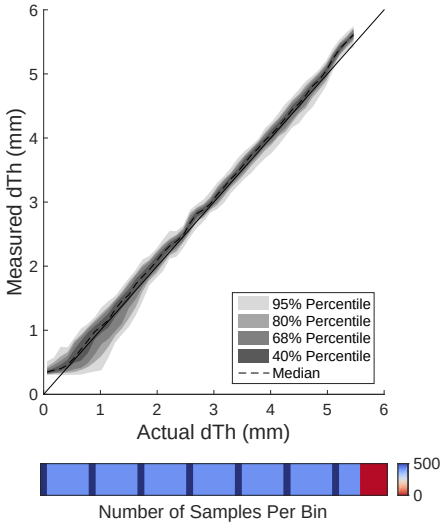
(b) Cortical bone thickness



(c) Mass surface density

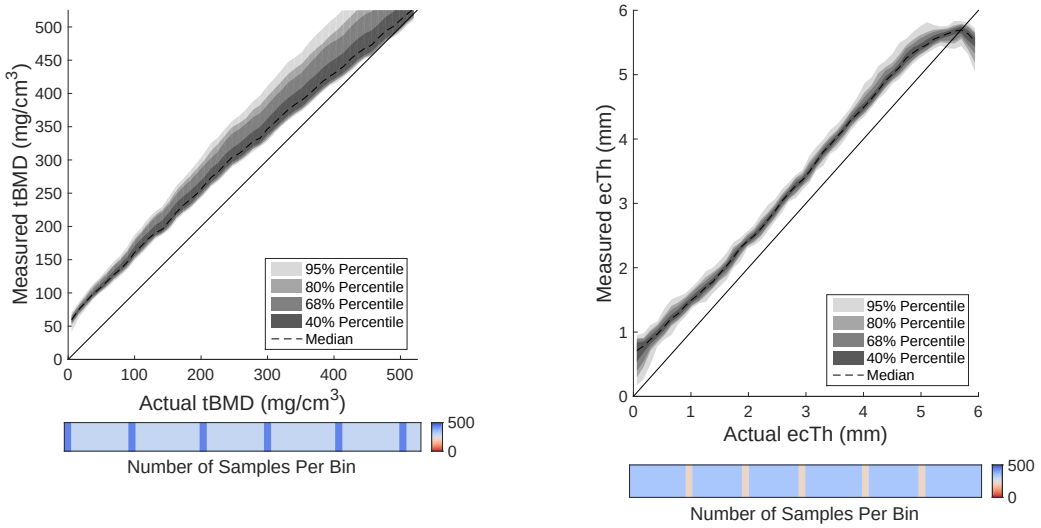


(d) Dense cortical bone density



(e) Dense cortical bone thickness

Fig. 2.37 CI plots of endocortical SSC-CBM measurements over the low noise cortical dataset.



(a) Trabecular bone density error

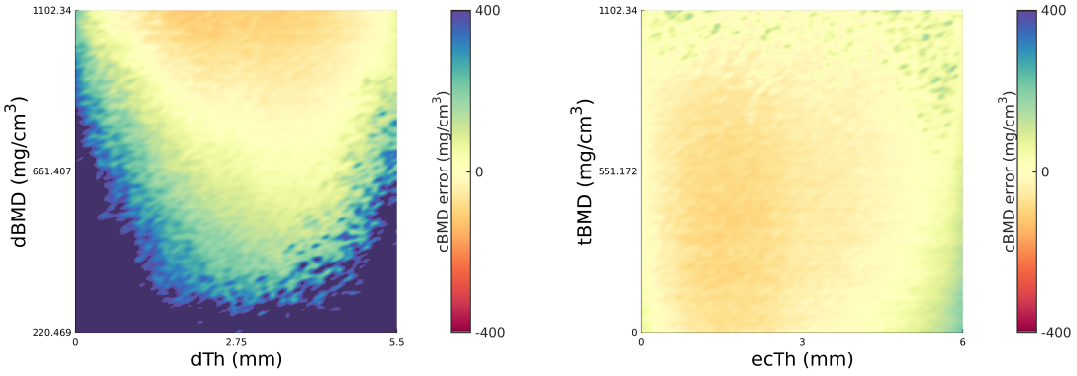
(b) Endocortical width error

Fig. 2.38 CI plots of endocortical SSC-CBM measurements over the low noise endocortical dataset.

that must be constrained. Despite this, the endocortical SSC-CBM method has a greater stability as reported in Table 2.3. This indicates that the smooth dBMD constraint combined with the cost function weightings selected in Section 2.5.2 improves the stability of the endocortical SSC-CBM method. Figures 2.39 and 2.40 show select measurement errors across the 70 mA · s QCT simulations of the cortical and endocortical synthetic datasets. The unsaturated areas of these colour-maps show that the range of architectures where the methods reliably provide low error measurements shrinks with increasing noise, so the larger it is in the 200 mA · s QCT simulations the better.

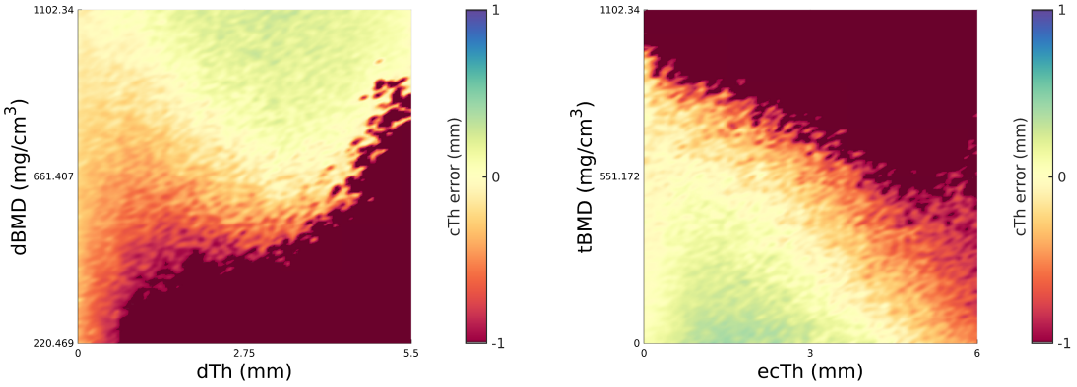
As discussed in Treece et al. (2015), the minimum cortical thickness above which cortical features can be measured unambiguously is a function of image blur [123]. This is reflected in the cortical thickness and cBMD errors over the synthetic cortical dataset, which are shown in Figures 2.41–2.42 for 0.25 mm and 0.75 mm QCT simulations, and Figures 2.32–2.33 for a 0.5 mm QCT simulation. These show that an improved simulation resolution leads to increased measurement quality where the dense cortical thickness is thin. In contrast, the measurement quality reduces with improved resolution where the dense cortical thickness is large as the optimiser increasingly fits the cortical model to local minima caused by noise along the dense cortical region. Figures 2.33 and 2.42 for the cortical datasets and Figures 2.35 and 2.44 for the endocortical datasets show the same general behaviour for endocortical measurements, but with a greater deterioration in measurement quality with a reduction in resolution. This reflects the lower stability of the endocortical thickness measurement.

The error maps show moderate cBMD errors over architectural regions with dense cortical thicknesses that should be large enough to eliminate PVE. This is predominantly due to calibration errors caused by differences in the level of beam hardening experienced within the



(a) Cortical bone density

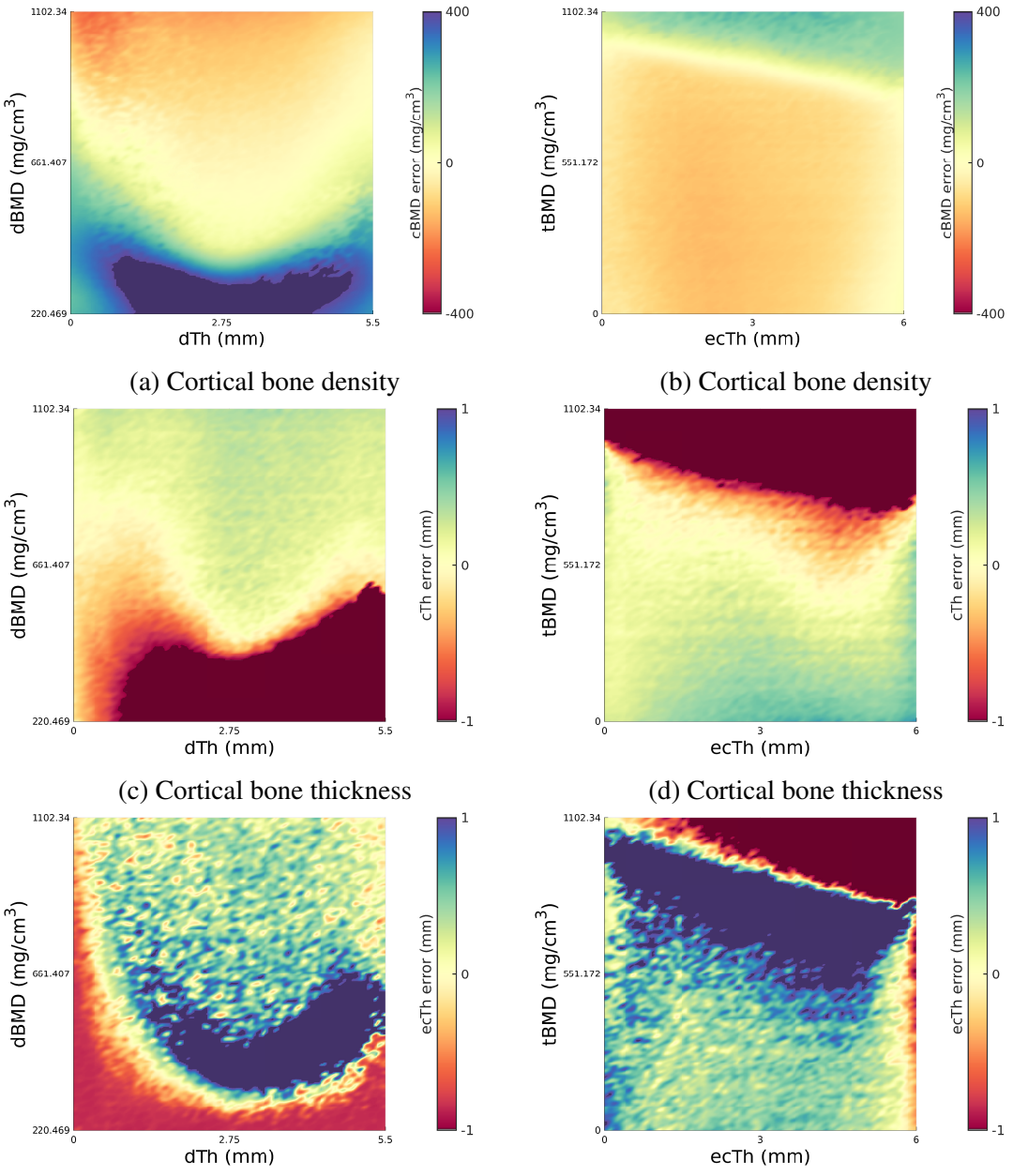
(b) Cortical bone density



(c) Cortical bone thickness

(d) Cortical bone thickness

Fig. 2.39 Rectangular CBMv2 measurement errors applied to moderate noise simulations of the cortical (a and c) and the endocortical datasets (b and d) with saturated colour-maps where the method is unable to provide measurements that are at all accurate.



(a) Cortical bone density

(b) Cortical bone density

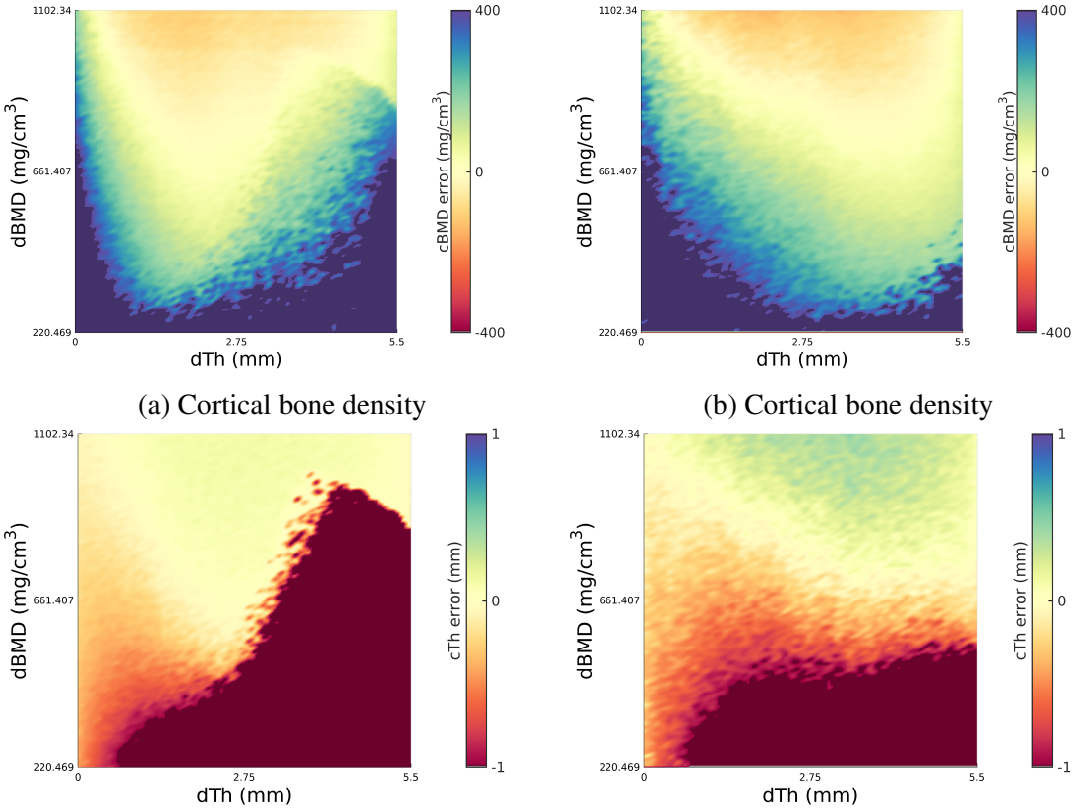
(c) Cortical bone thickness

(d) Cortical bone thickness

(e) Endocortical thickness

(f) Endocortical thickness

Fig. 2.40 Endocortical SSC-CBM measurement errors over the moderate noise simulations of the cortical (a, c and e) and the endocortical datasets (b, d and f) with saturated colour-maps where the method is unable to provide measurements that are at all accurate.



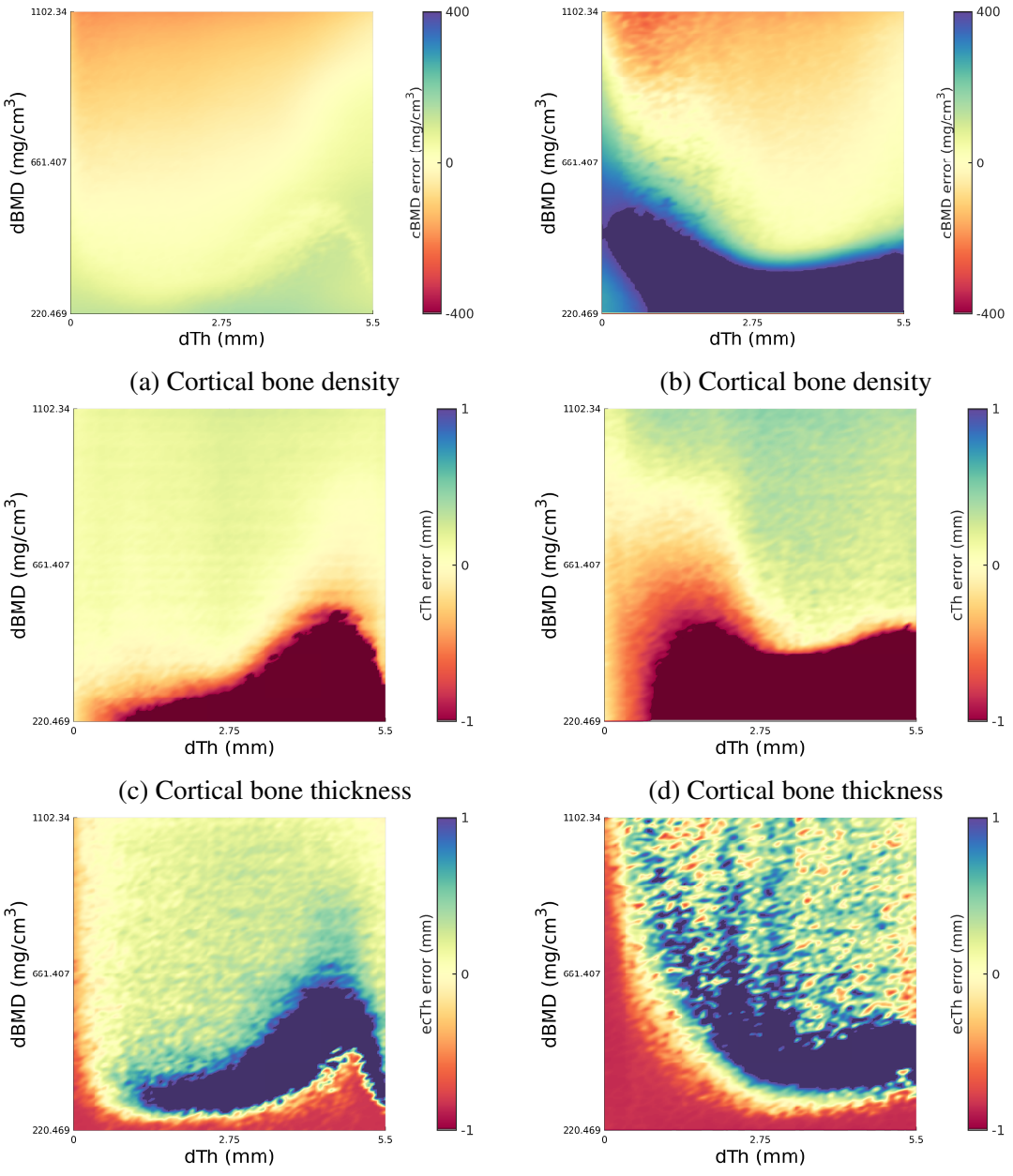
(a) Cortical bone density

(b) Cortical bone density

(c) Cortical bone thickness

(d) Cortical bone thickness

Fig. 2.41 Rectangular CBMv2 errors of high (a and c) and low (b and d) resolution simulations of the cortical dataset with saturated colour-maps where the method is unable to provide measurements that are at all accurate.



(a) Cortical bone density

(b) Cortical bone density

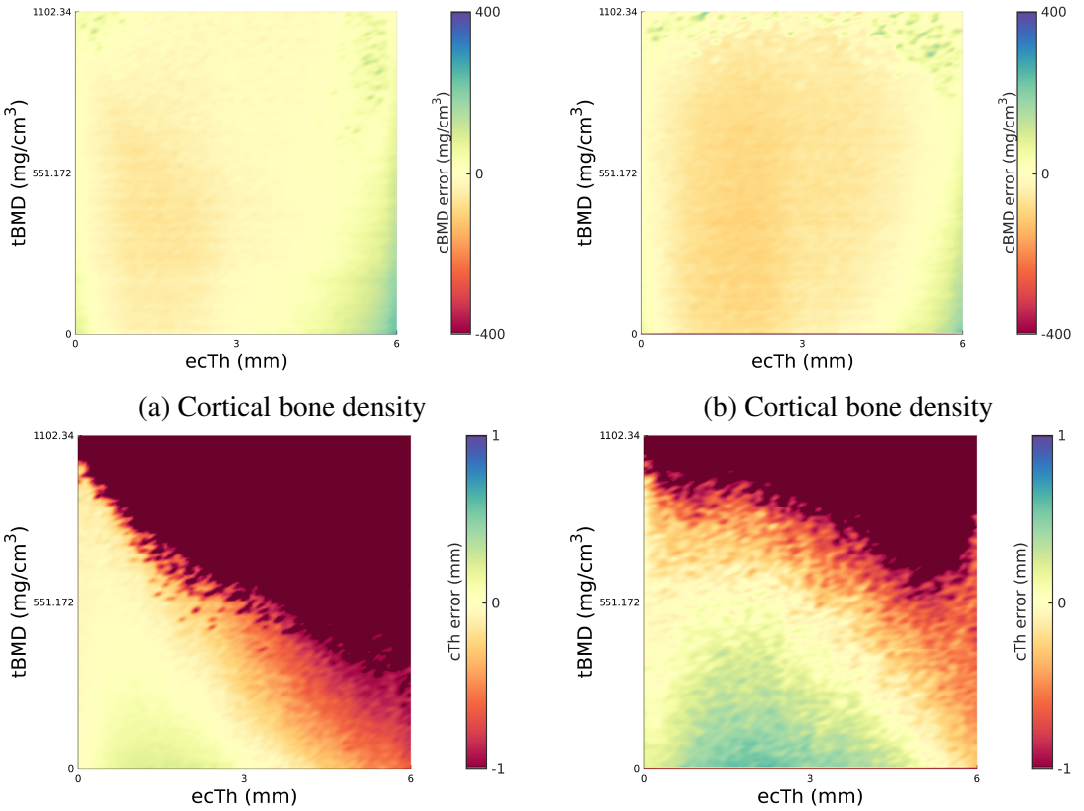
(c) Cortical bone thickness

(d) Cortical bone thickness

(e) Endocortical thickness

(f) Endocortical thickness

Fig. 2.42 Endocortical SSC-CBM errors of high (a, c and e) and low (b, d and f) resolution simulations of the cortical dataset with saturated colour-maps where the method is unable to provide measurements that are at all accurate.



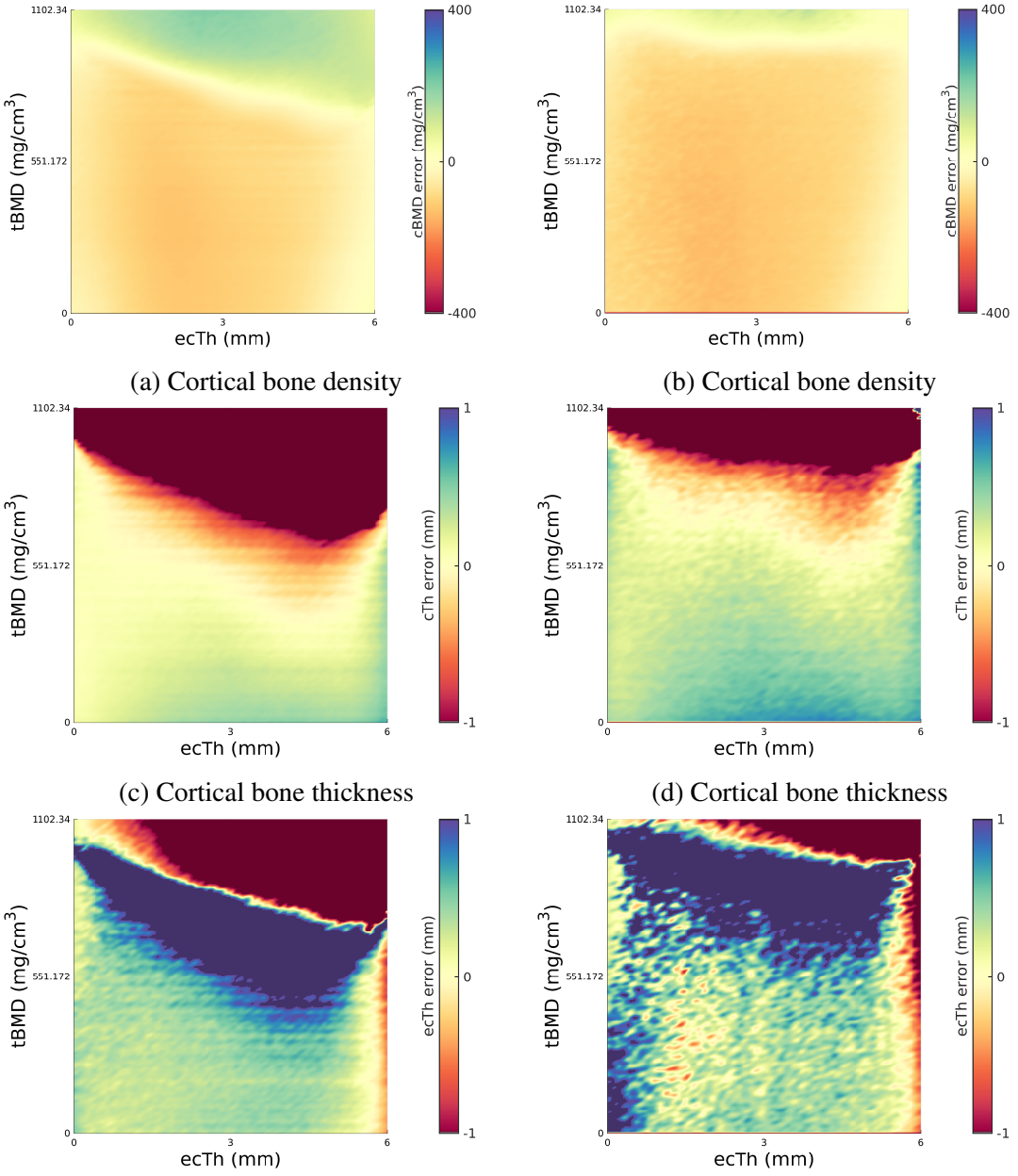
(a) Cortical bone density

(b) Cortical bone density

(c) Cortical bone thickness

(d) Cortical bone thickness

Fig. 2.43 Rectangular CBMv2 errors of high (a and c) and low (b and d) resolution simulations of the synthetic endocortical dataset.



(a) Cortical bone density

(b) Cortical bone density

(c) Cortical bone thickness

(d) Cortical bone thickness

(e) Endocortical thickness

(f) Endocortical thickness

Fig. 2.44 Endocortical SSC-CBM errors of high (a, c and e) and low (b, d and f) resolution simulations of the synthetic endocortical dataset.

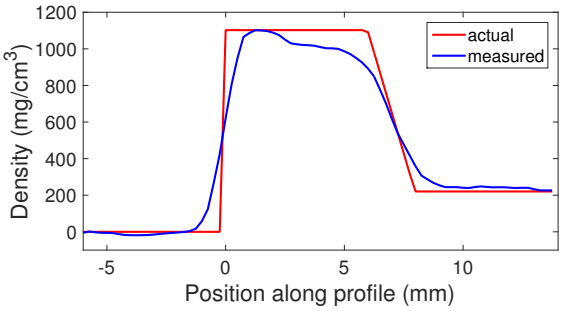
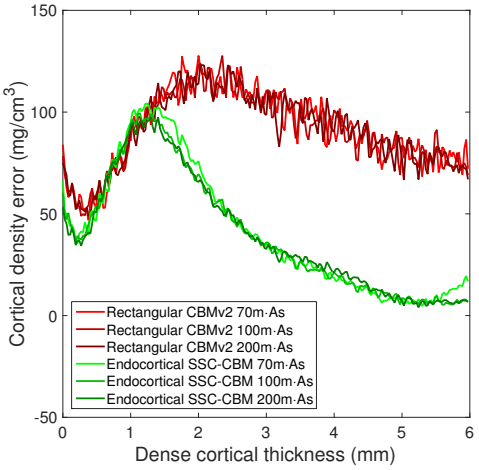
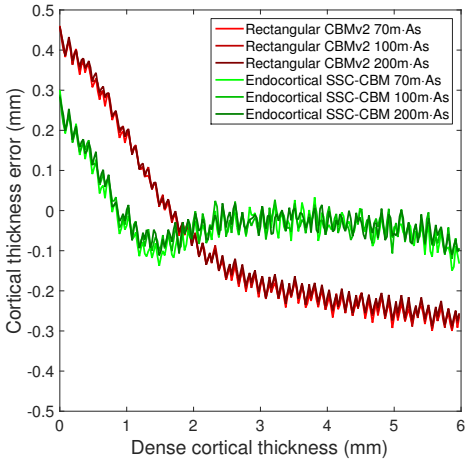


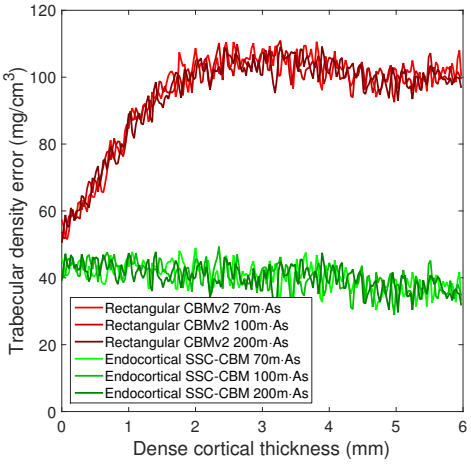
Fig. 2.45 The effect of beam hardening on the density of a cortical profile.



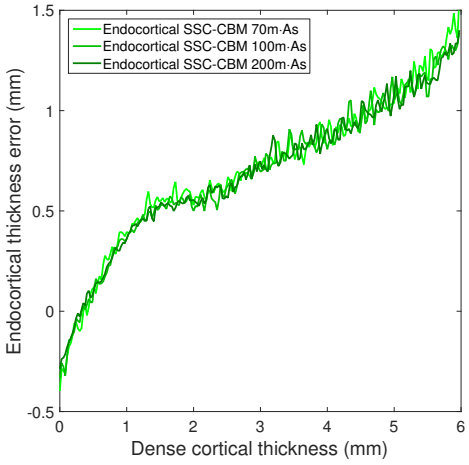
(a) cBMD error



(b) Cortical thickness error

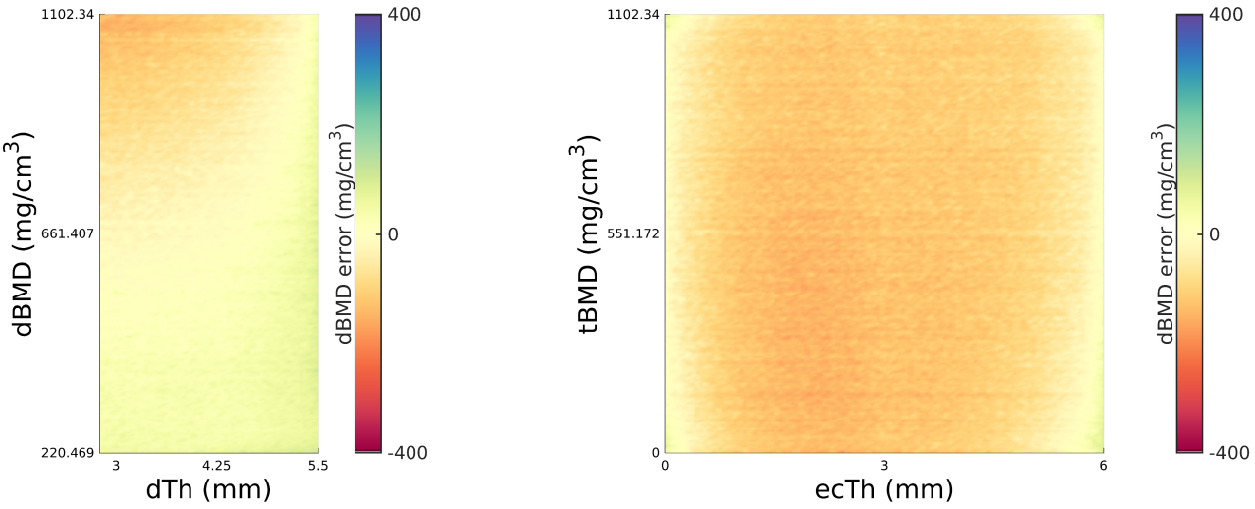


(c) tBMD error



(d) Endocortical thickness error

Fig. 2.46 The impact of increasing beam hardening on accuracy.

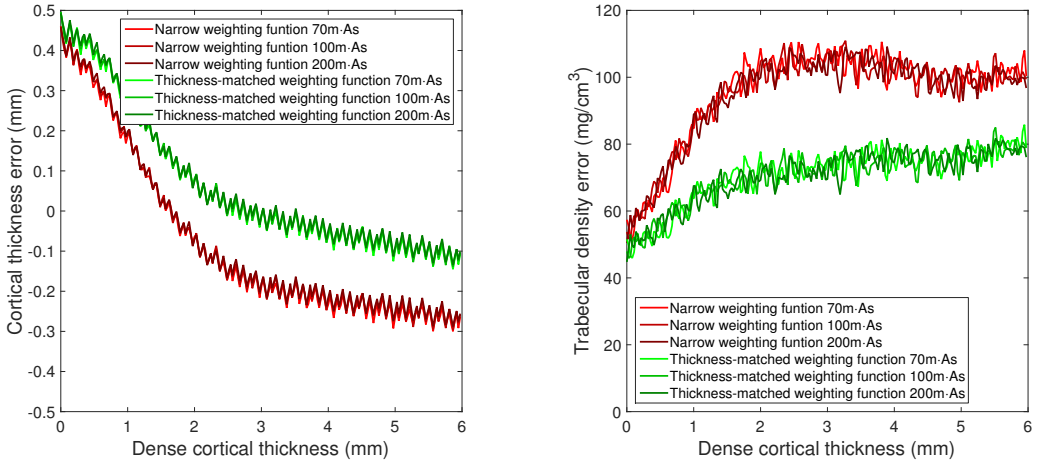


(a) Cortical dataset (b) Endocortical dataset

Fig. 2.47 The difference between the calibrated and actual dBMD values.

calibration cores and over the simulated datasets as discussed in Section 2.4.6. Figure 2.47 shows the difference between the maximum BMD value sampled along each profile and the actual dBMD value of the synthetic dataset at that profile location. The colour-map range is the same as the cBMD errors in Figures 2.32–2.33. The calibration error surface of the cortical dataset is trimmed to exclude architectures with a dense cortical thickness less than 3 mm to ensure the absence of PVE. These calibration errors are similar to the cBMD errors in Figures 2.32–2.33 over the subset of dBMD and tBMD values that commonly occur over the proximal femur and lumbar spine. This indicates that much of the dBMD and cBMD errors experienced over these architectural ranges are calibration errors caused by localised beam hardening effects. The increasing dBMD error with decreasing dense cortical thickness toward the centre of the measurement surfaces occurs as the amount of beam hardening experienced increases as the quantity of mineralised bone encountered by X-ray beams penetrating the sides of the synthetic dataset increases. The effect of beam hardening on BMD calibration cannot be eliminated, but its impact can be limited as discussed in Section 2.4.6.

The rectangular CBMv2 method is more affected by beam hardening than the endocortical SSC-CBM method as shown in Figure 2.46. Much of the improved performance of the endocortical SSC-CBM method is due to the use of a thickness adjusted weighting function as illustrated by Figure 2.48, which shows the cortical thickness and tBMD beam hardening errors for the rectangular CBMv2 method with both its original periosteal centred weighting function and with the thickness adjusted weighting function used in the endocortical SSC-CBM method. The measurements with the thickness adjusted weighting function are less affected by beam hardening, as increased weighting over the endocortical region means that compromising the endocortical edge and tBMD parameters to account for beam hardening effects over the cortical region receives a greater relative penalisation. Figure 2.46d shows that the endocortical thickness remains highly affected by beam hardening even with the



(a) Cortical thickness error

(b) tBMD error

Fig. 2.48 The impact of different weighting functions on the rectangular CBMv2 beam hardening errors.

cortical centred weighting functions. This is likely due to the similarities between an increase in the endocortical thickness and the progressive reduction in density caused by beam hardening. This may explain the constant bias in the endocortical thickness measurements of the endocortical synthetic data shown in Figure 2.38b. Figures 2.46 and 2.48 also show that the level of Poisson noise has little impact on the overall trend of the measurement errors and that the variability in measurement errors that it introduces is substantially less than the magnitude of the overall measurement errors. The regular oscillations in the rectangular CBMv2 cortical thickness errors are the result of the finite resolution of the synthetic datasets from which the simulations were generated. Finally, it should be noted that in both Figures 2.46 and 2.48 PVE appears to dominate the errors for dense cortical thicknesses below 1.5 mm.

2.7 Conclusions

The cortical, endocortical and beam hardening synthetic data were used to assess the performance of the rectangular CBMv2 and endocortical SSC-CBM methods over different architectures in detail. This shows that the endocortical SSC-CBM method allows for the extent of the endocortical region to be measured, while generally increasing the stability and precision of the rectangular CBMv2 architectural measurements. In addition, it is clear that the use of the thickness adjusted weighting function reduced the impact of beam hardening on cortical thickness and tBMD measurements.

Chapter 3

HR-pQCT cadaveric validation of CBM

3.1 Introduction

The simulated datasets used in Chapter 2 showed how the CBM methods perform over an idealised dataset with low noise, clean macro-structure and no micro-structure. Validation of QCT measurements over human bones is of critical importance as its complex multi-scale architecture is difficult to represent with simulated data, and any assumptions in its portrayal may have unforeseen consequences on the performance of QCT methods. HR-pQCT is a high resolution QCT technique with a nominal isotropic resolution of approximately 80–90 µm [17, 32, 48, 68]. This is much higher than the nominally 0.5 mm 'in-plane' pixel resolution typically offered by modern QCT scanners [45]. The higher resolution of HR-pQCT scans allow for them to be used to produce higher quality measurements for comparison with the QCT methods. In this chapter, the endocortical SSC-CBM method is compared against equivalent measurements extracted from HR-pQCT scans using cadaveric datasets of paired QCT and HR-pQCT scans from two different studies over the proximal femur and lumbar spine[1]. Datasets of the proximal femur and lumbar spine were used due to the high mortality and morbidity, increased disability, and relative frequency associated with fractures in these sites [23, 52, 64, 61, 135].

Morphological and model-based methods have both been used to extract feature measurements from HR-pQCT scans. Various methods have been used to measure some combination of cortical thickness, cortical density, cortical porosity, trabecular bone volume to total volume (BV/TV), trabecular number, trabecular spacing, and the width of the endocortical region. The endocortical SSC-CBM method requires a method capable of providing localised measurements of cortical thickness, cortical density, endocortical thickness and trabecular density measurements. No existing HR-pQCT techniques are capable of measuring all of these features, so a new method termed the high resolution tissue classifier (HRTC) was developed as described in Section 3.2.2 to allow for the performance of the endocortical SSC-CBM method to be critically assessed. Ideally the HRTC measurements would be used to validate the endocortical SSC-CBM method, but both the resolution of the HR-pQCT

[1]The high resolution lumbar spine scans are actually microCT, but throughout this thesis they are treated the same as HR-pQCT

scans and the rigour of the HRTC method validation were limited by the available datasets. Although, this means that the quality of the comparison is not high enough for it to be a true validation, the term validation is loosely used throughout the rest of this chapter to describe the comparison process.

3.1.1 Existing morphological HR-pQCT techniques

Morphological techniques are the standard tool for extracting architectural measurements from HR-pQCT [19, 73]. The resulting measurements define the mean architectural features over the analysed region. The standard morphological techniques involve the semi-automated segmentation of the periosteal edge over the image volume followed by Gaussian-based smoothing to remove trabecular structures and thresholding to identify mineralised cortical tissue. The cBMD is calculated as the mean of the thresholded voxels, and the cortical thickness is calculated as the area of the thresholded voxels divided by its periosteal perimeter [17, 19, 73].

Next the trabecular region is identified by subtracting the cortical region from a mask identifying the periosteal surface. The trabecular BV/TV is calculated as the mean density over the trabecular region divided by $1200\,\mathrm{mg/cm^3}$. This is based on the assumption that marrow has a density equivalent to $0\,\mathrm{mg/cm^3}$ and mineralised bone has a density of $1200\,\mathrm{mg/cm^3}$ [19, 48, 73]. Trabecular structure is extracted using a Laplace-Hamming filter followed by thresholding to identify regions of trabecular bone [19]. The trabecular number is then calculated as the mean three-dimensional distance between the trabecular mid-axes [55, 67]. From this, the mean trabecular thickness and spacing are derived using traditional plate model assumptions [93].

3.1.2 Existing model-based HR-pQCT techniques

Two model-based methods have also been employed to extract cortical and trabecular features from HR-pQCT scans: the FWHM method has been applied over the proximal femur to produce localised measurements of cortical thickness, cBMD and subsurface tBMD [123–125]; and the 'transitional zone segmentation' (TZS) method introduced in Zebaze et al. (2013) [138] has been used to measure the cortical and endocortical area over a slice.

As discussed in Section 2.1.4, the FWHM method estimates the density along a profile perpendicular to the periosteal surface of cortical bone, and provides localised estimates of the cortical thickness, cortical density and trabecular density at each measurement site. It has been validated against phantom data in several studies, which show the FWHM method begins overestimating cortical thickness and underestimating cortical density for structures with a thickness less than $1.5-3\times$ the FWHM of the image blur [33, 88, 102], but it has not been validated for HR-pQCT resolution scans of human cortical bone. This is a major limitation as the partially resolvable intracortical and trabecular micro-architecture can result in an ambiguous edge location as illustrated in Figure 3.1. In this instance, it is not clear if the first time the sampled density drops below the endosteal edge threshold it is encountering an intracortical or a trabecular pore.

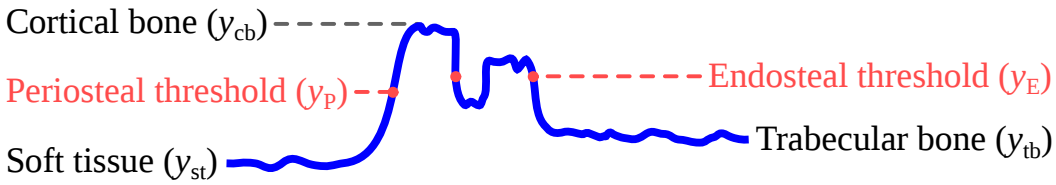
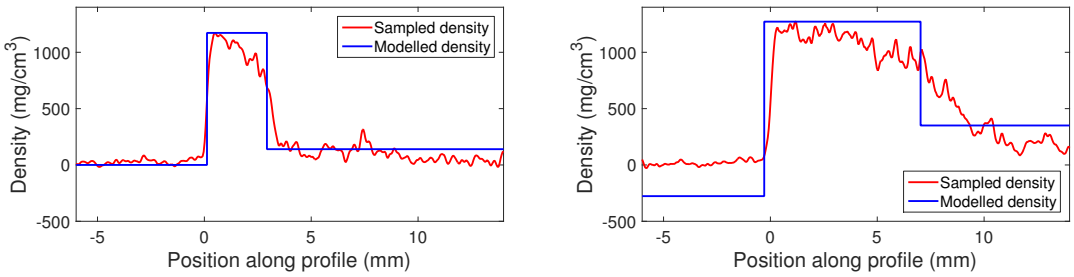


Fig. 3.1 A density profile with an ambiguous endosteal edge location.

The difficulty of selecting the endosteal edge location can be eased by using an optimiser to select the endosteal edge location based upon the sampled density profile as illustrated by Figure 3.2a. In addition, the mean of several closely spaced parallel density profiles can also be used to limit the effect of micro-structural variations parallel to the cortical bone surface as discussed in Section 3.1.3. Such optimiser-based implementations have previously been used in the validation of the CBM methods over paired QCT and HR-pQCT scans [123–125]. The complexity of the sampled density profile impacts the stability of the optimiser, which can result in poorly optimised FWHM models as shown in Figure 3.2b.



(a) Good fit

(b) Poor fit

Fig. 3.2 Two plots showing the variable quality of optimised FWHM models.

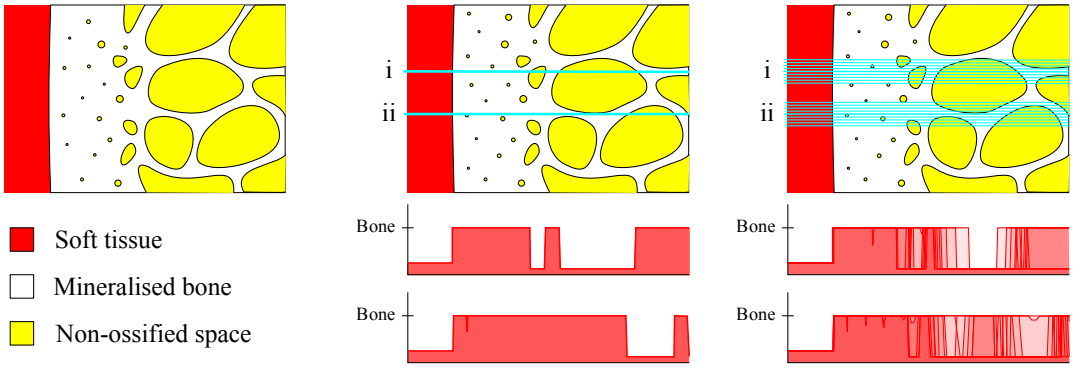
The TZS is a method of segmenting HR-pQCT scans into regions of cortical, endocortical and trabecular bone from which the mean cortical thickness, cortical density, cortical porosity and the endocortical thickness of a slice may be calculated [138]. The method defines the periosteal edge, the end of the compact appearing cortex, and the beginning of the trabecular bone along a series of sampled density profiles that project radially inwards to the centre of the bone roughly perpendicular to the cortical bone surface. As with the optimiser-based FWHM method, the mean of multiple closely space density profiles is used to reduce the effect of micro-architectural variations.

The periosteal edge is identified as the first major increase in density along the mean sampled profile. The TZS method defines the beginning of the cortical bone region as occurring near the top of the density increase across the periosteal edge, and the end of the cortical bone region as occurring at the beginning of the first endocortical pore. This is identified by the TZS method as the first pore after which the density samples do not return to that of the cortical bone near the periosteal edge. Finally, the TZS method defines the end of the endocortical region as the first pore with the same density as the trabecular bone at the end of the sampled profile [138]. The region boundaries that have been defined in each

profile are combined and used to segment each slice. This allows the area of each region in a slice to be calculated, which can be used to generate a mean measure of the endocortical thickness [138].

3.1.3 Macro-architectural and micro-architectural discrimination

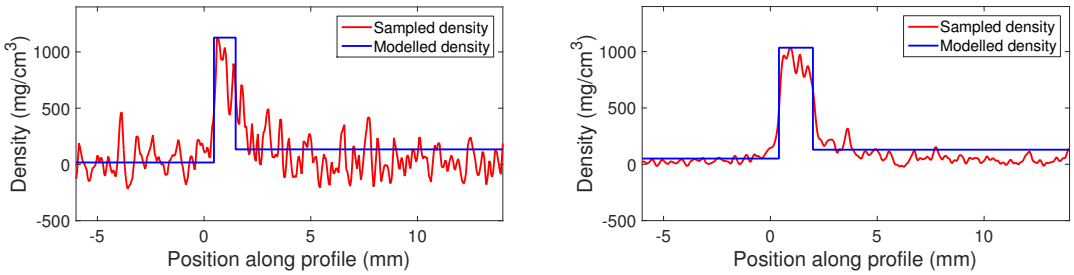
HR-pQCT scans include partially resolved micro-architectural features such as large intracortical pores, large trabecular pores and thick trabecular rods, which are located within the macro-architectural features. Extracting the macro-architectural features from amongst these partially resolved micro-architectural details for comparison with CBM method measurements is difficult, as illustrated by Figure 3.3. Consider the two sampled density profiles shown in Figure 3.3b, which are sampled at locations with similar macro-architecture. These profiles effectively contain binary density values, which either indicate mineralised bone or non-ossified tissue. The density of mineralised cortical and trabecular tissue is practically indistinguishable as they have similar DMB [4, 10, 39]. As a result, it is impossible to distinguish intracortical pores from trabecular pores and compact cortical bone from trabecular rods without considering the wider context around each sampled density profile.



(a) Architecture (b) Single profiles (c) Multiple profiles

Fig. 3.3 A two-dimensional representation of the three-dimensional bone architecture of cortical and subsurface trabecular bone and parallel profile sampling.

Multiple density profiles can be sampled in parallel and used to identify the underlying macro-architectural details as shown in Figure 3.3c. In the combined profiles shown in Figure 3.3c, the cortical bone can be identified as extending from the first instance of cortical bone to the next instance of non-ossified tissue occurring across most or all of the sampled profiles. The mean of these profiles gives the cortical and trabecular density at the macro-structural scale with minimal blurring across the cortical bone edges as shown in Figure 3.4. This mean density profile can be used directly by the FWHM method to model the cortical and subsurface trabecular values, but it cannot be used to measure the endocortical region as its extents are defined by micro-architectural features.



(a) Single HR-pQCT profile (b) Multiple HR-pQCT profile mean

Fig. 3.4 The effect of using multiple parallel profiles.

3.2 Implemented HR-pQCT techniques

The FWHM and other previously used techniques do not provide localised measurements of either the dense cortical or endocortical regions [124, 125, 140]. In addition, the use of only the mean of the sampled density profiles renders it more difficult to discriminate between macro- and micro-architectural causes of density variations. A new profile-based method of measuring cortical, endocortical and subsurface trabecular properties from HR-pQCT scans called the high resolution tissue classifier (HRTC) was developed to allow validation of the endocortical SSC-CBM method. The established FWHM method was used when considering the behaviour of this new method.

3.2.1 Full-width half-maximum (FWHM)

The FWHM has been used to provide HR-pQCT measurements for the validation of rectangular CBM measurements from QCT scans [124, 125, 140]. Previous validation studies use an optimiser to fit a rectangle model constrained with the local maximum density of each mean density profile. This is achieved within the CBM framework described in Section 2.2.1. The resulting model provides cortical thickness, cortical density and trabecular density measurements as discussed earlier. The increased resolution of the HR-pQCT scan means that PVE has less impact than in QCT scans.

3.2.2 High resolution tissue classifier (HRTC)

The HRTC was developed to provide local measurements of the cortical, dense cortical, endocortical, and subsurface trabecular regions from HR-pQCT scans. It can be used to define the cortical thickness, cortical density, endocortical thickness, and trabecular density for comparison with endocortical SSC-CBM measurements from a QCT scan of the same individual. The method stems from the challenges associated with identifying the endocortical region as its boundaries are defined by micro-architectural features. As a result, the HRTC method preserves the micro-architectural details within the individual density profiles sampled in parallel, and uses them together to identify the location of micro-structural features defining the transitions between each region, such as the pores marking the transition

to and from the endocortical region. Once the region boundaries are defined, the behaviour in each region is defined by the mean across all profiles. HRTC has four stages as shown in Figure 3.5: sampling; thresholding; sectioning; and modelling.

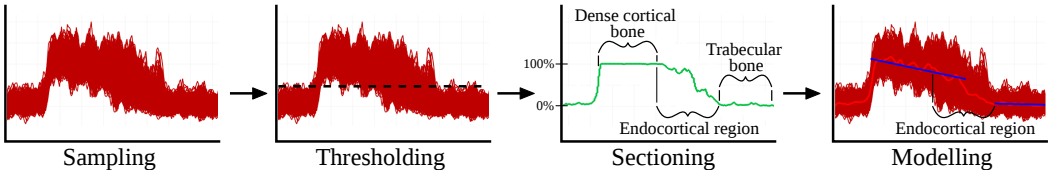


Fig. 3.5 The four stages of the High Resolution Tissue Classifier method: sampling parallel profiles at the image resolution; thresholding each density sample into ossified or non-ossified tissue; sectioning the bone percentage profile to define the extents of the different tissues along its length; and modelling the density of each tissue type.

Sampling

The sampling is performed along many parallel profiles which are arrayed around each measurement location. They are sampled on a regular grid at the Nyquist rate. A fixed radius is used to define the maximum extent of the profile grid.

Thresholding

At the HR-pQCT scale, cortical and trabecular bone both have similar BMD values due to the similar DMB in both tissues [4, 10, 39] coupled with the reduced impact of PVE in high resolution scans. This means that a single threshold value can accurately discriminate between mineralised cortical or trabecular bone, and various types of non-ossified tissue including soft tissue, skeletal muscle and marrow. The threshold value is important as poorly selected thresholds will cause the extent of the bone regions to be over- or underestimated. As non-ossified tissues all have relatively similar densities that are substantially lower than that of mineralised bone, a density threshold midway between the typical densities of mineralised bone and non-ossified tissue will be most reliable. The selection of such a value is described in Section 3.3.1. Each density sample is thresholded as either ‘mineralised bone’ or ‘non-ossified tissue’. The results can be used to define the percentage of the profiles passing through bone at each sample location, which is used to define the cross-sectional tissue structure along the length of the profile. The percentage profile provides a more reliable measure of the cross-sectional structure than the mean cross-sectional density as it is less affected by variations in the DMB and the density of the non-ossified tissue.

Sectioning

In the sectioning step the bone percentage profile from the thresholding step is classified into regions of four different tissue types: cortical bone; dense cortical bone; the endocortical region; and trabecular bone. Figure 3.5 shows the dense cortical, endocortical and trabecular

regions. The cortical bone region is not shown, but it extends from the first point where half the maximum percentage ($\%_{max}$) along the profile is reached, to the first point the percentage falls below the half-way point between the $\%_{max}$ and the mean trabecular percentage ($\%_{\bar{tb}}$). This aims to provide the same physical interpretation as the FWHM method. The start of the cortical region also marks the start of the dense cortical bone region. The dense cortical bone extends to the first endocortical pore, which is located where the percentage array last falls below $\%_{max}$. Physically the first endocortical pore marks the end of the last instance of the uniformly mineralised bone that exists in between the intracortical pores within the dense cortical bone. This physical interpretation is in line with previous methods [138]. The endocortical region extends from the first endocortical pore to the first minimum after the percentage array drops to the $\%_{\bar{tb}}$, which corresponds to the first marrow filled pore in the trabecular region. The trabecular region extends from the end of the endocortical region to the end of the profile.

Modelling

Each tissue region classified in the sectioning step can be categorised with a linear regression model fit to all the parallel density samples within that region. This gives slope, intercept, mean and STD values. The mean has a direct analogue to the density measurements in CBM methods and is used most, while the STD is an indication of the density variations caused by the micro-structural features in the region. The extents and mean densities of the HRTC dense cortical, endocortical and cortical regions combined with the mean density of the trabecular region can be directly compared with the endocortical SSC-CBM measurements.

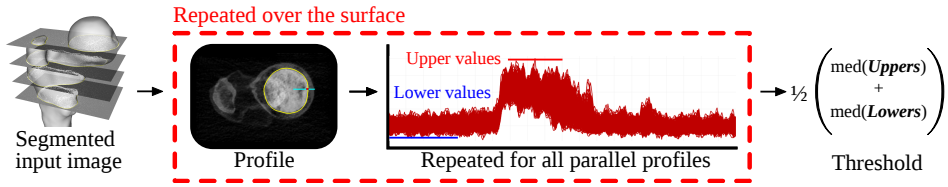
3.3 Adaptations for clinical and cadaveric datasets

Several adaptations were required before the profile based methods introduced in Sections 2.2 and 3.2 could be applied to clinical and cadaveric datasets.

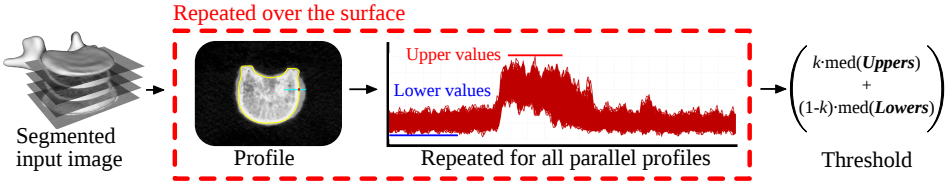
3.3.1 HRTC threshold selection

As discussed in Section 3.2.2, the optimal HRTC threshold is midway between the density of non-ossified tissue and mineralised bone. The pipeline shown in Figure 3.6a was developed to automatically calculate this threshold for tissue samples scanned in water, which should reduce the possibility of user introduced biases. At each measurement location over the mesh, the maximum and minimum density along each profile sampled in parallel are recorded. The threshold is calculated as the mean of the median of the maximum and minimum density sets. The median is used to protect against outliers.

The HR-pQCT scans in the spine dataset used in this thesis contain vertebra wrapped in plastic and scanned in air as discussed in Section 3.4.5. The density of plastic is similar to the density midway between air and mineralised bone, which can lead to its misclassification as mineralised bone when the mid-point threshold is used. This required the adaptation of the threshold selection pipeline so that it produces a threshold density that is approximately



(a) In water



(b) Plastic wrapped in air

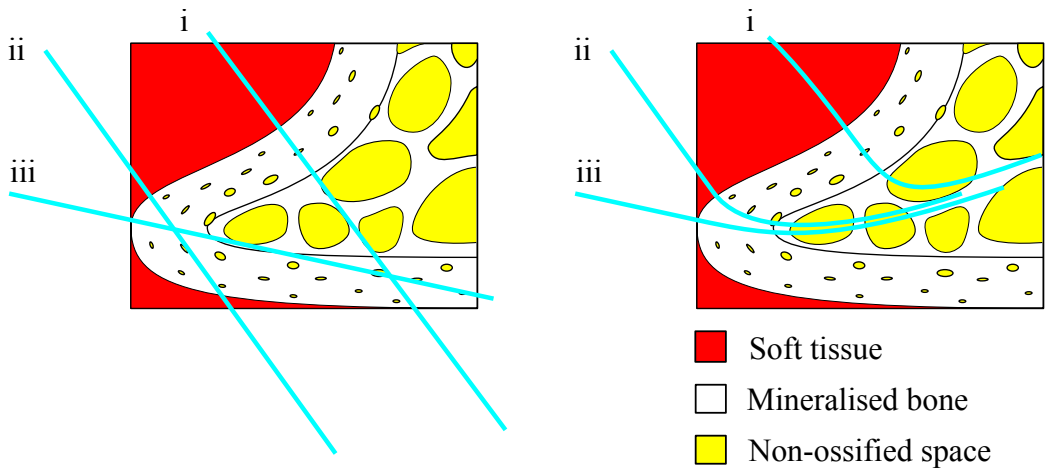
Fig. 3.6 Automated HRTC threshold selection

midway between mineralised bone and plastic. Histograms of the densities from a central slice of each scan were viewed, which showed that the distribution of the air, plastic and mineralised bone density values remained relatively constant between scans. This allowed the threshold density for each scan to be calculated as the unequally weighted mean of the maximum and minimum values as shown in Figure 3.6b where $k = 0.65$.

3.3.2 Curved profiles

Profile-based methods have traditionally used straight transects of sampled density. Figure 3.7 highlights three instances where straight profiles produce misleading results: (i) the profile passes through cortical bone twice, (ii) the profile passes through cortical bone and back into the surrounding soft tissue without passing through any trabecular bone, and (iii) the profile enters but never exits the cortical bone. All of these profiles will result in erroneous models with incorrect cBMD, cortical thickness and tBMD values. Such profiles do not generally occur in the proximal femur due to its geometry, but do occur in vertebrae where straight profiles routinely produce misleading densities in the vertebral processes and rims of the vertebral bodies.

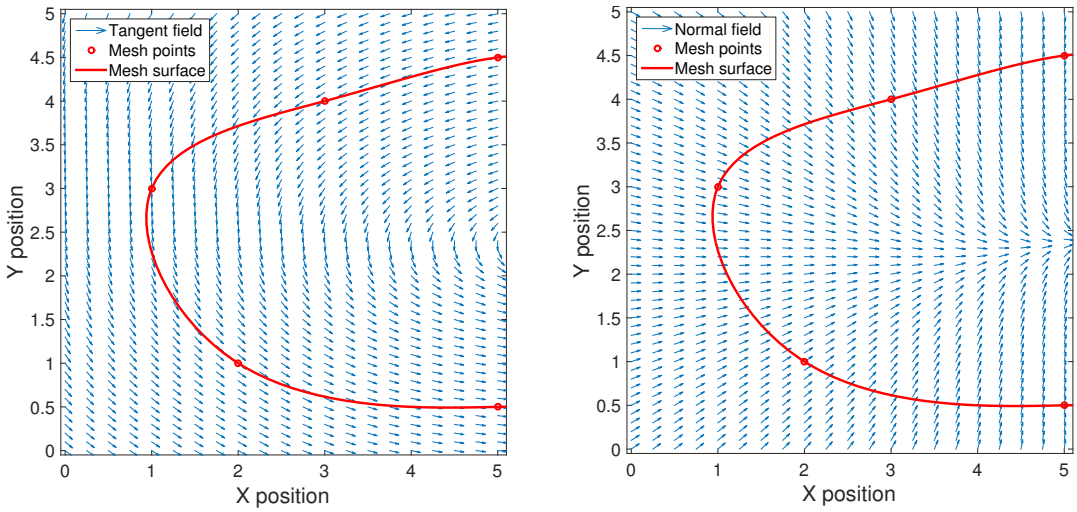
Straight profiles are based on the assumption that the laminar orientation of cortical bone remains constant with penetration depth, so a straight profile remains perpendicular to the cortical bone structure along its length. Unfortunately this is not a valid assumption where the surface curvature is non uniform as illustrated in two dimensions by the Gaussian distance-weighted tangent vector field of the surface in Figure 3.8a. This shows one possible manifestation of how the tangent direction to the cortical bone laminar structure may change with penetration depth. The normal to the laminar structure can be modelled by the same mechanism as shown in Figure 3.8b. This can be used to define curved CBM profile paths in three dimensions with the portion of each profile that extends into the bone defined by a streamline along the normal vector field, and the proportion of the profile that extends into the surrounding tissue defined by a streamline against the normal vector field. Where,



(a) Straight profiles (b) Curved profiles

Fig. 3.7 Acutely angled cortical structures requiring curved profiles. Straight profiles on the left and curved profiles on the right.

both streamlines originate at the same periosteal surface node. Tracing streamlines along a distance-weighted normal field is only one of several possible approaches. It was used as it directly expresses the physical nature of the problem.



(a) Tangent vector field (b) Normal vector field

Fig. 3.8 Variations in normal and tangent vector fields with penetration depth.

Many different techniques exist for tracing particles along vector fields with the fourth-order Runge-Kutta algorithm [54, 80] used in this thesis. Straight and Gaussian distance-weighted normal profiles have fundamentally different interpretations of the behaviour of laminar bone, which will lead to differences in the resulting thickness measurements. This is no great problem as long as only measurements produced by the same method are compared,

and any measurements produced by straight profiles over thin or acutely angled anatomical regions are discarded. Hence, the vertebral datasets were all analysed with curved profiles to ensure accurate measurements over the vertebral rims and the processes, while the proximal femur datasets were all analysed with straight profiles in line with previously published CBM-based analysis of this region [98, 99, 123–126, 132, 133].

3.4 Preparation of paired HR-pQCT to QCT datasets

The proximal femur and lumbar spine validation studies examined in this chapter required careful preparation to ensure equivalence between the QCT and HR-pQCT measurements before they could be used to assess the performance of the methods in Sections 3.5 and 3.6. Each validation dataset contains many paired QCT and HR-pQCT scans of individuals, which were used to compare clinical CBM techniques applied to QCT scans against the more precise ex-vivo HRTC technique applied to the higher resolution HR-pQCT scan. Equivalence in sampling, surface registration and density calibration was needed before these measurements could be compared. The proximal femur dataset is introduced in Section 3.4.4, and the lower spine dataset is introduced in Section 3.4.5.

3.4.1 Equivalent sampling

The QCT and HR-pQCT measurements need to be made over the same localised anatomic regions if they are to be equivalent. This requires alignment through some form of surface-based registration as discussed in Section 3.4.2, but also an equivalence of the extent of the localised region contributing to each sampled density profile. Equivalence of the sampling region can only be achieved by considering the differing effects of blur on the two modalities, which have different resolutions. Imperfections in the imaging process create blur as discussed in Section 2.1.2, in which each image value is a weighted mean of the values in the surrounding region. As the QCT blur is greater than the HR-pQCT blur, each QCT pixel is affected by values over a larger region than each HR-pQCT pixel, so HR-pQCT measurements should be generated from values sampled from several adjacent pixels. The size of the region from which the HR-pQCT measurements should be generated depends on the relative difference in the QCT and HR-pQCT blurs. In both datasets, the HR-pQCT region required for equivalent comparison with the QCT scans is large enough to allow discrimination between most macro- and micro-architectural features.

3.4.2 Surface registration

Registration is a process for aligning equivalent features between images [29]. In the context of validating QCT methods with HR-pQCT techniques, it is used to define anatomical correspondence between the measurement surfaces in each scan pair. This allows for direct comparisons to be made between measurement sites. In the broader context of CBM, registration is required before comparisons can be made between surface measurements

in different scans, either of the same individual taken at different times or with different modalities, or between images of different individuals in a cohort [42].

Rigid registration algorithms involve only translation and rotation of the surface. This is often extended to include similarity transforms in which each axis of the rotated and translated image can be independently scaled to account for slight calibration differences between scans [29]. Such registration algorithms are useful when a constant shape can be assumed between scans. This is generally the case between multiple scans of one individual. In this thesis, all paired QCT and HR-pQCT scans were aligned using a similarity transform calculated with an iterative closest point registration algorithm as described in Treece (2010) [124]. Measurement surfaces in scans of different individuals in a cohort can only be aligned with some form of 'non-rigid' registration, as the individuals in the cohort cannot be assumed to have bones with the same shape [29, 42]. Non-rigid registration involves some localised deformation of the measurement surface. All other registrations performed in this thesis were performed using locally affine deformation calculated by the iterative closest point registration algorithm (LAD) [38], which is one possible example of several similar algorithms that are commonly used for registration in medical imaging [38, 42, 108].

The paired QCT and HR-pQCT scans from each dataset were all registered to the same canonical surface using the process shown in Figure 3.9, where the canonical surfaces were selected to be anatomically representative of the populations analysed in this thesis. First, each scan was segmented using a combination of thresholding and manual intervention using the software tool Stradwin[2]. The increased noise and micro-architectural detail in the HR-pQCT scans meant it was harder to generate high quality surfaces. Next, the surfaces were each registered to a canonical surface using WxRegsurf[3]. The resulting surfaces were reviewed and manually re-registered as necessary. The canonical proximal femur was selected from an individual with a fairly average head size, neck angle and trochanter prominence, while the canonical vertebrae was selected as the mean shape from the dataset of 56 L1 vertebrae introduced later in Section 4.3. Ideally, different canonical vertebra would have been used for vertebra from different anatomical sites, but this was not possible with available data.

As shown in Figure 3.9, the QCT surfaces were registered directly to the canonical surface using a similarity transform followed by LAD, while the HR-pQCT scans were registered to the canonical surface via a similarity registration to the paired QCT surface. This effectively allowed the higher quality QCT generated surface to replace the HR-pQCT surface. Manual registration was particularly important for the vertebrae as the absence of processes made registration to a canonical particularly difficult. In response, alignment between mesh nodes with substantially different normals was ignored to avoid fitting to the segmented surface covering the removed processes.

[2] A freely available segmentation and surface extraction toolbox produced by the Cambridge University Medical Imaging Group: http://mi.eng.cam.ac.uk/~rwp/stradwin/

[3] A freely available registration toolbox produced by the Cambridge University Medical Imaging Group: http://mi.eng.cam.ac.uk/~ahg/wxRegSurf

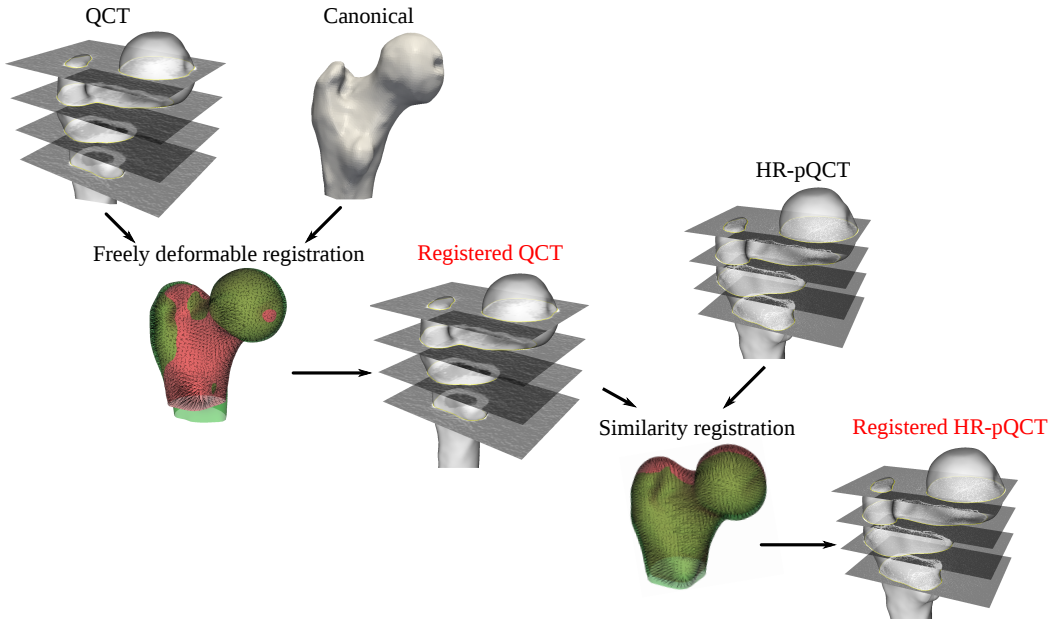
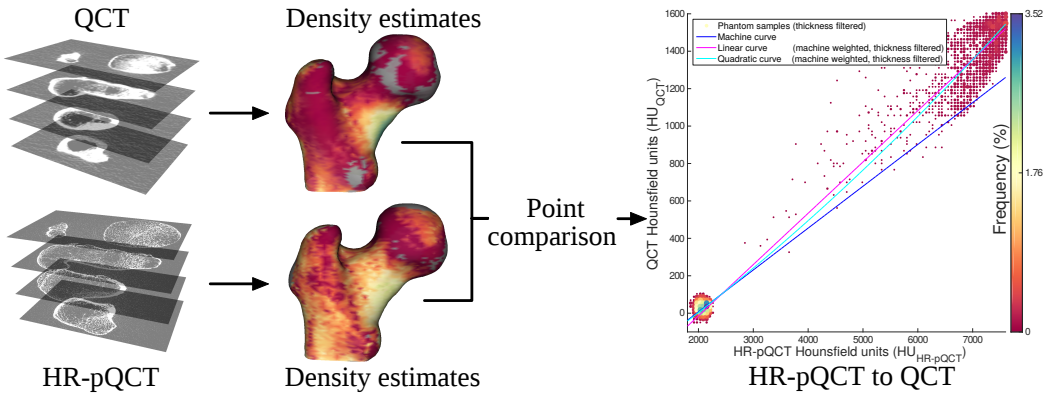


Fig. 3.9 The pipeline used to perform registration across a paired dataset of QCT and HR-pQCT scans.

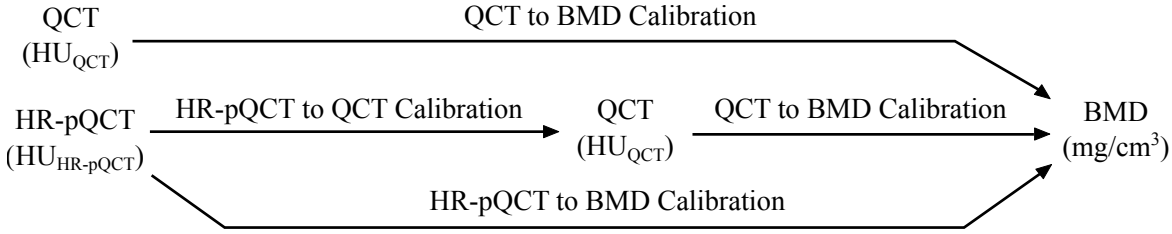
3.4.3 Density calibration

In both cadaveric validation studies, the QCT scans include a calibration phantom and the HR-pQCT scans do not. This is relatively common due to differing QCT and HR-pQCT scan protocols. The lack of a HR-pQCT calibration phantom means the HR-pQCT HU values cannot be readily converted to BMD values for comparison. Calibration can be based on machine settings in some instances, but this information is only available for some HR-pQCT scanners, and it does not provide BMD estimates with an acceptable level of certainty for comparison with QCT BMD measurements. As each QCT to HR-pQCT scan pair is of the same individual, direct density calibration between HR-pQCT HU and QCT HU values should be possible as shown in Figure 3.10. Any equivalent densities could be compared as long as they provide density measurements over a wide range of BMD values. Figure 3.10b shows how the QCT calibration curve generated with a calibration phantom can then be used to relate HR-pQCT HU to BMD.

The differing energy spectra of QCT and HR-pQCT scans mean they will experience different linear attenuations, which will introduce non-linearities in the relationship between QCT and HR-pQCT HU in the presence of beam hardening or non-linear variations in the material composition. Figure 3.11a shows the differing effect of beam hardening caused by the penetration of 3 mm of mineralised bone on QCT and HR-pQCT HUs, and Figure 3.11b shows the differing effect of the non-linear compositional changes with density in the proximal femur model introduced in Figure 2.23 on QCT and HR-pQCT HUs. Both figures show the non-linear affects are small.

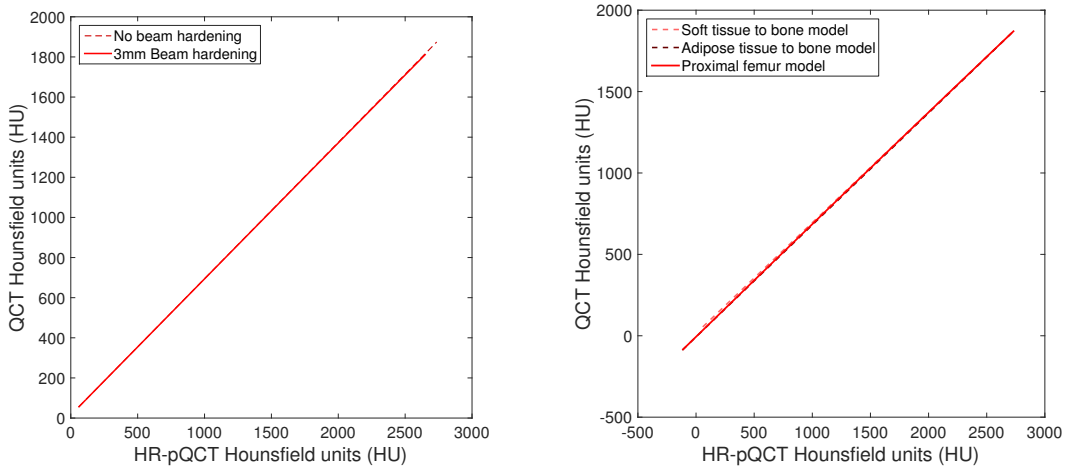


(a) The pipeline used to generate comparable QCT and HR-pQCT density estimates that can be used to produce an HR-pQCT HU to QCT HU calibration curve.



(b) The relationship between HR-pQCT HU, QCT HU and BMD.

Fig. 3.10 A generic calibration pipeline for paired HR-pQCT and QCT scans.



(a) Beam hardening

(b) Variable material composition

Fig. 3.11 Plots considering the relationship between HR-pQCT HU to QCT HU.

Differing material composition

Validation studies of paired QCT and HR-pQCT scans are ex-vivo as the limited chamber size of HR-pQCT scanners necessitates the scanned samples be removed from the cadaver. The samples can then be scanned in water or air. Water is preferable as its material composition, density and linear attenuation is more similar to that of the soft tissue and muscle that surrounds bone in-vivo, but in some instances the limited chamber size of the HR-pQCT scanner may necessitate scanning in air. Ideally when HR-pQCT scans are performed in air the QCT scanning protocol would be updated to reflect this change, but this is not always the case. This results in QCT and HR-pQCT scans of cadaveric samples with differing material compositions, which complicates the density calibration process. While the QCT and HR-pQCT scans of the proximal femur study were both performed in water, the lumbar spine validation study introduced in Section 3.4.5 contains HR-pQCT scans in air and QCT scans in water.

This frustrates the density calibration process as it becomes necessary to consider compositional changes, both internal and external to the scanned subject, when comparing the QCT and HR-pQCT scans. In the vertebral study, the internal compositional changes result from the displacement of varying quantities of hematopoietic marrow during the embalming process. This is mainly replaced with water for QCT scans and air for HR-pQCT scans. These compositional differences change the HR-pQCT HU to QCT HU relationship. Figure 3.12 shows two simple models that were used to investigate the possible nature of the HU relationship between scans taken in air and in water.

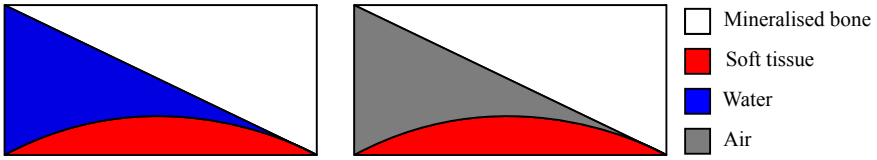


Fig. 3.12 A cross sectional image showing non-linearly varying tissue composition. QCT on the left, and HR-pQCT on the right.

These models are based on the assumption that the presence of mineralised bone structure increasingly protects marrow from removal during the embalming process, so in both models the quantity of mineralised bone increases linearly from left to right, while the amount of either water or air decreases from left to right. The QCT HU and HR-pQCT HU values were calculated for an idealised X-ray beam travelling vertically through the model using Equations 2.9 and 2.10, the tables of X-ray mass attenuation and densities provided on the NIST Standard Reference Database [57], and the X-ray sources introduced in Section 2.3.1.

Figure 3.13 shows the resulting relationship between the HR-pQCT and QCT HU values. The relationship is convex as the linear attenuation of air is substantially less than that of water and the soft tissue representing hematopoietic marrow, and the proportion of air to soft tissue decreases with increasing BMD in the model. A very similar relationship would result if stromal marrow were incorporated in the model so as to represent the proximal femur due to the relative similarity in the linear attenuation of stromal and hematopoietic marrow as compared to that of air. The convex relationship will remain as long as the

proportion of marrow removed by the embalming process decreases with the amount of the local mineralised bone.

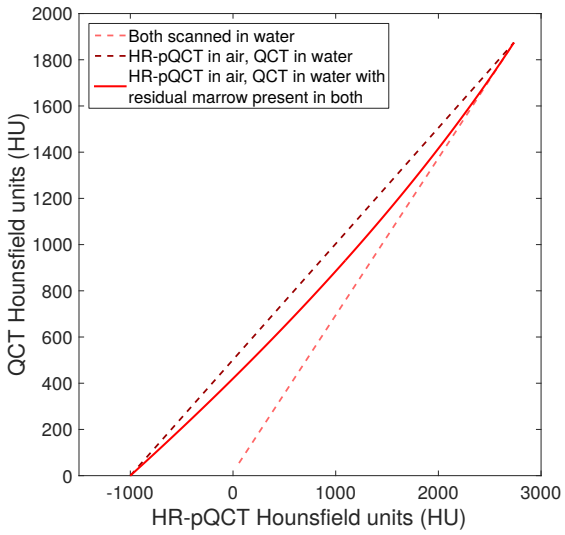


Fig. 3.13 The relationship between HR-pQCT HU scanned in air, and QCT HU scanned in water for tissue with linearly varying BMD and residual marrow.

Differing image blur

A final challenge results from the differences in QCT and HR-pQCT resolution, which means that QCT cortical bone HU measurements experience more PVE than the corresponding HR-pQCT measurements. As discussed in Section 2.1.2, PVE leads to an increasing underestimation of cortical bone HU as the blur extent becomes greater than its width, so the QCT cortical bone measurements will underestimate the actual cortical bone HU where it is thin. The simplest solution to this problem is to filter the paired HR-pQCT HU and QCT HU values by cortical thickness, and only consider value pairs that are both unaffected by the image blur. This is typically possible over the proximal femur where the shaft cortical thickness typically approaches or exceeds 6 mm [123]. Unfortunately this is not possible over the vertebral body, where cortical thickness typically ranges from 0.15–0.86 mm[35, 106, 37]. These will all be affected by PVE in a typical QCT scan.

Paired HU measurements

The paired HR-pQCT to QCT calibration pipeline shown in Figure 3.10, requires a large set of directly comparable HU density measurements representing a wide range of BMD values. Surface registration can be performed as described in Section 3.4.2 to provide aligned density profiles across each scan pair, but these must still be processed to extract equivalent density measurements given the differences in resolution between the QCT and HR-pQCT scans. Figures 3.14a and 3.14b show methods that are designed to extract equivalent estimates of

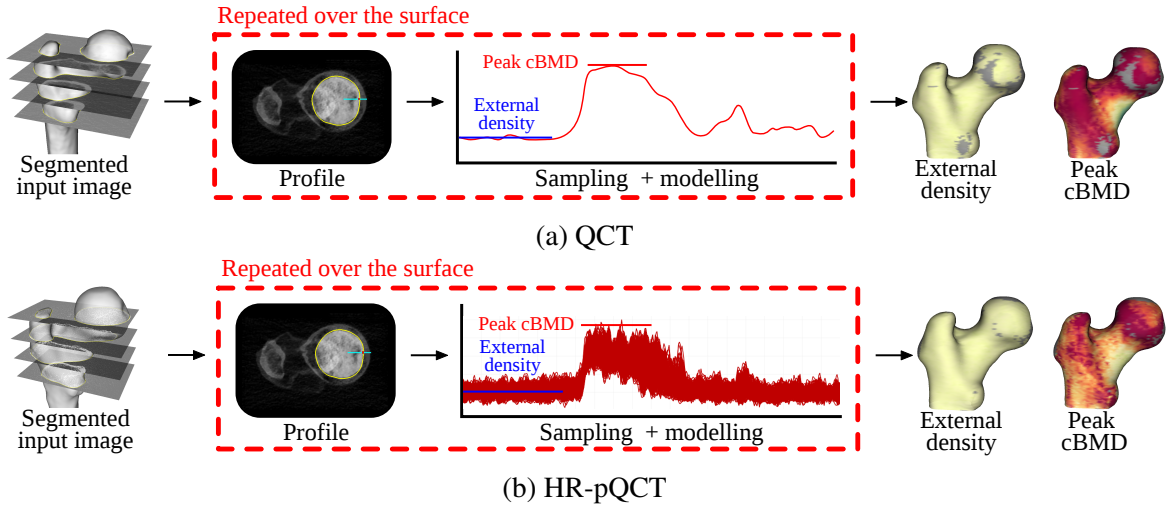


Fig. 3.14 The pipeline used to generate surface mapped values for calibrating HR-pQCT scans to QCT scans.

the peak cBMD and the external density from QCT and HR-pQCT scans. In QCT scans a single profile is sampled as shown in Figure 3.14a, while in HR-pQCT scans several parallel profiles are sampled as shown in Figure 3.14b to allow macro-architecture of a similar scan scale to be compared as discussed in Section 3.1.3.

The profiles are each aligned with the periosteal surface so that the first 30 % extends externally into the surrounding material and the remaining 70 % extends internally through the cortical bone and into the subsurface trabecular bone. The mean of the first 20 % of the sampled values along each profile is used to estimate the external tissue density. Only the first 20% of the profile is used to protect against inaccuracies due to blur effects and poor periosteal edge alignment. In addition, any profiles where the maximum value occurs in the first 20 % of the profile are ignored. The peak cBMD is estimated as the maximum QCT profile value, and the median of the maximum values across all of the parallel HR-pQCT profiles.

3.4.4 Proximal femur validation study

74 paired QCT and HR-pQCT scans of the proximal femur were obtained from an ethically approved study [30, 31]. The left and right femurs were removed from 18 female and 17 male cadavers. These were stripped of soft tissue, submerged in a saline solution and vacuum packed to remove air bubbles. The mean age of the individuals was 77 years with a range of $59-96$ years. The QCT data was acquired using a Brilliance64[4] scanner at 120 kV with a pixel sizing of $0.33 \times 0.33 \times 1.00$ mm, while the reference HR-pQCT data was captured using an XTremeCT[5] scanner at 70 kV with a pixel sizing of $0.082 \times 0.082 \times 0.082$ mm. This gave a total of 74 paired high- and low-resolution datasets from 37 patients. Two scan

[4] Philips, Germany
[5] Scanco Medical AG Switzerland

Table 3.1 The mean and STD of the linear QCT HU to density calibration coefficients and the quadratic HR-pQCT HU to density calibration coefficients.

QCT HU to BMD		HR-pQCT HU to BMD		
p_0	p_1	c_0	c_1	c_2
-6.30 ± 1.00	0.79 ± 0.01	-324.33 ± 28.15	0.15 ± 0.02	$7.77 \times 10^{-6} \pm 1.96 \times 10^{-6}$

pairs were excluded, one for high levels of ring artefact in the HR-pQCT scan, and one for a HR-pQCT scan imaged with a restricted FoV. Surface registration of the dataset was performed as detailed in Section 3.4.2.

A QRM-BDC calibration phantom was included with each QCT scan. The QRM-BDC phantom contains three 200 mm long cylindrical cores with 9 mm radii and known BMD values. The central 50 % of each core was sampled at the QCT resolution to produce approximately 50 samples per core per slice. The paired HU and BMD values were used to calibrate each scan. These curves and the underlying density measurements were visually inspected for each scan to ensure the calibration phantoms were correctly sampled.

The manufacturer provided a linear calibration curve to relate HR-pQCT HU to BMD based on the machine settings. This was not used directly as it produced poor alignment between equivalent QCT and HR-pQCT BMD profiles. Instead the manufacturer calibration curve was incorporated into an adaptation of the general paired calibration pipeline introduced in Figure 3.10a as shown in Figure 3.15. The density estimates included paired high and low QCT and HR-pQCT values that were generated using the methods shown in Figure 3.14. The low values should represent the density of the surrounding saline solution, and the high values should represent the density of cortical bone. The measurements were filtered to remove all high measurements where the cortical thickness is less than 3 mm to limit the impact of uneven levels of PVE between the QCT and HR-pQCT scans. The manufacturer's machine-settings-based HR-pQCT HU to BMD calibration curve was used to influence the calibration curve over the range of BMD densities covered by the QRM-BDC. This acted to constrain the calibration curve for BMD values of 100–200 $\mathrm{mg/cm^3}$, where it was otherwise poorly constrained.

Figure 3.16b shows a representative example of the high and low HU values generated from a QCT and HR-pQCT scan pair. These values were combined with paired HR-pQCT and QCT HUs generated from BMD values over 0–200 $\mathrm{mg/cm^3}$ using the HR-pQCT manufacturer calibration curve and the QCT QRM-BDC calibration curve. Together these were used to generate linear and quadratic calibration curves relating the HR-pQCT and QCT HUs also shown in Figure 3.16b. The quality of these two curves was manually checked at a subset of the measurement sites by comparing the alignment and dynamic range of sampled QCT BMD and HR-pQCT BMD density profiles. The quadratic curve produced better alignment and dynamic range as shown in Figure 3.17, so it was used to relate HR-pQCT HU to QCT HU. The QRM-BDC calibration curve could then be used to complete the calibration to BMD producing the HR-pQCT HU to BMD curve shown in Figure 3.16c. Table 3.1 shows the mean and STD of the calibration polynomials.

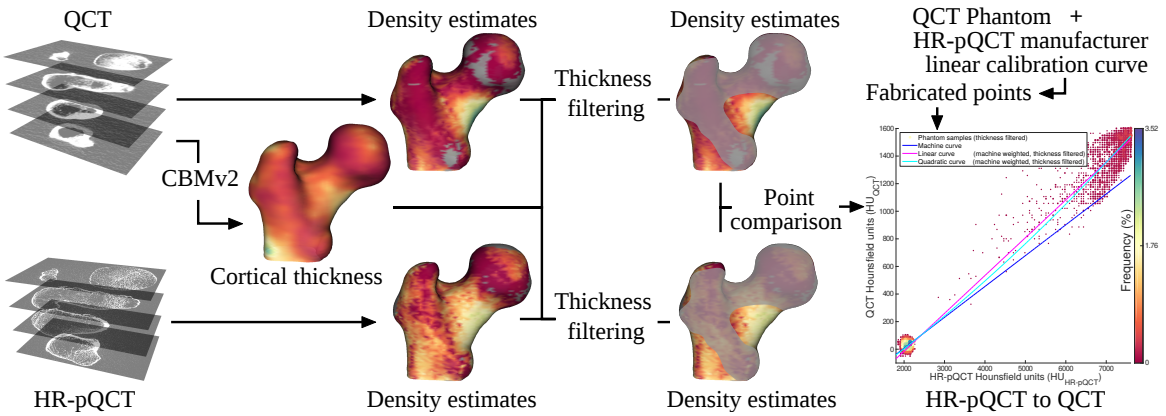


Fig. 3.15 The pipeline used to generate an HR-pQCT HU to QCT HU calibration curve over the proximal femur.

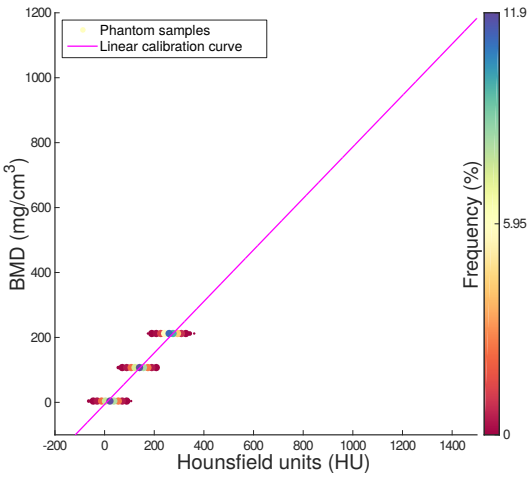
3.4.5 Vertebral bodies validation study

Vertebral specimens were obtained through the International Institute for the Advancement of Medicine[6] in an ethically approved study. The specimens were gathered in 2010 and 2012 from 6 donors who had been screened for standard infectious diseases. The lowest 3–4 vertebrae of the thoracic spine were extracted from each cadaver to produce 20 vertebrae. These were prepared for imaging by Luca Cristofolini[7] who stripped the vertebrae of soft tissue and removed the vertebral processes. The specimens were CT and microCT scanned by Enrico Schileo and Fabio Baruffaldi[8]. Surface registration of the dataset was performed as detailed in Section 3.4.2.

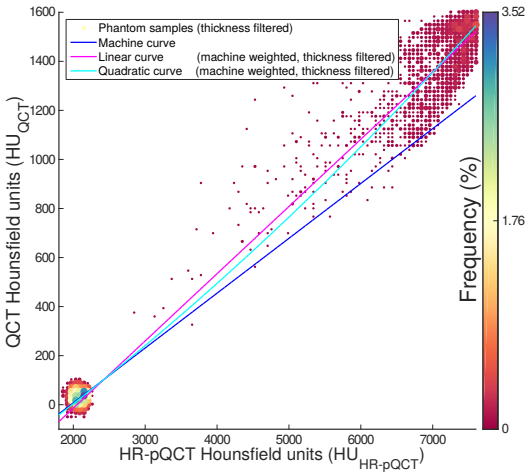
The stripped vertebral bodies were submerged in a saline solution and vacuum packed to remove air bubbles prior to QCT scanning. This improves the image quality by producing more realistic contrast between the cortical bone and surrounding material. The QCT data was acquired using a single Brilliance 16[9] scanner operating at 120 kV with a pixel spacing of $0.23\times0.23\times0.75$ mm. The QRM-ESP was also scanned in the same scanner with the same settings. The QRM-ESP is a calibration phantom that mimics the L1-L3 section of the spine. Each vertebra is modelled with three materials of different densities representing the cortex, processes, and spongious sections of the bone. In total the phantom contains materials with six different densities in the range of 0–800 $\mathrm{mg/cm^3}$. The phantom was sampled at the image resolution over twelve 4 mm squares centred at manually specified locations: one for each material in each vertebra, and three in the surrounding water-equivalent material. Each sample region was located in a single representative slice. Figure 3.19a shows the sampled calibration values and linear calibration curve.

The vacuum packed vertebral bodies did not fit in the HR-pQCT acquisition chamber. Instead, the vertebral bodies were removed from the saline solution and immediately wrapped

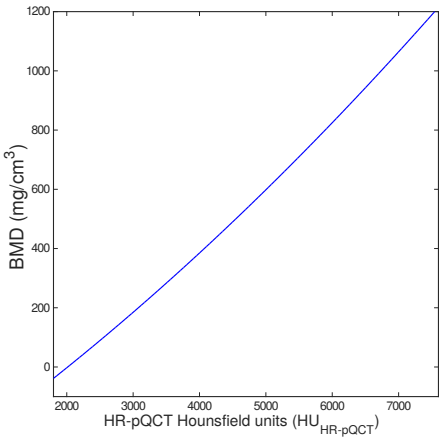
[6] International Institute for the Advancement of Medicine, Jessup, PA, US
[7] University of Bologna
[8] Istituto Ortopedico Rizzoli of Bologna
[9] Philips, Germany



(a) QCT

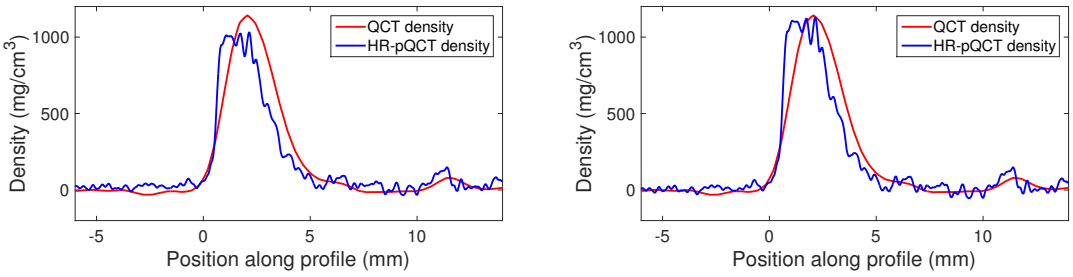


(b) Thickness filtered HR-pQCT to QCT



(c) HR-pQCT

Fig. 3.16 Calibration curves of a representative scan in the proximal femur validation study.



(a) Linear calibration

(b) Quadratic calibration

Fig. 3.17 The alignment of sampled QCT and HR-pQCT profiles after BMD calibration.

Table 3.2 The mean of the linear QCT HU to density calibration coefficients, and the mean and STD of the quadratic HR-pQCT HU to density calibration coefficients. No STD is included with the QCT density calibration coefficients as they were calibrated from a single QRM-ESP scan.

QCT HU to BMD		HR-pQCT HU to BMD		
p_0	p_0	c_0	c_1	c_2
6.02	0.86	-737.51 ± 133.28	0.04 ± 0.01	$4.68 \times 10^{-7} \pm 3.24 \times 10^{-7}$

in plastic to minimise the amount of air introduced to the medullary canal within the vertebral bodies. In spite of the plastic wrapping, most water drained from the vertebral bodies, and the HR-pQCT scans show large internal air cavities. The scans were performed in a single SkyScan 1176[10] microCT scanner at 80 kV with a pixel spacing of $0.035 \times 0.035 \times 0.035$ mm. No calibration phantom or manufacturer calibration curves were available making direct HR-pQCT to BMD calibration impossible. In addition, the differing material composition between the QCT and HR-pQCT scans caused by the HR-pQCT scan being performed in air means that the convex relationship between HR-pQCT HU and QCT HU measurements shown in Figure 3.13 can be expected.

Paired high and low HR-pQCT and QCT HU measurements were extracted from corresponding profiles using the pipeline introduced in Figure 3.14, where the low values represent the surrounding saline solution for the QCT scans, and air for the HR-pQCT scans. In contrast, the high values represent the peak cBMD in both scans, but the QCT measurements are more susceptible to PVE. Unfortunately none of the vertebral bodies included regions of thick cortex, so thickness filtering could not be employed.

Instead, the paired calibration pipeline introduced in Figure 3.10 was adapted as shown in Figure 3.18 to selectively weight the measurement pairs that experience less PVE. This should lead to an improvement in the quality of the HR-pQCT to QCT relationship. The selective weighting was employed based on the relative magnitude of the QCT to HR-pQCT measurements as a relatively large QCT to HR-pQCT ratio would only occur where the reduction in QCT magnitude due to the PVE is smaller than elsewhere. The selectively weighted calibration measurements and resulting quadratic calibration curve are shown in Figures 3.19b for a representative dataset. This dataset has a convex relationship between the HR-pQCT HU and QCT HU measurements as predicted by Figure 3.13. The differing levels of blur experienced by QCT and HR-pQCT scans meant that even with selective weighting some least squares quadratic curves still showed a concave relationship between HR-pQCT and QCT scans. This occurs when the curve is poorly constrained, so a linear calibration curve was used instead. The QRM-EPS calibration curve could then be used to complete the calibration to BMD. Table 3.2 shows the mean and STD of the calibration polynomials.

[10]Bruker microCT Belgium

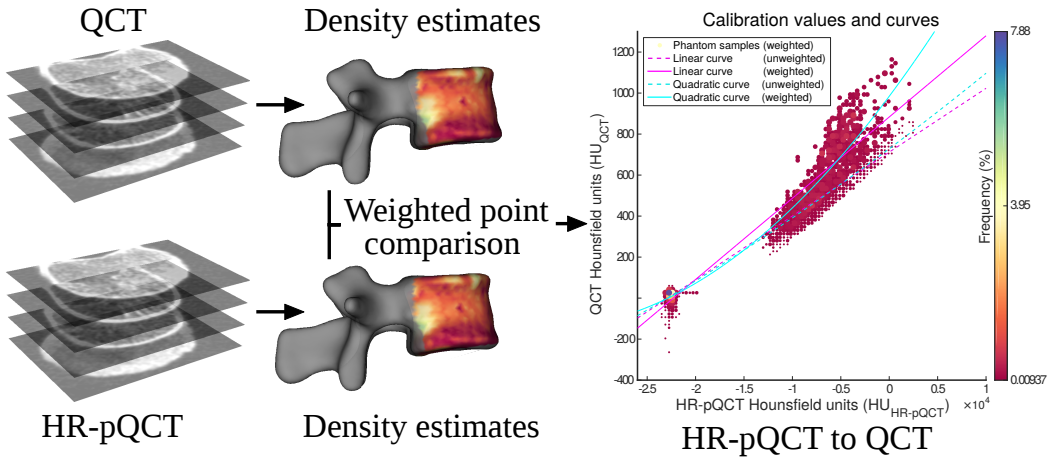


Fig. 3.18 The pipeline used to calibrate the HR-pQCT lumbar spine scans.

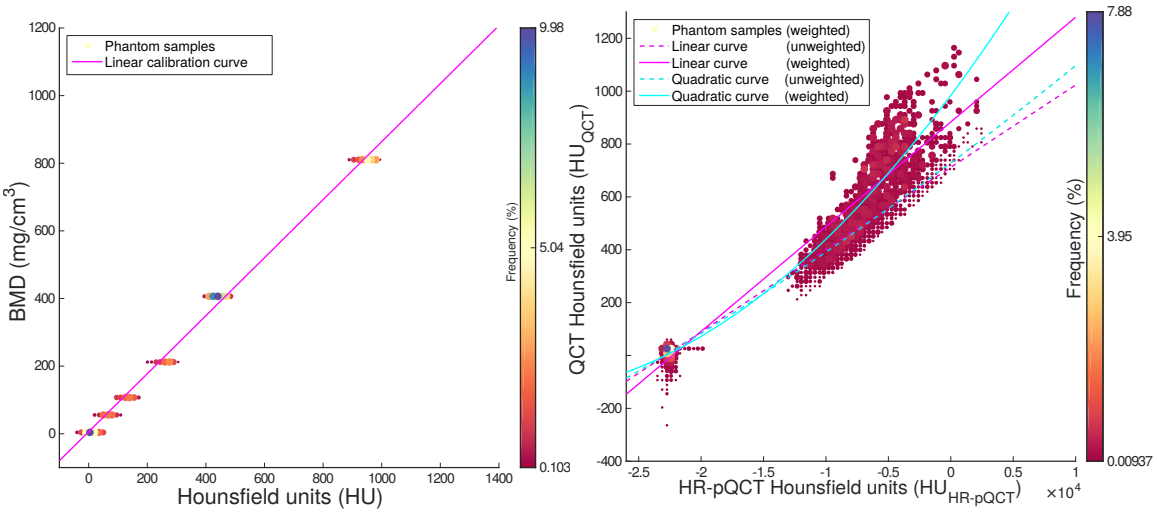
3.4.6 Discussion of paired density calibration

Paired QCT and HR-pQCT datasets allow for rigorous validation of QCT techniques over human bone, but this requires the BMD calibration of both scans. The direct density calibration technique developed in Section 3.4.3 for relating HR-pQCT HU to QCT HU was applied with minor dataset-specific adaptations to two validation datasets considered in this Chapter. The alignment of the BMD profiles between the calibrated QCT and HR-pQCT was spot-checked to provide a crude indication of the calibration quality. Both datasets show better alignment using a quadratic curve to relate HR-pQCT HU to QCT HU. In the lumbar spine dataset this is partially due to the convex non-linearities resulting from comparing a scan in air to a scan in water illustrated in Figure 3.4.3. In both datasets this may also partially result from the greater levels of PVE experienced in QCT scans. Lower density cortical bone has lower levels of DMB and higher levels of porosity. Cortical pores are typically smaller than the QCT resolution [104], so the QCT measurements will experience a greater level of PVE than the HR-pQCT measurements even with thickness filtering.

Calibration of HR-pQCT scans, through direct comparison with calibrated QCT scans, is viable and desirable as it allows for the validation of density measurements. However, care should be taken to account for the differing levels of image blur experienced by QCT and HR-pQCT, and equivalent QCT and HR-pQCT profiles should be visually inspected to ensure there is an acceptable alignment and similar dynamic range between them.

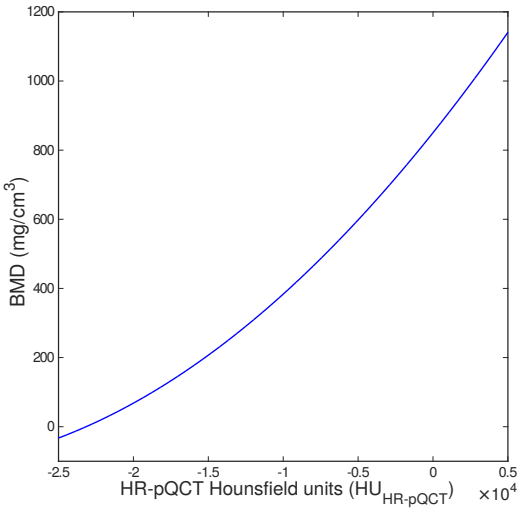
3.5 Cadaveric validation of the HRTC method

Validation of the HRTC method is difficult as no suitable phantoms nor clinical techniques exist for comparison. The HRTC method relies on micro-architectural information to classify bone tissue. The complex and regionally variable nature of bone micro-architecture makes it hard to model with synthetic data. Similarly, no μCT based techniques exist that are capable of producing suitable measurements for comparison with the HRTC measurements. Instead,



(a) QCT BMD calibration

(b) HR-pQCT to QCT HU calibration



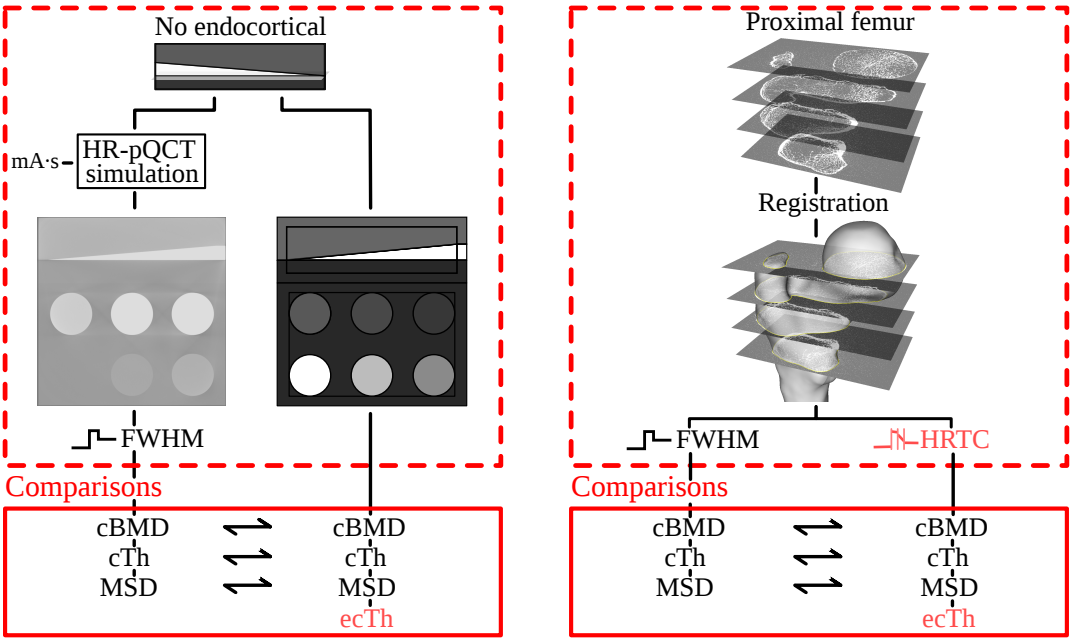
(c) HR-pQCT BMD calibration

Fig. 3.19 Calibration curves of a representative vertebra dataset from the spine validation study.

the established FWHM technique was validated against a HR-pQCT resolution simulation. Then the FWHM and equivalent HRTC measurements were compared over a cadaveric proximal femur dataset of HR-pQCT scans.

3.5.1 Experimental pipeline

Figure 3.1 shows the pipelines that were used to evaluate and compare the performance of both methods. Figure 3.20a was used to validate the FWHM measurements, and Figure 3.20b was used to compare the HRTC and FWHM methods. The FWHM measurements were smoothed with two iterations of mesh-connectivity-based precision smoothing, which were matched with two iterations of evenly weighted smoothing over the HRTC measurements. The FWHM precisions were extracted from the estimated model parameter variances of the final LMM optimisation at each measurement site. The smoothing was performed to remove outlier FWHM measurements resulting from very low precision optimisations.



(a) FWHM validation

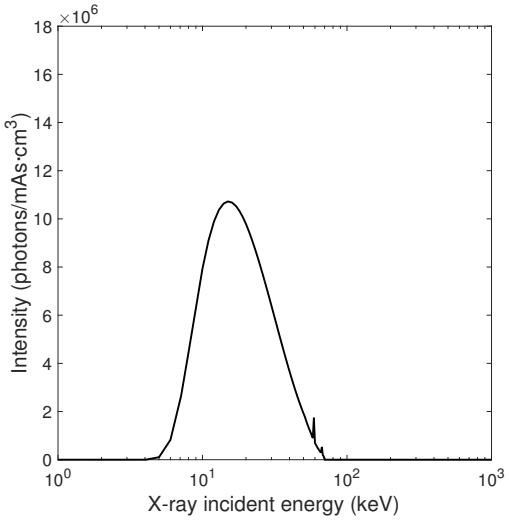
(b) HRTC and FWHM comparison

Fig. 3.20 The pipelines used to validate the FWHM method and to compare the performance of the FWHM and the HRTC methods.

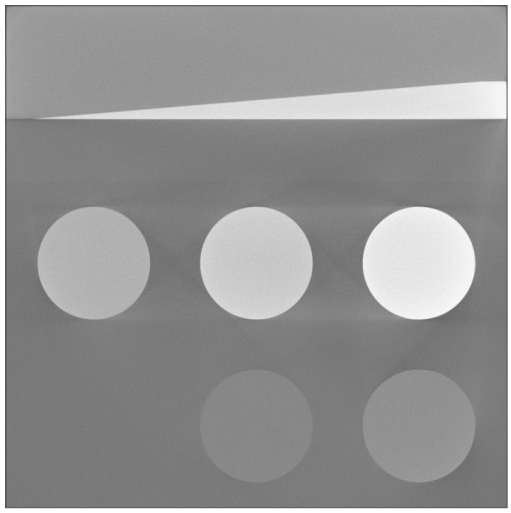
3.5.2 FWHM performance

The performance of the FWHM method was investigated at a resolution similar to HR-pQCT scans with the maximum cBMD slice of the synthetic cortical bone dataset with no endocortical region that was introduced in Section 2.4.1. 70 mA·s, 100 mA·s and 200 mA·s HR-pQCT simulations were generated at a resolution of 0.05 mm with the X-ray characteristic curve

shown in Figure 3.21a. The 200 mA·s simulation is shown in Figure 3.21b. The blur is almost imperceptible as the slice is 80×80 mm. The accuracy of the FWHM measurements was assessed over the full range of cortical thicknesses included within the slice as shown in Figure 3.22. These measurements were made at the pixel resolution along the periosteal surface.



(a) X-ray source



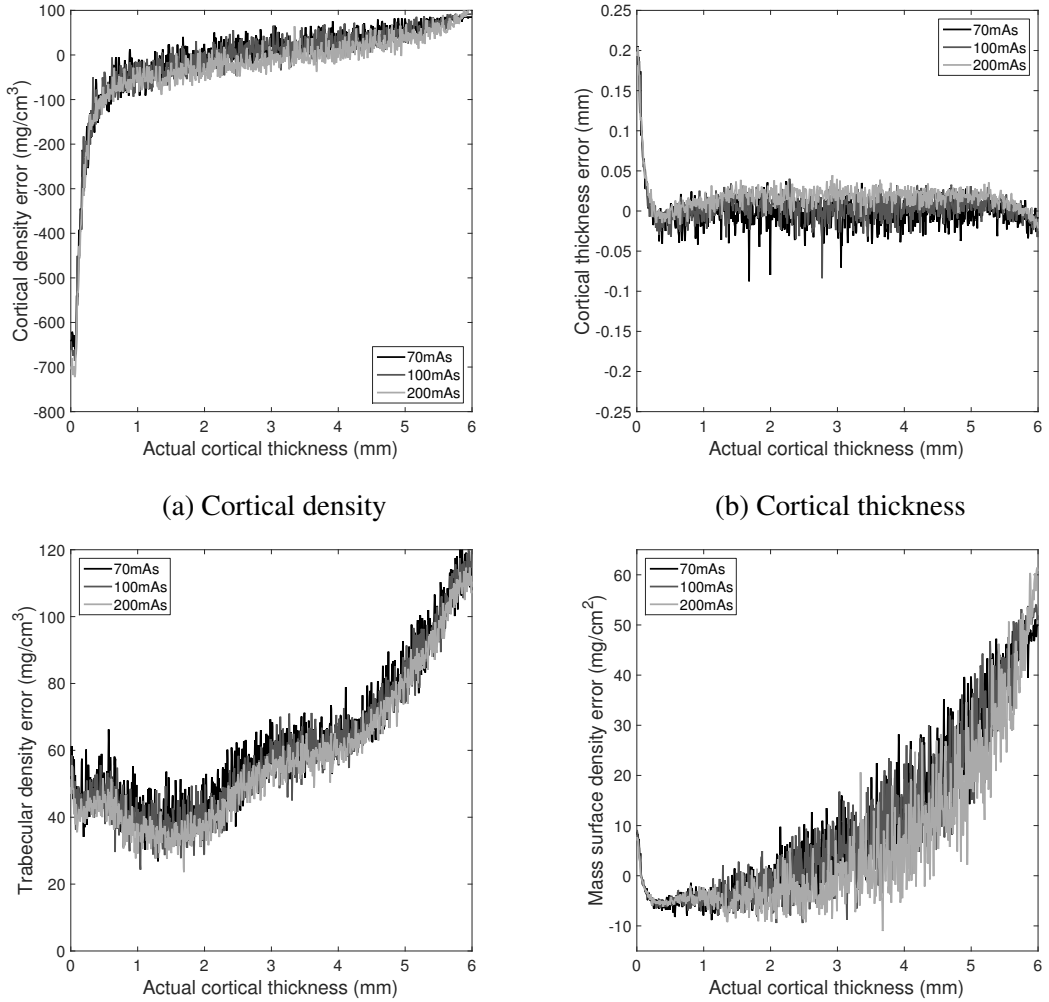
(b) Simulated scan

Fig. 3.21 The X-ray source and slice of the simulated HR-pQCT test data with no endocortical region.

3.5.3 FWHM performance discussion

Figure 3.22 shows that the FWHM method provides reasonably accurate cortical density and cortical thickness measurements until inaccuracies due to PVE dominate. This occurs when the cortical thickness drops below about 0.3 mm or approximately $6\times$ the slice resolution. This is larger than would be expected if the pixel resolution equalled the data resolution. This illustrates an important point: the pixel resolution is not necessarily the data resolution. In fact, it is typical for the image data resolution to be substantially less than the image pixel resolution.

The cortical and trabecular densities both show noticeable inaccuracies across all cortical thicknesses. These are due to variable beam hardening effects across the image. As discussed in Section 2.4.6, the calibration curve is generated from samples within the centre of the phantom cores. These samples experience beam hardening from the penetration of the core, but also from the material encountered on the way to the core. Given the densities and geometries of the calibration cores, the trabecular bone and the cortical bone; the trabecular bone experiences more beam hardening than the cores and the cortical bone experiences less beam hardening than the cores. This is the cause of the overestimation of the tBMD and the underestimation of the cBMD. The noticeable variations in tBMD with cortical thickness are



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

(d) Mass surface density

Fig. 3.22 FWHM measurement errors over an HR-pQCT simulation of the synthetic cortical bone dataset with no endocortical region.

due to the combined effect of increasing beam hardening caused by the cortical bone and differing amounts of beam hardening caused by the different density calibration cores.

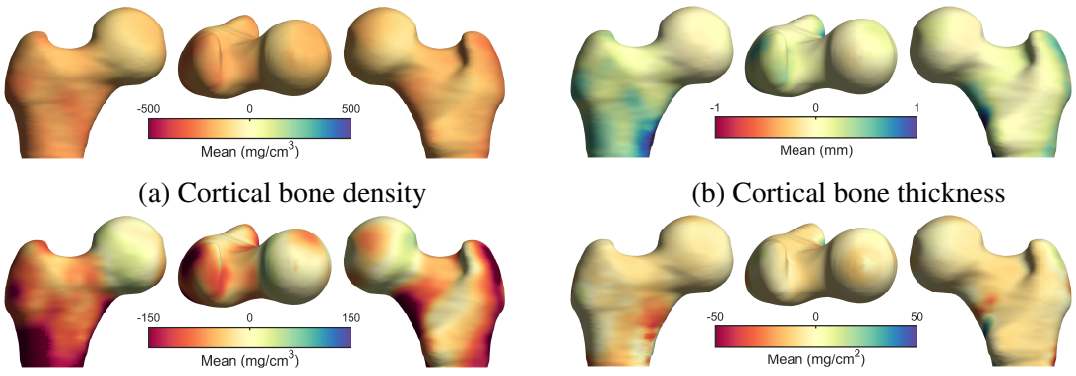
3.5.4 FWHM and HRTC performance comparison

The HR-pQCT scans from the proximal femur validation study introduced in Section 3.4.4 were used to compare FWHM and HRTC measurements. The dataset provided 72 HR-pQCT scans, which were all registered to the same canonical proximal femur. The HRTC threshold selection was performed using the mid-point method described in Section 3.3.1. The differences between the cBMD, cortical thickness, tBMD and mass surface density measurements are shown in Table 3.3. These are calculated over thickness ranges of 0.3–

1 mm, 1–3 mm and 3–6 mm to aid comparisons with previous publications [123–125]. The surface mapped mean of the differences between the FWHM and HRTC measurements across the dataset are shown in Figures 3.23. The CI plots comparing the HRTC and FWHM measurements are shown in Figure 3.24.

Table 3.3 Differences in the FWHM and HRTC measurements over a cadaveric HR-pQCT dataset of the proximal femur, where the difference is defined in relation to the FWHM measurements. Any sites without a valid FWHM or HRTC measurement were excluded.

Quantity	Thickness ranges (mm)	Difference between FWHM and HRTC
cBMD (mg/cm^3)	$0.3 \leq th_{cb} < 1.0$	-159 ± 64
	$1.0 \leq th_{cb} < 3.0$	-222 ± 54
	$3.0 \leq th_{cb} < 6.0$	-217 ± 40
cTh (mm)	$0.3 \leq th_{cb} < 1.0$	0.08 ± 0.24
	$1.0 \leq th_{cb} < 3.0$	0.36 ± 0.44
	$3.0 \leq th_{cb} < 6.0$	0.66 ± 0.73
tBMD (mg/cm^3)	$0.3 \leq th_{cb} < 1.0$	-36 ± 64
	$1.0 \leq th_{cb} < 3.0$	-123 ± 77
	$3.0 \leq th_{cb} < 6.0$	-137 ± 82
MSD (mg/cm^2)	$0.3 \leq th_{cb} < 1.0$	-6.1 ± 9.6
	$1.0 \leq th_{cb} < 3.0$	-5.9 ± 31.6
	$3.0 \leq th_{cb} < 6.0$	-20.8 ± 69.8



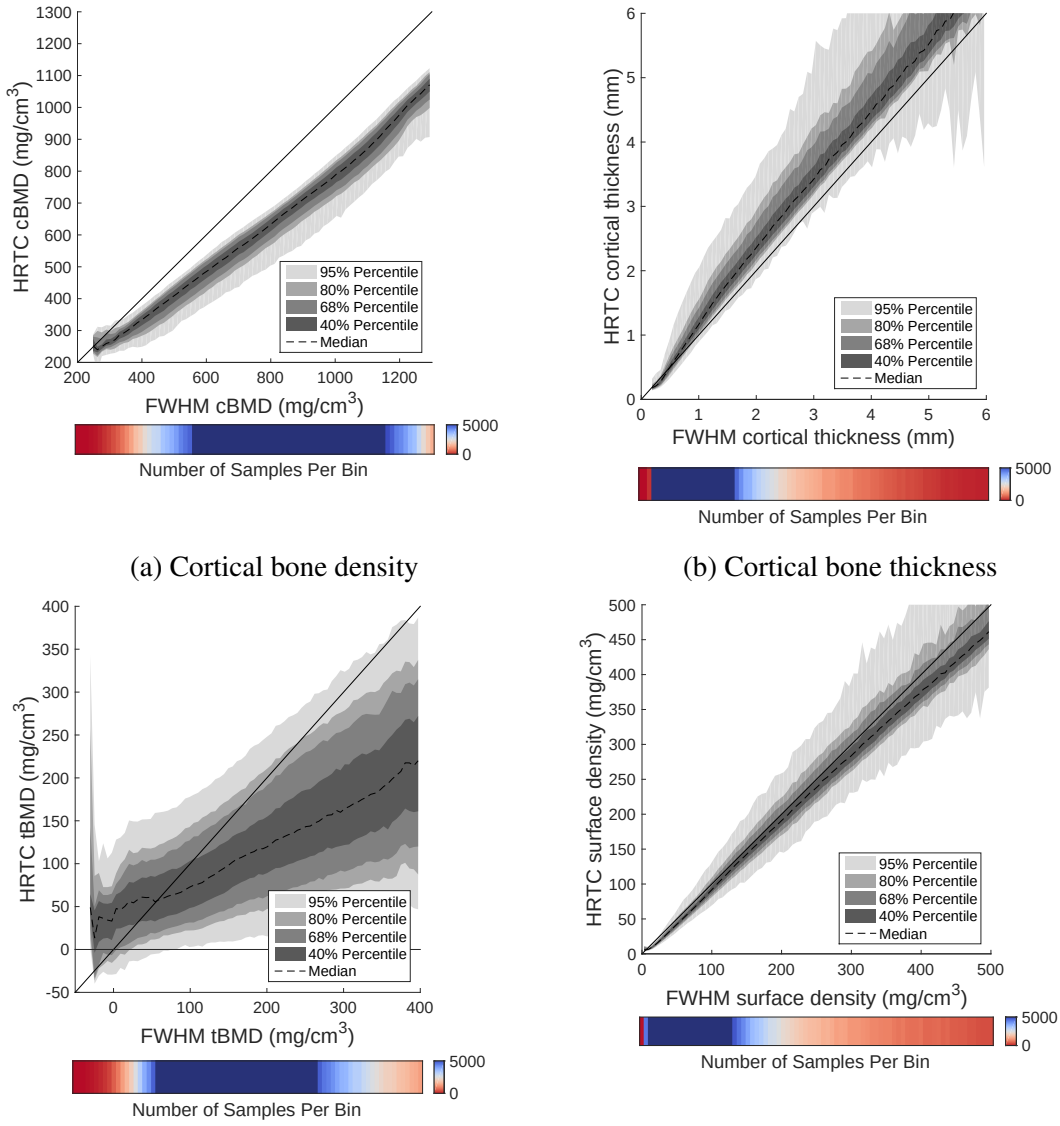
(a) Cortical bone density

(b) Cortical bone thickness

(c) Trabecular bone density

(d) Mass surface density

Fig. 3.23 Surface mapped differences between the rectangular FWHM and HRTC methods over the proximal femur, where the difference is defined in relation to the FWHM measurements.



(a) Cortical bone density

(b) Cortical bone thickness

(c) Trabecular bone density

(d) Mass surface density

Fig. 3.24 CI plots comparing the rectangular FWHM and HRTC methods applied over the proximal femur.

3.5.5 HRTC and FWHM comparison discussion

The CI plots comparing the HRTC and FWHM measurements in Figure 3.24 show that the FWHM cortical densities are highly correlated but not directly comparable with the HRTC density, and that the cortical thickness measurements are directly comparable but the HRTC thicknesses are slightly larger. The combined differences in these measurements is reflected in highly correlated mass surface density measurements.

Cortical bone density

The major differences in the cBMD measurements result as the HRTC measures the mean cBMD, while the FWHM measures the peak cBMD. The mean cBMD incorporates the effect of the intracortical DMB and porosity on the cBMD. A secondary cause of the differences is beam hardening. The HRTC cortical density is the unweighted mean over the cortical region, which does not account for the increasing underestimation of density with penetration depth due to beam hardening. In contrast, the maximum cortical density measured by FWHM generally occurs nearer to the periosteal edge, where the effects of beam hardening are less.

Cortical bone thickness

The HRTC cortical thickness measurement is consistently larger than the FWHM measure. The difference increases with thickness and is about 0.4 mm at a cortical thickness of 6 mm. As noted in Treece et al. (2015) [123], some of this is due to the difference in cortical bone density interpretation between the two methods. The peak density measured by the FWHM method is an overestimation of the mean cortical density as it does not account for the presence of intracortical pores and varying levels of DMB. This leads to an underestimation in the cortical thickness. Treece et al. (2015) [123] investigate the extent of this underestimation using another less stable optimiser-based technique; it was found it decreases from roughly 0 mm for a cortical thickness of 0.4 mm to roughly -0.15 mm for cortical thicknesses greater than 2 mm as shown by the correction values reported in Figure 3.25. This partially explains the differences observed in the HRTC and FWHM cortical thickness measurements.

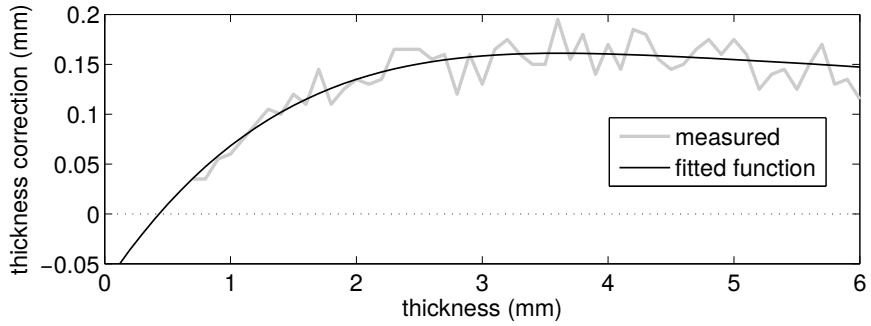
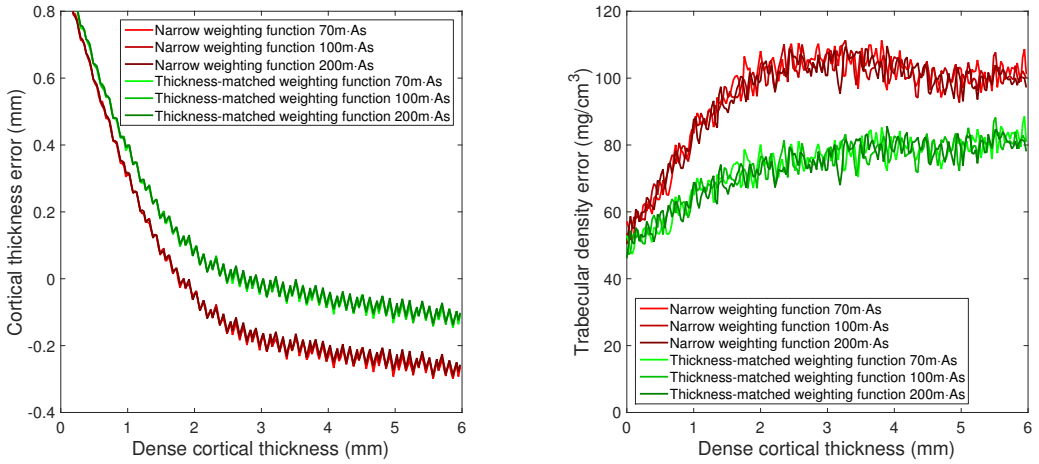


Fig. 3.25 The measured corrections for the skewed FWHM thickness measurements resulting from an overestimation of the cortical density in HR-pQCT proximal femur scans. This figure is copied from Treece et al. (2015) [123].

Some of the remaining difference can be explained by the impact of beam hardening on the optimiser CBM methods. Figure 3.26 shows the effect of beam hardening on the FWHM cortical thickness and trabecular density errors. These plots were generated using a 0.5 mm resolution simulated scan of the beam hardening phantom introduced in Section 2.4.3. The figure shows the impact of beam hardening on the accuracy of the cortical thickness and trabecular density measurements generated by the FWHM method for two different weighting functions. The narrow weighting function implementation is the one that has been

used in the FWHM validation of other CBM methods [123, 125] and was used in comparison with the HRTC method. The thickness-matched weighting function implementation uses the same weighting functions as the endocortical SSC-CBM method. The FWHM method errors are dominated by PVE for dense cortical thicknesses below 2 mm, but in a scan with a similar resolution to the HR-pQCT scans this would be closer to 0.5 mm. Above this, beam hardening results in a gradually increasing underestimation of cortical thickness to around 0.2 mm. This fits with the unaccounted discrepancies between the HRTC and FWHM thickness measurements. The thickness-matched weighting function makes the optimiser-based FWHM method more resistant to over-fitting the endosteal and trabecular portion of the model to account for beam hardening effects.



(a) Cortical thickness

(b) Trabecular density

Fig. 3.26 Reducing the FWHM errors over the simulated QCT resolution variable beam hardening dataset by using thickness adjusted weighting functions.

As noted previously, no suitable phantoms were developed for the direct performance assessment of the HRTC method. It is still worth considering its susceptibility to beam hardening effects. As shown in Figure 2.45, beam hardening can be expected to reduce the measured cortical density by about 10% over a 6 mm section of cortical bone. As the HRTC method defines each section of bone using a straight threshold that should be set 50% of the way between mineralised bone and typical non-ossified tissue, the threshold and the resulting thickness measurements should be minimally affected by the 10% change in density due to beam hardening. In contrast, the density measurements are the unweighted mean of each section, which will be directly affected by beam hardening.

Trabecular bone density

As shown in Figure 3.24c, the HRTC and FWHM trabecular density measurements are correlated, but not directly comparable. As with the cortical density measurements, some of the differences in the trabecular density measurements can be explained by considering

what each method is measuring. The HRTC trabecular density measurement is the mean of the trabecular region, which is defined from the end of the endocortical region to the end of the sampled density profile. In contrast, the FWHM trabecular density measurement is the optimal trabecular density estimate generated from a periosteal edge weighted error function. This means that differences between the trabecular density estimate and the sampled density profile are selectively weighted to penalise the differences in the density samples closest to the endocortical edge. This results in an endocortical weighted FWHM trabecular density that is generally larger than the unweighted HRTC trabecular density measurement. The remaining discrepancies can be explained by beam hardening. As with the cortical density measurements, the HRTC method will underestimate the trabecular density due to beam hardening. In contrast, the periosteal weighted optimiser will over-fit the endosteal edge and trabecular density of the model producing an overestimation in trabecular density as shown in Figure 3.26b. This is why the FWHM method with a thickness-matched weighting function is more resistant to increased trabecular density errors with increased beam hardening.

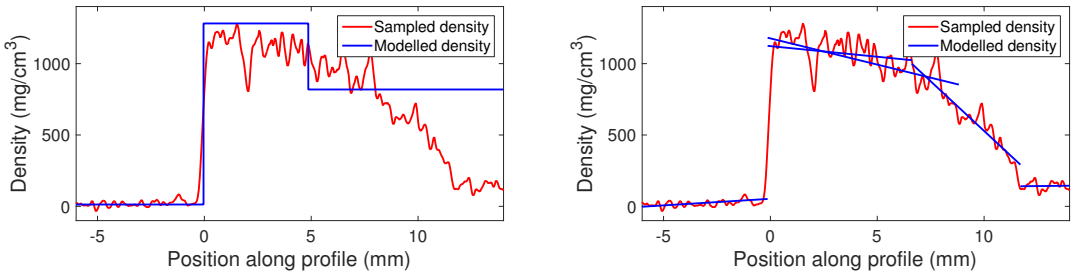
Spatial distribution

The surface mapped differences between the HRTC and FWHM surfaces in Figure 3.23 show that the differences in the cortical density measurement are broadly uniform, while the differences in the other three measurements differ spatially. The largest cortical thickness differences are clustered medially and laterally below the greater trochanter to around the lesser trochanter. The trabecular density differences have a similar spatial distribution to the cortical thickness but in the opposite direction. The HRTC cortical thicknesses are greater than FWHM measurements, and the HRTC trabecular densities are smaller. The spatial component and inverse relationship between the cortical thickness and trabecular density measurements indicate that the methods may be interpreting the dominant macro-structure in these regions differently.

Figure 3.27 shows the HRTC and FWHM models of a typical profile from the medial proximal femur beside the lesser trochanter, which shows that the FWHM method can struggle to identify the transition from cortical to trabecular bone and instead confuse intracortical pores for the beginning of the trabecular region. This results in an underestimation in cortical thickness and an overestimation in trabecular density, and it seems to explain most of the differences between these measurements and the HRTC measurements.

Conclusions

Although both methods have some susceptibility to measurement errors due to beam hardening, the HRTC method is generally more robust to this error. This is particularly true of the thickness measurements, which will typically be more affected by PVE. The FWHM measurement of cortical density may be less impacted by beam hardening, but the HRTC cortical density measurement is more clinically relevant as it includes an indirect measure of porosity, which has been linked to bone strength and fracture risk [2, 5]. As such, the HRTC method is used to validate all CBM measurements in Section 3.6.



(a) FWHM (b) HRTC

Fig. 3.27 Typical FWHM and HRTC models of a density profile from below the lesser trochanter. The line segments in (b) define the linear fit to the density samples in the cortical, dense cortical, endocortical and trabecular bone regions.

3.6 Cadaveric validation of CBM

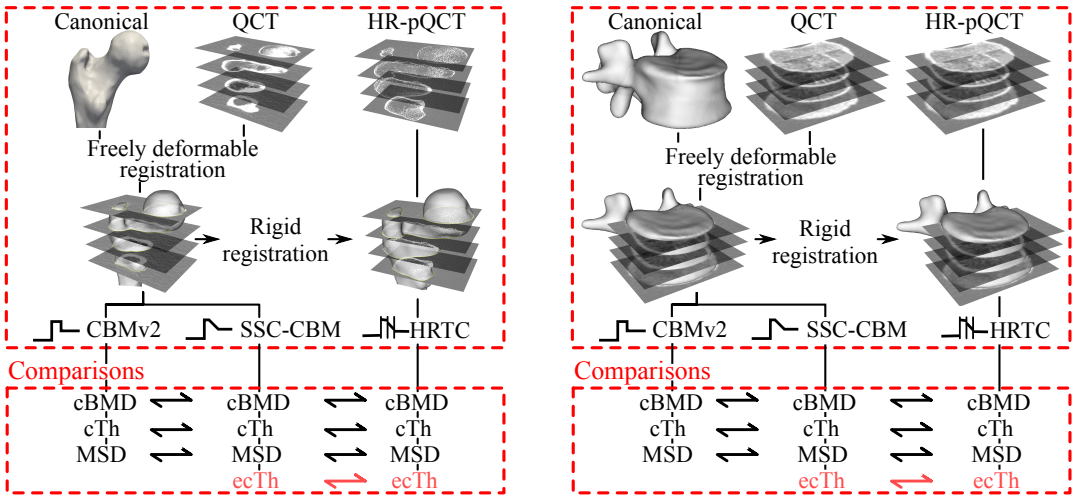
The performance of the endocortical SSC-CBM method was assessed over the proximal femur and vertebral body in the two cadaveric validation studies of paired HR-pQCT and QCT scans introduced in Section 3.5.

3.6.1 Experimental pipeline

The pipelines shown in Figure 3.28 were used to assess the performance of the endocortical SSC-CBM method. In both datasets, the scans were segmented and registered and calibrated as described in Section 3.4 before the CBM and HRTC methods were applied to the QCT and HR-pQCT scans. The HR-pQCT were processed using multiple density profiles sampled in parallel as discussed in Section 3.1.3. Sampling radii of 0.7 mm and 0.4 mm were used over the proximal femur and lumbar spine, respectively. The measurements were then processed to remove outlier CBM measurements using two iterations of mesh-connectivity-based smoothing as discussed in Section 2.5.1 for the CBM measurements and Section 3.5.1 for the HRTC measurements. The smoothed CBM measurements were compared directly with the smoothed HRTC measurements. The mean bias and precision of each measurement was considered along with its mean error across the measurement surface and its error across the range of different HRTC values.

3.6.2 Proximal femur results

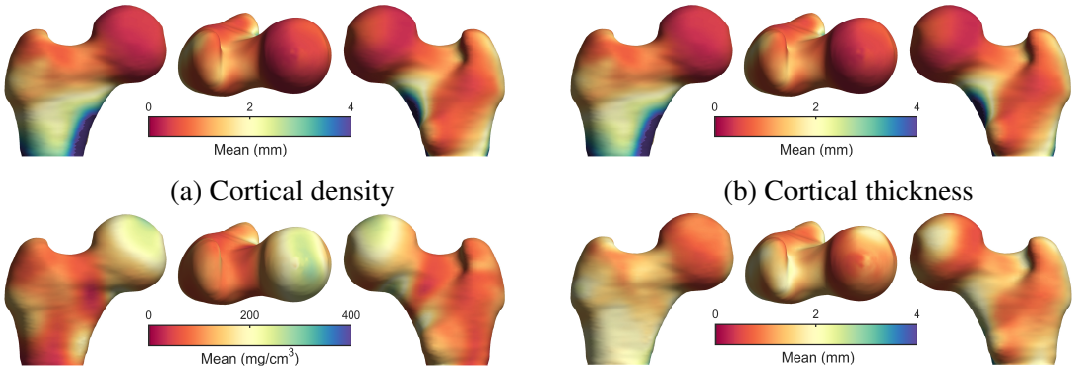
Figure 3.29 shows a selection of the mean cortical, endocortical and subsurface trabecular features across the proximal femur as measured by the HRTC method. Table 3.4 and Figures 3.30–3.34 show the performance of both CBM methods compared to the HRTC measurements. The bias and precision of each measurement, and the overall stability of each method is shown in the table for 0.3–1 mm, 1–3 mm, and 3–6 mm cortical thickness ranges. Figures 3.30, and 3.32–3.33 show the CI plots of each method, while Figures 3.31 and 3.34 show the spatial distribution of errors produced by each method over the proximal femur.



(a) The proximal femur dataset (b) The vertebral body dataset

Fig. 3.28 The pipelines used to evaluate the proposed endocortical SSC-CBM against both the rectangular CBMv2 and HRTC over the proximal femur and lumbar spine.

The endocortical SSC-CBM CI plots are split between two figures with the additional dBMD, dense cortical thickness and endocortical thickness plots shown in Figure 3.33.



(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Endocortical thickness

Fig. 3.29 Surface mapped HRTC measurements over the proximal femur.

3.6.3 Proximal femur discussion

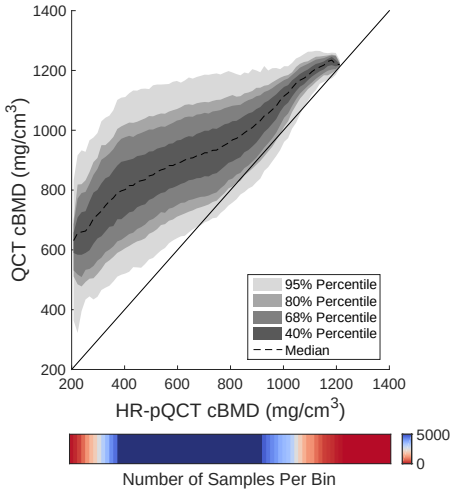
As shown by the CI plots in Figures 3.30 and 3.32, both the rectangular CBMv2 and endocortical SSC-CBM methods provide fairly accurate cortical thickness, mass surface density and trabecular density measurements, but while both methods are able to adapt to changing cortical densities only the endocortical SSC-CBM method can lay any claim to measuring it. The endocortical method also provides fairly accurate measurements of the dBMD, dense cortical thickness and endocortical thickness.

Table 3.4 The biases, precisions and overall stability of the two CBM methods over the proximal femur. The stability of each method is the % of physically valid models out of the set of all measurement locations where a valid HRTC model exists.

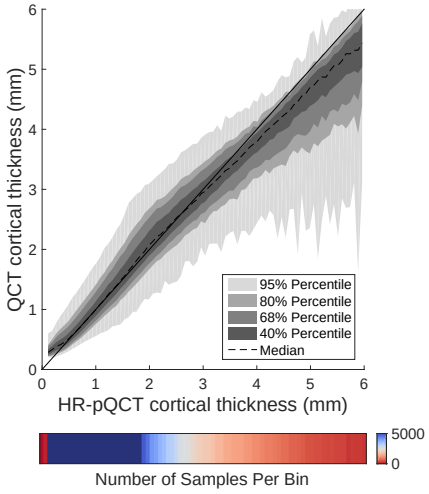
Quantity	Thickness Ranges (mm)	Rectangle CBMv2	Endocortical SSC-CBM
Stability (%)	$0.0 \leq t < 6.0$	99.6	100.0
cBMD (mg/cm^3)	$0.3 \leq th_{cb} < 1.0$	321±153	173±148
	$1.0 \leq th_{cb} < 3.0$	164±127	208±152
	$3.0 \leq th_{cb} < 6.0$	129±88	118±121
cTh (mm)	$0.3 \leq th_{cb} < 1.0$	0.06±0.23	0.10±0.24
	$1.0 \leq th_{cb} < 3.0$	0.07±0.41	−0.21±0.37
	$3.0 \leq th_{cb} < 6.0$	−0.28±0.60	−0.34±0.61
tBMD (mg/cm^3)	$0.3 \leq th_{cb} < 1.0$	−10±42	−15±42
	$1.0 \leq th_{cb} < 3.0$	42±49	22±44
	$3.0 \leq th_{cb} < 6.0$	80±77	−17±58
MSD (mg/cm^2)	$0.3 \leq th_{cb} < 1.0$	20.7±11.0	17.2±12.1
	$1.0 \leq th_{cb} < 3.0$	26.4±22.9	12.4±23.4
	$3.0 \leq th_{cb} < 6.0$	20.3±44.0	10.1±44.7
dBMD (mg/cm^3)	$0.3 \leq th_{cb} < 1.0$	-	233±154
	$1.0 \leq th_{cb} < 3.0$	-	306±163
	$3.0 \leq th_{cb} < 6.0$	-	170±142
dTh (mm)	$0.3 \leq th_{cb} < 1.0$	-	0.08±0.43
	$1.0 \leq th_{cb} < 3.0$	-	−0.56±0.36
	$3.0 \leq th_{cb} < 6.0$	-	−0.83±0.72
ecTh (mm)	$0.3 \leq th_{cb} < 1.0$	-	−0.47±0.43
	$1.0 \leq th_{cb} < 3.0$	-	0.12±0.74
	$3.0 \leq th_{cb} < 6.0$	-	0.58±0.93

Cortical bone density

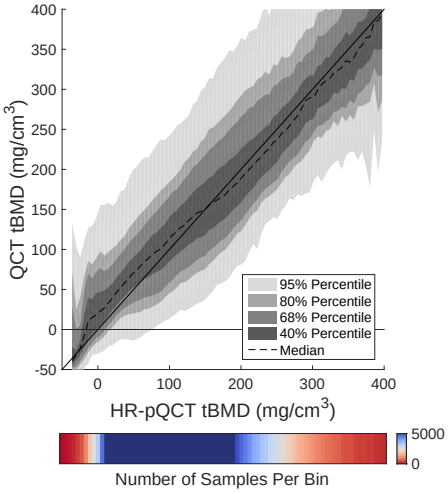
Both methods overestimate the cortical bone density. Considering the CI plots, the rectangular CBMv2 errors range from 0–420 mg/cm^3, while the endocortical SSC-CBM errors range from −20–300 mg/cm^3 excluding an increase to 450 mg/cm^3 for cBMD below 300 mg/cm^3. In part this is likely due to an underestimation in the HRTC cBMD measurement caused by beam hardening as discussed in Section 3.5.5. In the rectangular CBMv2 method, the bias will also partially be due to its optimiser weighting function, which selectively weights errors near the periosteal edge. This will produce a model that is biased to measure the cBMD near the periosteal edge, which is often where it is most dense leading to an overestimation of the



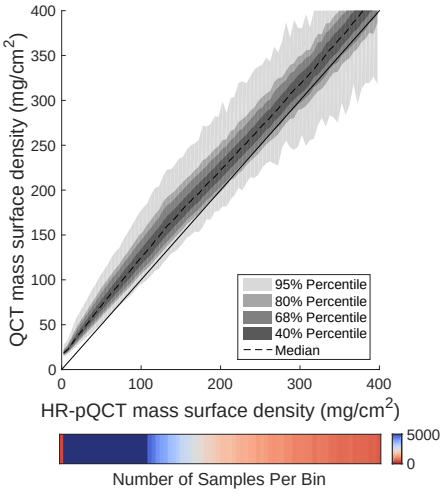
(a) Cortical density



(b) Cortical thickness



(c) Trabecular density



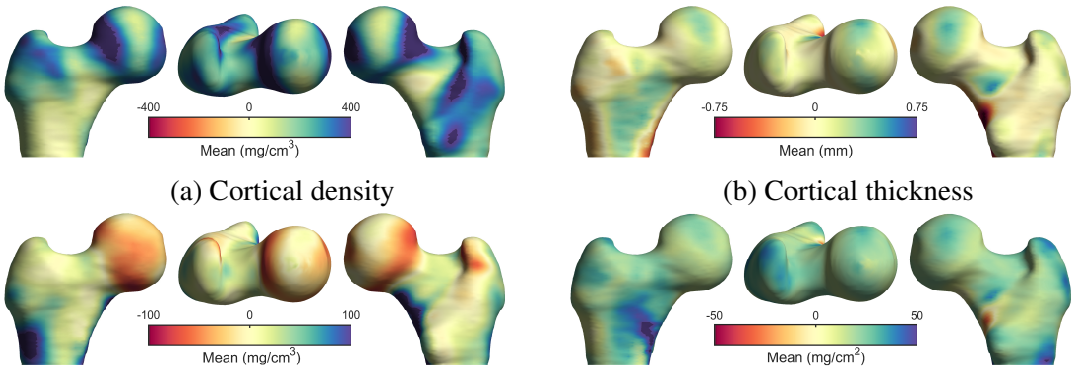
(d) Mass surface density

Fig. 3.30 CI validation plots of the rectangular CBMv2 method against the HRTC method over the proximal femur.

mean cBMD. The remaining differences are an indication of the limitations in each method's ability to adapt to local variations in the cortical density.

Cortical thickness

The rectangular CBMv2 method accurately estimates cortical thicknesses that are less than 3 mm, but increasingly underestimates larger cortical thicknesses. This is similar to its performance over the variable beam hardening simulation data analysed in Section 2.5.7. The endocortical SSC-CBM method also underestimates cortical thickness above 3 mm, but unlike the rectangular CBMv2 method its bias is approximately constant from 1.5–6 mm. Although this persistent bias limits its ability to measure cortical thickness, its constant nature



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

(d) Mass surface density

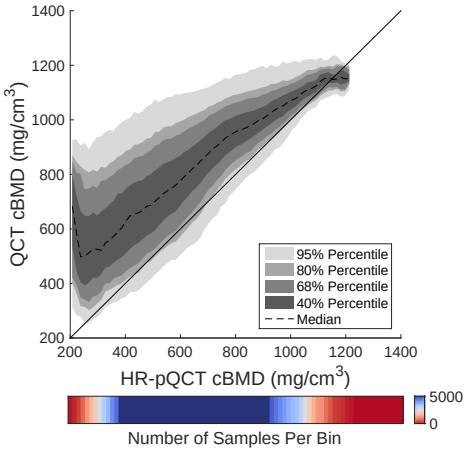
Fig. 3.31 Surface mapped rectangular CBMv2 errors over the proximal femur.

indicates that the endocortical SSC-CBM method appears to be robust to beam hardening effects over the proximal femur. This gives further evidence that the thickness-matched weighting function, that was used in Section 2.6 to improve the resilience of the rectangular CBMv2 method to the beam hardening in the simulated beam hardening dataset, may also improve its robustness over the proximal femur.

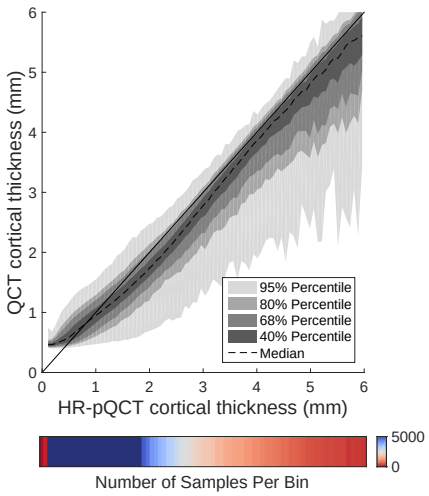
The endocortical SSC-CBM cortical thickness is calculated using Equation 2.6. It is overestimated for cortical thicknesses that are less than 1 mm, and underestimated for larger thicknesses. Both the equation, and the dense cortical and endocortical thickness measurements may be sources of the errors in the cortical thickness measurements. Using Equation 2.6 to estimate cortical thickness from the dense cortical and endocortical HRTC thickness measurements shows that Equation 2.6 consistently overestimates cortical thickness by approximately 0.2 mm as shown in Figure 3.35. As such, the errors in the cortical thickness measurement can be attributed to limitations in the physical assumptions governing Equation 2.6, and limitations in the ability of the method to accurately measure the dense cortical and endocortical thickness. These biases are larger than those experienced over the simulated data examined in Section 2.5.4.

Trabecular density

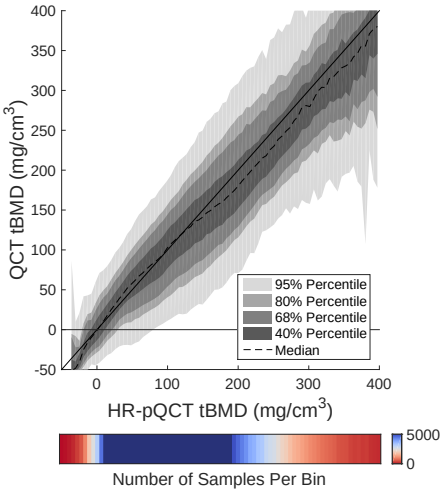
Both CBM methods provide fairly low bias tBMD estimates, but underestimate the trabecular densities above 150 mg/cm^3. The spatial distribution of these errors is reported in Figures 3.31c and 3.34c, and it show that both methods underestimate the tBMD around the sides of the femoral head just above the neck. It is possible this is because these profiles will encounter the 'principle compressive system', which is a region of dense trabecular bone that stretches from the medial cortex of the head into the femoral neck, near the end of the density profile. This will be almost entirely ignored by the periosteal and cortical weighted CBM methods, but given an equal weighting in the HRTC method. The rectangular CBMv2 method overestimates the trabecular densities below 150 mg/cm^3. Figures 3.31c and 3.29b show that these errors mostly occur where the cortical thickness is large. This indicates



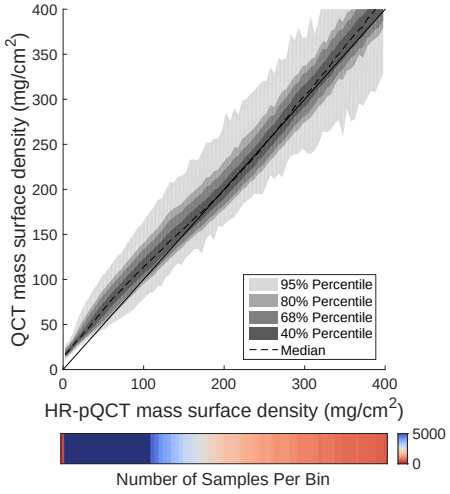
(a) Cortical density



(b) Cortical thickness



(c) Trabecular density



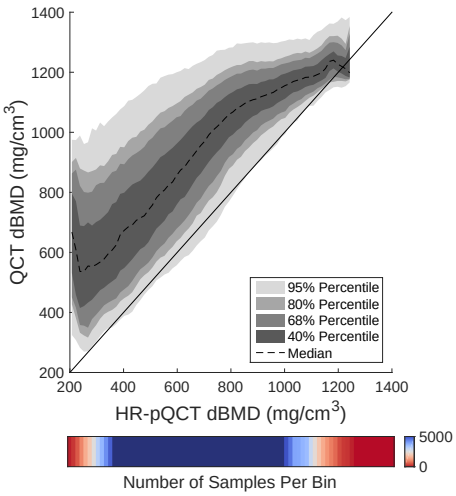
(d) Mass surface density

Fig. 3.32 CI validation plots of the endocortical SSC-CBM method against the HRTC method over the proximal femur.

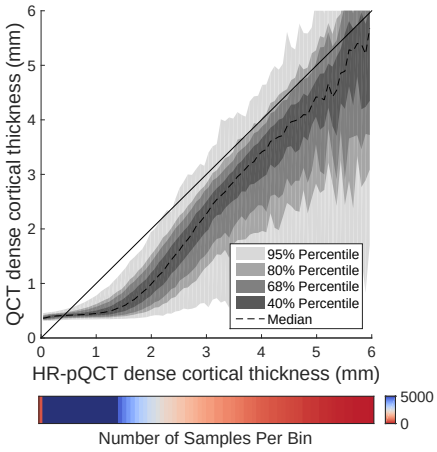
that much of the overestimations can be attributed to the optimiser over-fitting the model to account for beam hardening effects as discussed in section 3.5.5.

Endocortical thickness

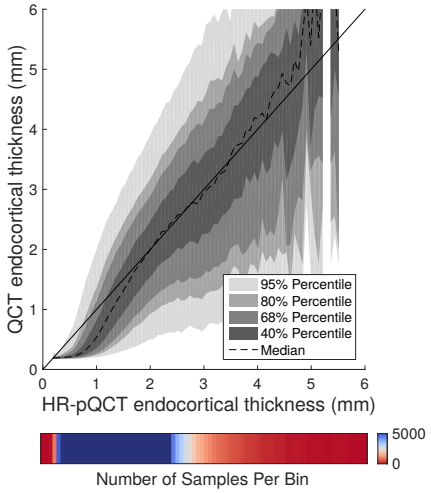
Figure 3.33c shows the endocortical SSC-CBM method provides a fairly unbiased relatively precise measure of the endocortical thickness over the range of 0–5 mm. Although, it should be noted it slightly underestimates endocortical thicknesses in the range of 0.3–1.7 mm, which encompasses most endocortical thicknesses. This is generally reflected by the spatial distribution of mean endocortical thickness errors in Figure 3.34g, which shows large regions where the endocortical thickness is underestimated. However, it also shows three regions



(a) Dense cortical density



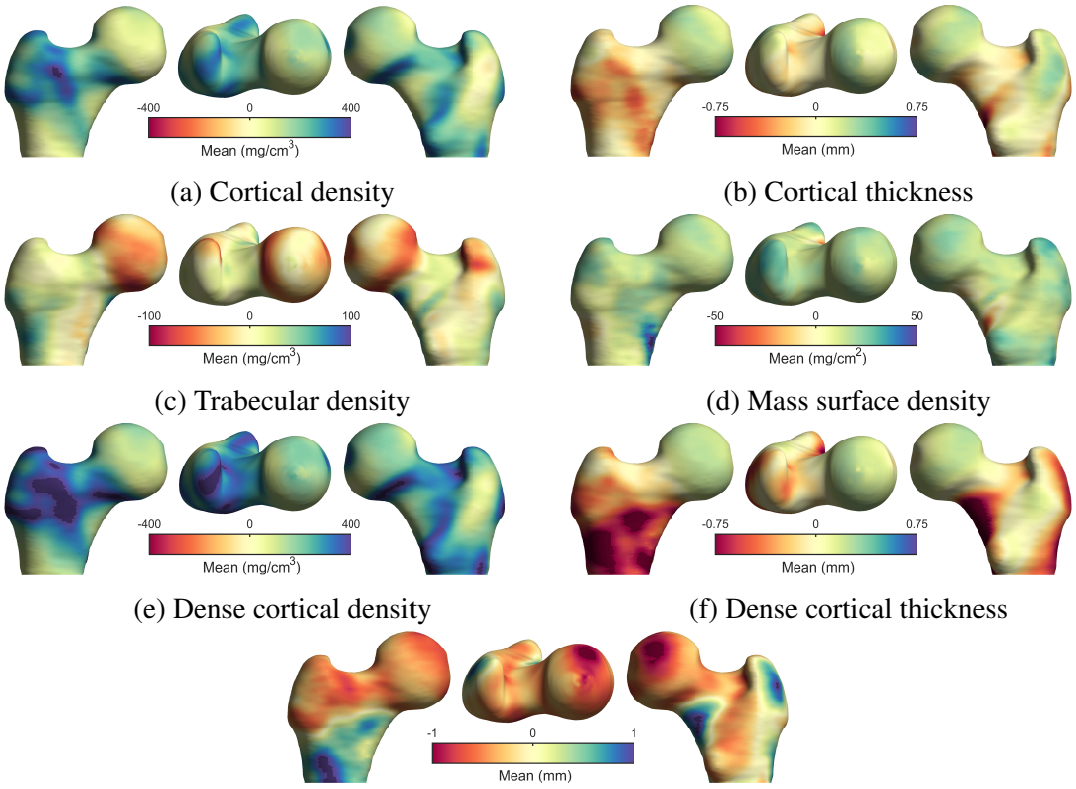
(b) Dense cortical thickness



(c) Endocortical thickness

Fig. 3.33 CI validation plots of the new endocortical SSC-CBM measurements against the HRTC method over the proximal femur.

of considerable area where the endocortical thickness is overestimated. These areas appear to occur where there are substantial subsurface trabecular structures that are successfully discriminated as trabecular bone by the HRTC, but confounded as an extension of the endocortical region by the endocortical SSC-CBM method at the lower QCT resolution. This leads to relatively infrequent but large regionally localised overestimations in the endocortical thickness, which have a greater impact on the mean shown in Figure 3.34g than the median values shown in Figure 3.33c.



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

(d) Mass surface density

(e) Dense cortical density

(f) Dense cortical thickness

(g) Endocortical thickness

Fig. 3.34 Surface mapped errors of the endocortical SSC-CBM measurements against the HRTC method over the proximal femur.

Conclusions

The validation of the CBM methods with the HRTC method shows that both methods are able to provide cortical thickness, trabecular density and mass surface density measurements with a low level of bias over the proximal femur as highlighted by the values reported in Table 3.4. Figures 3.30b and 3.32b show that the rectangular CBMv2 method is more susceptible to cortical thickness errors caused by beam hardening than the endocortical SSC-CBM method. A general comparison of the cortical thickness, cBMD, tBMD and MSD measurements over the proximal femur shows that the rectangular CBMv2 method remains better able to provide low bias cortical thickness measurements in the sub-millimetre range, while the endocortical SSC-CBM method generally provides lower bias cBMD, tBMD and MSD measurements. In addition, the endocortical SSC-CBM method is able to measure the endocortical thickness to a reasonable level of accuracy and precision. This indicates that unless the cortical thickness is the sole measurement of interest the endocortical SSC-CBM method is more versatile.

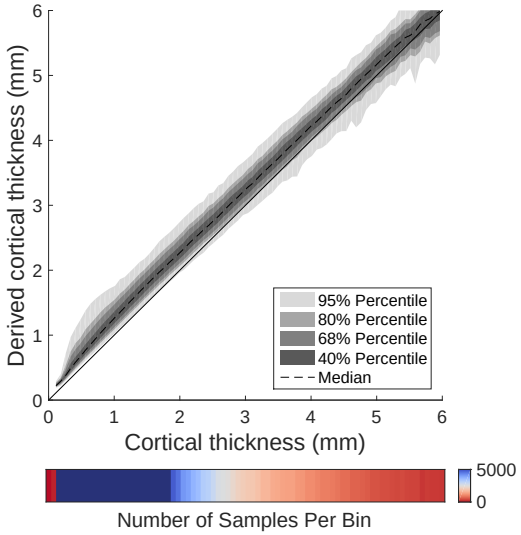
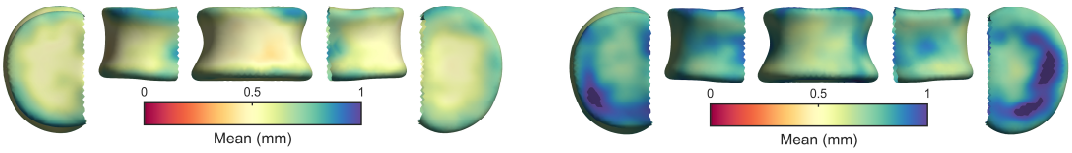


Fig. 3.35 The relationship between HRTC cortical thickness and the thickness derived from the dense cortical and endocortical thicknesses over the proximal femur.

3.6.4 Vertebral bodies results

Figure 3.36 shows the mean cortical and endocortical thickness measurements across the vertebral bodies as measured by the HRTC method. The validation results are shown in Table 3.5, and Figures 3.37–3.41 using the same general layout as the proximal femur results. The bias and precision values in Table 3.5 are only calculated over the 0.3–1 mm and 1–3 mm cortical thickness ranges as the vertebral bodies lack any thick cortical bone. The CI plots were also displayed over smaller thickness and density ranges to reflect the more porous and thinner cortical and trabecular structures [35, 106, 37], and the number of CI bins was reduced to 30 to reflect the smaller size of the dataset. The colour-map ranges were adjusted to better reflect the range of errors observed over the lumbar spine except for the cBMD and dBMD colour-maps; these were limited to -400–$400\,\mathrm{mg/cm^3}$ or 40 % of the maximum cBMD and dBMD measurements recorded by the HRTC method.



(a) Cortical thickness (b) Endocortical thickness

Fig. 3.36 Select surface mapped HRTC measurements over the lumbar spine.

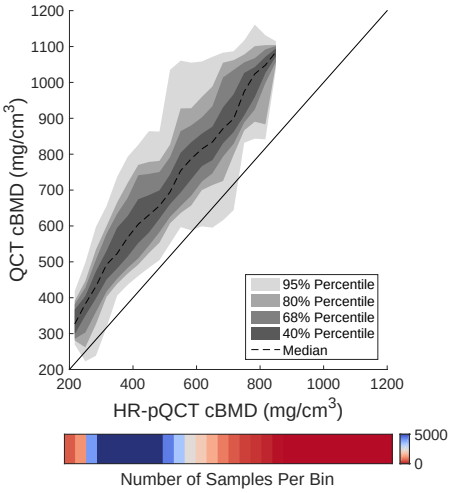
Table 3.5 A comparison of the biases, precisions and overall stability of the two CBM methods over the lumbar spine. The biases and precisions are calculated for each physical feature measured by each CBM method. The stability of each method is the % of successfully fit valid models out of the set of all measurement locations where a valid HRTC model exists.

Quantity	Thickness Ranges (mm)	Rectangle CBMv2	Endocortical SSC-CBM
Stability (%)	$0.0 \leq t < 6.0$	98.6	100.0
cBMD (mg/cm^3)	$0.3 \leq th_{cb} < 1.0$ $1.0 \leq th_{cb} < 3.0$	187 ± 93 231 ± 104	428 ± 128 343 ± 164
cTh (mm)	$0.3 \leq th_{cb} < 1.0$ $1.0 \leq th_{cb} < 3.0$	0.50 ± 0.34 0.49 ± 0.73	0.08 ± 0.23 0.24 ± 0.76
tBMD (mg/cm^3)	$0.3 \leq th_{cb} < 1.0$ $1.0 \leq th_{cb} < 3.0$	-20 ± 45 37 ± 57	-24 ± 46 14 ± 58
MSD (mg/cm^2)	$0.3 \leq th_{cb} < 1.0$ $1.0 \leq th_{cb} < 3.0$	39.6 ± 17.4 62.7 ± 51.3	33.3 ± 15.9 60.2 ± 54.2
dBMD (mg/cm^3)	$0.3 \leq th_{cb} < 1.0$ $1.0 \leq th_{cb} < 3.0$	- -	470 ± 135 433 ± 166
dTh (mm)	$0.3 \leq th_{cb} < 1.0$ $1.0 \leq th_{cb} < 3.0$	- -	0.01 ± 0.15 -0.35 ± 0.36
ecTh (mm)	$0.3 \leq th_{cb} < 1.0$ $1.0 \leq th_{cb} < 3.0$	- -	-0.26 ± 0.39 0.69 ± 1.15

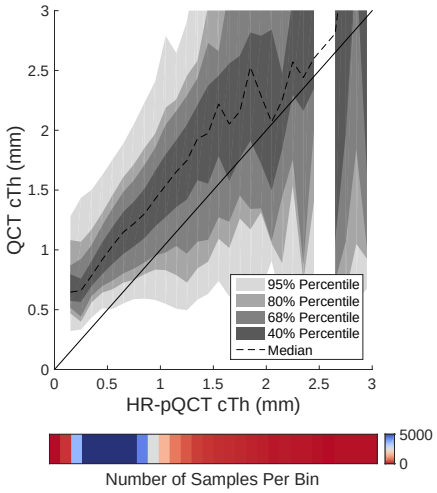
3.6.5 Vertebral bodies discussion

The density measurements of the vertebral validation study are challenging due to inconsistencies in the scanning protocol between the QCT and HR-pQCT scans as discussed in Section 3.4.5. This is a likely cause of the offset observed in both mass surface density CI plots. Despite these challenges, the CI plots of both methods show the cBMD and tBMD measurements are strongly correlated. It is worth noting that the rectangular CBMv2 cBMD measurements are actually better correlated over this dataset then over the proximal femur.

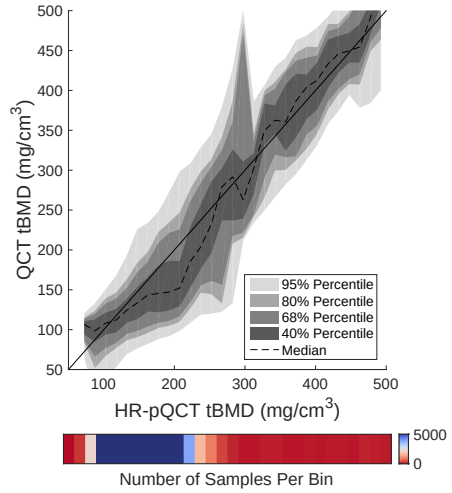
Both methods struggle to provide accurate cortical thickness measurements with each substantially overestimating thicknesses in the range of 1–3 mm, where both methods are able to provide measurements with low bias in the proximal femur. This cannot be explained by the difficulties in providing an accurate density calibration between the QCT and HR-pQCT scans. It may perhaps be due to the substantially greater proportion of bone that is endocortical as compared to the proximal femur. In the proximal femur the endocortical



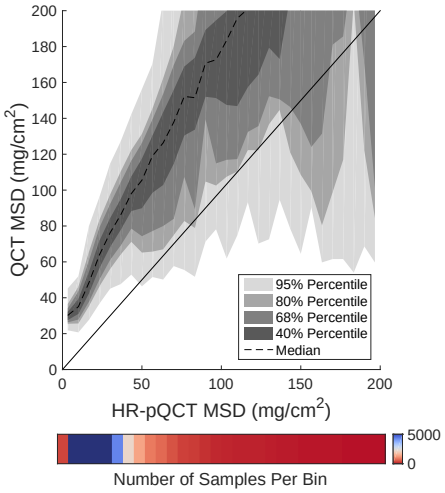
(a) Cortical density



(b) Cortical thickness



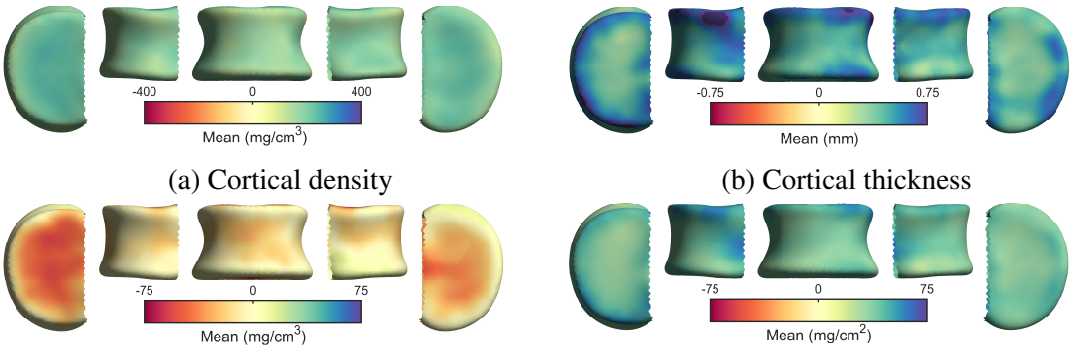
(c) Trabecular density



(d) Mass surface density

Fig. 3.37 CI validation plots of the rectangular CBMv2 method against the HRTC method over the lumbar spine.

thickness is similar to the cortical thickness where the cortical thickness is small, but it is smaller where the cortical thickness is large as shown by a comparison of Figures 3.29b and 3.29d. In contrast, the endocortical thickness is typically larger than the cortical thickness over the vertebral bodies as shown in Figures 3.36a and 3.36b. It may be that both methods are confounding the increased prominence of the endocortical region for an increase in the cortical thickness. The endocortical SSC-CBM method still provides a measure of the endocortical thickness, but like the cortical thickness measurement it is of a substantially lower quality than over the proximal femur.



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

(d) Mass surface density

Fig. 3.38 Surface mapped rectangular CBMv2 measurement errors over the lumbar spine.

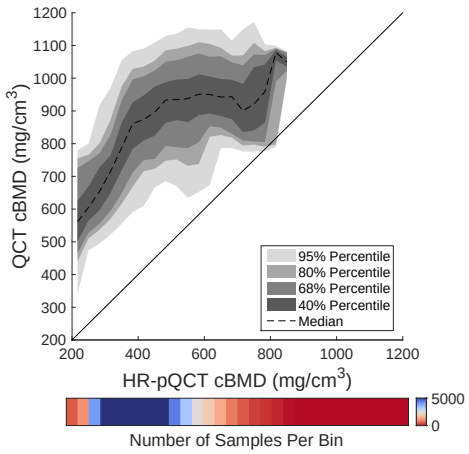
Conclusions

Validation of CBM methods over the spine is challenging due to its micro-structure, which means the cortical bone is substantially thinner than over most of the proximal femur [35, 37, 106]. This makes the estimation of both cortical density and thickness more difficult. These challenges are confounded by the size, age distribution and differing scanning QCT and HR-pQCT conditions in the vertebral bodies validation dataset. Despite this, both CBM methods are able to provide highly correlated cBMD, cortical thickness, tBMD and mass surface density results with precisions approaching those of measurements over the proximal femur. The endocortical SSC-CBM method is also able to provide highly correlated measurements of the endocortical thickness.

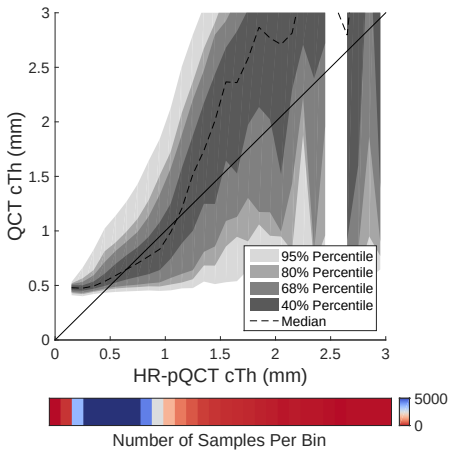
3.7 Conclusions

The HRTC method measures a wider range of features over the cortical and subsurface trabecular region than the FWHM method. It is also less susceptible to errors in thickness measurements caused by beam hardening effects. As such, it offers a favourable alternative to the FWHM method for generating gold standard measurements for comparison with CBM measurements from QCT scans.

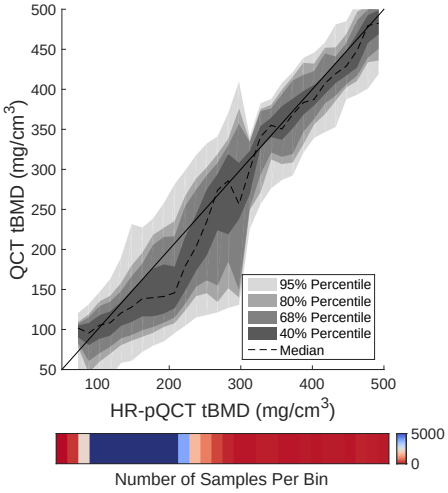
Comparisons with the HRTC method, show that the endocortical SSC-CBM method is able to measure the endocortical region with a low bias over the proximal femur. It also provides an improved measure of the cortical density, and is still able to measure the cortical thickness, trabecular density and mass surface density. However, its sub-millimetre cortical thickness measurements are less accurate than the rectangular CBMv2 method, which should be used to supplement the endocortical SSC-CBM measurements when cortical thickness is of interest. The results shown in Section 3.6.4 mark the first time that CBM methods have been validated over the lumbar spine. Both methods are able to measure cortical and subsurface trabecular features over the vertebral bodies in the lower spine, and the endocortical SSC-CBM method is also able to measure the endocortical thickness. These



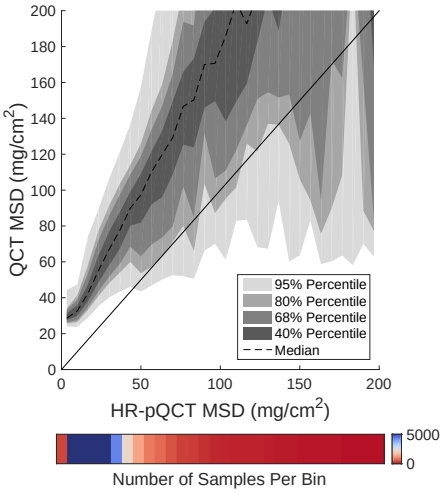
(a) Cortical density



(b) Cortical thickness



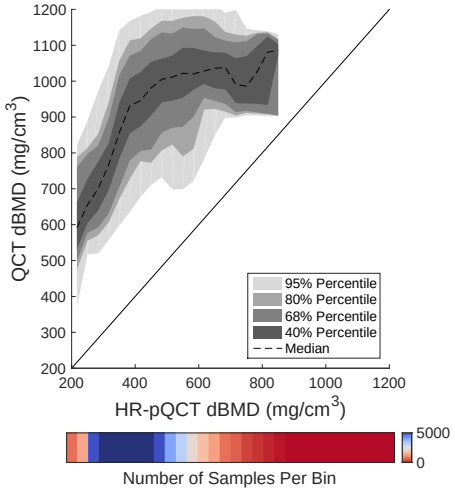
(c) Trabecular density



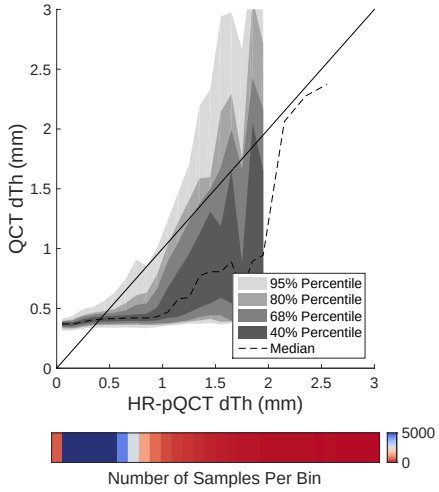
(d) Mass surface density

Fig. 3.39 CI validation plots of the endocortical SSC-CBM method against the HRTC method over the lumbar spine.

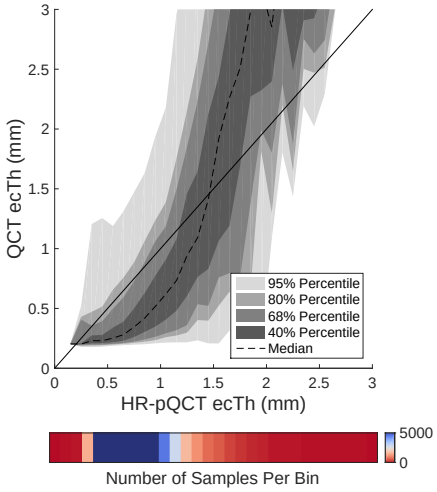
measurements have similar precisions to the measurements over the proximal femur, but higher biases.



(a) Dense cortical density

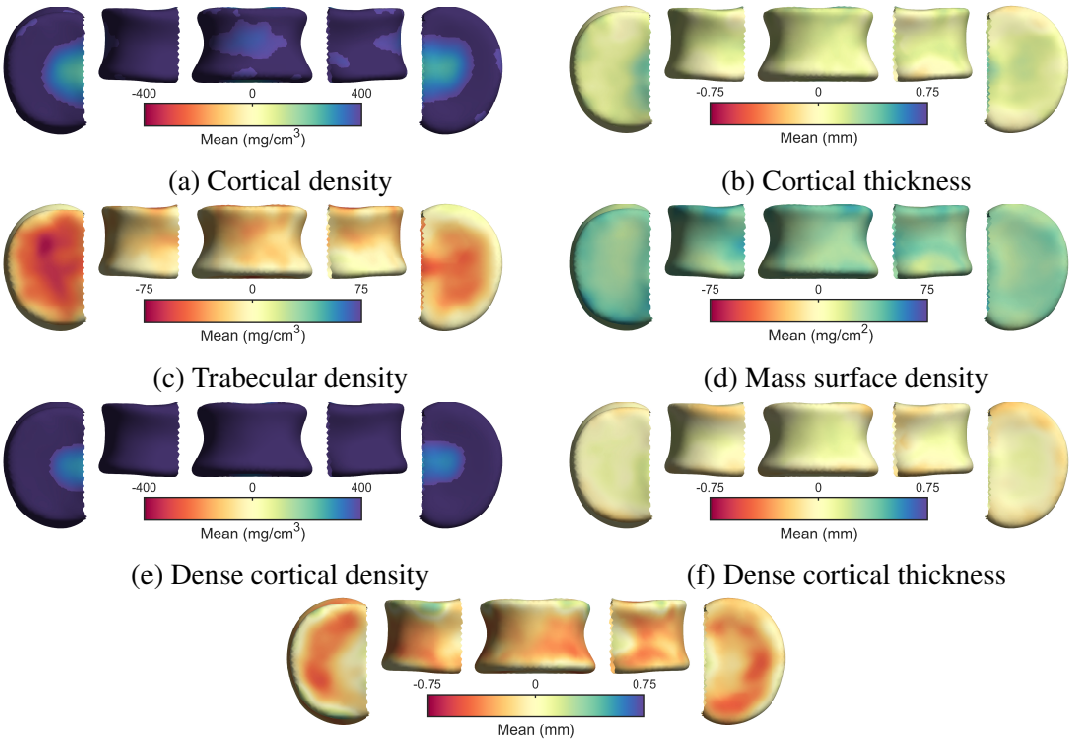


(b) Dense cortical thickness



(c) Endocortical thickness

Fig. 3.40 CI validation plots of the new endocortical SSC-CBM measurements against the HRTC method over the lumbar spine.



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

(d) Mass surface density

(e) Dense cortical density

(f) Dense cortical thickness

(g) Endocortical thicknes

Fig. 3.41 Surface mapped endocortical SSC-CBM errors over the lumbar spine.

Chapter 4

Clinical applications

4.1 Introduction

The main motivator for the development of the endocortical SSC-CBM method is to improve on existing techniques for clinical osteoporosis research. In this chapter, I investigate how well it is able to both detect physical changes in the cortical and subsurface trabecular bone architecture that occur in response to pharmaceutical therapies, and how well it is able to identify architectural differences in individuals with an elevated fracture risk. QCT scans from three clinical trials are examined. The first two are used to examine the effect of teriparatide treatment on the bone architecture of the proximal femur and lumbar spine, while the third allows for the changes in bone architecture associated with increased fracture risk to be examined.

4.1.1 Teriparatide

Teriparatide (human parathyroid hormone) is an established anabolic therapy for osteoporosis treatment [28, 77, 133], which induces bone formation and resorption [87]. Daily subcutaneous injections of teriparatide have been associated with a significant increase in aBMD in the lumbar spine and femoral neck [28, 34, 87], and an associated decrease in vertebral and non-vertebral fractures [87]. Chen et al. (2006) examined the contribution of the increased aBMD to the reduction in fracture risk over the lumbar spine, and found that it only accounted for 30–40 % of the observed change. The rest was attributed to changes in bone architecture and other non-BMD related determinants in bone strength [25]. This finding prompted extensive study into the impact of teriparatide on bone architecture [11, 43, 47, 59, 77, 83, 100, 127, 128, 133].

Hirano et al. (1999) and Burr et al. (2001) used animal models with similar Haversian based intracortical remodelling to humans to examine the impact of teriparatide therapy on bone architecture and strength using detailed but invasive histomorphic measurements and mechanical strength tests [20, 56]. Both studies showed a significant increase in periosteal and endosteal bone formation and intracortical remodelling, which were accompanied by a significant increase in cortical porosity [20, 56]. Although both were significant, the rate

of endosteal apposition was substantially greater than the rate of periosteal apposition. The studies also showed that the increase in cortical porosity associated with teriparatide had no impact on the bone strength, and it was hypothesised that this was due to an overall increase in cortical bone mass resulting from the increased periosteal and endosteal bone formation. Burr et al. (2001) also noted that the increase in cortical porosity, although significant throughout the cortex, was concentrated near the endocortical surface, where it would have the least impact on the bending strength of the bone [20].

HR-pQCT scans have been used to examine the impact of teriparatide on bone micro-architecture and to estimate bone strength at peripheral sites in humans [77, 127, 128]. Tsai et al. (2015) studied the effect of teriparatide therapy on the distal radius and distal tibia. They showed that 12 months of teriparatide therapy coincided with an increase in tBMD, trabecular thickness, trabecular separation and cortical porosity; no change in total BMD and cortical thickness; and a decrease in cBMD and the trabecular number. It could be argued that these changes indicate a reduction in bone quality, but importantly they were not associated with a decrease in the estimated strength at either site [127]. Two additional studies showed that bone quality in the distal radius and distal tibia continued to deteriorate with continued teriparatide therapy. Macdonald et al. (2011) and Tsai et al. (2016) both showed a significant decrease in total BMD, cBMD, a significant increase in cortical porosity and a decrease in cortical thickness in both sites after 18 and 24 months respectively [77, 128]. The only positive response to treatment was a continued increase of the tBMD. Despite the marked decrease in bone quality, both studies reported no significant change in bone strength.

Although the proximal femur cannot be scanned at the resolutions possible in HR-pQCT [44], several studies have been able to extract some information about the changes in bone architecture and strength in response to teriparatide therapy [11, 47, 59, 100, 133]. Analysis of the proximal femur using the rectangular CBMv2 method has shown that teriparatide therapy leads to widespread increases in trabecular density and cortical thickness, a widespread decrease in cortical density, and little change in the cortical mass surface density [100, 133]. The observed changes in cBMD and tBMD were statistically significant over most of the proximal femur, while the increase in cortical thickness was only statistically significant in regions of high stress [100, 133]. Borggrefe et al. (2010) and Ito et al (2014) both used the commercially available CTXA Hip Exam Analysis protocol by Mindways[1], which can determine trabecular and cortical contributions to overall BMD [21, 65]. Borggrefe et al. (2010) shows 24 months of teriparatide therapy leads to a significant increase in buckling strength and an increase in bending strength due in part to an increase in cortical thickness over the femoral neck due to endosteal deposition of cortical bone [11, 59]. Ito et al (2014) reports similar changes in the femoral neck, shaft and inter-trochanter in a 16 month placebo controlled trial. Although these studies have provided additional insight into the effect of teriparatide therapy on the bone architecture of the proximal femur, none have been able to comment on changes in the level of intracortical remodelling or endosteal deposition over the wider proximal femur.

The lumbar spine is another region of research interest with several studies examining changes in the bone architecture and strength [25, 34, 43, 46, 47, 83, 131]. All have shown

[1]CTXA-Hip™ by Mindways Software Inc. http://qct.com/home/products/cxta-hip/

a significant increase in aBMD, and all that considered tBMD also showed it significantly increased [43, 46, 47, 83, 131]. Graeff et al. (2007, 2009) considered the impact of 12 and 24 months of teriparatide therapy on trabecular bone structure and estimated strength and found it produced a significant increase in BV/TV, trabecular number, trabecular thickness, bending strength, compressive strength and a significant decrease in trabecular separation. Cortical bone was excluded from this analysis [46, 47]. Genant et al. (2016) and Whitmarsh et al. (2014) did include cortical bone in their analysis of the effect of 12 months of teriparatide treatment. Genant et al. (2016) used the MIAF-Spine software to compartmentalise the lumbar spine into cortical and trabecular bone and found an increase in aBMD, cBMD, cortical thickness and tBMD [43], but the significance of these changes was not reported. Whitmarsh et al. (2014) used the rectangular CBMv2 method to show regionally significant increases in tBMD, mass surface density and cortical thickness combined with a significant increase in the overall cortical thickness, tBMD and mass surface density [131]. As with the proximal femur, changes in the level of intracortical remodelling and endosteal deposition have not been considered over the lumbar spine.

4.1.2 Fracture risk assessment

Fracture risk prediction allows for the targeted treatment of high risk individuals. As discussed in Section 1.3.1, aBMD defined osteoporosis is the current clinical standard for identifying individuals with an elevated fracture risk [48, 135]. Unfortunately fewer than half of fragility fractures occur in individuals with aBMD defined osteoporosis [58, 110], so the identification of measurements that allow for improved prediction of fracture risk is a key area of research [96]. CBM methods can be used to identity focal regions where architectural measurements differ significantly between fracture and control groups [96, 99, 126].

In several recent studies, the tBMD and MSD measurements were incorporated into DXA-based fracture prediction models. The measurements were included from regions that are significantly different in individuals with elevated fracture risk. This has led to a small improvement in the model's predictive ability [96], which was significant over the large Osteoporotic Fractures in Men (MrOS) study [126]. It is hoped that measuring the endocortical thickness and dBMD may allow for further improvements in fracture prediction, but before this can be directly investigated, the ability of the endocortical SSC-CBM method to measure regions where elevated fracture risk is associated with significant changes in endocortical thickness or dBMD must be established.

4.1.3 Statistical parametric mapping (SPM)

CBM methods produce localised measurements of bone architecture over a surface. The statistical significance of the regional variations in these measurements either with time or across different population groups can be established using statistical parametric mapping (SPM). This is an established technique in the neuro-imaging community that allows statistical inferences to be drawn from dense spatially correlated data [40]. As each CBM model across the measurement surface is fit independently, it cannot be assumed to be be

spatially correlated. Spacial normalisation is typically achieved with some form of mesh connectivity based smoothing that is applied across each measurement surface [98–100]. An architectural feature measured across a population can be represented by a general linear model (GLM), which relates the spatially correlated measurements to characteristics that may have impacted the measurements in each individual such as the individual's age, weight, sex or treatment profile. Once generated, the GLM may then be used to implement univariate statistical analyses, from which SPM can be used to identify significant regional deviations from the norm using random field theory [40].

4.1.4 Density calibration

In each clinical QCT study analysed in this chapter, most or all of the scans contained a solid Mindways calibration phantom that was used to perform density calibration. The scans containing Mindways solid phantoms were sampled along the phantom cores to produce at least 20,000 BMD to HU value pairs. A least squares fit was used to produce a linear calibration curve. The calibration value pairs and linear calibration curves were visually checked for each calibration phantom with no suspicious value pairs or calibration curves detected. The fracture discrimination study included some scans that did not include a solid Mindways calibration phantom. These were only included for further analysis if they could be calibrated from a Mindways phantom scanned in the same machine with the same settings.

Solid Mindways calibration phantom

The solid Mindways calibration phantom includes five evenly spaced cylindrical calibration cores, each with a diameter of 19 mm. The manufacturer provides the BMD values of each core as well as additional correction values for beam hardening due to differences in material composition between the calibration cores [84]. In each scan, the phantom location was manually specified before the centre 50 % of each core was sampled at the image pixel resolution in every slice containing the phantom. The calibration cores of the phantom contain differing but known amounts of K_2HPO_4 and water equivalent plastic, where K_2HPO_4 is the dominant mineral in bone. This means that the relationship between the HU measured in each core and the density of K_2HPO_4 can only be established after accounting for the density of the water equivalent plastic also present in each core. This can be achieved using Equations 4.1.

$$HU_{\text{core } K_2HPO_4} = m_{K_2HPO_4}\rho_{K_2HPO_4} + b_{K_2HPO_4} \tag{4.1a}$$

$$HU_{\text{core}} = \rho_{H_2O \text{ eqv.}} + m_{\text{core}}\rho_{K_2HPO_4} + b_{\text{core}} \tag{4.1b}$$

$$m_{K_2HPO_4} = m_{\text{core}} - 0.2174 \tag{4.1c}$$

$$b_{K_2HPO_4} = b_{\text{core}} + 999.6 \tag{4.1d}$$

Equation 4.1a defines the relationship between the density of the K_2HPO_4 in the cores and the measured HU values. The $HU_{\text{core } K_2HPO_4}$ values cannot be measured directly due

to the varying quantities of water equivalent plastic in each core. Instead the overall HU of the cores, HU_{core}, can be measured as shown in Equation 4.1b. m_{core} and b_{core} are the slope and intercept of the linear regression model of the combined K_2HPO_4 and water equivalent plastic HU measurements. The effect of the water equivalent plastic can be removed from the slope and intercept values using equations 4.1c and 4.1d, which are provided by the manufacturer [84]. This gives the slope and intercept relating HU to BMD.

4.2 Teriparatide over the proximal femur

QCT scans of the proximal femur from the EUROFORS trial were re-analysed using the endocortical SSC-CBM method to investigate the impact of treatment on intracortical remodelling activity and endosteal apposition of cortical bone.

4.2.1 Data acquisition

EUROFORS[2] (The EUROpean Study of FORSteo) was a multinational randomised, open label controlled, prospective two year clinical trial into the impact of one year of teriparatide treatment followed by an additional year of either teriparatide or an anti-resorptive treatment [34]. EUROFORS was conducted in 95 centres across 10 European countries on post-menopausal ambulatory women with a history of fragility fractures who were otherwise healthy. Eligible women had to be $\geq$55 years old, at least two years post-menopausal, with a lumbar spine BMD T-score of -2.5 or less, who had suffered at least one documented fragility fracture in the last three years. Prospective participants were checked to ensure they had normal baseline levels of serum PTH, alkaline phosphatase, and calcium, and were not suffering any severe or chronically disabling diseases aside from osteoporosis. Women were excluded who had severe or chronically disabling diseases or took pharmaceutical treatments that are known to cause secondary forms of osteoporosis [34] . A subgroup of participants were QCT scanned at baseline, 6, 12 and 24 months. Only QCT scans from trial participants who received two years of teriparatide therapy and attended all check-ups were included for analysis by the endocortical SSC-CBM method, and where possible both left and right hips were included. Scans were excluded for high levels of artefacts and one instance of a corrupted file. This resulted in 236 scans from 63 individuals.

4.2.2 Experimental pipeline

The ability of the endocortical SSC-CBM method to provide a more detailed interpretation of the architectural changes occurring over the proximal femur in response to teriparatide treatment was examined using the pipeline shown in Figure 4.1. All scans contained Mindways solid phantoms and were calibrated using the approach described in Section 4.1.4, then segmented and registered to the same canonical surface using the process described in Section 3.4.2. The rectangular CBMv2 and endocortical SSC-CBM methods were applied

[2]ClinicalTrials.gov number NCT00191425

to the baseline and 24 month scans, and the resulting measurements were compared to give the change in bone architecture in response to 24 months of teriparatide therapy. SPM was used to examine the regional significance of changes in these measurements using the GLMs discussed in Section 4.2.3. A mask was used to exclude measurements where the acetabulum approaches the femoral neck as the modelling of closely adjacent cortices was not supported in the CBM implementation developed for this thesis. Similar masks were used in the other clinical datasets considered in this chapter.

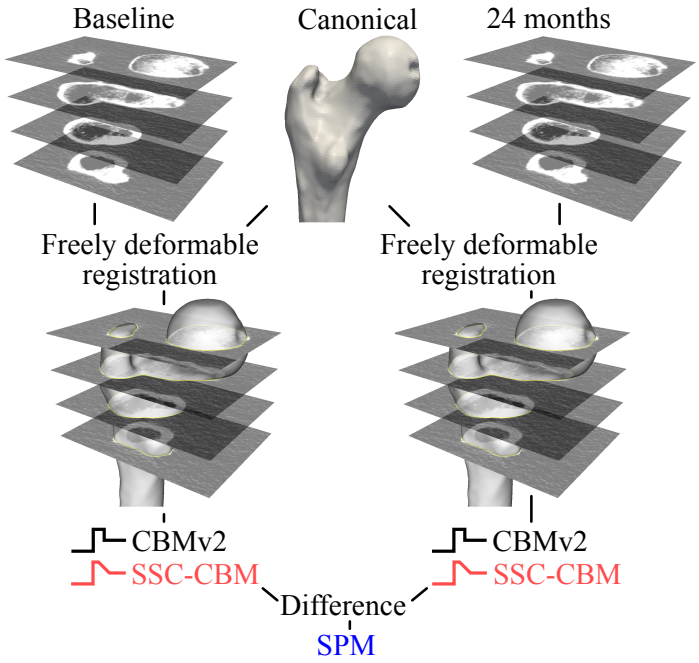


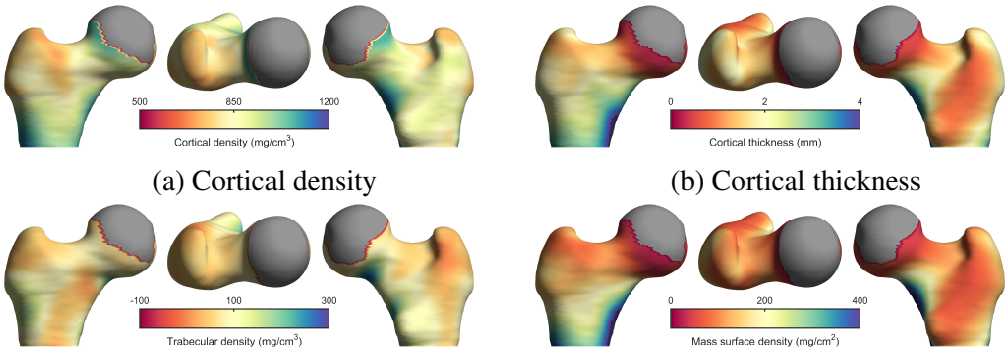
Fig. 4.1 The pipeline used to investigate the effect of teriparatide therapy on bone macrostructure over the proximal femur.

4.2.3 Statistical parametric mapping (SPM)

SPM was used to identify regions where the bone architecture changed significantly during the 24 months of teriparatide therapy. Spatial correlation between each subject's measurements was achieved with six iterations of mesh connectivity based precision smoothing to give an equivalent filter FWHM of 10 mm. As this is a longitudinal study, the baseline measurements were subtracted from the follow-up measurements to give the change in bone architecture during treatment across 118 proximal femurs. Any left and right femurs from the same individuals were combined to give the mean change across both femurs. This gave the change in bone architecture with treatment for 63 individuals [100]. A constant GLM was then compared with a null model to give F-statistics across the proximal femur, from which p-values across the surface were determined using random field theory. P-values of < 0.05 were used to define regions where an architectural measure had significantly changed.

Table 4.1 Architectural measurements of the baseline population and its changes with treatment made by the rectangular CBMv2 and endocortical SSC-CBM methods. The significance of the architectural changes with treatment is calculated using a paired t-test (bias) at $p < 0.01$. Significance is indicated with an *.

		Baseline		Teriparatide 24 months	
		Rectangle CBMv2	**Endocortical SSC-CBM**	**Rectangle CBMv2**	**Endocortical SSC-CBM**
cBMD	(mg/cm^3)	890±128	869±128	-29^*±52	-37^*±48
cTh	(mm)	1.72±0.96	1.52±0.92	0.06^*±0.18	0.07^*±0.14
tBMD	(mg/cm^3)	84±65	66±59	11^*±19	9^*±17
MSD	(mg/cm^2)	150.1±103.6	137.0±99.9	0.0±12.9	0.4^*±12.5
dBMD	(mg/cm^3)	-	997±136	-	-42^*±52
dTh	(mm)	-	0.63±0.61	-	0.02^*±0.12
eTh	(mm)	-	1.78±1.01	-	0.11^*±0.27



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

(d) Mass surface density

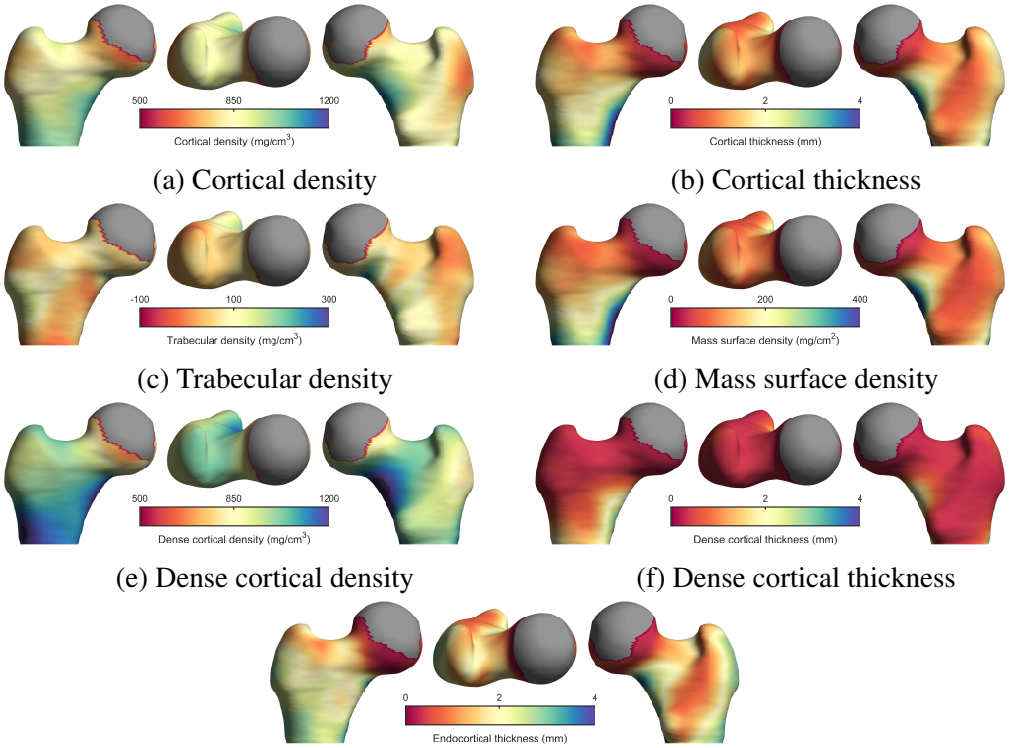
Fig. 4.2 Surface maps of the proximal femur showing baseline features measured by the rectangular CBMv2 method before teriparatide treatment.

4.2.4 Results

Figures 4.2 and 4.3 show the baseline features measured by the two CBM methods. Table 4.1 shows the mean changes in bone architecture in response to teriparatide therapy measured by each method, and Figures 4.4 and 4.5 show the mean regional changes in bone architecture. The regions in which these changes are significant are highlighted.

4.2.5 Discussion

The QCT scans from the 24 month teriparatide subgroup of the EUROFORS study have been analysed twice before [100, 133] using independent implementations of the rectangular CBM method, but never before with the endocortical SSC-CBM method. The CBMv2 results shown in Table 4.1 and Figure 4.4 broadly agree with these previous publications, and were included to allow comparison with the endocortical SSC-CBM results. Both CBM methods show teriparatide is associated with a decrease in cortical density, and an increase in



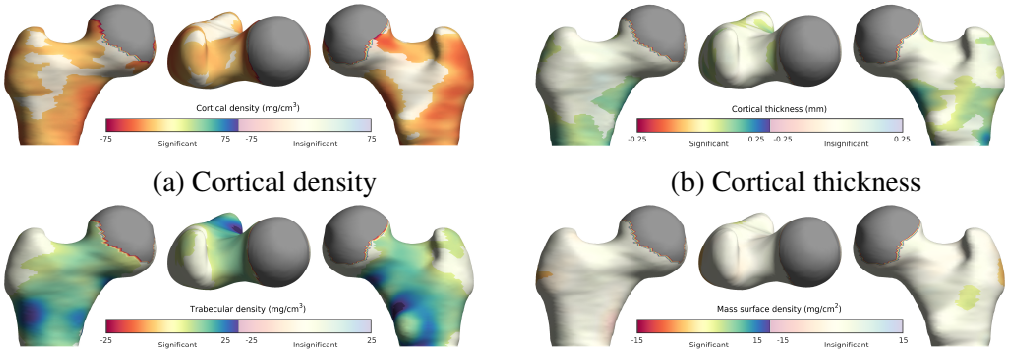
(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Mass surface density

(e) Dense cortical density (f) Dense cortical thickness

(g) Endocortical thickness

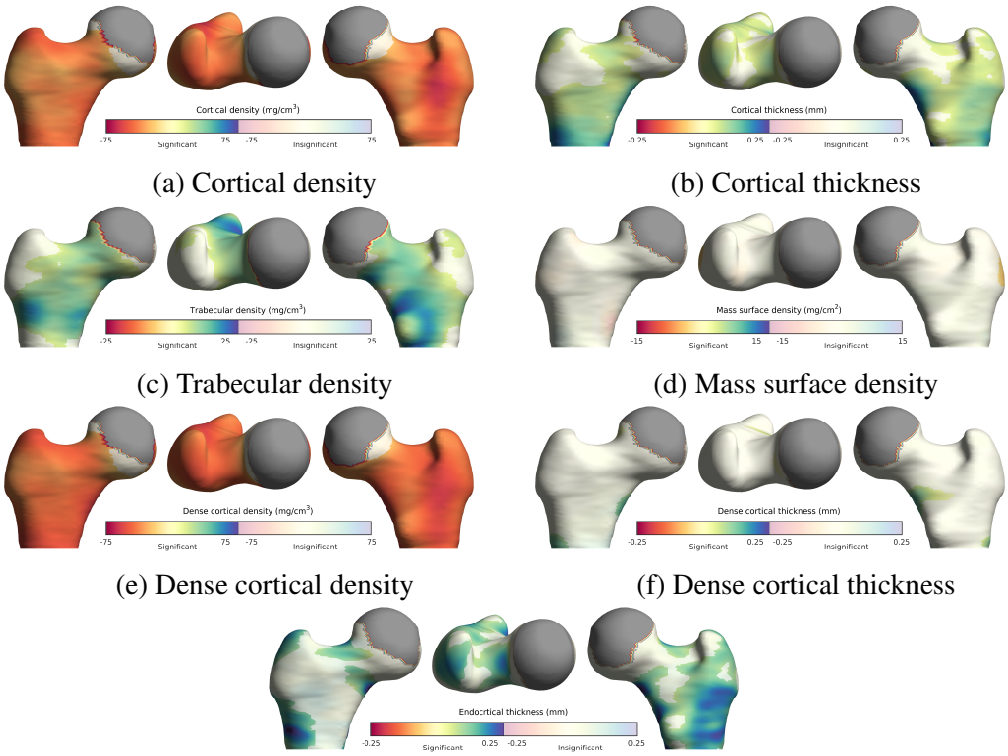
Fig. 4.3 Surface maps of the proximal femur showing baseline features measured by the endocortical SSC-CBM method before teriparatide treatment.



(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Mass surface density

Fig. 4.4 Surface maps of the proximal femur showing areas of significant change in the rectangular CBMv2 measurements due to teriparatide treatment.



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

(d) Mass surface density

(e) Dense cortical density

(f) Dense cortical thickness

(g) Endocortical thickness

Fig. 4.5 Surface maps of the proximal femur showing areas of significant change in the endocortical SSC-CBM measurements due to teriparatide treatment.

cortical thickness and trabecular density. The endocortical SSC-CBM also shows a decrease in dBMD and an increase in mass surface density, dense cortical thickness and endocortical thickness. The mean changes in bone architecture during treatment are all significant, except for the mass surface density measured by the rectangular CBMv2 method which does not change perceptibly as shown in Table 4.1.

The regional changes in bone architecture in Figures 4.4 and 4.5 show that both methods measure a widespread decrease in cBMD and widespread increases in tBMD and cortical thickness combined with focal increases and decreases in mass surface density. The endocortical SSC-CBM method measured a larger magnitude of decrease in cBMD, and greater regions of significance for all features measured aside from the mass surface density. The endocortical SSC-CBM method also showed significant regional decreases in dBMD and increases in endocortical thickness, and the dense cortical thickness. The changes in both the dBMD and endocortical thickness are significant over much of the proximal femur, while the significant increases in dense cortical thickness were confined to the base of the calcar femorale and a small section of the lateral aspect of the femoral shaft. Unfortunately, no placebo group was included in the EUROFORS study. As a result, the observed changes should be viewed with some caution as they will include changes due to ageing, but as ageing is typically associated with decreased cBMD, cortical thickness and tBMD [60, 113], it is unlikely to change the broad conclusions especially regarding the observed increases in cortical thickness and tBMD.

The endocortical SSC-CBM method allows changes in intracortical remodelling and endosteal apposition in response to teriparatide therapy to be considered for the first time over the proximal femur. The dBMD and dense cortical thickness measurements show that most of the decrease in cBMD is due to a decrease in dBMD with little change in the dense cortical thickness. This suggests an increase in intracortical remodelling with a resultant increase in cortical porosity and decrease in dBMD as observed over the distal radius and distal tibia in several studies of humans [20, 56]. The large increase in endocortical thickness compared with dense cortical thickness suggests that teriparatide therapy leads to an increase in endosteal apposition, but that most of this new bone does not have a chance to mature into dense bone during continued teriparatide therapy. As noted in Poole et al. (2011), most of the significant increases in cortical thickness are centred about several regions of high stress including the inferomedial cortex and the calcar femorale, and the insertion sites for several large muscles involved in locomotion [99]. The endocortical thickness also significantly increases over these regions, while the dense cortical thickness only increases significantly around the base of the calcar femorale.

4.2.6 Conclusions

For the first time, the endocortical SSC-CBM method allows the increase in cortical thickness over the proximal femur to be largely attributed to an increase in endosteal apposition, and much of the decrease in cortical density to be attributed to intracortical remodelling. These findings are in line with those observed previously both over the proximal femur, and in other regions in the human body. Some caution should be reserved for the attribution of the

decrease in dBMD to intracortical remodelling as some is likely also due to natural ageing. A placebo controlled trial is needed.

4.3 The lumbar spine

The QCT scans from the placebo and teriparatide groups of a double-blind clinical trial were analysed using the endocortical SSC-CBM method to investigate the impact of teriparatide therapy on the cortical architecture of the lumbar spine. This trial also included several romosozumab groups, which were not included in this analysis to avoid distracting from the comparative effects of teriparatide on the proximal femur and lumbar spine.

4.3.1 Data acquisition

The analysed placebo and teriparatide groups were part of a multicentre phase two randomised, placebo-controlled clinical trial designed to examine the efficacy and safety of the pharmaceutical, romosozumab, over a 12-month period. The trial included a total of 419 women enrolled in eight parallel groups: six with different romosozumab dosages; one receiving teriparatide; and one receiving a placebo. The placebo and teriparatide groups yielded 20 and 19 participants respectively. Eligible trial participants were ambulatory postmenopausal women between the ages of 55 and 85 who had low BMD with a T-score between -2.5 and -3.5 at the lumbar spine, total hip and femoral neck. Participants were excluded for: vertebral fractures or fragility fractures of the wrist, humerus, hip or pelvis suffered after age 50; a history of metabolic bone disease; a serum level of 25-hydroxy; vitamin D of less than $20\,\mathrm{ng/ml}$; untreated hyperthyroidism or hypothyroidism; substantially impaired renal function (estimated creatinine clearance $30\,\mathrm{ml/min}$, as assessed by the Modification of Diet in Renal Disease equation); current hypercalcemia or hypocalcemia; cancer; or prior use of agents affecting bone metabolism. Details of the study design and study participants have been previously published [43, 83].

4.3.2 Experimental pipeline

The baseline and 12 month scans of the teriparatide and placebo group contained Mindways solid calibration phantoms and were calibrated using the approach described in Section 4.1.4. The L1 vertebra was segmented and registered in each scan using the process described in Section 4.3.3, before the rectangular CBMv2 and endocortical SSC-CBM methods were applied using the pipeline shown in Figure 4.6. The measured changes between baseline and 12 months were calculated for each group, then the changes in the teriparatide group were compared against those in the placebo groups. The regions with significant architectural changes were identified as discussed in Section 4.3.4. A mask was used to exclude measurements where the adjacent vertebrae approach the measurement surface as the modelling of closely adjacent cortices was not supported in this thesis as discussed in Section 4.2.2.

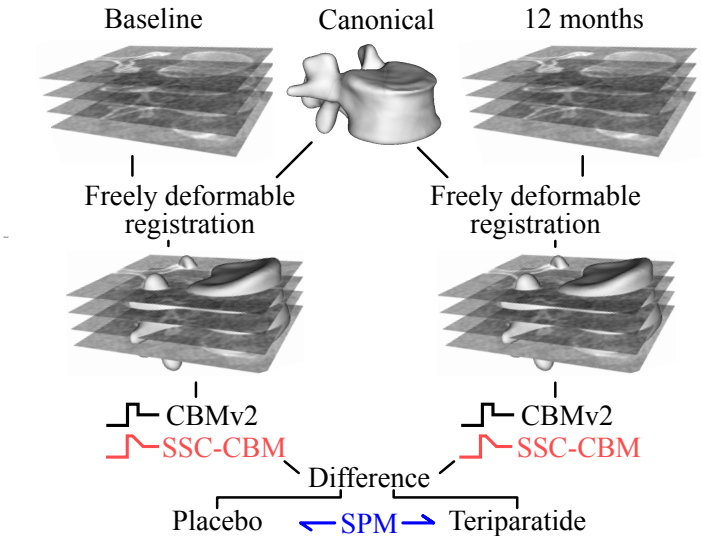


Fig. 4.6 The pipeline used to investigate the effect of teriparatide therapy on the bone architecture of the lumbar spine.

4.3.3 Registration

Registration was broadly performed as described in Section 3.4.2, but due to the difficulty in registering vertebral processes a slightly different protocol was used. The scans were roughly aligned using a similarity transform, and then registered using the freely deformable LAD algorithm. During the LAD a large search radius was used to align the processes, but this search radius was reduced with each iteration to avoid over-fitting. Also in the initial iterations, normal-aligned but widely separated points were matched in an attempt to align processes before the search radius was reduced using the protocol described in the WxRegsurf help[3].

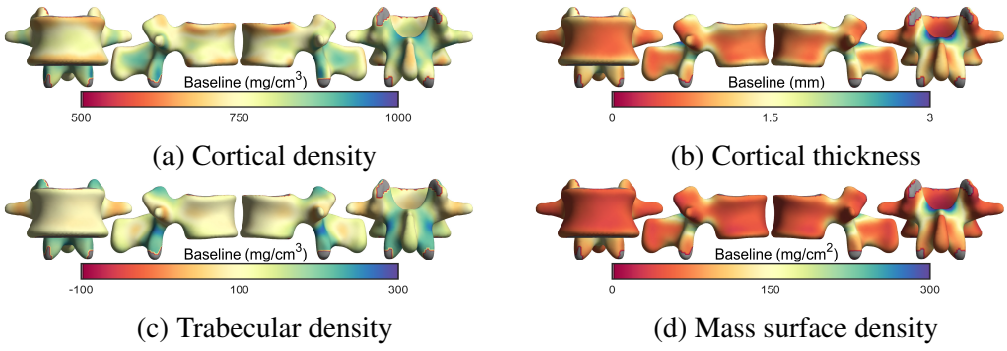
4.3.4 Statistical parametric mapping (SPM)

SPM was used to identify regions where the bone architecture changed significantly during the 12 months of teriparatide therapy, while controlling for changes resulting from ageing through comparison with the placebo group. As with the proximal femur study, all measurements were made over the same canonical surface, and spatial correlation between each subject's measurements was achieved with six iterations of mesh connectivity based precision smoothing to give an equivalent filter FWHM of 5 mm. The significance of the architectural changes in the placebo group over the 12 months was assessed using a constant GLM compared with a null model, while the significance of the architectural changes in the teriparatide group compared with the placebo group was assessed using a constant GLM with a random-effect term for the group compared with a constant GLM. In both cases, this gives F-statistics across the vertebrae, from which p-values across the surface were determined

[3]http://mi.eng.cam.ac.uk/~ahg/wxRegSurf/vertebra.html

Table 4.2 The mean baseline measurements of the placebo and the teriparatide group as measured by the rectangular CBMv2 and endocortical SSC-CBM methods with an * included after the baseline teriparatide measurements that differ significantly from the placebo populations. Significance values are calculated using an unpaired t-test at $p < 0.01$.

	Placebo (n = 20)		Teriparatide (n = 19)	
	Rectangle CBMv2	**Endocortical SSC-CBM**	**Rectangle CBMv2**	**Endocortical SSC-CBM**
Age (years)	67.2±6.1	67.2±6.1	65.6±6.1	65.6±6.1
cBMD (mg/cm^3)	779±107	755±56	786*±111	752*±60
cTh (mm)	1.07±0.64	0.99±0.69	1.06±0.59	0.99±0.63
tBMD (mg/cm^3)	128±64	120±60	130*±63	122*±60
MSD (mg/cm^2)	80.6±58.5	74.3±54.8	80.0*±52.9	73.6*±49.5
dBMD (mg/cm^3)	-	830±54	-	829*±57
dTh (mm)	-	0.53±0.49	-	0.51*±0.44
eTh (mm)	-	0.92±0.66	-	0.94*±0.65



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

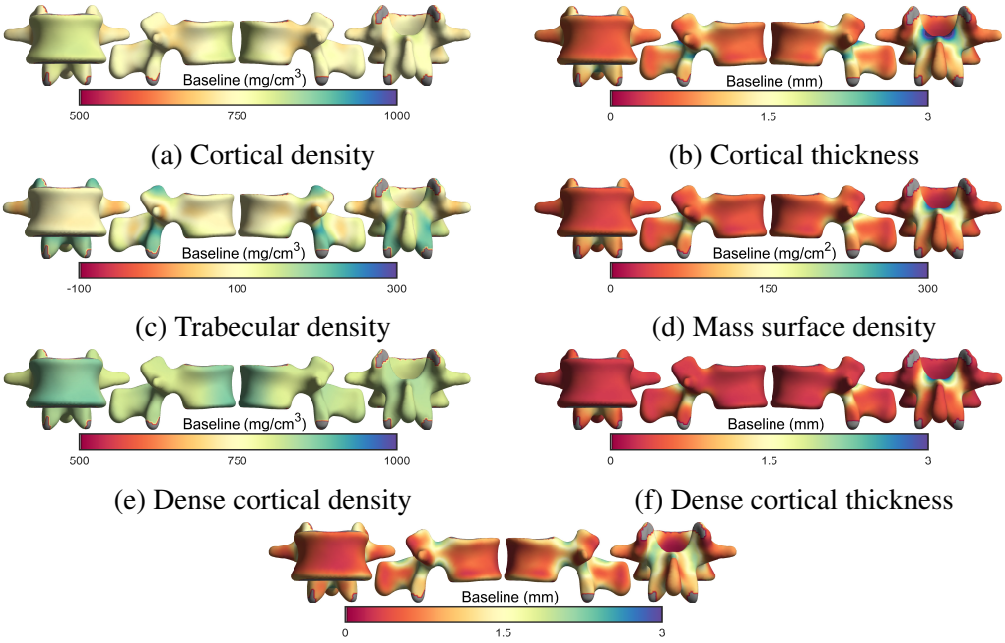
(d) Mass surface density

Fig. 4.7 Surface mapped baseline rectangular CBMv2 measurements over the lumbar spine before teriparatide treatment.

using random field theory. P-values of < 0.05 were used to define regions of significant change for each measurement.

4.3.5 Results

Table 4.2 shows the mean measurements made by each method across the baseline scans in both groups, and Figures 4.7 and 4.8 show the mean baseline measurements by each method across the vertebrae. The mean changes in the bone architecture of each group during the 12 months measured by the two CBM methods are shown Table 4.3. Figures 4.9 and 4.10 show the impact of 12 months of ageing on the bone architecture of the placebo group, while Figures 4.11 and 4.12 show the mean regional change in bone architecture due to teriparatide treatment that are beyond those measured over the placebo group due to ageing. The regions where the architectural changes are significantly different from those in the placebo group are highlighted.



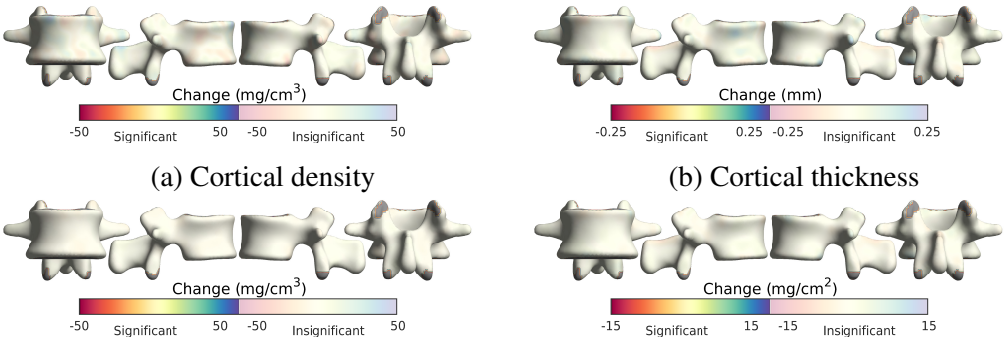
(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Mass surface density

(e) Dense cortical density (f) Dense cortical thickness

(g) Endocortical thickness

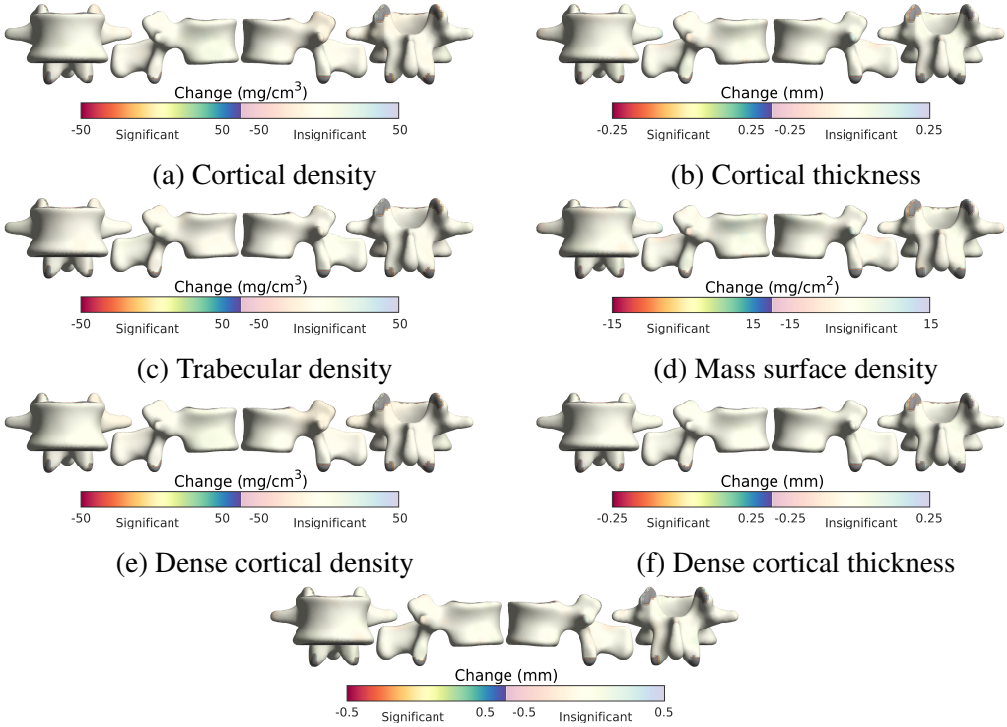
Fig. 4.8 Surface mapped baseline endocortical SSC-CBM measurements over the lumbar spine before teriparatide treatment.



(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Mass surface density

Fig. 4.9 Surface mapped architectural changes in rectangular CBMv2 measurements during one year of natural ageing over the lumbar spine.



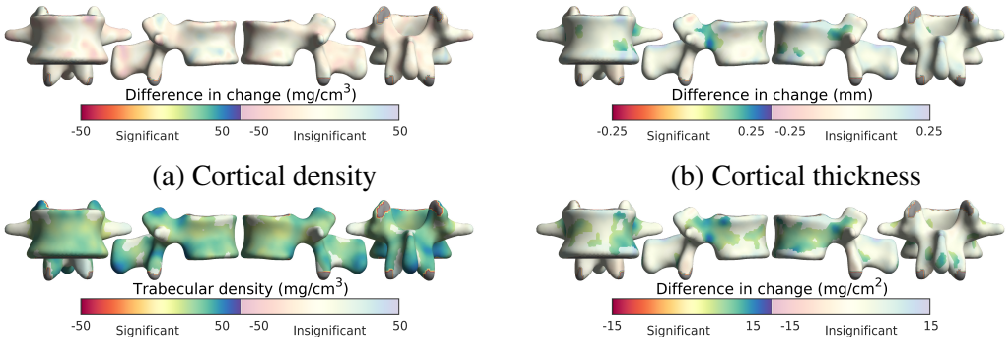
(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Mass surface density

(e) Dense cortical density (f) Dense cortical thickness

(g) Endocortical thickness

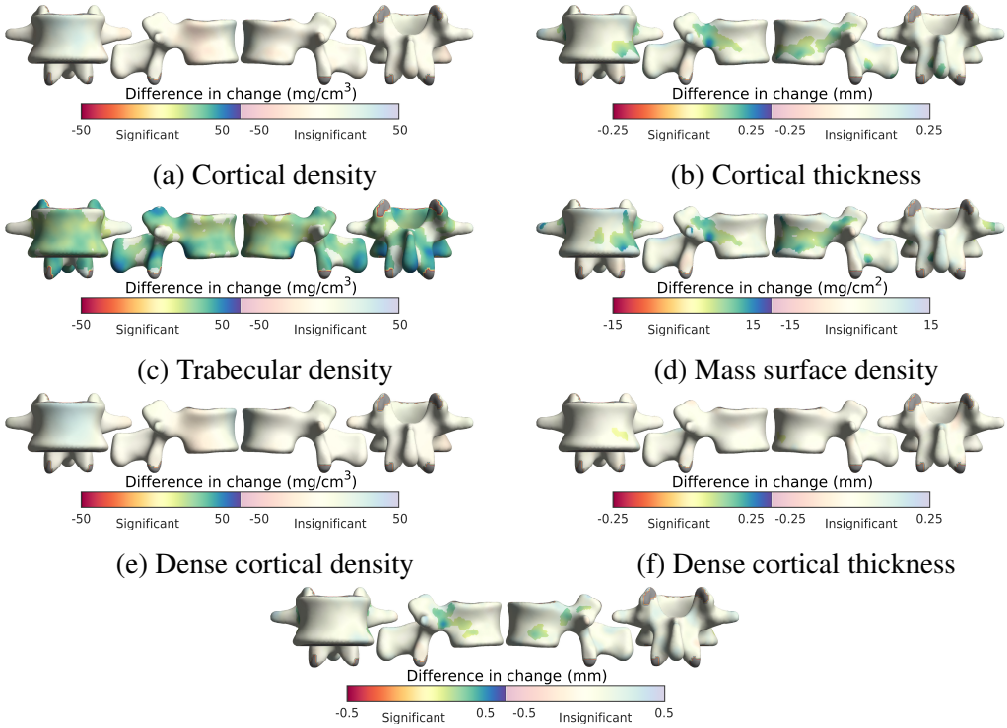
Fig. 4.10 Surface mapped architectural changes in endocortical SSC-CBM measurements during one year of natural ageing over the lumbar spine.



(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Mass surface density

Fig. 4.11 Surface maps showing significant changes in rectangular CBMv2 measurements over the lumbar spine due to teriparatide treatment compared with natural ageing.



(a) Cortical density

(b) Cortical thickness

(c) Trabecular density

(d) Mass surface density

(e) Dense cortical density

(f) Dense cortical thickness

(g) Endocortical thickness

Fig. 4.12 Surface maps showing significant changes in endocortical SSC-CBM measurements over the lumbar spine due to teriparatide treatment compared with natural ageing.

Table 4.3 The mean change from baseline in each treatment group as measured by the rectangular CBMv2 and endocortical SSC-CBM methods. An * is included after measurements that changed significantly during the 12 months of treatment as calculated with a paired t-test at $p < 0.01$. A † is included after teriparatide measurements that changed significantly as compared to the placebo population based on an unpaired t-test at $p < 0.01$. All comparisons were made between measurements by the same method.

	Placebo (n = 20)		Teriparatide (n = 19)	
	Rectangle CBMv2	**Endocortical SSC-CBM**	**Rectangle CBMv2**	**Endocortical SSC-CBM**
cBMD (mg/cm^3)	0±55	$-2^{*}\pm23$	$-8^{*\dagger}\pm43$	$-1^{*\dagger}\pm31$
cTh (mm)	$0.01^{*}\pm0.20$	0.01±0.12	$0.07^{*\dagger}\pm0.17$	$0.06^{*\dagger}\pm0.12$
tBMD (mg/cm^3)	$-2^{*}\pm15$	$-2^{*}\pm15$	$21^{*\dagger}\pm19$	$20^{*\dagger}\pm19$
MSD (mg/cm^2)	0.2±8.2	0.2±8.5	$4.0^{*\dagger}\pm8.4$	$3.8^{*\dagger}\pm7.8$
dBMD (mg/cm^3)	-	$-2^{*}\pm22$	-	$0^{\dagger}\pm32$
dTh (mm)	-	0.00±0.09	-	$0.01^{*\dagger}\pm0.08$
eTh (mm)	-	0.01±0.24	-	$0.10^{*\dagger}\pm0.25$

4.3.6 Discussion

The QCT scans from the placebo and teriparatide subgroups have been analysed twice before [43, 131] to give the mean change in tBMD, cBMD and cortical thickness [43, 131], and localised regionally significant changes in bone architecture as measured with an independent implementation of the rectangular CBMv2 method [131]. The baseline measurements in Table 4.2 and Figures 4.7–4.8 both show broadly similar results with a thin vertebral body and a thicker, denser vertebral ring. The slight differences in thickness compared with previous analysis is likely due to the use of curved profiles instead of straight profiles. Figures 4.9–4.10 and Table 4.3 show the changes in the placebo group due to the 12 months of ageing. Both methods report a significant decrease in the overall tBMD of the vertebrae. The rectangular CBMv2 method also reports a small but significant increase in the overall cortical thickness, while the endocortical SSC-CBM method reports a small but significant decrease in the cBMD and dBMD. The methods report similar insignificant changes in mass surface density, and neither method detects any significant regional changes in bone architecture.

Figures 4.11–4.12 and Table 4.3 show the changes in bone architecture due to 12 months of teriparatide therapy. Significances are reported in comparison with both the baseline measurements and the change in the placebo measurements. Both methods show a clear increase in tBMD, cortical thickness and mass surface density, and the endocortical SSC-CBM method also shows an increase in the endocortical thickness. As discussed in Section 3.6.5, the thin nature of cortical bone over the vertebrae makes it challenging to produce high quality rectangular CBMv2 and endocortical SSC-CBM measurements, so it will be more difficult to detect significant regional changes in the bone architecture. The increase in tBMD is significant over most of the vertebra in both methods, while the increases in cortical thickness, mass surface density and endocortical thickness are concentrated around the vertebral

ring and the lateral sides of the vertebral body. The dense cortical thickness measured by the endocortical SSC-CBM method significantly increases over a single small region in the medial aspect of the vertebral body.

The large increase in endocortical thickness and small increase in dense cortical thickness measured by the endocortical SSC-CBM method during teriparatide treatment suggest an increase in endosteal apposition combined with the transformation of some endocortical bone into dense cortical bone. This leads to an increase in the cortical thickness, which was measured by both methods. Teriparatide treatment is also associated with an increase in MSD measured by both methods, which suggests a positive remodelling balance unlike the proximal femur. This would lead to endocortical and intracortical pore filling, which is consistent with the increase in cortical thickness and MSD measured by both methods, but not the decrease in cBMD. Whitmarsh et al. (2015) argue that endosteal apposition due to pore infill can be associated with a decrease in cBMD due to the lower density of the new bone. The same mechanism could affect dBMD, and may explain the absence of any increase in the dBMD measured by the endocortical SSC-CBM method.

4.3.7 Conclusions

As in the proximal femur, the endocortical SSC-CBM method allows most of the increase in cortical thickness to be attributed to increased endosteal apposition for the first time. However in contrast to the proximal femur, this does not seem to coincide with an increase in cortical porosity. Instead, the increased endocortical apposition coincides with an increase in mass surface density over much of the vertebra. The endocortical SSC-CBM method also measured a slight increase in dense cortical thickness. The increases in endosteal apposition and positive remodelling balance could indicate that teriparatide therapy can lead some endocortical bone to be transformed into dense cortical bone.

4.4 Fracture discrimination

The QCT scans from the fracture and control groups of a cross-sectional fracture discrimination trial were analysed with the endocortical SSC-CBM method to investigate differences in bone architecture between both groups.

4.4.1 Data acquisition

The analysed fracture and control groups were part of the FEMCO study, which was adopted onto the UK National Institute for Health Research (NIHR) portfolio and performed in several centres throughout the UK. All participants provided informed consent prior to participation in the study. Fracture participants had to have sustained a femoral neck or trochanteric fracture for which they were awaiting surgical fixation. Exclusions included cognitive impairment, unconsciousness, terminal illness, metastatic cancer, previous hip replacement, previous hip fracture, osteomyelitis, bone tumour, currently taking oral corticosteroids, prior hemiplegia, and prior treatment with teriparatide or strontium ranelate. The controls were sex

and minimum age matched, but not fracture risk matched. Clinical and CTXA measurements were made over both hips in the control patients and the intact hip for fracture patients before surgical repair. All fractures were classified using the Miller AO criteria by a consultant radiologist as femoral neck or trochanteric fracture types, which were considered together in the present analysis. The study yielded 53 female fracture cases (40 femoral neck, 13 trochanteric), 20 male fracture cases (14 femoral neck, 6 trochanteric), 50 female controls and 24 male controls. Details of the study design and study participants have been published previously [96, 99].

4.4.2 Experimental pipeline

Most scans contained Mindways solid phantoms with the subject, but in some the calibration phantom was scanned separately. Calibration was performed using the approach described in Section 4.1.4. The proximal femur was segmented and registered in each scan using the process described in Section 3.4.2, before the rectangular CBMv2 and endocortical SSC-CBM methods were applied using the pipeline shown in Figure 4.13. The mean architectural features of each group were calculated and compared, before regions with significant architectural differences were identified as discussed in Section 4.4.3. A mask was used to exclude measurements where the acetabulum approaches the femoral head as discussed in Section 4.2.2.

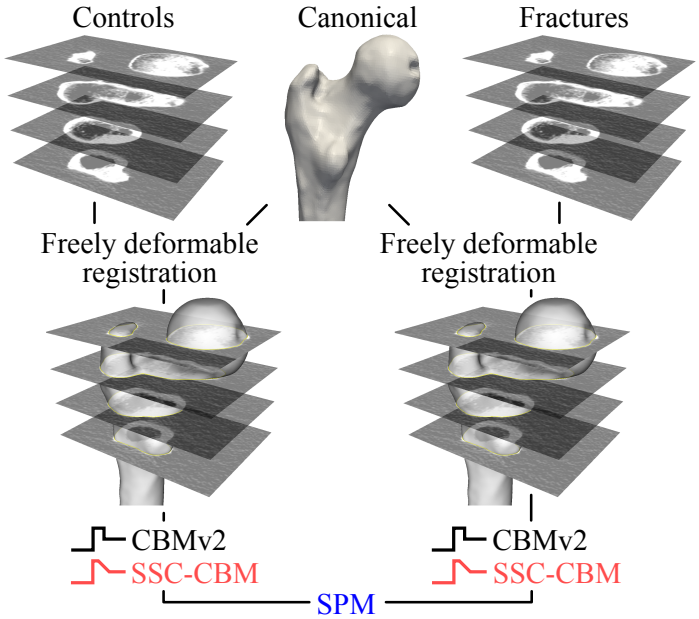


Fig. 4.13 The pipeline used to investigate differences in bone macro-structure between fracture patients and healthy age and sex matched controls.

Table 4.4 The mean population attributes of the fracture and control groups, where an * is included after the fracture group attribute if it is significantly different to the control population. Significance values are calculated using an unpaired t-test $p < 0.01$.

	Controls (n = 74)	Fractures (n = 73)
Age (years)	74.1±9.2	78.3*±9.5
Weight (kg)	71.5±13.6	62.3*±13.1
Height (m)	1.64±0.10	1.64±0.09

4.4.3 Statistical parametric mapping (SPM)

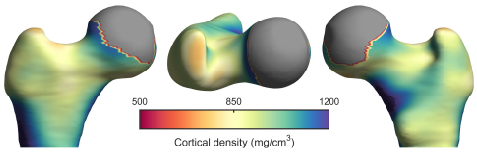
SPM was used to identify regions where the bone architecture significantly differed between the fracture and control groups, while controlling for non fracture related differences between the groups. All measurements were made over the same canonical surface, and spatial correlation between each subject's measurements was achieved with six iterations of mesh connectivity based precision smoothing to give an equivalent filter FWHM of 10 mm. No individuals were excluded. The significance of the differences between the fracture and control groups was assessed using a constant GLM with fixed effect terms for the age, weight, sex, shape modes 2–5 and group, compared with a constant GLM with fixed effect terms for the age, weight and shape modes 2–5. Shape modes were included to protect against false inferences arising from systematic misregistration errors [42]. The shape mode 1 and height were not included in the model to avoid correlated regressors and the danger of missing significant effects as these covariates were correlated with group [96, 136]. The GLMs were used to give F-statistics, from which the p-values across the surface were determined using random field theory. P-values of < 0.05 were used to define regions of significant change for each measurement.

4.4.4 Results

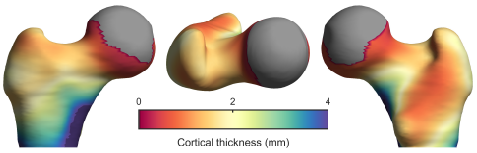
Table 4.4 shows the mean height, weight and age of each group. Table 4.5 shows the mean measurements made by each method across the two groups, and the difference between the two groups. Figures 4.14 and 4.15 show the mean measurements by each CBM method across the control group over the proximal femur. The regional variations in the mean differences in the bone architecture between the fracture and control groups are shown in Figures 4.16 and 4.17. Regions with significantly different bone architecture are highlighted.

4.4.5 Discussion

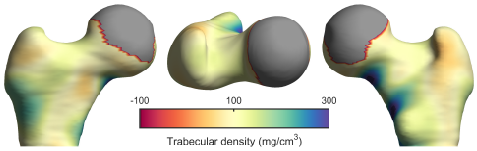
The QCT scans from the FEMCO study have been analysed once before using an independent implementation of the rectangular CBMv2 method [96], with the publication reporting the regions where the tBMD and mass surface density significantly differed between the fracture and control groups. The rectangular CBMv2 measurements are reported in Figures 4.14



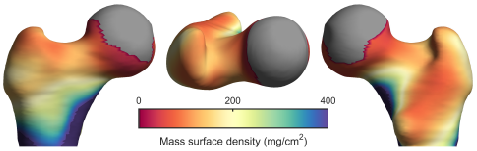
(a) Cortical density



(b) Cortical thickness

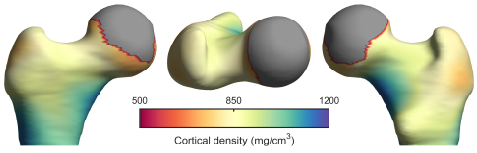


(c) Trabecular density

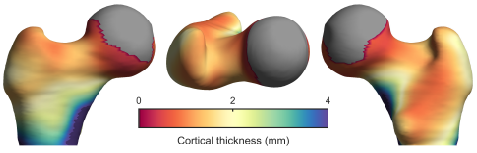


(d) Mass surface density

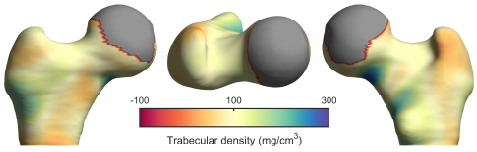
Fig. 4.14 Surface maps of the proximal femur showing rectangular CBMv2 measurements of the control group.



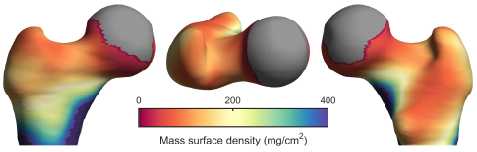
(a) Cortical density



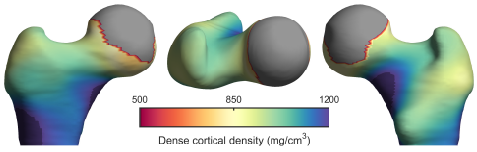
(b) Cortical thickness



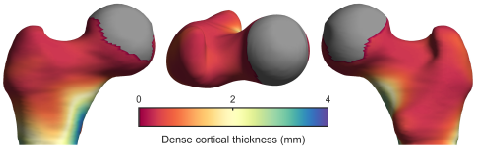
(c) Trabecular density



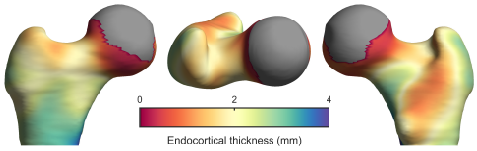
(d) Mass surface density



(e) Dense cortical density

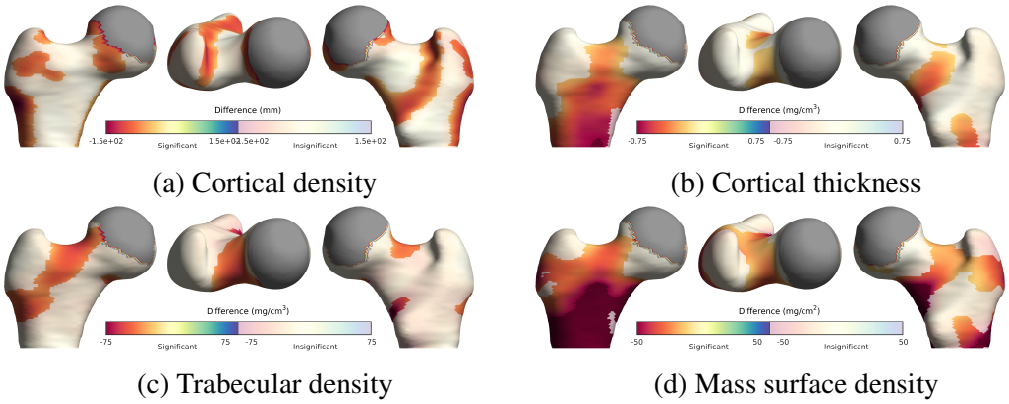


(f) Dense cortical thickness



(g) Endocortical thickness

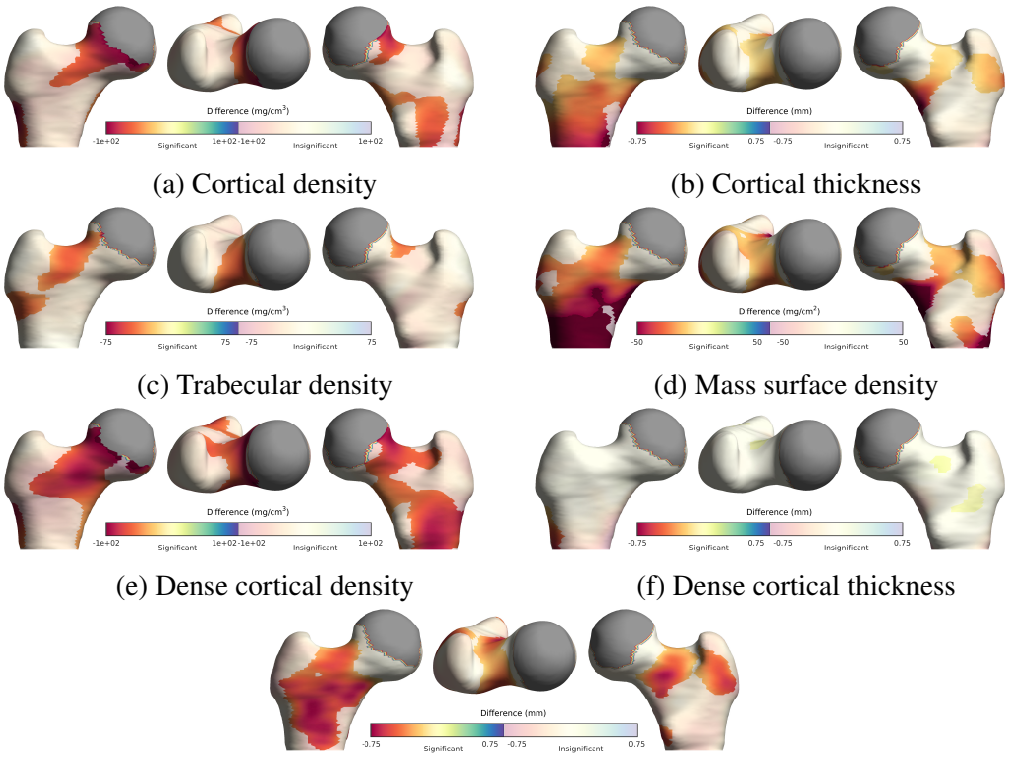
Fig. 4.15 Surface maps of the proximal femur showing endocortical SSC-CBM measurements of the control group.



(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Mass surface density

Fig. 4.16 The areas over the proximal femur where fracture incidence is associated with significant changes in the rectangular CBMv2 measurements.



(a) Cortical density (b) Cortical thickness

(c) Trabecular density (d) Mass surface density

(e) Dense cortical density (f) Dense cortical thickness

(g) Endocortical thickness

Fig. 4.17 The areas over the proximal femur where fracture incidence is associated with significant changes in the endocortical SSC-CBM measurements.

Table 4.5 The rectangular CBMv2 and endocortical SSC-CBM bone architecture measurements of both groups and the difference between them. An * is included after all difference measurements where the fracture and control measurements differ significantly. Significance values are calculated using an unpaired at t-test at $p < 0.01$.

	Controls (n = 74)		Fractures (n = 73)		Difference	
	Rectangle CBMv2	**Endocortical SSC-CBM**	**Rectangle CBMv2**	**Endocortical SSC-CBM**	**Rectangle CBMv2**	**Endocortical SSC-CBM**
cBMD (mg/cm^3)	996±138	911±137	938±143	864±145	−57*	−47*
cTh (mm)	1.83±1.14	1.75±1.14	1.57±0.99	1.49±0.97	−0.25*	−0.26*
tBMD (mg/cm^3)	126±80	102±71	99±72	79±64	−28*	−23*
MSD (mg/cm^2)	180.7±132.6	167.8±131.0	147.0±111.3	136.3±108.5	−33.6*	−31.5*
dBMD (mg/cm^3)	-	1040±143	-	981±156	-	−59*
dTh (mm)	-	0.78±0.83	-	0.69±0.70	-	−0.09*
eTh (mm)	-	1.92±1.03	-	1.59±0.93	-	−0.33*

and 4.16 and Table 4.5. The tBMD and mass surface density regions of significance broadly agree with the previously published results, although the extents of the significant regions are slightly greater in this analysis [96]. The endocortical SSC-CBM measurements of the control group are shown in Figure 4.15. These are broadly similar to the rectangular CBMv2 measurements, although the cBMD and tBMD measurements do differ visibly in some regions. The large cBMD values measured by the rectangular CBMv2 method at the superior femoral head neck junction are believed to be artefacts resulting from the method's low density failure mode that has been discussed in Sections 2.2.4 and 3.6.3, while the larger tBMD values measured using the rectangular CBMv2 method were largely attributed to beam hardening errors as discussed in Section 2.6.

The architectural changes associated with elevated fracture risk are shown in Figures 4.16–4.17. These broadly agree between the two methods, although the extent of the statistically significant regions differ for the cBMD and cortical thickness measurements. The surface error plots from the ex-vivo validation of the methods over the proximal femur in Section 3.6.2 show that the major differences in the extents occur where the endocortical SSC-CBM cBMD measure is substantially more accurate. This suggests that the endocortical SSC-CBM extents will better reflect the underlying changes in bone architecture associated with elevated fracture risk. The endocortical SSC-CBM method also shows that dBMD and endocortical thickness are significantly lower in the fracture group over much of the femoral neck and anterior femoral shaft. This indicates that it may be viable to include endocortical thickness and dBMD in future fracture prediction models.

The independent consideration of the cortical, dense cortical and endocortical regions by the endocortical SSC-CBM method allows for the comparison of equivalent measurements in each region. The distribution of density changes within the cortical region has not been considered before. The dBMD measured by the endocortical SSC-CBM method significantly decreases over a greater region than the accompanying cBMD. The measurements indicate that the architectural changes causing the decrease in density are concentrated in the dense

cortical bone. In contrast, a comparison of the significant changes in dense cortical thickness, cortical thickness and endocortical thickness show that the significant decreases in cortical thickness are almost exclusively due to significant decreases in the endocortical thickness with very little caused by localised decreases in the dense cortical thickness.

4.4.6 Conclusions

The endocortical SSC-CBM method allows the endocortical thickness and dBMD in individuals with elevated fracture risk to be considered over the proximal femur for the first time. This shows that both are significantly lower than in healthy age and sex matched controls over much of the femoral neck and anterior femoral shaft. This gives added insight into the architectural changes associated with increased fracture risk, and provides new measures that can be incorporated into fracture prediction models.

4.5 Conclusions

The endocortical SSC-CBM method allows for a more detailed consideration of the changes in cortical architecture in response to teriparatide therapy. It specifically allows for the changes in intracortical remodelling and endosteal apposition to be considered for the first time from QCT scans. It should be possible to use the same method to measure the response of bone architecture to other pharmaceutical therapies or exercise routines. The endocortical SSC-CBM method can also be applied to fracture discrimination studies to identity regions of significantly decreased dBMD and endocortical thickness in individuals with an elevated fracture risk. It will be possible to include the dBMD and endocortical thickness measurements over these regions in future fracture prediction models.

Chapter 5

Conclusions

5.1 Introduction

The endocortical SSC-CBM method provides a model of the cortical and subsurface trabecular region from which the cortical thickness, cBMD, tBMD and endocortical width can be measured at many locations. Its ability to measure the thickness of the endocortical region and to more accurately measure the cBMD distinguishes it from previously published QCT based techniques. It also provides more accurate tBMD and mass surface density measurements, while maintaining a similar standard of cortical thickness measurement.

In order to validate this technique, I have also developed a novel HR-pQCT technique which allows for the localised measurement of the endocortical thickness. This was used to validate the endocortical SSC-CBM method over the proximal femur and lumbar spine. Once validated, the endocortical SSC-CBM method was successfully applied to several clinical datasets to examine its ability to detect changes in bone architecture in response to treatment therapies, and to detect differences in bone architecture associated with elevated fracture risk.

The major contributions of this thesis are:

- The localised measurement of endocortical thickness and improved localised measurement of cBMD from QCT scans with an examination of its resilience to the effects of beam hardening, and ex-vivo validation over the proximal femur and lumbar spine.
- The development of an HR-pQCT method capable of producing localised endocortical thickness, cBMD, tBMD, cortical thickness and mass surface density measurements independently of the micro-architectural detail included in HR-pQCT scans.
- The confirmation of a significant increase in intracortical remodelling over the proximal femur, and endosteal deposition over the proximal femur and lumbar spine in response to teriparatide therapy, and the identification of regions of significantly decreased cBMD and endocortical thickness over the proximal femur in individuals with elevated fracture risk.

5.2 Beam hardening

Three simulated QCT phantoms with differing bone architecture were used to assess the performance of the endocortical SSC-CBM method. This allowed for a detailed study into the effect of beam hardening on CBM performance for the first time. This showed that increased levels of beam hardening causes CBM cortical thickness and tBMD measurements to be increasingly under and overestimated. In Section 2.6, I demonstrate that the effect of beam hardening on these measurements can be reduced through the use of an optimiser weighting function that reflects the local cortical thickness. Significantly, this can be applied to existing optimiser-based methods including the rectangular CBMv2 and FWHM method to improve their accuracy during high levels of beam hardening.

5.3 High-resolution tissue-classification (HRTC)

A new technique for analysing HR-pQCT scans was developed for this thesis: high resolution tissue classification (HRTC). HRTC acts to extract the macro-architectural cortical, endocortical, and sub-surface trabecular bone detail from the micro-architectural detail included in the HR-pQCT scan. It was used in place of FWHM and other established HR-pQCT techniques for its ability to provide localised measurements of the endocortical region. It further distinguishes itself from the FWHM method for its ability to measure the mean cBMD and its greater resistance to beam hardening effects. Notably, the mean cBMD has a greater clinical relevance as it includes the impact of porosity, which is not included in the peak cBMD measured by the FWHM method.

5.4 Cadaveric validation

5.4.1 Proximal femur

A large ex-vivo dataset of paired QCT and HR-pQCT scans allowed for the validation of the endocortical SSC-CBM method over the proximal femur. This showed it is able to measure the endocortical thickness to (-0.15 ± 0.71) mm and cBMD down to densities of $300\,\mathrm{mg/cm^3}$, while still measuring the cortical thickness, tBMD and mass surface density. The extra information included in the endocortical model makes it more difficult to accurately fit it to a sampled density profile. Despite this, the endocortical SSC-CBM method provides measurements that have a similar level of precision to the rectangular CBMv2 method. This attests to the improved stability afforded by modelling the dBMD as a smoothly varying property of cortical bone.

5.4.2 Lumbar spine

An ex-vivo dataset of paired QCT and HR-pQCT scans was used to validate the rectangular CBMv2 and endocortical SSC-CBM method over the lumbar spine, which marked the

first time any CBM method has been validated over this clinically important region. The validation dataset was challenging with differences in material composition between the QCT and HR-pQCT making it difficult to provide an accurate density calibration relating the HR-pQCT and QCT densities. In addition, the dataset included only 20 vertebral bodies from 6 individuals with a mean age of 71. Finally, the architecture of the lumbar spine is challenging. Its thin cortex makes it difficult to provide high quality estimates of the dBMD. Despite this, both methods were able to provide reasonable CBM measurements although generally at a lower accuracy than is possible over the proximal femur. Importantly, the endocortical SSC-CBM method was still able to measure the endocortical region with an accuracy (-0.20 ± 0.53) mm.

5.5 Treatment effectiveness

Two clinical trials involving teriparatide were used to examine the ability of the endocortical SSC-CBM method to detect changes in bone architecture over the proximal femur and lumbar spine.

5.5.1 Proximal femur

In the analysis of the first trial, the endocortical SSC-CBM method was used to detect changes in the bone architecture over the proximal femur in response to teriparatide therapy. The measurement of the endocortical thickness, dense cortical thickness and dBMD allowed the impact of teriparatide therapy on intracortical remodelling and endosteal apposition to be considered for the first time over the proximal femur. The measurements showed an increase in both. This indicates that the commonly observed decrease in cBMD over the proximal femur is largely due to an increase in intracortical remodelling. This is significant as previous QCT-based analyses of teriparatide have only been able to speculated about the cause of this decrease in cBMD over the proximal femur.

5.5.2 Lumbar spine

The second clinical trial was used to examine the ability of the endocortical SSC-CBM method to investigate changes in bone architecture over the lumbar spine. As noted in Section 5.4.2, the architecture of the vertebrae makes it difficult to provide high quality dBMD estimates, which impacts the quality of the endocortical SSC-CBM measurements. This has a larger impact on the endocortical SSC-CBM method than the rectangular CBMv2 method given the challenges associated with the more complex endocortical model discussed in Section 5.4.1. Despite this, the endocortical SSC-CBM is able to provide similar measurements of significant regional changes in cortical thickness, tBMD and mass surface density to the rectangular CBMv2 method. In addition, it is also able to provide high enough quality endocortical thickness measurements to show that it increases significantly over a similar but slightly smaller region than the cortical thickness.

5.6 Fracture discrimination

The third clinical trial was a cross-sectional fracture discrimination study of individuals who had recently suffered a hip fracture, and healthy sex and minimum age matched controls. This allowed for the differences in the dBMD, dense cortical thickness and endocortical thickness associated with elevated fracture risk to be considered for the first time, which showed that the dBMD and endocortical thickness, in particular, vary significantly over clearly defined regions of the proximal femur. This is important as it clearly indicates that changes in the dBMD and endocortical thickness of the proximal femur are associated with increased fracture risk, and it shows that the endocortical SSC-CBM method can be used to define dBMD and endocortical thickness measurements that can be incorporated into fracture prediction models.

5.7 Discussion and conclusions

The aim of my PhD was to improve on existing methods for clinical osteoporosis research with fracture risk prediction, treatment assessment, and the monitoring of osteoporosis disease progression all areas of research interest. The endocortical SSC-CBM method developed in this thesis allows for the separate consideration of the cortical, dense cortical and endocortical regions from clinical QCT. This is powerful as it allows for clinically measured differences in cortical thickness and cBMD measurements to be attributed to differences in intracortical remodelling and endosteal apposition. This allows for changes in intracortical remodelling and endosteal apposition to be considered over the lumbar spine and proximal femur for the first time.

The endocortical SSC-CBM technique was validated over the proximal femur and lumbar spine, before it was used to perform an initial analysis of several clinical datasets. The validation showed that it provides improved cBMD measurements compared to other existing QCT techniques, and can measure the thickness of the endocortical region for the first time from QCT scans. The analysis of the clinical datasets indicated that the endocortical SSC-CBM method: can be an effective tool for assessing the changes in bone architecture associated with treatment therapies; and, is able to measure new regions of significantly different endocortical and cBMD values between fracture and control groups, which can be added to future fracture prediction models.

References

[1] Ahmed, L. A., Schirmer, H., Bjørnerem, Å., Emaus, N., Jørgensen, L., Størmer, J., and Joakimsen, R. (2009). The gender- and age-specific 10-year and lifetime absolute fracture risk in Tromsø, Norway. *European Journal of Epidemiology*, 24(8):441–448.

[2] Ammann, P. and Rizzoli, R. (2003). Bone strength and its determinants. *Osteoporosis International*, 14(3):13–18.

[3] Augat, P. and Schorlemmer, S. (2006). The role of cortical bone and its microstructure in bone strength. *Age and Ageing*, 35(Supplement 2):ii27–ii31.

[4] Bala, Y., Farlay, D., Delmas, P. D., Meunier, P. J., and Boivin, G. (2010). Time sequence of secondary mineralization and microhardness in cortical and cancellous bone from ewes. *Bone*, 46(4):1204 – 1212.

[5] Bala, Y., Zebaze, R., Ghasem-Zadeh, A., Atkinson, E. J., Iuliano, S., Peterson, J. M., Amin, S., Bjørnerem, Å., Melton, L. J., Johansson, H., Kanis, J. A., Khosla, S., and Seeman, E. (2014). Cortical porosity identifies women with osteopenia at increased risk for forearm fractures. *Journal of Bone and Mineral Research*, 29(6):1356–1362.

[6] Baxter, B. S. and Sorenson, J. A. (1981). Factors affecting the measurement of size and CT number in computed tomography. *Investigative Radiology*, 16:337 – 341.

[7] Benson, H. (1996). *University Physics*. New York : Wiley, revised (2nd) edition.

[8] Boden, S. D., Goodenough, D. J., Stockham, C. D., Jacobs, E., Dina, T., and Allman, R. M. (1989). Precise measurement of vertebral bone density using computed tomography without the use of an external reference phantom. *Journal of Digital Imaging*, 2(1):31–38.

[9] Boivin, G., Bala, Y., Doublier, A., Farlay, D., Ste-Marie, L., Meunier, P., and Delmas, P. (2008). The role of mineralization and organic matrix in the microhardness of bone tissue from controls and osteoporotic patients. *Bone*, 43(3):532 – 538.

[10] Boivin, G., Chavassieux, P., Santora, A., Yates, J., and Meunier, P. (2000). Alendronate increases bone strength by increasing the mean degree of mineralization of bone tissue in osteoporotic women. *Bone*, 27(5):687 – 694.

[11] Borggrefe, J., Graeff, C., Nickelsen, T. N., Marin, F., and Glüer, C. C. (2010). Quantitative computed tomographic assessment of the effects of 24 months of teriparatide treatment on 3D femoral neck bone distribution, geometry, and bone strength: Results from the EUROFORS study. *Journal of Bone and Mineral Research*, 25(3):472–481.

[12] Boutroy, S., Bouxsein, M. L., Munoz, F., and Delmas, P. D. (2005). In vivo assessment of trabecular bone microarchitecture by high-resolution peripheral quantitative computed tomography. *The Journal of Clinical Endocrinology & Metabolism*, 90(12):6508–6515.

[13] Bouxsein, M. and Karasik, D. (2006). Bone geometry and skeletal fragility. *Current Osteoporosis Reports*, 4(2):49–56.

[14] Brandi, M. L. (2009). Microarchitecture, the key to bone quality. *Rheumatology*, 48(Supplement 4):iv3–iv8.

[15] Brown, J. (1966). *X-rays and their applications*. London : Iliffe Books.

[16] Buie, H. R., Campbell, G. M., Klinck, R. J., MacNeil, J. A., and Boyd, S. K. (2007). Automatic segmentation of cortical and trabecular compartments based on a dual threshold technique for in vivo micro-CT bone analysis. *Bone*, 41(4):505 – 515.

[17] Burghardt, A. J., Buie, H. R., Laib, A., Majumdar, S., and Boyd, S. K. (2010a). Reproducibility of direct quantitative measures of cortical bone microarchitecture of the distal radius and tibia by HR-pQCT. *Bone*, 47:519–528.

[18] Burghardt, A. J., Kazakia, G. J., Ramachandran, S., Link, T. M., and Majumdar, S. (2010b). Age- and gender-related differences in the geometric properties and biomechanical significance of intracortical porosity in the distal radius and tibia. *Journal of Bone and Mineral Research*, 25(5):983–993.

[19] Burghardt, A. J., Kazakia, G. J., Sode, M., de Papp, A. E., Link, T. M., and Majumdar, S. (2010c). A longitudinal HR-pQCT study of alendronate treatment in postmenopausal women with low bone density: Relations among density, cortical and trabecular microarchitecture, biomechanics, and bone turnover. *Journal of Bone and Mineral Research*, 25:2558–71.

[20] Burr, D. B., Hirano, T., Turner, C. H., Hotchkiss, C., Brommage, R., and Hock, J. M. (2001). Intermittently administered human parathyroid hormone (1–34) treatment increases intracortical bone turnover and porosity without reducing bone strength in the humerus of ovariectomized cynomolgus monkeys. *Journal of Bone and Mineral Research*, 16(1):157–165.

[21] Cann, C. E., Adams, J. E., Brown, J. K., and Brett, A. D. (2014). CTXA hip - an extension of classical DXA measurements using quantitative CT. *PLOS ONE*, 9(3):1–9.

[22] Carter, D. R. (2009). *Skeletal function and form : mechanobiology of skeletal development, aging, and regeneration*. Cambridge : Cambridge University Press.

[23] Cauley, J. A. (2013). Public health impact of osteoporosis. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*.

[24] Chappard, D., Baslé, M., Legrand, E., and Audran, M. (2011). New laboratory tools in the assessment of bone quality. *Osteoporosis International*, 22(8):2225–2240.

[25] Chen, P., Miller, P. D., Delmas, P. D., Misurski, D. A., and Krege, J. H. (2006). Change in lumbar spine BMD and vertebral fracture risk reduction in teriparatide-treated postmenopausal women with osteoporosis. *Journal of Bone and Mineral Research*, 21(11):1785–1790.

[26] Cheung, A. M., Adachi, J. D., Hanley, D. A., Kendler, D. L., Davison, K. S., Josse, R., Brown, J. P., Ste-Marie, L.-G., Kremer, R., Erlandson, M. C., Dian, L., Burghardt, A. J., and Boyd, S. K. (2013). High-resolution peripheral quantitative computed tomography for the assessment of bone strength and structure: A review by the Canadian bone strength working group. *Current Osteoporosis Reports*, 11(2):136–146.

[27] Consensus development conference (1993). Consensus development conference: Diagnosis, prophylaxis, and treatment of osteoporosis. *The American Journal of Medicine*, 94(6):646 – 650.

[28] Cranney, A., Papaioannou, A., Zytaruk, N., Hanley, D., Adachi, J., Goltzman, D., Murray, T., and Hodsman, A. (2006). Parathyroid hormone for the treatment of osteoporosis: a systematic review. *Canadian Medical Association Journal*, 175:52–9.

[29] Crum, W. R., Hartkens, T., and Hill, D. L. G. (2004). Non-rigid image registration: theory and practice. *The British Journal of Radiology*, 77(Supplement 2):S140–S153.

[30] Dall'Ara, E., Luisier, B., Schmidt, R., Kainberger, F., Zysset, P., and Pahr, D. (2013a). A nonlinear QCT-based finite element model validation study for the human femur tested in two configurations in vitro. *Bone*, 52(1):27 – 38.

[31] Dall'Ara, E., Luisier, B., Schmidt, R., Pretterklieber, M., Kainberger, F., Zysset, P., and Pahr, D. (2013b). DXA predictions of human femoral mechanical properties depend on the load configuration. *Medical Engineering & Physics*, 35(11):1564 – 1572.

[32] Donnelly, E. (2011). Methods for assessing bone quality: A review. *Clinical Orthopaedics and Related Research*, 469:2128–2138.

[33] Dougherty, G. and Newman, D. (1999). Measurement of thickness and density of thin structures by computed tomography: A simulation study. *Medical Physics*, 26(7):1341–1348.

[34] Eastell, R., Nickelsen, T., Marin, F., Barker, C., Hadji, P., Farrerons, J., Audran, M., Boonen, S., Brixen, K., Gomes, J. M., Obermayer-Pietsch, B., Avramidis, A., Sigurdsson, G., and Glüer, C. C. (2009). Sequential treatment of severe postmenopausal osteoporosis after teriparatide: Final results of the randomized, controlled European study of forsteo (EUROFORS). *Journal of Bone and Mineral Research*, 24(4):726–736.

[35] Edwards, T., Zheng, Y., Ferrara, L. A., and Yuan, H. A. (2001). Structural features and thickness of the vertebral cortex in the thoracolumbar spine. *Spine*, 26(2):218–225.

[36] Eriksen, E. F., Hodgson, S. F., Eastell, R., RIGGS, B. L., Cedel, S. L., and O'Fallon, W. M. (1990). Cancellous bone remodeling in type I (postmenopausal) osteoporosis: Quantitative assessment of rates of formation, resorption, and bone loss at tissue and cellular levels. *Journal of Bone and Mineral Research*, 5(4):311–319.

[37] Eswaran, S. K., Gupta, A., Adams, M. F., and Keaveny, T. M. (2006). Cortical and trabecular load sharing in the human vertebral body. *Journal of Bone and Mineral Research*, 21(2):307–314.

[38] Feldmar, J. and Ayache, N. (1996). Rigid, affine and locally affine registration of free-form surfaces. *International Journal of Computer Vision*, 18(2):99–119.

[39] Follet, H., Boivin, G., Rumelhart, C., and Meunier, P. (2004). The degree of mineralization is a determinant of bone strength: a study on human calcanei. *Bone*, 34(5):783 – 789.

[40] Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4):189–210.

[41] Garnero, P., Sornay-Rendu, E., Chapuy, M.-C., and Delmas, P. D. (1996). Increased bone turnover in late postmenopausal women is a major determinant of osteoporosis. *Journal of Bone and Mineral Research*, 11(3):337–349.

[42] Gee, A. H. and Treece, G. M. (2014). Systematic misregistration and the statistical analysis of surface data. *Medical Image Analysis*, 18(2):385 – 393.

[43] Genant, H. K., Engelke, K., Bolognese, M. A., Mautalen, C., Brown, J. P., Recknor, C., Goemaere, S., Fuerst, T., Yang, Y.-C., Grauer, A., and Libanati, C. (2017). Effects of romosozumab compared with teriparatide on bone density and mass at the spine and hip in postmenopausal women with low bone mass. *Journal of Bone and Mineral Research*, 32(1):181–187.

[44] Genant, H. K., Engelke, K., Fuerst, T., Glüer, C.-C., Grampp, S., Harris, S. T., Jergas, M., Lang, T., Lu, Y., Majumdar, S., Mathur, A., and Takada, M. (1996). Noninvasive assessment of bone mineral and structure: State of the art. *J Bone Miner Res*, 11(6):707–730.

[45] Genant, H. K., Engelke, K., and Prevrhal, S. (2008). Advanced CT bone imaging in osteoporosis. *Rheumatology*, 47.

[46] Graeff, C., Chevalier, Y., Charlebois, M., Varga, P., Pahr, D., Nickelsen, T. N., Morlock, M. M., Glüer, C. C., and Zysset, P. K. (2009). Improvements in vertebral body strength under teriparatide treatment assessed in vivo by finite element analysis: Results from the EUROFORS study. *Journal of Bone and Mineral Research*, 24(10):1672–1680.

[47] Graeff, C., Timm, W., Nickelsen, T. N., Farrerons, J., Marín, F., Barker, C., and Güer, C. C. (2007). Monitoring teriparatide-associated changes in vertebral microstructure by high-resolution CT in vivo: Results from the EUROFORS study. *Journal of Bone and Mineral Research*, 22(9):1426–1433.

[48] Griffith, J. F. and Genant, H. K. (2012). New advances in imaging osteoporosis and its complications. *Endocrine*, 42(1):39–51.

[49] Hangartner., T. N. (2007). Thresholding technique for accurate analysis of density and geometry in QCT, pQCT and μCT images. *Journal of Musculoskeletal and Neuronal Interactions*, 7(1):9–16.

[50] Hangartner, T. N. and Gilsanz, V. (1996). Evaluation of cortical bone by computed tomography. *Journal of Bone and Mineral Research*, 11(10):1518–1525.

[51] Hangartner, T. N. and Short, D. F. (2007). Accurate quantification of width and density of bone structures by computed tomography. *Medical Physics*, 34(10):3777–3784.

[52] Harvey, N., Dennison, E., and Cooper, C. (2010). Osteoporosis: impact on health and economics. *Nature Reviews Rheumatology*, pages 99–105.

[53] Herman, G. T. (2009). *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Springer Publishing Company, Incorporated, 2nd edition.

[54] Hildebrand, F. B. (1956). *Introduction to numerical analysis*. New York ; London : McGraw-Hil.

[55] Hildebrand, T. and Rüegsegger, P. (1997). A new method for the model-independent assessment of thickness in three-dimensional images. *Journal of Microscopy*, 185(1):67–75.

[56] Hirano, T., Burr, D. B., Turner, C. H., Sato, M., Cain, R. L., and Hock, J. M. (1999). Anabolic effects of human biosynthetic parathyroid hormone fragment (1–34), LY333334, on remodeling and mechanical properties of cortical bone in rabbits. *Journal of Bone and Mineral Research*, 14(4):536–545.

[57] Hubbell, J. and Seltzer, S. (1996). Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients from 1 keV to 20 MeV for elements Z = 1 to 92 and 48 additional substances of dosimetric interest.

[58] Ito, M., Ikeda, K., Nishiguchi, M., Shindo, H., Uetani, M., Hosoi, T., and Orimo, H. (2005). Multi-detector row CT imaging of vertebral microstructure for evaluation of fracture risk. *Journal of Bone and Mineral Research*, 20(10):1828–1836.

[59] Ito, M., Oishi, R., Fukunaga, M., Sone, T., Sugimoto, T., Shiraki, M., Nishizawa, Y., and Nakamura, T. (2014). The effects of once-weekly teriparatide on hip structure and biomechanical properties assessed by CT. *Osteoporosis International*, 25(3):1163–1172.

[60] Johannesdottir, F., Turmezei, T., and Poole, K. E. (2014). Cortical bone assessed with clinical computed tomography at the proximal femur. *Journal of Bone and Mineral Research*, 29(4):771–783.

[61] Kanis, J., McCloskey, E., Johansson, H., Cooper, C., Rizzoli, R., and Reginster, J.-Y. (2013). European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporosis International*, 24(1):23–57.

[62] Kanis, J. and WHO. (1994). Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: Synopsis of a WHO report. *Osteoporosis International*, 4(6):368–381.

[63] Kanis, J. A., Johnell, O., Oden, A., Dawson, A., De Laet, C., and Jonsson, B. (2001). Ten year probabilities of osteoporotic fractures according to BMD and diagnostic thresholds. *Osteoporosis International*, 12(12):989–995.

[64] Kanis, J. A., Johnell, O., Oden, A., Sernbo, I., Redlund-Johnell, I., Dawson, A., De Laet, C., and Jonsson, B. (2000). Long-term risk of osteoporotic fracture in Malmö. *Osteoporosis International*, 11(8):669–674.

[65] Khoo, B. C. C., Brown, K., Cann, C., Zhu, K., Henzell, S., Low, V., Gustafsson, S., Price, R. I., and Prince, R. L. (2009). Comparison of QCT-derived and DXA-derived areal bone mineral density and T scores. *Osteoporosis International*, 20(9):1539–1545.

[66] Laib, A., Häuselmann, H., and Rüegsegger, P. (1998). In vivo high resolution 3D-QCT of the human forearm. *Technology and health care : official journal of the European Society for Engineering and Medicine*, 6(5-6):329—337.

[67] Laib, A., Hildebrand, T., Häuselmann, H., and Rüegsegger, P. (1997). Ridge number density: A new parameter for in vivo bone structure analysis. *Bone*, 21(6):541 – 546.

[68] Lala, D., Cheung, A. M., Gordon, C., and Giangregorio, L. (2012). Comparison of cortical bone measurements between pQCT and HR-pQCT. *Journal of Clinical Densitometry*, 15(3):275 – 281.

[69] Lee, J. K. T. (1998). *Computed body tomography with MRI correlation*. Philadelphia : Lippincott-Raven.

[70] Link, T. (2010). The founder's lecture 2009: advances in imaging of osteoporosis and osteoarthritis. *Skeletal Radiology*, 39(10):943–955.

[71] Lippuner, K., Johansson, H., Kanis, J., and Rizzoli, R. (2009). Remaining lifetime and absolute 10-year probabilities of osteoporotic fracture in Swiss men and women. *Osteoporosis International*, 20(7):1131–1140.

[72] Liu, X. S., Cohen, A., Shane, E., Stein, E., Rogers, H., Kokolus, S. L., Yin, P. T., McMahon, D. J., Lappe, J. M., Recker, R. R., and Guo, X. E. (2010a). Individual trabeculae segmentation (ITS)-based morphological analysis of high-resolution peripheral quantitative computed tomography images detects abnormal trabecular plate and rod microarchitecture in premenopausal women with idiopathic osteoporosis. *Journal of Bone and Mineral Research*, 25(7):1496–1505.

[73] Liu, X. S., Zhang, X. H., Sekhon, K. K., Adams, M. F., McMahon, D. J., Bilezikian, J. P., Shane, E., and Guo, X. E. (2010b). High-resolution peripheral quantitative computed tomography can assess microstructural and mechanical properties of human distal tibial bone. *Journal of Bone and Mineral Research*, 25(4):746–756.

[74] Lonsdale, K. (1948). *Crystals and X-rays*. London : G. Bell and Sons.

[75] Louis, O., Willnecker, J., Soykens, S., Van den Winke, P., and Osteaux, M. (1995). Cortical thickness assessed by peripheral quantitative computed tomography: Accuracy evaluated on radius specimens. *Osteoporosis International*, 5(6):446–449.

[76] Ma, P. X. (2014). *Biomaterials and regenerative medicine*. Cambridge : Cambridge University Press.

[77] Macdonald, H., Nishiyama, K., Hanley, D., and Boyd, S. (2011). Changes in trabecular and cortical bone microarchitecture at peripheral sites associated with 18 months of teriparatide therapy in postmenopausal women with osteoporosis. *Osteoporosis International*, 22(1):357–362.

[78] Magnusson, A. (1987). Object size determination at computed tomography. *Uppsala Journal of Medical Sciences*, 92(3):277–286.

[79] Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–41.

[80] Mayers, D. F. (2014). *An introduction to numerical analysis*. Cambridge : Cambridge University Press.

[81] Mayhew, P. M., Thomas, C. D., Clement, J. G., Loveridge, N., Beck, T. J., Bonfield, W., Burgoyne, C. J., and Reeve, J. (2005). Relation between age, femoral neck cortical stability, and hip fracture risk. *The Lancet*, 366(9480):129 – 135.

[82] McCalden, R. W., McGeough, J. A., Barker, M. B., and Court-Brown, C. M. (1993). Age-related changes in the tensile properties of cortical bone. The relative importance of changes in porosity, mineralization, and microstructure. *The Journal of Bone & Joint Surgery*, 75(8):1193–1205.

[83] McClung, M. R., Grauer, A., Boonen, S., Bolognese, M. A., Brown, J. P., Diez-Perez, A., Langdahl, B. L., Reginster, J.-Y., Zanchetta, J. R., Wasserman, S. M., Katz, L., Maddox, J., Yang, Y.-C., Libanati, C., and Bone, H. G. (2014). Romosozumab in postmenopausal women with low bone mineral density. *New England Journal of Medicine*, 370(5):412–420.

[84] Mindways Software, Inc (2011). *CT Calibration Phantom : User's Guide*. Mindways.

[85] Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. In Watson, G. A., editor, *Numerical Analysis: Proceedings of the Biennial Conference Held at Dundee, June 28–July 1, 1977*, pages 105–116. Springer Berlin Heidelberg, Berlin, Heidelberg.

[86] Muller, R., Hildebrand, T., and Ruegsegger, P. (1994). Non-invasive bone biopsy: a new method to analyse and display the three-dimensional structure of trabecular bone. *Physics in Medicine & Biology*, 39(1):145.

[87] Neer, R. M., Arnaud, C. D., Zanchetta, J. R., Prince, R., Gaich, G. A., Reginster, J.-Y., Hodsman, A. B., Eriksen, E. F., Ish-Shalom, S., Genant, H. K., Wang, O., Mellström, D., Oefjord, E. S., Marcinowska-Suchowierska, E., Salmi, J., Mulder, H., Halse, J., Sawicki, A. Z., and Mitlak, B. H. (2001). Effect of parathyroid hormone (1-34) on fractures and bone mineral density in postmenopausal women with osteoporosis. *New England Journal of Medicine*, 344(19):1434–1441.

[88] Newman, D. L., Dougherty, G., Obaid, A. A., and Hajrasy, H. A. (1998). Limitations of clinical CT in assessing cortical thickness and density. *Physics in Medicine and Biology*, 43(3):619.

[89] NIH Consensus Development Panel on Osteoporosis Prevention and Diagnosis, and Therapy (2001). Osteoporosis prevention, diagnosis, and therapy. *JAMA*, 285(6):785–795.

[90] Ohnaru, K., Sone, T., Tanaka, K., Akagi, K., Ju, Y.-I., Choi, H.-J., Tomomitsu, T., and Fukunaga, M. (2013). Hip structural analysis: a comparison of DXA with CT in postmenopausal Japanese women. *SpringerPlus*, 2(1):331.

[91] Pakdel, A., Robert, N., Fialkov, J., Maloul, A., and Whyne, C. (2012). Generalized method for computation of true thickness and X-ray intensity information in highly blurred sub-millimeter bone features in clinical CT images. *Physics in Medicine and Biology*, 57(23):8099.

[92] Parfitt, A., Villanueva, A., Foldes, J., and Rao, D. S. (1995). Relations between histologic indices of bone formation: Implications for the pathogenesis of spinal osteoporosis. *Journal of Bone and Mineral Research*, 10(3):466–473.

[93] Parfitt, A. M., Drezner, M. K., Glorieux, F. H., Kanis, J. A., Malluche, H., Meunier, P. J., Ott, S. M., and Recker, R. R. (1987). Bone histomorphometry: Standardization of nomenclature, symbols, and units: Report of the ASBMR histomorphometry nomenclature committee. *Journal of Bone and Mineral Research*, 2(6):595–610.

[94] Pearson, R. A. and Treece, G. M. (2017). Measurement of the bone endocortical region using clinical CT. *Medical Image Analysis*, Accepted.

[95] Pistoia, W., van Rietbergen, B., and Rüegsegger, P. (2003). Mechanical consequences of different scenarios for simulated bone atrophy and recovery in the distal radius. *Bone*, 33(6):937 – 945.

[96] Poole, K., Skingle1, L., Gee, A., Turmezei, T., Johannesdottir, F., Blesic, K., Rose, C., Vindlacheruvu, M., Donell, S., Vaculiz, J., Dungl, P., Horak, M., Stepan, J., Reeve, J., and Treece, G. (2016). Focal osteoporotic defects play a key role in hip fracture. *Bone*, 94:124–134.

[97] Poole, K. E., Mayhew, P. M., Rose, C. M., Brown, J. K., Bearcroft, P. J., Loveridge, N., and Reeve, J. (2010). Changing structure of the femoral neck across the adult female lifespan. *Journal of Bone and Mineral Research*, 25(3):482–491.

[98] Poole, K. E., Treece, G. M., Gee, A. H., Brown, J. P., McClung, M. R., Wang, A., and Libanati, C. (2015). Denosumab rapidly increases cortical bone in key locations of the femur: A 3D bone mapping study in women with osteoporosis. *Journal of Bone and Mineral Research*, 30(1):46–54.

[99] Poole, K. E. S., Treece, G. M., Mayhew, P. M., Vaculík, J., Dungl, P., Horák, M., Štěpán, J. J., and Gee, A. H. (2012). Cortical thickness mapping to identify focal osteoporosis in patients with hip fracture. *PLoS ONE*, 7(6):e38466.

[100] Poole, K. E. S., Treece, G. M., Ridgway, G. R., Mayhew, P. M., Borggrefe, J., and Gee, A. H. (2011). Targeted regeneration of bone in the osteoporotic human femur. *PLoS ONE*, 6(1):e16190.

[101] Powell, M. J. (1978). A fast algorithm for nonlinearly constrained optimization calculations. In *Numerical analysis*, pages 144–157. Springer.

[102] Prevrhal, S., Engelke, K., and Kalender, W. (1999). Accuracy limits for the determination of cortical width and density: The influence of object size and CT imaging parameters. *Physics in Medicine and Biology*, 44(3):751–764.

[103] Prevrhal, S., Fox, J. C., Shepherd, J. A., and Genant, H. K. (2003). Accuracy of CT-based thickness measurement of thin structures: Modeling of limited spatial resolution in all three dimensions. *Medical Physics*, 30(1):1–8.

[104] Qiu, S., Fyhrie, D. P., Palnitkar, S., and Rao, D. S. (2003). Histomorphometric assessment of haversian canal and osteocyte lacunae in different-sized osteons in human rib. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology*, 272A(2):520–525.

[105] Rittweger, J., Michaelis, I., Giehl, M., Wüsecke, P., and Felsenberg, D. (2004). Adjusting for the partial volume effect in cortical bone analyses of pQCT images. *J Musculoskel Neuron Interact*, 4(4):436–441.

[106] Ritzel, H., Amling, M., Pösl, M., Hahn, M., and Delling, G. (1997). The thickness of human vertebral cortical bone and its changes in aging and osteoporosis: A histomorphometric analysis of the complete spinal column from thirty-seven autopsy specimens. *Journal of Bone and Mineral Research*, 12(1):89–95.

[107] Rucci, N. (2008). Molecular biology of bone remodelling. *Clinical Cases in Mineral and Bone Metabolism*, 5(1):49–56.

[108] Rueckert, D., Frangi, A. F., and Schnabel, J. A. (2003). Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Transactions on Medical Imaging*, 22(8):1014–1025.

[109] Schaffler, M. B. and Burr, D. B. (1988). Stiffness of compact bone: Effects of porosity and density. *Journal of Biomechanics*, 21(1):13 – 16.

[110] Schuit, S., van der Klift, M., Weel, A., de Laet, C., Burger, H., Seeman, E., Hofman, A., Uitterlinden, A., van Leeuwen, J., and Pols, H. (2004). Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam study. *Bone*, 34(1):195 – 202.

[111] Seeman, E. and Delmas, P. D. (2006). Bone quality — the material and structural basis of bone strength and fragility. *New England Journal of Medicine*, 354(21):2250–2261.

[112] Seibert, C., Barnes, J., Dreisbach, J., Swanson, W., and Heck, R. (1981). Accurate CT measurement of the spinal cord using metrizamide: physical factors. *American Journal of Roentgenology*, 136.

[113] Sigurdsson, G., Aspelund, T., Chang, M., Jonsdottir, B., Sigurdsson, S., Eiriksdottir, G., Gudmundsson, A., Harris, T. B., Gudnason, V., and Lang, T. F. (2006). Increasing sex difference in bone strength in old age: The age, gene/environment susceptibility-Reykjavik study (AGES-Reykjavik). *Bone*, 39(3):644 – 651.

[114] Silva, M., Wang, C., Keaveny, T., and Hayes, W. (1994). Direct and computed tomography thickness measurements of the human, lumbar vertebral shell and endplate. *Bone*, 15(4):409 – 414.

[115] Spoor, F. C., Zonneveld, F. W., and Macho, G. A. (1993). Linear measurements of cortical bone and dental enamel by computed tomography: applications and problems. *Am. J. Phys. Anthropol.*, 91:409 – 414.

[116] Stein, M. S., Feik, S. A., Thomas, C. D. L., Clement, J. G., and Wark, J. D. (1999). An automated analysis of intracortical porosity in human femoral bone across age. *Journal of Bone and Mineral Research*, 14(4):624–632.

[117] Stiehl, J., Jacobson, D., and Carrera, G. (2007). Morphological analysis of the proximal femur using quantitative computed tomography. *International Orthopaedics*, 31(3):287–292.

[118] Streekstra, G. J., Strackee, S. D., Maas, M., ter Wee, R., and Venema, H. W. (2007). Model-based cartilage thickness measurement in the submillimeter range. *Medical Physics*, 34(9):3562–3570.

[119] Styner, M., Brechbuhler, C., Szckely, G., and Gerig, G. (2000). Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Transactions on Medical Imaging*, 19(3):153–165.

[120] Styner, M. and Gerig, G. (1997). Evaluation of 2D/3D bias correction with 1+ 1ES-optimization. *Rapport de recherche*, 179.

[121] Suetens, P. (2009). *Fundamentals of medical imaging*. Cambridge : Cambridge University Press, 2nd edition.

[122] Szulc, P., Seeman, E., Duboeuf, F., Sornay-Rendu, E., and Delmas, P. D. (2006). Bone fragility: Failure of periosteal apposition to compensate for increased endocortical resorption in postmenopausal women. *Journal of Bone and Mineral Research*, 21(12):1856–1863.

[123] Treece, G. and Gee, A. (2015). Independent measurement of femoral cortical thickness and cortical bone density using clinical CT. *Medical Image Analysis*, 20(1):249 – 264.

[124] Treece, G., Gee, A., Mayhew, P., and Poole, K. (2010). High resolution cortical bone thickness measurement from clinical CT data. *Medical Image Analysis*, 14(3):276 – 290.

[125] Treece, G., Poole, K., and Gee, A. (2012). Imaging the femoral cortex: Thickness, density and mass from clinical CT. *Medical Image Analysis*, 16(5):952 – 965.

[126] Treece, G. M., Gee, A. H., Tonkin, C., Ewing, S. K., Cawthon, P. M., Black, D. M., and Poole, K. E. (2015). Predicting hip fracture type with cortical bone mapping (CBM) in the osteoporotic fractures in men (MrOS) study. *Journal of Bone and Mineral Research*, 30(11):2067–2077.

[127] Tsai, J. N., Uihlein, A. V., Burnett-Bowie, S.-A. M., Neer, R. M., Zhu, Y., Derrico, N., Lee, H., Bouxsein, M. L., and Leder, B. Z. (2015). Comparative effects of teriparatide, denosumab, and combination therapy on peripheral compartmental bone density, microarchitecture, and estimated strength: the DATA-HRpQCT study. *Journal of Bone and Mineral Research*, 30(1):39–45.

[128] Tsai, J. N., Uihlein, A. V., Burnett-Bowie, S. M., Neer, R. M., Derrico, N. P., Lee, H., Bouxsein, M. L., and Leder, B. Z. (2016). Effects of two years of teriparatide, denosumab, or both on bone microarchitecture and strength (DATA-HRpQCT study). *The Journal of Clinical Endocrinology & Metabolism*, 101(5):2023.

[129] Turner, C. H. (2002). Biomechanics of bone: Determinants of skeletal fragility and bone quality. *Osteoporosis International*, 13(2):97–104.

[130] Van Rietbergen, B., Huiskes, R., Eckstein, F., and Rüegsegger, P. (2003). Trabecular bone tissue strains in the healthy and osteoporotic human femur. *Journal of Bone and Mineral Research*, 18(10):1781–1788.

[131] Whitmarsh, T., Treece, G., Gee, A., Bolognese, M., Brown, J., Goemaere, S., Grauer, A., Hanley, D., Mautalen, C., Recknor, C., Yang, Y.-C., Libanati, C., and Poole, K. (2014). Romosozumab and teriparatide effects on vertebral cortical mass, thickness and density in postmenopausal women with low bone mineral density (BMD). In *Proceedings of the American Society for Bone and Mineral Research*, Houston, USA. ASBMR.

[132] Whitmarsh, T., Treece, G. M., Gee, A. H., and Poole, K. E. S. (2015). 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*, 30(7):1309–1318.

[133] Whitmarsh, T., Treece, G. M., Gee, A. H., and Poole, K. E. S. (2016). The effects on the femoral cortex of a 24 month treatment compared to an 18 month treatment with teriparatide: A multi-trial retrospective analysis. *PLoS ONE*, 11(2):1–9.

[134] Williams, D. F. (2014). *Essential biomaterials science*. Cambridge : Cambridge University Press.

[135] World Health Organisation (2007). *WHO Scientific Group on the Assessment of Osteoporosis at Primary Health Care Level*, Summary Meeting Report., Geneva. World Health Organisation.

[136] Worsley, K., Taylor, J., Carbonell, F., and et al. (2009). Surfstat: a Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. page 47–S102. NeuroImage Organization for Human Brain Mapping.

[137] Yang, L., Burton, A. C., Bradburn, M., Nielson, C. M., Orwoll, E. S., Eastell, R., and for the Osteoporotic Fractures in Men (MrOS) Study Group (2012). Distribution of bone density in the proximal femur and its association with hip fracture risk in older men: The osteoporotic fractures in men (MrOS) study. *Journal of Bone and Mineral Research*, 27(11):2314–2324.

[138] Zebaze, R., Ghasem-Zadeh, A., Mbala, A., and Seeman, E. (2013). A new method of segmentation of compact-appearing, transitional and trabecular compartments and quantification of cortical porosity from high resolution peripheral quantitative computed tomographic images. *Bone*, 54(1):8 – 20.

[139] Zebaze, R. M., Ghasem-Zadeh, A., Bohte, A., Iuliano-Burns, S., Mirams, M., Price, R. I., Mackie, E. J., and Seeman, E. (2010). Intracortical remodelling and porosity in the distal radius and post-mortem femurs of women: a cross-sectional study. *The Lancet*, 375(9727):1729 – 1736.

[140] Zebaze, R. M., Libanati, C., Austin, M., Ghasem-Zadeh, A., Hanley, D. A., Zanchetta, J. R., Thomas, T., Boutroy, S., Bogado, C. E., Bilezikian, J. P., and Seeman, E. (2014). Differing effects of denosumab and alendronate on cortical and trabecular bone. *Bone*, 59(0):173 – 179.