TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Tomas Cervinka

Bone Structural Analysis with pQCT:
Image Preprocessing and Cortical Bone Segmentation

Julkaisu 1236 • Publication 1236

Tampere 2014

Tampereen teknillinen yliopisto. Julkaisu 1236
Tampere University of Technology. Publication 1236

Tomas Cervinka

Bone Structural Analysis with pQCT:
Image Preprocessing and Cortical Bone Segmentation

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 5th of September 2014, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2014

Supervisors

Professor Jari Hyttinen
Department of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland

Docent Harri Sievänen, ScD.
Bone Research Group, UKK Institute, Tampere, Finland
Department of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland

ISBN 978-952-15-3340-2 (printed)
ISBN 978-952-15-3353-2 (PDF)
ISSN 1459-2045

To my grandma

Dum spīro, spēro

Acknowledgements

This thesis is a product of a research project conducted at the Department of Electronics and Communications Engineering at Tampere University of Technology in co-operation with the Bone Research Group of the UKK Institute in Tampere and the Institute of Aerospace Medicine at the German Aerospace Center.

I wish to express my gratitude to my thesis supervisors Professor Jari Hyttinen and Docent Harri Sievänen. I am grateful for their guidance, support and motivational support during the course of the study, as it has proved to be priceless. I am very thankful for the medical insight and visionary ideas of Professor Jörn Rittweger who enabled the testing of developed methods on relevant clinical data. Co-operation with him was very productive and valuable. I would also like to thank Professor Jari Viik and Docent Prasun Dastidar for their valuable guidance and co-operation when needed.

I also wish to give a special thanks to my colleagues Markus Hannula, Nathaniel Narra, Baran Aydogan, Michelangelo Paci, Florentino Caetano dos Santos, Jarno Tanskanen, Soile Nymark, Antti Ahola, Atte Joutsen, Emre Kapucu, Narayan Puthanmadam Subramaniyam, Ville-Pekka Seppä, Javier Gracia Tabuenca and Edite Areias Figueiras for their valuable comments and co-operation and for creating a friendly and fruitful atmosphere at our work place. Furthermore, I owe my best thanks to the personnel of Department of Electronics and Communication Engineering but especially to Soile Lönnqvist and Pasi Kauppinen whose motivational support and guidance throughout my studies was limitless. I would also like to express my gratitude to all my friends from TAKO rowing club for making my life a bit more diverse than just sitting in front of my computer.

I would like to thank Professor Ulla Ruotsalainen, (Department of Signal Processing, Tampere University of Technology, Finland) who gave me valuable comments on the introductory chapters of this thesis. I also wish to thank Dr. Jodi Dowthwaite (Department of Orthopedic Surgery, SUNY Upstate Medical University, USA) and Dr. James Johnston (Department of Mechanical Engineering, University of Saskatchewan, Canada) for their constructive criticism and advice as examiners of this thesis. I also thank Peter Heath for carefully revising the English of my thesis.

I am grateful for the financial support received from the Finnish Cultural Foundation and the Graduate School of Tampere University of Technology and my supervisor Professor Jari Hyttinen.

Importantly, I would like to thank my family. I owe them the deepest gratitude for their constant interest, understanding, help and support during the long years of my studies.

Finally, my warmest thanks go to my wife Marketa for her love and support during this study, and for the first stage proofreading of my manuscripts. Without her, this thesis would not have been possible. Last but not least, my thanks go to my first born son Matias for his many smiles that gave me energy to go on.

Tampere, 2014

Tomas Cervinka

Abstract

Bone fragility, or conversely, its strength is largely determined by structural properties (e.g., cross-sectional geometry, cortical thickness and trabecular architecture). Accordingly, a proper structural analysis of bones might help identify those individuals who are susceptible to fragility fractures. In this respect, peripheral quantitative computed tomography (pQCT) provides a reasonable option to assess bone cross-sectional geometry and separate it into the trabecular and the cortical compartments. Although the commonly used pQCT systems lack sufficient spatial resolution to capture fine structural traits compared with present high-resolution pQCT (HR-pQCT) systems, they are cheaper and still widely used among bone researchers. Current techniques used for pQCT image analysis (e.g., median filtering, threshold-based segmentation), however, may not be optimally suited for characterizing bone geometry and strength.

Therefore, in this work, image preprocessing based on a statistical approach, and a fast and reliable segmentation method were developed with the aim of improving the *in vivo* precision of pQCT image analyses, *in vivo* assessment of tibial bone geometry and strength. The segmentation method is based on the outer boundary detection and subsequent shrinking (OBS) procedure. These methods were applied on various phantom and *in vivo* pQCT image data and their performance was compared with commonly used analysis approaches and, in the case of segmentation, with a more advanced level-set based method. The manual segmentation of cortical bone, done by 3 independent evaluators, was considered as a gold standard.

The findings demonstrate that the developed preprocessing method provided the most consistent results and improved the reliability of the analysis of the cortical bone area by approximately 30%. However, the preprocessing failed to improve the assessment of trabecular density. Nevertheless, some coarse structural patterns could be seen in the preprocessed images in contrast to a disperse distribution of density levels in the raw images. Further, the developed OBS method showed nearly 50% less variation in error compared with threshold-based analysis and, in conjunction with the developed preprocessing method, agreed well with the results obtained from manual segmentation. The level-set based segmentation resulted in a consistent ~15% mean overestimation of all geometrical traits with a similar variation of data as that obtained from the OBS method.

The results of this research show that the preprocessing method can enhance cortical bone analysis. Further, the OBS method (i) improved assessment of cortical geometry, (ii) enhanced cortical bone analysis of pQCT images and (iii) can, in conjunction with the developed preprocessing, form a simple and applicable pQCT image analysis tool for clinical bone research applications.

CONTENTS

List of abbreviations

µ-CT	micro computed tomography
2D	two-dimensional
3D	three-dimensional
95%CI	95% confidence intervals
aBMD	areal bone mineral density
BBR	Berlin Bed Rest
BMC	bone mineral content
BMD	bone mineral density
BSI	bone strength index
BV	bone volume
BV/TV	bone volume to total volume fraction
C1/P2	threshold segmentation implemented in the Stratec analysis software (contour mode 1 and peeling mode 2)
C2/P2	threshold based segmentation implemented in the Stratec analysis software (contour mode 2 and peeling mode 2)
CoA	cortical cross-sectional bone area
CoD	volumetric density of the cortical bone
CoTh	mean cortical wall thickness
CSMI	cross-sectional area moment of inertia
DRLSE	distance regularized level set evolution
DXA	dual-energy X-ray absorptiometry
FEA	finite element analysis
FEM	finite element modelling
FIR	finite impulse response
GaussW	field with Gaussian weight coefficient distribution
GHT	generalized Hough transform
HR-CT	high-resolution computed tomography
HR-MRI	high-resolution magnetic resonance imaging
HR-pQCT	high-resolution peripheral quantitative computed tomography
k-NN	k – nearest neighborhood classifiers
LinW	field with linear weight coefficient distribution

LTBR	Long Term Bed Rest
OBS	outer boundary detection and shrinking
OWT-UCM	ultrametric contour mapping algorithm
pQCT	peripheral quantitative computed tomography
PSF	point spread function
PSNR	peak signal-to-noise ratio
QCT	quantitative computed tomography
SMF	standard median filter
SMSE	signal to mean squared error
SNR	signal-to-noise ratio
SSI	stress strain index
Tb.N	trabecular number
Tb.Sp	trabecular separation
Tb.Th	trabecular thickness
TIST	threshold-independent segmentation tool introduced by Valentinitsch et al. [2012]
ToA	total cross-sectional bone area
TrA	trabecular cross-sectional bone area
TrD	volumetric density of the trabecular bone
ULLS	Unilateral Lover Limb Suspension
VOI	volume of interest
WT	wavelet transform
Z	section modulus

List of original publications

This thesis is base on the following articles, referred to in the text by Roman numerals.

I. CERVINKA T., HYTTINEN J., SIEVANEN H.: Enhanced bone structural analysis through pQCT image preprocessing. *Medical Engineering & Physics,* 2010; 32(4): 398-406.

II. CERVINKA T., RITTWEGER J., HYTTINEN J., FELSENBERG D., SIEVANEN H.: Anatomical sector analysis of load-bearing tibial bone structure during 90-day bed rest and 1-year recovery. *Clinical Physiology and Functional* Imaging, 2011; 31(4): 249-257.

III. CERVINKA T., HYTTINEN J., SIEVANEN H.: Threshold-Free Automatic Detection of Cortical Bone Geometry by Peripheral Quantitative Computed Tomography. *Journal of Clinical Densitometry*, 2012; 15(4): 413-421.

IV. CERVINKA T., SIEVANEN H., HYTTINEN J., RITTWEGER J.: Bone loss patterns in cortical, subcortical and trabecular compartments during simulated microgravity. *Journal of Applied Physiology*, submitted January 2014.

The author's contribution

The author has developed an image processing and segmentation algorithm, performed image analysis and statistical analysis, has taken a part in the conception of studies and their design and has been mainly responsible for the preparation of all manuscripts.

1. Introduction

The main goal of this study was to develop an image analysis tool for performing reliable bone structural analysis in pQCT image data that would be feasible for use in clinical bone research. Osteoporosis and associated fragility fractures are common health problems in aging populations and are a considerable socioeconomic burden on health care systems [Christensen et al. 2010]. Osteoporotic fragility fractures are attributable both to interaction between external loading (mostly fall-induced) and bone fragility. Bone fragility, in turn, is largely determined by the bone structure and its particulars [Järvinen et al., 2005].

Improvements in medical imaging and image preprocessing during the past 3 decades have enabled deeper *in vivo* analysis of bone macro- and micro-structure, increased the overall knowledge and understanding of bone anatomy, physiology and assessment of osteoporotic processes [Ito, 2011; Griffith & Genant, 2012]. Nevertheless, the present diagnostic assessment of bone fragility and fracture risk still rests largely on areal bone mineral density (aBMD) measured by dual-energy X-ray absorptiometry (DXA) [Kanis et al., 2005]. The aBMD approach is limited not only by the inherent inaccuracy of DXA caused by the violation of the 2-component (homogeneous soft tissue and bone) assumption in individual patients [Bolotin & Sievänen, 2001] but also by the inability of planar DXA to yield reliable information on actual bone geometry and structure that is necessary for proper assessment of bone strength [Sievänen et al., 2007]. Therefore, the aBMD approach has been recognized to be insufficient in describing bone fragility and individual fracture risk [Stone et al., 2003; Kanis et al., 2004].

Peripheral quantitative computed tomography (pQCT) offers a reasonable option to assess bone cross-sectional geometry and to separate it into trabecular and cortical compartments [Sievänen et al., 1998]. Although commonly used pQCT systems lack sufficient spatial resolution to capture specific micro-structural traits compared with present high-resolution

pQCT (HR-pQCT) systems, they have been commonly used in many recent studies [Lochmüller et al., 2002; Rittweger et al., 2005; Rittweger et al., 2009; Rittweger et al., 2010; Nikander et al., 2006; Szabo et al., 2011; Shedd-Wise et al., 2011]. Due to obvious technical limitations, pQCT cannot be applied to clinically relevant vertebral and proximal femur sites. Nonetheless pQCT does provide similar information on apparent structural traits from appendicular bones, to data provided by clinical quantitative computed tomography (QCT) at the proximal femur and lumbar spine [Lang et al., 2006; Melton et al., 2007; Melton et al., 2010].

The evaluation of bone fragility and fracture risk is technically challenging, because trade-offs between parameters such as image resolution, radiation exposure, signal-to-noise ratio (SNR) and acquisition time are limiting factors for accurate bone image analysis [Genant et al., 2008]. The lower the resolution, the less accurate is the determination of the bone structures obtained because the partial volume effect distorts the real borders of bone structures. Higher resolution can be reached by increasing radiation exposure. Obviously, safety concerns and ethical issues limit larger dosages in clinical applications. However, by increasing the acquisition time, the resolution and the SNR level can be improved, but longer scan times increase risk of image distortion due to movement artefacts [Beutel et al., 2000].

Hence, identifying the proper imaging method for reasonable assessment of bone structure is needed for early diagnosis of osteoporosis. Successful management of osteoporosis requires detection of subtle changes in bone mass, structure and tissue mechanical properties for accurate prediction of bone fragility [Lochmüller et al., 2003; Kazakia & Majumdar, 2006; Griffith & Genant, 2012]. Whether this kind of structural assessment will outperform conventional aBMD for fragility fracture prediction and identification of individuals-at-risk will be a crucial issue [Sievänen, 2010a]. Before that, proper analysis algorithms need to be developed; the present study is considered as a step towards the ultimate objective of predicting bone fragility and revealing relevant changes in bone structure that affect whole bone strength, making it more susceptible to fracture.

Current image processing and analysis algorithms used in clinical research are mostly limited to the use of median filtering and application of density thresholds [Veitch et al., 2004; Ward et al., 2005; Hangartner, 2007]. This practice comes not only from technical simplicity but also from the need for reproducible and tangible results in clinical settings. Such needs, however, are not always met by current threshold-based analyses [Ward et al., 2005; Ashe et al., 2006; Kontulainen et al., 2007], mainly due to partial volume effects, a relatively low signal to noise ratio and the presence of movement artefacts. All these obstacles cause blurred, indistinct or discontinued edges in analysed images and result in a non-uniform density values along these edges. Therefore, the use of a median filter, which causes an additional blur, together with variable threshold values substantially modulate the results of the threshold-based analysis [Ward et al., 2005; Ashe et al., 2006; Kontulainen et al., 2007] and result in discontinued segmented areas in analysed images.

This considerably affects an assessment of bone strength indexes (e.g., area moment of inertia [CSMI]) that are largely dependent on the precise definition of the cortical compartment. Obviously, the accuracy and precision of measures of relevant bone structural changes depends on image quality of the used imaging modality [Lochmüller et al., 2003]. Subsequent image processing cannot overcome inherent limitations of the image data due to limited resolution, for example. Many sophisticated algorithms such as level set-based segmentation [Osher & Paragios, 2003] are also not considered suitable for clinical research because of the need for large operator involvement (increasing the risk of decreased reproducibility of the results) or the high sensitivity to noise. However, any improvement in the accurate and precise description of the structure of the given bone is essential for the appropriate assessment of whole bone strength [Sievänen et al., 2007].

2. Literature review

2.1. Bone structure

Bones and bone structures have been an object of interest among researchers during the past centuries. Recently, the interest has grown extensively as osteoporosis and related fragility fractures have become common health problems in aging populations. Basically, human bone structure reflects the overall environment where each bone is and the purpose for which they are developed: to support body weight, to work together with muscles, to generate body movements and to provide protection for vital organs. Demands for the arrangement of bone structures are high considering the occasionally high forces that have to be coped with while, at the same time being sufficiently light to cause no excess metabolic cost or movement restrictions. Bones are comprised of a cortical shell, with trabecular network primarily limited to the long bone ends, vertebrae, flat and small bones. A thick cortical shell is primarily found in the shaft of long bones to resist bending forces from a limited range of directions. A flexible, lightweight trabecular network is located near joints to resist compression forces and to absorb energy coming from many different directions, as seen during locomotion for example. Figure 1 illustrates an example of a bone with a trabecular bone framework and a relatively thin cortical shell. The comprehensive description of bone structures can be found in many anatomy books [Marieb, 2004; Martini et al., 2006].

Bone tissue is distinguished between two types of osseous tissue**:** cortical bone, also known as a compact bone; and trabecular bone, also called cancellous or spongy bone. As the terms already suggest, the porosity of cortical bone is low and ranges from 5-10%. Cortical bone is composed of a hierarchical structure that forms a series of concentric rings around the central canal. Trabecular bone has a porosity of between 50-90% and is relatively light. Trabecular bone is composed of a plate and strut-like structure that forms

an open spongy framework and helps to reduce bone weight. A more detailed comparison between the general features of cortical and trabecular bone can be found in the work of Jee [1983].

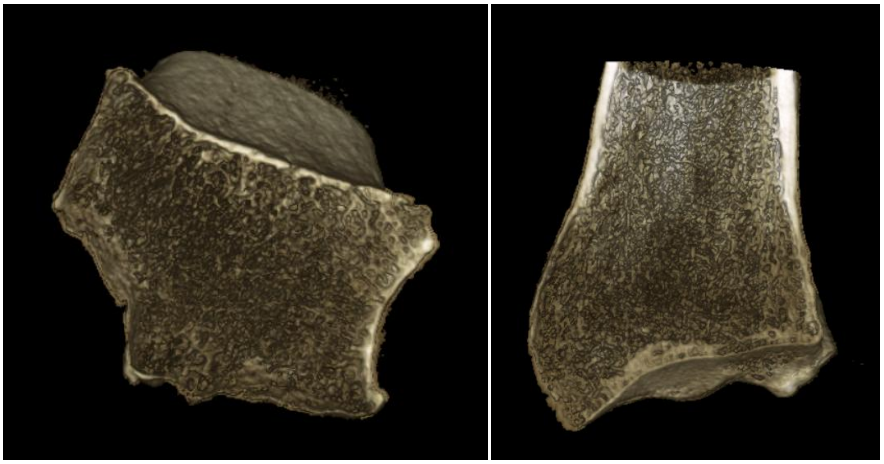


Figure 1: Example of the distal radius. The image was obtained from a cadaver bone by using a high resolution computed tomography (HR-CT) scanner with an isotropic voxel size of 0.1 mm.

2.2. Bone imaging methods in vivo

The main purpose of bone densitometric techniques is to determine parameters that accurately assess bone strength, predict bone fragility and associated fractures, and assess aging processes and endocrine and physical conditions in bones. In this respect, DXA derived aBMD was among the first quantitative measurements [Engelke et al., 2008]. Areal BMD does not provide a detailed picture of whole bone strength. More parameters to measure bone structure and tissue mechanical properties of both trabecular and cortical bone are needed to improve the prediction of fragility fractures [Järvinen et al., 2005]. However, in clinical practice, the 25-year old DXA-measured aBMD is still the most used and recommended parameter [Kanis et al., 2005; Lang, 2010]. Selected bone structural indicators (traits) that are commonly used for the estimation of whole bone strength and their definitions are shown in Table 1.

Spatial resolution of imaging – determined by the size of the smallest possible feature that can be detected – is typically represented by a point spread function (PSF) for each imaging system. The PSF describes the response of an imaging system to a small point object (e.g., small bead), with the spread of the point object in the image characterizing the PSF. However, voxel size is commonly used to indicate differences in spatial resolutions between different imaging modalities, with the assumption that the voxel size is greater than the PSF. It is important to note that voxel size is not equivalent with PSF. In the following pages of this chapter the term voxel size will be used to indicate spatial resolutions for different types of imaging modalities

2.2.1. In plane methods

Digital radiographs

As the first method used for the detection of osteoporotic bones, plain film radiography has been used since the 1940s. Although radiography scanners have been radically developed during the past decades and digital detectors have replaced the film medium, the main principle of plain radiography has remained the same since its introduction by Wilhelm Conrad Roentgen in 1895. Radiography is based on the measurement of the linear attenuation of x-ray beams that penetrate the body and visualize the skeleton.

Radiography is very simple, cheap and fast (imaging time < 10 s), is readily available, does not require a large radiation dose and is capable of providing sufficient bone details [Griffith & Genant, 2011]. There are, however, some limitations with radiography that should be mentioned. The use of a lower radiation dose restricts radiography to detecting osteoporosis mainly in the appendicular skeleton, due to the large attenuation of x-rays by soft tissues**.** Even though some studies have developed methods to detect semi-quantitative indices such as trabecular quantification [Seeley et al., 1991] or cortical thickness quantification [Gärdsell et al., 1993], conventional radiographs represent an overall projection of the bone structure and are usually not able to show individual trabeculae [Link, 2010]. Furthermore, there needs to be about 20% bone mineral loss before the bone loss can be detected from radiographs [Ardran, 1951]. Therefore, the sensitivity and reproducibility are very poor and, as a result, radiography has been mainly used to identify fractures rather than to predict bone fragility. However, there have been some recent studies on human cadaver specimens [Pulkkinen et al., 2008] and postmenopausal women [Rachidi et al., 2008; Pulkkinen et al., 2011] that suggest the possible prediction of hip fracture load and the accurate separation between subjects with fractures and controls independently of BMD. Nevertheless, in these studies, a high-resolution scanner with an in-plane voxel size of 50 µm was used. This type of scanner is a refined version of the standard radiographs used for clinical examinations and further studies are needed to reveal the usability of these approaches for the early detection of osteoporosis.

Dual-energy X-ray Absorptiometry

Dual-energy x-ray absorptiometry (DXA) and DXA-measured areal bone mineral density (aBMD in g/cm^2) has been in clinical use since 1987. In 1994, the World Health Organization (WHO) proclaimed aBMD to be the primary measure to assess bone strength because of its relatively high correlation with actual bone strength and the sufficient ability to predict relative fracture risk [Kanis, 1994]. Further advantages of DXA include the following: small radiation dose, short screening times (< 1 min.), large availability, simple scanner operation and lower costs in comparison with computed tomography. Areal BMD is determined by the measurement of attenuation of x-ray beams applied on clinically relevant sites of the body. Hard tissues such as bones attenuate more x-ray beams than soft

Table 1: Selected bone structural indicators and their definitions

Measure types		Indicator	Definition
Macro-structures	Bulk	Bone mineral content (BMC)	a measure of the amount of bone mineral (calciumhydroxyapatite) contained in a certain volume of bone
		Areal bone mineral density (aBMD)	the amount of bone mineral per square centimetre of bone
	Geometry	Total cross-sectional bone area (ToA)	the area of bone enclosed by the outer periosteal boundary
		Trabecular cross-sectional bone area (TrA)	the area between the outer periosteal and the endosteal boundary
		Cortical cross-sectional bone area (CoA)	the area between the outer periosteal and the endosteal boundary
		Mean cortical wall thickness (CoTh)	the mean value of distances between the endosteal and the periosteal boundary measured along their contours
		Cross-sectional area moment of inertia (CSMI)	a measure of cortical bone distribution around the neutral axes, determines the ability of bone to resist bending or torsion relative to specific neutral axes
		Section modulus (Z)	the product of CSMI and the maximum distance from the periosteum to a particular bending or torsional axis, determines the ability of bone to resist bending or torsion in a given plane
	Strength	Bone strength index (BSI)	the density weighted polar section modulus of a given bone cross-section
		Stress strain index (SSI)	the product of the polar- or bending- section modulus and the normalized volumetric cortical density value of each pixel in a selected cortical bone cross-section, essentially a weighted CSMI
Micro-structures	Architecture	Bone volume (BV)	the total volume of bone within the volume of interest (VOI) that corresponds to the detected outer periosteal boundary
		Bone volume to total volume fraction (BV/TV)	the proportion of the VOI occupied by bone
		Porosity	the ratio of aggregated pore space to the volume of the entire bone mass, equal to (1 – BV/TV)
	Geometry	Trabecular thickness (Tb.Th)	the trabecular bone area divided by one-half of the perimeter or the trabecular bone area divided by the length of the skeleton network
		Trabecular separation (Tb.Sp)	the average distance between adjacent trabeculae
		Trabecular number (Tb.N)	the reciprocal of the distance between the centres of adjacent trabeculae

tissues (i.e. muscles, bone marrow and fat) at a given x-ray energy. DXA scanners use two disparate x-ray energies and evaluate their relative differences in x-ray attenuation.

Today, the evaluation of osteoporosis related bone fragility and fracture is one of the most commonly performed clinical measurements [Kanis et al., 2005]. In 1994, the WHO set the reference diagnostic value (T-score) for the definition of osteoporosis to -2.5 or less of the standard deviation from the young adult mean aBMD level. The reliability of aBMD can, however, be compromised by inherent inaccuracies and ambiguities in interpretation [Sievänen, 2000; Bolotin & Sievänen, 2001; Bolotin, 2004]. The aBMD does not take into account bone depth and, subsequently, a larger bone might seem to be more dense than a smaller one [Griffith & Genant, 2012]. Furthermore, aBMD is derived under an oversimplified assumption that the region of interest contains only hard tissues (i.e., bone) and homogeneous soft tissues. Evidently, the latter assumption cannot be met in practice [Bolotin & Sievänen, 2001]. Hence, use of the reference T-score value does not disclose the majority of so-called osteoporotic fractures that appear to occur mainly above this value [Stone et al., 2003]. Further, DXA's inherent inability to directly assess cortical or trabecular bone properties (since DXA cannot differentiate between these two tissues) illustrates the limited usability of DXA for an accurate assessment of an individual's bone strength [Sievänen, 2010a] and associated fracture risk [Delmas & Seeman, 2004; Li et al., 2004].

Notwithstanding the previously mentioned limitations, DXA-measured aBMD is one of the best available measures of osteoporosis [Griffith & Genant, 2012] due to high precision (1-2%) [Sievänen et al., 1992; Glüer et al., 1993] and simple usability (e.g., the lower the aBMD value is, the more likely the patient has osteoporosis, and the more likely the patient will benefit from medical treatment).

2.2.2. Volumetric methods

Quantitative Computed Tomography

Quantitative computed tomography (QCT) for bone densitometry measurements was developed at the end of the 1970s and the beginning of the 1980s, and allowed the measurement of apparent volumetric BMD (in mg/cm^3) [Link, 2010]. QCT uses a fan of x-rays and a detector-array that rotates around the patient and continuously provides a distribution of linear attenuation coefficients within the patient (radiographic projections). An image is then reconstructed from these radiographic projections by using Radon projection theory. In the image, particular intensity values represent local x-ray linear attenuation. The BMD value is determined by these linear attenuations of x-ray beams. The BMD measurement is, however, derived by the linear calibration of the CT number to BMD using a bone-density equivalent phantom. The CT number is derived from the attenuation coefficient by calibration to the x-ray attenuation of distilled water at room temperature. Hence, all QCT scanners are identically calibrated. Further, recent QCT, based on a spiral system, allows the reconstruction of the 3D structure and brings a number

of advantages: reasonable *in vivo* assessment of bone geometry, the separate assessment of the cortical and trabecular bone compartment, the measurement of the apparent densities at clinically relevant sites (lumbar spine and proximal femur) and the measurement of a volumetric BMD independent of bone size [Riggs et al., 2004; Carpenter et al., 2005; Riggs et al., 2006; Black et al., 2008; Yang et al., 2008].

The two main limitations of QCT-based densitometry are (i) the need for a dedicated QCT scanner (cost) that is not portable and (ii) the relatively high radiation dose involved in measurements that may pose safety and ethical issues. In addition, despite the use of a larger radiation dose compared with DXA, the systems lack sufficient spatial resolution to differentiate between individual trabeculae and are, therefore, unable to assess micro-structural traits (e.g. trabecular number, separation, thickness, connectivity) [Engelke et al., 2008].

However, a relatively small in-plane voxel size (below 300 µm [Ramamurthi et al., 2012]) allows separate discrimination of the cortical and trabecular bone compartments. This advantage makes it possible to selectively assess trabecular bone and its BMD where observed age, treatment and disuse-related changes are greater [Sievänen, 2010b]. Consequently, bone fracture prediction is improved in comparison to DXA-measured aBMD [Yu et al., 1995; Ito et al., 1997]. Nevertheless, the above mentioned in-plane voxel size prevents the correct determination of cortical thickness in the lumbar spine region [Prevrhal et al., 1999] and, therefore, due to partial volume effects, cortical BMD always contains some portion of sub-cortical bone (i.e., transitional zone between compact-appearing cortex and trabecular bone) and adjacent tissues (e.g., tendon, muscle, fascia).

Peripheral QCT

The term peripheral quantitative computed tomography (pQCT) denotes a special low-cost and low-dose application of QCT scanners developed at the turn of the 1980s and the 1990s for quantitative determination of BMD at appendicular sites such as upper and lower limbs [Griffith & Genant, 2011]. Although, pQCT cannot be applied to an axial skeleton or hip region, its data acquisition and reconstruction methods are similar to the second generation of QCT scanners. Therefore, this imaging method provides similar information on macro-structural traits in the forearm and the distal tibia [Sievänen et al., 1998] but also inherits the resolution related benefits and limitations of QCT-based densitometry. Moreover, the use of low power x-ray sources results in a longer imaging time (about 4 min.) and, therefore, the appearance of motion artefacts is more frequent when compared with QCT scanners (screening time for a single image slice ~ 30 s) [Engelke, Libanati, et al., 2009].

The advantages of pQCT scanners are their smaller size and thus higher mobility, lower price compared with whole body clinical QCT scanners and, most importantly, a lower radiation dose. Although pQCT has somewhat larger in-plane voxel size (up to 400 µm), its precision for BMD measurements is comparable to QCT [Sievänen et al., 1998].

Furthermore, due to their application to perpendicular sites where x-rays are not largely attenuated by surrounding soft tissues, pQCT can determine cortical thicknesses in forearms and lower limbs more accurately than QCT can in the lumbar spine. With regards to a recent study by Kontulainen et al. [2008], it is known that the geometry-based parameters (e.g. stress strain index [SSI] and bone strength index [BSI]) can substantially improve the prediction of whole bone strength and bone failure loads at the tibial epiphysis and diaphysis. Therefore, an improvement in the accurate detection of the cortical bone can improve the prediction of bone failure in this respect. The high association of geometry-based parameters with whole bone strength has been reported in some pertinent studies [Lochmüller et al., 2002; Muller et al., 2003; Kontulainen et al., 2008].

High-resolution pQCT

Recently, a new high-resolution imaging system for appendicular sites (XtremeCT, Scanco Medical AG, Bruttisellen, Swizterland) has become available on the market. The imaging system is the same as any other computed tomography system in that it is based on the measurement of the attenuation of x-ray beams passing through the test subject. The system is the result of the convergence of classical clinical CT with features of micro-CT and is designed for the assessment of micro-architecture and morphometric analysis [Boutroy et al., 2005; MacNeil & Boyd, 2007a]. The scanner achieves isotropic voxel sizes as small as 82 µm [Krug et al., 2010]. With smaller voxel sizes, micro-architectural properties such as trabecular thickness, trabecular separation, trabecular number, porosity and connectivity can be estimated. Despite smaller voxel sizes, the effective radiation dose is relatively small and comparable to standard pQCT or DXA measurements, and several orders smaller when compared with QCT. The screening time of high-resolution pQCT (HR-pQCT) is shorter in comparison with standard pQCT scanners (~3 min.). Therefore, the occurrence of movement artefacts is somewhat reduced. The size of the voxels provides a good basis for finite element modelling (FEM) and consequently for various simulations of bone strength in different loading conditions [MacNeil & Boyd, 2008; Liu et al., 2010] or relative load distributions [MacNeil & Boyd, 2007b]. However, it has not yet been demonstrated that FEM improves fracture predictions more than the combination of densitometric and geometric parameters [Engelke et al., 2008].

HR-pQCT is a great step forward toward better understanding and assessment of bone fragility and fractures. However, the system does have some limitations that need to be recognized. At present, the system is made by only one manufacturer. As a result, its use is still rather limited and the devices are mainly installed in major research institutions [Krug et al., 2010]. Further, the system is limited to appendicular sites (wrist, ankle) due to gantry size limitation and radiation dose restriction. The imaging of clinically relevant sites such as the proximal femur and the spine (located deeper in the body) would need a higher radiation dose in order to attain reasonable image quality. Finally, but probably most importantly, even with these smaller voxel sizes, the achievable voxel size is only close to the thickness of individual trabeculae (about 50 – 150 µm). Therefore, the HR-pQCT

image analysis is still affected by partial volume effects and derived micro-architectural properties are overestimated compared to histomorphometry results [Boutroy et al., 2011].

Nevertheless, HR-pQCT proved to be able to show differences in several structural parameters in test subjects, in spite of similar aBMD values [Li et al., 2010; Seeman et al., 2010]. The precision for HR-pQCT assessment of changes in cortical and trabecular bone is about 2-5% [Boutroy et al., 2005]. The values of approximate radiation exposure during different densitometry measurements are shown in Table 2.

Table 2: Approximate radiation exposures during densitometry measurements

Imaging modality	Approx. effective dose (mSv)
Radiography	< 0.01
DXA	~ 0.01 – 0.05
Single-slice QCT	< 0.06 – 0.3
3D QCT scan	~ 1 – 1.5
pQCT	< 0.003
HR-pQCT	< 0.005

High-resolution Magnetic Resonance Imaging

High-resolution magnetic resonance imaging (HR-MRI) systems are becoming more popular among bone researchers. Magnetic resonance imaging uses a strong magnetic field to align the spin orientation of hydrogen proton atoms in the imaged tissue, then radio frequency pulse sequences are applied to change the proton spin orientation; the change of their spin orientation depends on tissue type. This change and return of the spin orientation to a stable state in the magnetic field is detected by the scanner to produce image data. HR-MRI systems cannot directly measure BMD values. However, they can estimate the bone structure information as topology, cortical thickness, cortical area, trabecular network type and trabecular orientation [Krug et al., 2010]. The two biggest advantages of HR-MRI systems are direct multiplanar acquisition and absence of radiation exposure. Next, by virtue of recent developments in the optimization of the pulse sequences and coils used, the systems can be easily used for imaging clinically relevant sites (lumbar spine, proximal femur) [Krug et al., 2005]. In addition, MRI-based systems allow the physiological aspects of bones to be determined beyond the mineralized bone component (marrow fat content, water content, marrow perfusion and marrow diffusion) [Griffith & Genant, 2012].

Nevertheless, when compared with HR-pQCT scanners, HR-MRI has a significantly lower SNR and larger in-plane voxel size (typically ~300 μm [Krug et al., 2010]) due to their "trade-off" character (reducing voxel size reduces SNR and vice versa). Further, imaging

times are relatively long (about 10-12 min., depending on a system type) [Krug et al., 2008; Krug et al., 2010]. Because patients must lie still in a long narrow tube, this imaging modality is not feasible for all cases.

In a recent study by Krug et al. [2008], the use of pulse sequences optimized for the imaging of trabecular bone structures, allowed in-plane voxel sizes of the HR-MRI scanners to reach close to 150x150 μm. However, this is still similar to the thickness of individual trabeculae. Therefore the partial volume effect arises, and is compensated for by increasing the image slice thickness (usually three times larger than in-plane voxel size) to gain sufficient SNR [Krug et al., 2008].

Typical characteristics of bone imaging modalities are shown in Table 3.

2.2.3. Image quality

In the end, reliable estimation of whole bone strength also depends on image quality yielded by the above-mentioned modalities. Image quality depends on factors such as: detector and source characteristics, screening time (longer screening time increases risk of patient motion artefacts) and artefactual noise. In x-ray-based modalities, artefactual noise can be cause by, e.g., beam hardening or beam polychromaticity. In MRI-based modalities, sources of artefactual noise can be low-frequency phase shifts due to magnetic field inhomogeneity or localized differences in magnetic susceptibility [Beutel et al., 2000; Dhawan et al., 2008].

X-ray based imaging modalities may be adversely affected by factors that reduce images resolution, increase noise level or occurrence of image artefacts. Image resolution may be reduced by photon scattering during the process of conversion of x-rays to light in a detector. Photon scattering results in the blurring of an image, however, modern systems have been refined to minimize this problem. Further, detector element size may also result in image blurring, as larger detector elements can fail to discriminate the transition between two tissues that fall within a single pixel/voxel (partial volume effect).

Further, the main noise sources in radiographic images are the x-ray quantum noise, the non-uniform intensity of the x-ray beam, the internal noise of the system, scattered radiation, quantization noise, patient related motion artefacts and structural noise [Hanson, 1981; Kak & Slaney, 1988; Williams et al., 2007; Boas & Fleischmann, 2012]. The x-ray quantum noise causes of random streaks artefacts usually in direction of larger attenuation. The non-uniform intensity of the x-ray beam is a statistical noise that reduces the contrast sensitivity. This noise arises from detection of only a finite number of x-ray quanta. The internal noise of the system is an electronic noise corresponding to properties of particular detectors (e.g., variation in sensitivity can cause ring artefacts). Scattered radiation causes a reduction in the dynamic range of x-ray intensities and an increase in noise as it contains no usable signal. Quantization noise is a truncation error caused by the digitalization process. Patient motion during imaging procedure causes blurring, double images or long

Table 2: Comparison of several modalities used for bone analysis

Modality	In plane voxel sizes	Slice thickness	Precision			Time	Site of application
			Assessing BMD	Assessing structural parameters	Reference		
X-ray	up to 50 µm	–	less than 2%	1-14%†	Pulkkinen et al. 2008	10 s	spine, hip, calcaneus, distal radius
DXA	up to 500 µm	–	less than 2%	3-9%#	Sievanen et al. 1992, Hind et al. 2012	< 1 min.	Spine, hip, forearm, calcaneus
QCT	234 µm	up to 250 µm	2-4%	8%*	Braillon 2005, Engelke et al 2009	30 s	Spine, hip
pQCT	up to 400 µm	2.5 mm	less than 3%	7%*	Sievanen et al. 1998	4 min.	Forearm, tibia
HR-pQCT	82 µm	82 µm	–	2-5%¤	Boutroy et al. 2005, Griffith et al. 2011	3 min.	Forearm, tibia
HR-MRI	< 200 µm	500 µm	–	2-7%¤	McKay et al. 2004, Krug et al. 2010	12 min.	Spine, hip, forearm, tibia, calcaneus, whole body

† macro-structural-related parameters (trabecular bone area, the Euler number, homogeneity index and trabecular main orientation)

macro-structural parameters derived from 2-dimensional DXA images (sectional modulus, cross-sectional moment of inertia and cross-sectional area, buckling ratio and femoral strength index)

* volumetric macro-structural parameters (total bone cross-sectional area, trabecular bone cross-sectional area, cortical bone cross-sectional area and mean cortical wall thickness)

¤ micro-structural parameters (bone volume, bone volume to total volume fraction, porosity, trabecular thickness, trabecular separation, trabecular number)

range streaks (motion artefacts). Structural noise is caused by the superposition of various types of overlapping tissues and results in false determination of transition of particular tissues.

Furthermore, the quality of images yielded by CT-based imaging modalities can be compromised by high-pass filtration, the back projection of noise, the back projection reconstruction algorithm itself and noise artefacts caused by variations in the translational and rotational motion of the x-ray source and detector assembly during the screening procedure [Hanson, 1981; Kak & Slaney, 1988; Boas & Fleischmann, 2012]. The back projection of noise in the projections causes streaks artefacts if the maximum frequency presented in the image data exceeds the maximum recordable at given sampling rate. The back projection algorithm can cause Gibbs phenomenon, streaks and Moire patterns. These artefacts appear due to incorrectly selected number of samples per projections (under sampling), a number of views, a reconstruction grid or filter.

MRI noise sources can be divided into three categories: sources that are related to the use of the MRI scanner (technician/user-defined settings), MRI scanner-related sources and patient-related sources [Gudbjartsson & Patz, 1995; Beutel et al., 2000].

Technician/user-defined settings involve compromises in scanner settings (noise optimization), so the final noise level must be considered before actual imaging. Primary, the noise level is affected by the strength of the magnetic field. Stronger field increases SNR, approximately twice with twofold increase in a field strength but also quadruplicates specific absorption rate (SAR). Increase in SAR can result in increased tissue heating and therefore longer screening times are used to reduce tissue heating. This, however, leads to higher occurrence of motion artefacts. Further, the noise level depends on the selected field of view. The larger the field is, the more protons contribute to the outcome signal. This is, however, at the cost of decreased resolution. Next, the noise level is affected by the number of image acquisitions. Larger number of image acquisitions reduces uncorrelated noise (e.g., termal noise), Next, the noise level depends on the sample bandwidth (a narrower bandwidth decreases the noise digitalization) and the quality of the receiving coil. The optimal setting of the coil to the tissue of interest gives better SNR. Moreover, noise artefacts that can occur in MRI images due to the incorrect setting of pulse sequences are called the Gibbs phenomenon, aliasing, chemical shift and the partial volume effect (interface between two kinds of tissues captured within a pixel/voxel). The Gibbs phenomenon appears in images as ringing artefact that is caused by data truncation and undersampling. Aliasing causes the misplacement of anatomical structures or striping. Chemical shift artificially decreases or increases signal intensity coming from particular tissue.

In addition, quality of MRI images can be compromised by appearance of image artefacts due to MRI scanner-related sources. These artefacts are caused by magnetic field variations shifting tissue contrast to the wrong places or gradient nonlinearity producing rectangular sizes of pixels at the edges of the field of view. Further, artefact can appear due

to local field variations in consequence of the presence of ferromagnetic material. This is called a susceptibility effect and causes incorrect placement of detected image information. Next, MRI image artefacts can be caused by non-uniform response of the imaging coils (fall-off/decreased intensity in the image) or radio frequency system. Leaked radio frequency pulse into the imaging room causes image distortions.

Furthermore, the MRI images can also be compromised by patient-related noise artefacts. Sources of this type noise are the motions of the patient during the imaging and the thermal conductivity of the patient. Motion artefacts are not only caused by blinking, sneezing or speaking during acquisition but are also caused by uncontrollable movements such as respiratory or blood flow motion. Such movements lead to the partial miss-positioning of image information, the so called ghost effects [Beutel et al., 2000].

Since the focus and purpose of this study is to enhance the quality of image analysis for the assessment of bone structures in radiological images, the following chapters will focus on image processing techniques applicable to QCT, pQCT, HR-pQCT images. The description of the image processing techniques used for planar radiographic, DXA or HR-MRI measurements is considered irrelevant and will not be discussed in the following chapters.

2.3. Preprocessing - Denoising Techniques

The purpose of image preprocessing-denoising procedures is to emphasise the desired details in images and to reduce noise generated during imaging procedures. The accurate categorization of noise removal techniques is quite difficult as most of the recent methods are a hybrid of multiple methods. For simplicity, we can divide noise removal techniques into two main groups: linear and non-linear noise removal methods. Linear noise removal methods are not spatially adaptive and are based on the direct change of observed values according to filter settings (e.g. Wiener filter [Rangayyan & Neuman, 2005]). Linear methods radically reduce the amount of noise in the image, but also over smooth and thus lose fine image detail (information from high frequency spectral components – similar to noise). Therefore, the main focus of this study will be on non-linear noise removal techniques that search the neighbourhoods of the observed values to determine the most appropriate values. These techniques, have greater potential to deal with the non-additive nature of noise in medical images [Gravel et al., 2004]. The evaluation of image denoising techniques is relatively straightforward with the use of technical parameters such as Peak Signal-to-Noise Ratio (PSNR) and Signal to Mean Squared Error (SMSE). However, use of these parameters to evaluate medical images is limited, as it is rather difficult, if not impossible, to distinguish the true (relevant) signal value from signal affected by the noise. Therefore, raw images that come directly from clinical imaging modalities are usually preferred due to the possible loss of relevant detail during denoising procedures [Bedi & Goyal, 2010].

To date, hundreds of denoising methods with numerous modifications have been developed and their review is outside the scope of this study. Therefore, in the following sections there will only be an overview of commonly used image denoising methods: Order-statistic techniques, Diffusion-based techniques, and Transform-based techniques.

Order-statistic techniques

The class of order-statistic filters is relatively large. The concept of this class of filters is based on the assessment (ranking-order) of pixel values in the determined neighbourhood of the processed pixel from the minimum to the maximum. It includes not only the median filter, which is probably the most commonly used clinical noise removal technique, likely due to its well understood nature (see Table 4 [Schalkoff, 1989]) but also additional filters such as: Max filters, Min filters, Max/Min filters, α-trimmed filters and L-filters. Max filters remove low-valued impulsive noise whereas Min filters remove high-valued impulsive noise. Max/Min filters combine characteristics of both previously mentioned filters. The α-trimmed filters are hybrid of the mean and median filters. They discard the most atypical pixel values (based on alpha parameter) and calculate mean value from the rest. At last, L-filters comprise weighted combination of all previous filters. However, the preference and broad usability of median filters is mainly due to their simplicity of implementation and their ability to provide relatively reasonable results.

Table 4: Properties of the standard median filter

The median filter reduces the variance of image intensities
Intensity oscillations with a period less than the window width are smoothed
Median filters will change the image intensity mean value if the spatial noise distribution in the image is not symmetrical within the window
Median filters preserve certain edges shapes
Median filters preserve the location of edges if the window is symmetrical
Median filters do not generate new gray-level values
The shape of the median filter window may affect processing results

The standard median filter (SMF) provides reasonable results while removing the impulsive type of noise. However, its performance declines when the noise is signal-dependent. Hence, several modifications to the standard SMF have been introduced since its introduction for time series analysis [Tukey, 1977] and its extension to 2D space [Pratt, 1978]. Between the first variations, the techniques based on threshold decomposition and stack filters [Fitch et al., 1984; Lin et al., 1990], separation of dimension [Nodes & Gallagher, 1983], weighted windows [Brownrigg, 1984; Arce, 1991], Max/median filtering [Arce & McLoughlin, 1987], FIR-median hybrid filtering [Nieminen et al., 1987] and a multistage (preserving details in horizontal vertical and diagonal directions separately)

approach [Arce & Foster, 1989; Arce, 1991] have been developed. These techniques more preserve desired image details and, in the case of weighted median filters, allow control of the filtering performance at the expense of somewhat lower noise suppression. The desire to preserve fine image details while still using powerful median filtering characteristics led to the introduction of switching-based median filters. Based on defined selection criteria (noise level estimator), these filters first detect pixels corrupted by noise and then apply the SMF to the pixels while preserving the rest of the image pixels [Luo, 2006]. More recently, switching median filters have been enhanced by using a fuzzy statistic-based approach [Eng & Ma, 2001; Toh & Isa, 2010].

Diffusion-based techniques

Diffusion-based denoising methods were introduced by Perona and Malik [1990]. They adapted the isotropic heat diffusion equation for image denoising and for preserving important image features. In an iterative process, the values of particular image pixels are "diffused" with pixels from their neighbourhood. However, the diffusion-based approach is limited by assessment of a diffusion conductance. The conductance is only controlled by local value gradients in each of the particular neighbourhoods and does not provide any image-dependent estimation for selection of the optimal gradient diffusion flow magnitudes. Furthermore, the original approach does not use any morphological or structural information to control the diffusion in different locations of the processed images (fine structures or smooth regions).

Over the last two decades, researchers have tried to overcome the limitations of the original seminal approach by introducing a number of modifications. For instance, the diffusion conductance was controlled by: local neighbourhood value gradients estimated from the smoothed image [Catte et al., 1992], magnitude and direction of gradients [Weickert, 1999], use of a modified optical flow technique model with generalized intensity constraint [Monteil & Beghdadi, 1999], or use of a robust estimator to preserve sharper boundaries [Black et al., 1998]. Furthermore, additional modifications were introduced to control the level of the diffusion, which is different for edge areas and the inside of the smoother regions. These modifications comprise introduction of a measure of local scale described by: the minimum reliable scale of Gaussian kernel above the required critical value (defined threshold) [Elder & Zucker, 1998], or the estimation of rough object size using so-called digital hyperball or hyperellipsoid definition [Saha & Udupa, 2001]. More recently, the "ball" scale of the hyperball scaling approach was further modified into generalized scale object size estimation that allows effective diffusion along edges and elongated structures [Souza et al., 2008]. In another study, Gilboa et al. [2004] modified the standard diffusion approach to a more general framework by introducing extensive analysis in a complex domain called a nonlinear complex diffusion. In the nonlinear complex diffusion, the imaginary part acts as a robust edge detector and the real part behaves as a real nonlinear diffusion process. This modified approach clearly outperformed the original standard diffusion approach [Salinas and Fernandez, 2007].

Transform-based techniques

The transformation of an image from space domain to the frequency domain became available after the introduction of the Fourier transform and the Fast Fourier Transform algorithm [Cooley et al., 1969]. Fast Fourier Transform is an efficient algorithm for computation of the frequency domain representation for large data sets. The idea of using the frequency domain comes from the natural behaviour of natural images where step changes (high frequencies) are usually scarce while scenes that vary slowly and smoothly across the image space (low frequencies) are present in abundance. Therefore, frequency domain-based filtering (i.e., removal of high frequency components) may provide better results in some applications (e.g., mimicking natural bone geometry and structure), ideally with guidance from *a priori* information about noise characteristic frequency components [Rangayyan & Neuman, 2005]. However, that is not entirely true and the simple suppression of high frequency components may lead to the removal of some of the desirable image components along with the noise components. For the construction of low-pass (smoothing effect), high-pass (sharpening and extraction of edges), selective band pass or comb filters (periodic artefact removal), the Butterworth or Wiener designs are usually used [Rangayyan & Neuman, 2005; Dhawan et al., 2008].

Although the Fourier transform has proved to be a very useful tool to study signals and images in the frequency domain, it lacks the possibility of space localization. Therefore, a theory for multiresolution signal decomposition, known as Wavelet transform (WT), was introduced, allowing space localization of image spectral components [Mallat, 1989]. The main advantage of the wavelet approach is sampling interval variation. Through sampling interval variation, WT can perform multiresolution analysis for event localization in data over time or space, with respect to all frequency components. This proved to be very useful for the effective definition of specific image features and image classification [Dhawan et al., 2008].

Denoising based on WT was used and studied more extensively after the introduction of the soft-thresholding approach by Donoho [1995]. Soft-thresholding uses a defined threshold value at coarser scales of decomposed image data to reduce the values of the wavelet coefficients obtained after wavelet decomposition of the original image. Soft-thresholding provides a simple and relatively successful noise removal tool. However, the setting of the optimal threshold value remains challenging unless noise characteristics are known. Therefore, researchers introduced methods for automatic noise level estimation and adaptive threshold setting based on image characteristics, often modelled by Bayesian framework [Chang et al., 2000; Sanches et al., 2008]. These techniques treat wavelets coefficients as random variables and try to estimate noise free images with the use of various prior models in adjacent neighbourhoods from the wavelet coefficients of given (noisy) images (e.g. generalized Gaussian priors [Gupta et al., 2005], generalized Laplacian priors [Pizurica & Philips, 2006], Markov Random Fields [Malfait & Roose, 1997], non-Gaussian priors [Sendur & Selesnick, 2002] or scale mixtures of Gaussians

[Portilla et al., 2003]). However, these methods are usually time consuming and require high computational power when dealing with minimization of a maximum a posteriori optimization function. More recently, WT- based denoising was enhanced by substitution of dependence on statistical estimation of noise-free images for dependence only on estimation of noisy data [Luisier et al., 2007]. Other refinements were the introduction of the modification of the WT itself [Luisier & Blu, 2008] and so-called Fractal-based denoising approaches (scaling and copying similar blocks of wavelet coefficient to the lower decomposition levels) [Zhong & Ning, 2005; Ghazel et al., 2006].

2.4. Segmentation Techniques

Image segmentation is defined as the process of identifying and delineating objects in images [Udupa et al., 2006]. In other words, image segmentation analyses the image data and then subdivides the regions/objects of interest in the images so that parts with similar/cohesive properties are grouped together. Image segmentation is an important part of any image analysis as proper segmentation process is crucial for following assessment of the image data (e.g. recognition, classification, registration and 3D visualization). However, despite extensive research during the last decades, accurate segmentation remains a difficult task [Sharma & Aggarwal, 2010]. The accurate categorization of image segmentation techniques is as challenging as that of denoising methods, and different authors divide segmentation techniques in different ways [Heinonen, 1999; Withey & Koles, 2007; Sharma & Aggarwal, 2010]. We will follow the line proposed by Heinonen [1999] and divide image segmentation techniques according to their operational principle. This study will consider region-based techniques, boundary-based techniques, mathematical morphology-based techniques, active contours-based techniques and classification-based techniques. In reality, however, the majority of the techniques used in medical image analysis are hybrid to significantly improve the segmentation performance and robustness of approach used [Yoo, 2004]. The evaluation and quantitative comparison of different image segmentation algorithms is difficult because of the number of parameters and factors that have to be taken into account. Nevertheless, a few researchers have suggested a practical evaluation framework for segmentation algorithms that considers the precision, accuracy and efficiency [Udupa et al., 2006].

Threshold-based techniques

Thresholding is the most used segmentation method for clinical images nowadays. Its wide popularity is partially due to its simplicity, effectiveness and clear interpretation of segmentation outcomes. The technique is based on the use of the pixel gray-level values of the image data to find a group of data with similar pixel values [Sahoo et al., 1988; Sezgin & Sankur, 2004]. Such a procedure can either use a simple single-threshold approach that creates binary images or, more often, a multiple-threshold approach that clusters image data into more classes defined by the number of thresholds used.

Despite its simplicity, finding the optimal threshold values, to accurately differentiate regions of interest, is not a simple task. When applied to noisy data, improperly set thresholds can cause larger image assessment errors. Therefore, over the past decades, many approaches have been developed to replace operator intervention and automatically estimate the optimal threshold values for particular image data sets [Sezgin & Sankur, 2004]. These approaches can be assigned into the following categories. Image histogram shape-based techniques determine the optimal threshold setting by analyzing of the shape properties of the histogram (detection of peaks and valleys or concavities) [Tsai, 1995; Guo & Pandit, 1998]. Clustering-based techniques cluster divided data into two classes that comprise foreground and background pixels. The threshold value is set according to analysis of the mean variances of these two classes [Leung & Lam, 1996], the weighted sum of the inner class variances [Liao et al., 2001], the minimum error of misclassification [Moser & Serpico, 2006] or variances with the use of the fuzzy clustering approach [Jawahar et al., 1997]. In Entropy-based techniques, the optimal threshold is found when the entropy of a segmented image is maximized [Yen et al., 1995] or the cross entropy between the original and segmented image is minimized [Yin, 2007]. Further, the fuzzy classification entropy approach considers inter- and intra-class variance to determined the optimal threshold selection [Liu et al., 2006]. The spatial probability-based techniques take into account local characteristics of an image and select the optimal thresholds based on an analysis of the local histograms [Li et al., 1997; Yang et al., 2005] or the nearest-neighbour statistics of each particular pixel in an image [Sauvola & Pietikäinen, 2000; Vasilic & Wehrli, 2005]. More recently, Batenburg and Sijbers [2009] introduced an approach for tomography-based images where the optimal thresholds are computed by minimizing the distance between the forward projection of the segmented image and the measured projection data. However, in spite of decades of research on this topic, threshold-based segmentation still remains vulnerable to image inhomogeneity, artefacts and noise present in the image data.

Boundary-based techniques

The boundary-based approaches use the *a priori* information that abrupt changes in pixel intensity values occur at boundaries between regions in the image data. First the algorithm detects these localized discontinuities (edge detection), commonly found with different gradient operators, and provides candidates for possible edges of objects of interest. These operators usually use the first-order differentiation (e.g., Sobel or Roberts operators) or the second-order differentiation (e.g., Laplacian operator) [Sonka & Fitzpatrick, 2000], although some more sophisticated techniques based on Fuzzy, Genetic and Neural Network approaches were recently introduced [Senthilkumaran & Rajesh, 2009]. Second, a thresholding procedure is applied to suppress spurious and enhance coherent edges. Finally, segmentation is completed by assembling the enhanced edges into closed and meaningful object boundaries [Cheng et al., 2001]. During the past decade, several algorithms have been developed for boundary assembly. These algorithms are based on

border tracing [Yao et al., 2005], graph searching [Sanfeliu et al., 2002], shortest spanning trees [Kwok et al., 2004], or the EdgeFlow approach [Ma & Manjunath, 2000]. Nevertheless, mainly due to the common presence of noise and the inhomogeneity in natural image data, completing fully closed boundaries is still a difficult task and methods tend to result in spurious or broken edges [Rangayyan & Neuman, 2005]. As a result, the edge detection algorithms are usually combined with region-based segmentation approaches to reduce their susceptibility to noise and to provide additional spatial information about the object of interest in the image data [Sonka & Fitzpatrick, 2000; Cheng et al., 2001].

Region-based techniques

In comparison to threshold based segmentation, region-based techniques use information about neighbourhood characteristics and try to create connected regions from pixels that correspond to the same object in an image. These characteristics can comprise a similarity of pixel intensity values, spatial proximity and connectedness. However, real images rarely include only homogenous objects. Therefore, the standard approach was extended with the use of statistical properties (e.g. mean, variance) or homogeneity criteria [Sonka & Fitzpatrick, 2000; Yoo, 2004; Rangayyan & Neuman, 2005]. Region-based techniques can be divided into two groups including region growing-based approaches and region splitting/merging algorithms.

Region growing segmentation, proposed in its simplest form by Adams and Bischof [1994], starts from the selection of a seed point (usually defined by the operator) that is located in a region of interest. The nearest neighbourhood is then searched. If neighbourhood pixels satisfy the defined criteria of homogeneity (similar pixel intensity values), the region grows by including these pixels into the region. The correct setting of seed points, the proper homogeneity criteria and the order in which neighbourhood pixels are explored are the main challenges of this approach. Therefore, since the original work where region growing relies only on absolute pixel intensity values, the approach has been further improved by integrating boundary information [Fan et al., 2001], the use of so called fuzzy-connectedness [Udupa et al., 2002], Bayesian probability [Fan & Xia, 2001] or Markov random fields [Qin & Clausi, 2010] into homogeneity criteria, or by methods for automatic seed selections [Stewart et al., 2002; Fan et al., 2005].

Region splitting and merging segmentation algorithms are based on two successive steps. The first step involves the splitting of the image data into square sub-regions that contain only pixels fulfilling the criterion of homogeneity. In the second step, the neighbourhood sub-regions with similar characteristics can be merged, if the second criterion of homogeneity is reached. The main problem with this approach is that the image segmentation can result in a blocky (square) representation [Sonka & Fitzpatrick, 2000]. This problem can be reduced by splitting the regions into substantially smaller sub-regions [Cheng et al., 2001] or through the use of Delaunay triangulation instead of the standard

square-block approach [Gevers, 2002]. This is, however, at the cost of largely increased computational expenses.

Mathematical morphology-based techniques

Mathematical morphology-based techniques are used for geometrical structure analysis in image data. The analysis is performed by using a collection of morphological segmentation operators that are based, in principle, on a type of distance transform and the chosen structuring elements. The main operators are morphological erosion and dilation. In combination with each other, these two main operators can further provide additional morphological transformations such as morphological opening (erosion followed by dilation) and closing (dilation followed by erosion). These operators have been originally defined in binary images. However, the approach was later extended for gray-scale applications [Sonka & Fitzpatrick, 2000; Pesaresi & Benediktsson, 2001]. Morphological operators cannot be applied directly as a segmentation tool. However, in combination with edge detection or region-based approaches, they can provide an effective approach for image segmentation, the so called Watershed transformation [Vincent & Beucher, 1989].

The Watershed transformation is based on the detection of border-lines that separate distinct regions in the image data. Usually, this approach is not applied directly to image data, but it employs distance transform or gradient operators that create a "topographic surface" of the image data and then detect potential borders between segmented regions [Vincent & Soille, 1991]. The topographic surface corresponds to the measure of local elevation (slope) of the pixel intensity values. Then, the segmentation starts by "flooding" this topographic surface from the regional minima, called basins, to the highest (peak) pixel intensity value. The algorithm further raises the "water level" uniformly over the topographic surface. When the "water" rises to a curtain level, neighbourhood basins start to merge and a dam is built to prevent it. The segmentation process ends at the moment when no other dam needs to be constructed, the topographic image is fully flooded, and only the tops of the dams, called watershed lines, are visible [Sonka & Fitzpatrick, 2000].

The original approach to watershed-based segmentation that uses gradient information is usually slow and leads to over-segmentation. This is mainly caused by the high sensitivity of the gradient operators to noise level or image texture that produce large amounts of non-relevant regional minima and maxima in the topographic image. To overcome these problems, common filtering (e.g. low-pass filter) can be applied to image data before using the gradient operators. As a result, the number of segmented regions is lowered. However, some relevant information, presented by high-frequencies, might be lost [Pesaresi & Benediktsson, 2001]. These problems led to the introduction of a computationally efficient watershed algorithm [Vincent & Soille, 1991]. This algorithm employs *a priori* knowledge about possible patterns or structures presented in the image data and the use of so called markers that determine regions of interest [Vincent & Beucher, 1989; Beucher, 1992; Gao et al., 2001]. In practice, choosing the proper region markers remains a difficult task. To

address this challenge, new watershed-based algorithms have been developed, including: the “fuzzy approach” [Bouchet et al., 2007], algorithms based on fast region merging using graph representations [Haris et al., 1998], the integration of the watershed transform with an energy minimalization technique (Watersnakes) [Nguyen et al., 2003] or with ultrametric contour mapping (OWT-UCM algorithm) [Arbelaez et al., 2011].

Deformable models

The parametric deformable model-based segmentation approach was firstly introduced by Kass et al. [1988]. The approach uses the evolution of closed smooth spline curves, called snakes, within an image so that the desired object of interest in the given image can be detected. The evolution of these original spline curves is based on energy minimization controlled by: 1) internal forces to keep desired smooth curves, and 2) by external forces derived from imaged data to attract curve evolution towards desired object boundaries. This unique approach opened a new way to image segmentation, but suffers from some inherent limitations [McInerney & Terzopoulos, 1996; Cremers et al., 2007]: First, due to the use of image gradient information as an external force, snakes are very sensitive to initialization that must be close to the desired object; otherwise, the contour can be attracted towards undesired local edges. Second, due to internal forces keeping contour stiffness, there is a limited contour evolution into boundary concavities, especially in noisy images. Third, the original approach does not allow segmentation of multiple objects in an image and thus is not able to follow topological changes in image data. Finally, the original approach is computationally expensive.

Over the past two decades, several authors have introduced new approaches to solve problems with the necessary initialization close to the desired object and the difficult progression into boundary concavities. These new approaches include: methods based on the implementation of pressure forces [Cohen & Cohen, 1993], control points [Davatzikos & Prince, 1995], directional attraction [Abrantes & Marques, 1996] or the Gradient Vector Flow approach [Xu & Prince, 1998]. However, multiple object segmentation still remains difficult for the original approach in the absence of *a priori* statistical knowledge on the shape of the desired object [Cremers et al., 2002].

Therefore, to address the multiple segmentation problem, Osher and Sethian [1988] introduced the geometrical active contour model, so called level set. Level sets employ implicit geodesics or minimal distance curves where contour propagation is controlled by its curvature. In comparison to snakes, this approach is less sensitive to initialization and local noise. In addition, level sets naturally allow the following of topological changes and the use of probabilistic interpretations to the segmentation approach (including criteria as colour, texture, motion, etc.) [Suri et al., 2002; Cremers et al., 2007]. However, level sets also have drawbacks: possible segmentation failure if an object is embedded in another one; possible evolving contour “leaks” due to object boundary gaps; and segmentation failure caused by curtain singularities (e.g. protrusion, indentation) [Kimia et al., 1994].

Since their introduction, many level sets variations have been developed to incorporate stopping criteria (e.g. gradient-, edge-, area minimization-, curvature-based), clustering, Bayesian statistic, shape prior or image statistic [Suri et al., 2002; Cremers et al., 2007]. Recently, such algorithms have evolved to provide better performance while dealing with: noisy and topologically complicated images [Chan & Vese, 2001; Chen, 2010; Shen et al., 2011], images containing intersecting objects and occlusion [Ali & Madabhushi, 2012], images with different texture, colour or shape patterns [Karoui et al., 2010]. Furthermore, a new algorithm have been developed to eliminate the need for reinitialization and to radically reduce computational costs [Li et al., 2010].

Classification-based techniques

Classification-based segmentation techniques are typically used for the segmentation or pattern recognition of multispectral image data or image data from where multiple features such as pixel intensity, gradient, spatial properties or texture measures can be derived [Sonka & Fitzpatrick, 2000]. These multiple features are grouped into feature vectors in the feature space that is related to the original data. The classification procedure analyses the feature space and assigns labels to all feature vectors that characterize particular classes/subgroups (e.g. particular tissues) in the image data in the best way [Bezdek et al., 1993; Sonka & Fitzpatrick, 2000]. Classification techniques can be categorized as supervised and unsupervised (so called clustering techniques.

Supervised classification techniques are initialized via user identification and labelling of samples of the physical classes (tissues) that need to be segmented. First, these samples are used for the algorithm training and then the trained algorithm performs automatic classification of all the remaining image data. Unsupervised classification techniques group data with similar features vectors together and create natural structures (subgroups) in the image data without the need for initialization. However, at the end of the unsupervised classification, the user must identify particular subgroups and assign desired labels (e.g., corresponding to particular tissues) [Bezdek et al., 1993; Sonka & Fitzpatrick, 2000].

Supervised classification techniques can be further divided according to parametric or non-parametric design. The supervised parametric methods are based on Bayesian design with the use of maximum likelihood estimation [Lei & Udupa, 2003]. These classifiers use a mixed population (weighted summation) of a statistical distribution that contains prior probabilities and conditional probability density functions corresponding to each of the selected classes.

The advantage of non-parametric methods is that they do not require any prior knowledge about the statistical properties of the image data. The group of non-parametric methods include: a simple minimum distance classifier, the well known Parzen window, k – nearest neighbourhood classifiers (k-NN) or classifiers based on decision tree theory. The simple minimum distance classifier searches for the minimum Euclidean distance of a given

feature vector from the selected class mean vectors [Sonka & Fitzpatrick, 2000]. The Parzen window directly estimates a posteriori probabilities from samples in the given window [Fukunaga & Hayes, 1989]. The k-NN classifiers perform direct classification depending on the majority of the occurring classes in the nearest neighbourhood [Goin, 1984]. At last, classifiers based on decision tree theory use the Fisher's linear discriminant to classify given image data [Ouyang et al., 2009].

Despite their high classification accuracy, non-parametric supervised classification methods usually require large computational times to find the differences between the training sets and the tested patterns, are vulnerable to noise and need good quality training sets in order to perform reasonably well. To overcome these limitations, the popular k-NN algorithm has been modified recently in order to improve classification effectiveness by incorporating similarity criteria [Chen et al., 2009; Liu et al., 2013] or by decreasing the need for large training sets [Huang, 2006; Triguero et al., 2012].

Unsupervised classification (clustering) is used in many situations where accurate labelling of the test samples and algorithm training are difficult or computationally expensive [Jain et al., 2000; Xu & Wunsch II, 2005]. The main goal of the clustering techniques is to divide an image into a finite number of natural subgroups (clusters) that are sufficiently homogeneous, according to the defined measure of similarity without operator involvement. The desired result is that the similarity is greater between objects within the same cluster than between objects from different clusters. The clustering is difficult, as the process can produce clusters with many different sizes and shapes, and fritter regions with holes or containing a single pixel as its own region. However, the advantages of speed, reliability and consistency in large data organization were the driving forces behind the recent development in this field [Jain et al., 2000]. The group of clustering classifiers comprises a large number of methods and their combinations that are based on squared error criteria [Kanungo et al., 2002; Chang et al., 2009], graph theory [Silva & Zhao, 2012; Peng et al., 2013], fuzzy theory [Cai et al., 2007; Yu et al., 2010], the use of Markov Random Fields [Wu & Chung, 2007; Lin et al., 2011], the neural networks approach [Ou & Murphey, 2007; Sowmya & Rani, 2011], the kernel approach [Tuia et al., 2011; Huang et al., 2012] and/or the use of rough sets [Mitra et al., 2006; Hassanien et al., 2009]. A detailed description of the particular techniques is beyond the scope of this study; however, more information can be found in large survey studies [Bezdek et al., 1993; Jain et al., 2000; Xu & Wunsch II, 2005].

2.5. Bone Analysis

2.5.1. Current status

Accurate assessment of bone structural traits (BMD, cross-sectional geometry, cortical thickness, trabecular architecture, etc.) from medical images is challenging due to limitations imposed by image resolution, noise and movement artefacts. For clinical

settings, the results must be sufficiently reproducible and tangible. Hence, many sophisticated image-denoising algorithms have not been applied in bone clinical research due to their complexity, ambiguity of results and possible loss of fine detail, resulting in analysis distortion. Although some new denoising methods for different clinical applications based on statistical approaches have been developed [De Pierro & Yamagishi, 2001; Villain et al., 2003; Qi, 2005; Afonso et al., 2010], the previously mentioned issues limit their use in clinical imaging. Therefore, order-statistic, linear and local linear filters, whose influence on image data is well known and understood, are extensively used in clinical bone research [Bedi & Goyal, 2010; Krug et al., 2010].

Image segmentation is just as challenging as denoising. Limited image resolution level, presence of noise and usually unknown ground truth limits the performance of many of the previously mentioned techniques (see section 3.4). Therefore, hybrid techniques (e.g., [Kang et al., 2003; Zoroofi et al., 2003]) are used to improve the–segmentation performance [Sonka & Fitzpatrick, 2000; Yoo, 2004; Rangayyan & Neuman, 2005; Sharma & Aggarwal, 2010].

Plane radiography

Over the past decade, several methods have been developed for bone analysis in x-ray based modalities. Behiels et al. [2002] developed a method for accurate bone boundary segmentation in digital radiographs. The Behiels method is based on fitting active shape models (snakes) that are driven by a least square fit of the contour shape towards the bone boundary position. The boundary position is determined by statistical analysis of the image intensity in the neighbourhood of each point within the contour. The contour is further updated by a step-wise smoothing criteria to correct contrast or inhomogeneity in the bone border-line. Furthermore, three other approaches use variations of active contour methods for accurate bone segmentation in plane radiology.

Ballerini and Bocchi [2003] introduced a method for the automatic segmentation of hand radiographs to simplify bone age determination. Given *a priori* knowledge about hand structure, the method uses a multiple genetic snakes approach that segments the hand bone borders without the need for manual initialization and optimalization. McInerney [2008] introduced a method for fast and interactive 2D segmentation applicable in plane radiology and MRI images. The method uses curve snakes with sketch-line initialization. Although the method is apparently very fast, it requires time consuming user initialization close to the borders of segmented objects, the setting of various control parameters and occasionally some editing after the segmentation procedure. Diop and Burdin [2013] developed segmentation method for bi-planar radiographs that uses variational geometrical active contours with shape priors (i.e., shapes derived from prior rough segmentations) to reveal bone shape deformation and to improve 2D/3D reconstruction. The shape priors are used to improve the segmentation procedure in regions with poor contrast or blurred bone

border-lines. However, this guided approach still fails to correctly segment regions near joints where occlusion phenomena can appear.

Computed tomography – outer boundary detection algorithms

Yao et al. [2005], developed an algorithm for bone edge detection based on multiscale estimation and correction of the normal direction of bone edges. The method uses windowing function and weighted edge strength calculations to determine the initial seeds corresponding to the strong edge points. With these points, the local normal direction of bone is estimated and this information is used in the modification of a simple Canny's edge detector to improve the detection process. The authors report a relatively high accuracy level (~95%) for their automatic segmentation method compared to manual segmentation. However, the comparison was performed on only 9 CT images of the human pelvis. Furthermore, the authors reported that the algorithm could fail in the presence of a high noise level in the image or complicated bone structures.

Klinder et al. [2009] proposed a fully automatic, model based, method for identification and segmentation of the vertebral column. First, the method detects the spinal canal using a progressive adaptation of tube-shape segments into the spine area. Second, individual vertebrae are detected using a curved planar reformation and a generalized Hough transform models of all vertebra. Third, individual vertebrae are identified using rigid registration of vertebra shape models to detected vertebrae based on similarity criteria. In the last step, the final segmentation of the vertebral column is performed by modification of the shape models using a modified shape-constrained deformable model approach. This approach incorporates spatial relations between particular vertebral bodies and corrects inaccuracies caused by shape model fitting. The technique is fast and can segment the whole vertebral column in arbitrary CT images. However, the technique requires the training of models prior to the segmentation procedure and the method tends to overestimate segmentation results (the mean point-to-surface segmentation error ~1.2 ± 1 mm).

Zhang et al. [2010] introduced a fast segmentation method that uses 3D adaptive thresholding. Initially, standard thresholding is applied on 3D CT data. Then, inaccurate initial segmentation (holes in the bone structure, etc.) is corrected by applying a modified k-means clustering algorithm to 2D slices, then, the locally adaptive 3D thresholding is applied using a 10-voxel neighbourhood structure. The authors stated that the proposed algorithm overcomes the simplified method developed by Kang et al. [2003] with accuracy reaching nearly 100% for the correct vertebral segmentation. However, the authors have only used one set of data and only one radiologist performed manual segmentation (taken as a gold standard in this study). Therefore, the accuracy level reached is questionable.

Computed tomography – 3D volume segmentation

Kang et al. [2003] describe a method for accurate segmentation of skeletal structures such as the proximal femur, the knee and the skull, as well as the separation of cortical and trabecular bone. First a rough segmentation is performed by a region growing procedure that uses local adaptive thresholds with correction based on local neighbourhood information. Second, boundary discontinuities are closed with the use of morphological closing with a spherical element. Third, anatomically oriented boundary adjustment is performed to refine small, previously smoothed, structuring elements by application of a threshold set to 50% of the maximum bone border intensity profile (the periosteal line) in a particular image slice

Fourth, the endosteal border is determined using a 50% threshold of the endosteal intensity profile. In the last step, the final refinement of segmented structures is performed using morphological closing and hole-filling operations. The method needs the selection of a seed point for every selected bone for segmentation, specific separation of regions the process should avoid and occasionally, manual intervention for some specific cases. The authors stated high accuracy levels for the estimation of geometric parameters up to one voxel.

In contrast with the previous study, Zoroofi et al. [2003] introduced an automated method for particular segmentation of the pelvis and the femur. The method uses four steps. The first step includes operator selection of the femur-pelvis area, re-sampling of voxel dimension, smoothing of image data by 3D Gaussian filter and histogram-based thresholding using Otsu threshold criteria, binarizing and the application of morphological closing and filling holes operations. In the second step, the rough estimation of the joint space is performed with a model-based approach. The third step employs Hessian filtering for the image enhancement of the hip joint space. In the fourth step, this information is further used for refining bone borders with a Moving disk technique. The authors stated that their approach is of clinical use. Nevertheless, the proposed approach fails to properly segment (according to the authors' selected criteria of accuracy) up to 50% of cases where the femoral head did not have an ellipse-like shape or it was difficult to differentiate between the femur and the pelvis due to limited cartilage area, unevenly distributed bone tissue or poor quality of image data.

In a study by Treece et al. [2010], the authors describe a method that should be capable of the unbiased segmentation of cortical bone from low resolution clinical CT data into a sub-pixel range. The method uses a mathematical model of bone anatomy and an imaging system that is derived from the work of Prevrhal et al. [2003]. The method takes into account the in-plane point spread function (PSF) of the imaging system and the blurring caused by interaction between the CT slice thickness and the orientation of cortical bone. According to the authors, the method can segment cortical bone in low resolution data (~ 0.6 mm) up to a minimum thickness of 0.3 mm with estimated errors -0.01 ± 0.58 mm (mean ± 1 std.deviation) in comparison to reference high-resolution data. These results are

comparable with the work of Kang et al. [2003]. However, the method has a number of drawbacks such as using a standard thresholding procedure for the determination of cortical bone contour and; necessary correction of segmented contours due to errors caused by the thresholding procedure (thus more time needed for analysis). Also, usage of a fixed cortical CT over the whole image slice is unlikely realistic and could be the cause of largely overestimated cortical thickness (over 4 mm) in some test cases.

Mastmeyer et al. [2006] proposed a hierarchical 3D technique for the segmentation of vertebral bodies. The method includes the separation of particular vertebrae based on markers' placement followed by elliptic cylinder cuts. Then, the level set method (balloon model) is applied for the coarse segmentation of vertebral bodies in the selected cuts followed by the locally adaptive volume growing and morphological refinement of the bone border-line. Finally, the separation of the cortical shell is performed with the same technique as described by Kang et al. [2003], followed by optional peeling for the correction of inaccuracies in automatic segmentation that cannot separate the sub-cortical bone from the cortex. The method allows the accurate selection of the standard geometrical volumes of interest (VOI) used for analyses of trabecular bone. However, the method depends on initial operator initialization and the correction of selected regions of the vertebral bodies (necessary in up to 20% of cases with a normal data set; up to 40% with a complicated data set such as cases with tumours, fractures, etc.) The precision errors for segmented volume, BMD, and coordinate system position were below 2.0%, 0.6%, and 0.7%, respectively. Nevertheless, the accuracy of the correct separation of the cortical and trabecular bone compartment is dependent on a reconstructed field of view [Engelke et al., 2009].

High-resolution CT

In the work of Elmoutaouakkil et al. [2002], the authors presented an algorithm designed for the segmentation and separation of cortical and trabecular bone of the lumbar vertebra. The method includes the detection of a skeleton from the crest lines of the bone architecture using the gray-scale watershed transform. The skeleton is then iteratively expanded (perpendicularly to the skeleton) in order to find the remaining pixels in an image, corresponding to bone, under selected contrast criteria. In the next step, the cortical and trabecular compartment is separated with a clustering process that uses *a priori* information that at least two neighbourhood pixels of the examined pixel, specified in 3x3 neighbourhood, belong to the same class (cortical or trabecular bone). The authors reported relatively high accuracy of the segmentation when applied on a vertebral phantom image (~95%) and mentioned the higher stability of the results in comparison with commonly used segmentation methods (threshold-based techniques). Nevertheless, the relative error in the estimation of bone micro-structure traits was up to 20%. Moreover, the method requires an adjustment of two parameters for its optimal performance.

A method proposed by Buie et al. [2007] deals with the automatic segmentation of cortical and trabecular bone in HR-pQCT images of the distal tibia and the radius. The method is based on the use of a dual threshold technique that requires the setting of two threshold levels for detecting periosteal and endosteal surfaces. The detected area above the higher threshold is closed by morphological operations with the utilization of a connectivity filter and introduced as a mask of the bone volume. The detected area below the second threshold within the bone volume is considered to be a trabecular region mask and is processed similarly to the previous step. The final mask for the separation of the cortical and trabecular compartment is then constructed with a combination of these two created masks (bone volume and trabecular mask). The results of the proposed method were validated in comparison to a semi-automated hand contouring approach. The proposed method provided a uniform surface of the cortical bone border-lines and was able to fully extract thin cortices. However, it needs an experienced operator to initiate the segmentation process correctly.

Lublinsky et al. [2007] introduced a segmentation method for the separation of cortical and trabecular bone in micro-CT images. First, the operator measures the maximum and minimum cortical thickness to restrain the following analysis. This measurement is typically performed only once per data set. Second, the bone mask in the region of interest is created. The creation of the mask is based on the following steps: (i) narrow box-shaped Gaussian kernel, (ii) pre-set high threshold level, (iii) expansion of the segmented bone area, (iv) Gaussian smoothing (with window size equal to the minimal cortical bone thickness) and finally (v) application of the second threshold with a level reduced by 7% of the pre-set level. In the next step, the bone mask is further refined by filling bone cavities and with a peeling procedure for the correct determination of the periosteal border-line. The final cortical bone extraction is based on application of a Gaussian filter with a large window size, global thresholding set to a value 6% larger than in the previous step, and peeling or filling to geometrically restrict the cortical bone area where thicknesses of the segmented cortex exceeded or failed to reach operator-measured cortical bone thickness (measured in the first step). The accuracy of the proposed method was tested in comparison with results of manual segmentations yielded from 8 raters. The authors reported that the mean difference between the data produced by the proposed algorithm and the mean of the results from 8 raters was within 2%, which was far less than the variation of individual raters (~9%).

Recently, Burghardt et al. [2010], extended Buie's segmentation technique by incorporating the detection of intra-cortical porosity. The method uses the original binarized image (i.e, image after application of a dual threshold technique) together with a mask of the cortical compartment, generated by Buie's segmentation technique, for an initial estimate of the intra-cortical pores (voxels corresponding to background that are surrounded by bone). Then, a 2D connectivity criterion is applied to select the high-confidence pore voxels used as seeds for the initialization of the hysteresis region growing

process in z-plane. Region growing searches for low-confidence pore voxels that are longitudinally connected to the previously detected, high-confidence pore voxels and marked as high-confidence pore voxels. In the final stage, all the detected high-confidence voxels are combined with the initial cortical bone mask for accurate segmentation of the cortical compartment. The authors stated that their method is superior to the standard segmentation provided by the scanner manufacturer by reporting an *in vivo* precision for the determination of the bone structural parameters to be close to 1% (excluding cortical porosity precision of ~10%) in 90% of the segmented cases. In the remaining cases, the manual correction of the bone border-lines was required, due mainly to failed cortex definition caused by incorporation of an adjacent, dense trabecular network into the cortex.

More recently, Valenthinitxh et al. [2012] developed a new automated threshold-free algorithm for cortical bone segmentation (TIST) in HR-pQCT images. This method uses supervised classification of the cortical and trabecular compartment based on modelling appearance characteristics from manually annotated cases. For the extraction of the feature vectors used in the following classification process, a 3D gray-level co-occurrence matrix, a local structure tensor and the Hessian matrix calculations were performed on 15x15x15 voxel neighbourhood for each voxel in the image data. First, training sets are created by training the classification algorithm, which consists of assigning calculated feature vectors to classes determined by manual segmentation (cortical bone, trabecular bone, background). Then, the algorithm classifies every voxel in the given image data into either cortical, trabecular or background classes based on similarity with feature vectors from the learned training sets. The results of the TIST were compared with the manual segmentation of images obtained from the same samples at a twice higher resolution level and two additional segmentation methods (standard segmentation provided by the scanner manufacturer and the method proposed by Burghardt et al. [2010]). The results show that the TIST performed similarly in comparison to the method of Burghard et al and was superior to the standard segmentation method, with a mean error ~10% in the analysis of Ct.Th and cortical volume in comparison with manual segmentation. However, the analysis using the proposed TIST algorithm was performed only on a limited number of samples (14 cadaveric samples of the distal radius) where 13 of them were always used as a training-set for the segmentation of the remaining sample. This, together with need for properly defined training sets, clearly limits the study and more research should be done to validate the performance of the TIST algorithm, especially for *in vivo* studies where various types of artefact can appear.

2.5.2. Commonly used approaches for pQCT image analysis

Since the introduction of the pQCT scanner in 1973, researchers have tried to determine the correct way to analyse images yielded by this modality. The common approach, provided by the manufacturer, is based on the use of simple thresholding and peeling procedures [Stratec, 2004a; Stratec, 2004b]. However, the correct setting of threshold levels is crucial for achieving reasonable accuracy and precision of the performed

measurements. Therefore, there have been a number of studies that investigated the relationship between threshold setting and the accuracy of the yielded results. Ward et al. [2005] and Hangartner [2007] investigated the relationship between the threshold levels used and the accuracy of the cortical bone area (CoA), mean cortical wall thickness (CoTh) and cortical density (CoD) traits with the use of phantom data. They found out that the optimal threshold value for the accurate estimation of geometric traits in CoA and CoTh is ~50% of the difference of the density between the adjacent tissues. For CoD, it is recommended that 95% of the maximum density value of the bone is used in order to prevent errors due to the presence of partial volume effect in the image data.

Since the problem with partial volume effect has been recognized, Rittweger et al. [2004], and Hangartner and Short [2007] tried to develop proper adjustment for the correction of the measured CoD trait. Rittweger et al. identified, based on phantom data, the relationship between the used threshold value and CoD and stated that with the use of their correcting equations, the error in measurements of CoD was lower than 2%. Hangartner and Short selected a different approach. They developed a correction factor of the CoD trait based on the measured width of the cortical bone. However, in the end, they reported similar errors of measurements as Rittweger et al. (<2%).

In other studies, Veitch et al. [2004] and Kontulainen et al. [2007] investigated the influence of different analysis modes, incorporated in Norland/Stratec XCT software [Stratec, 2004b], on cross-sectional area and BMD analysis. In the study of Veitch et al., the authors used manually segmented images obtained from Adobe Photoshop and ash weight, respectively for the evaluation of the accuracy of CoA and BMD analysis. They reported that the most accurate mode was contour mode 2 in combination with peel mode 5 (both modes are based on a search of the steepest gradient within a profile in the region of interest). However, their analysis resulted in only moderate accuracy for CoA (8%-20%) and for cortical BMD (10%-22%) measurements (depending on the selected region). In contrast, Kontulainen et al. used a comparison of the measured cross-sectional areas to data obtained by histomorphometry. Their results showed that the smallest mean error in total cross-sectional area (-1%) was obtained using contour mode 3 (outer threshold value set to 169 mg/cm^3), in CoA (0.1%) using separation mode 4 (outer threshold value set to 200 mg/cm^3 and inner threshold value set to 661 mg/cm^3 that was derived from the use of the inflection point method). The inflection point method is an operator-independent and defines threshold value based on determination of the greatest change in bone mineral density at the endosteal surface.

Recently, two software applications (as plugins for ImageJ software) have been developed to improve the automated analysis of pQCT image data. Laskey et al. [2010] developed a plugin for the enhanced characterization of cortical bone. The program offers semi-automatic segmentation of cortical bone based on thresholding, manual contour placing and a locally adaptive region-growing algorithm. After segmentation of the cortical bone region, the program automatically calculates cortical bone distribution relative to the centre

of mass and provides the values of the maximum, minimum and mean cortical thickness. In addition, the program quantifies deviation of the analysed cortical bone cross-section from an ideal circle and/or ellipse. The second software plugin was developed by Rantalainen et al. [2011]. The program uses thresholding for the segmentation of CoA followed by regional analysis of cortical BMD. Cortical bone is also subdivided into three concentric cortical divisions and in 36 cortical sectors originating from the centroid of the bone. This regional division is of particular interest as it has been recently recognized that the age-related bone loss in cortical bone has non-uniform distribution and most of the bone is lost from the endosteal border [Zebaze et al., 2010].

3. Aims of the study

This thesis has both technological and clinical objectives. The technological goal is to develop image preprocessing and segmentation algorithms to improve the assessment of apparent structural traits from appendicular bones yielded by pQCT. In a broader perspective, these tools are expected to lead towards a more accurate estimation of whole bone strength for various clinical research applications. The main objectives of the present study and their links to the publications of this study are listed below:

1) To develop an image preprocessing method to enhance the analysis of pQCT images and to reduce the noise level, while preserving important details in the image data:
    a. In publication I, the aim was to evaluate whether the preprocessing of clinical pQCT images could improve the analysis of bone macro-structural traits in comparison to the analysis of raw, unprocessed images, and if so, to what extent. The other objective was to investigate the signal to noise ratio (SNR) and concomitant gray-scale resolution of pQCT images.
    b. In publication II, the aim was to assess whether (i) the bone responses to bed rest were different at anatomically different sectors of the distal tibia and tibial shaft in different subgroups and (ii) whether the sector-specific responses were different during the recovery after the bed rest. The additional objective in this study was to use the developed image preprocessing method [publication I] that permitted more consistent detection of the outer and inner cortical boundaries compared with commonly used median filtering.

2) To develop an algorithm for cortical bone segmentation in pQCT images that enables robust, fast and reproducible cortical bone segmentation and automatic analysis without involvement of the operator.

    a. In publication III, the aim was to address (i) the *in vivo* accuracy of the developed algorithm in the analysis of raw pQCT images using the manual segmentation of cortical bone as the gold standard, (ii) whether the preprocessing of pQCT images could enhance the cortical analysis compared with the analysis based on raw images and (iii) to what extent the new segmentation algorithm improves the analysis of real pQCT images compared to simple density threshold-based analyses and a recently proposed advanced analysis based on a new variational level set formulation.

    b. In publication IV, the aim was to carry out a density-threshold independent analysis of pQCT-measured cortical, subcortical, and trabecular bone mineral content (BMCc, BMCsub, BMCt) at the distal tibia by applying a robust, previously developed [publication III], segmentation algorithm on the same bone areas using data pooled from four prospective immobilization trials, each with a different duration and design.

4. Materials and Methods

The rationale for the development of the particular methods described in this study was to develop reliable, simple, fast and interpretable algorithms for use in clinical bone research. In addition, based on experience from the methods mentioned in section 2, the algorithms needed to be fully automatic in such a way that no operator involvement would be needed to find the optimal threshold values or to perform manual measurements. Furthermore, the operator would need no training in the use of the algorithm used.

During this study, various image preprocessing and segmentation methods for pQCT image bone analysis were developed. First, a preprocessing method for pQCT images based on gray-level transformation of image intensity histograms was developed; this method reduces noise level and generates simple datasets for following segmentation [I]. The resultant method was tested by reanalysing data from the LTBR study [II]. Second, a threshold-free segmentation algorithm (OBS) was developed to enable accurate cortical bone segmentation without the limitations inherent to common threshold-based approaches [III]. The developed image preprocessing and segmentation algorithms were further applied on pooled data from 4 immobilization studies [IV].

4.1. Phantom and In vivo data

The data used in this study were obtained from pQCT scanners (XCT 2000 and XCT 3000, Stratec Medizintechnik GmbH, Pforzheim, Germany). Images from four phantoms made of small plastic containers containing three known concentrations of homogenous K_2HPO_4 solutions (50, 100, 250 mg/cm^3) and tap water were used for the evaluation of the noise in pQCT images [I]. Next, data from 25 healthy volunteers taken from repeated images of the distal tibia (5% of the estimated tibial length from the distal endplate) were used to test the performance of the preprocessing and segmentation methods [I, III]. For both data sets, the

pixel size was 0.5 x 0.5 mm, the slice thickness was 2.5 mm and the translational scan speed was 30 mm/s. An example of an image from the four phantoms and an unprocessed pQCT image from the distal tibia is shown in Figure 2. In addition, the developed methods were tested in various clinical applications. The preprocessing was applied to data obtained at the 4% distal site from 23 healthy volunteers that underwent the LTBR (90 days of immobilization with 6-degrees head tilt down [Rittweger et al., 2005]) [II]. Furthermore, both the preprocessing together with the developed image segmentation algorithm were applied to data from the distal tibia (4% of the estimated tibial length from the distal endplate) from the control groups of 4 immobilization studies: LTBR – 10 volunteers underwent 90 days of immobilization with -6-degrees head tilt down [Rittweger et al., 2005]; BBR – 20 volunteers underwent 56 days of immobilization with 0 degree position of the test subject [Rittweger et al., 2010]; Valdoltra – 10 volunteers underwent 35 days of immobilization with 0 degree position [Rittweger et al., 2009]; and ULLS – 8 volunteers underwent 24 days of unilateral limb suspension [Rittweger et al., 2006] [IV]. Details of the study protocols of these four immobilization studies can be found in IV:Table 1. The pixel size of the pQCT image was 0.4 x 0.4 mm in the LTBR study and 0.5 x 0.5 mm in the BBR, Valdoltra and ULLS studies. The translational scan speed was 30 mm/s and the slice thickness was 2.5 mm in all studies. A summary of the pQCT image data used in this study is provided in Table 5.

Table 5: Summary of image data used in this study

pQCT data	Publication I	Publication II	Publication III	Publication IV
pQCT images of phantoms [Sievänen et al., 1998]	x			
Repeated *in vivo* pQCT images [Sievänen et al., 1998]	x		x	
In vivo pQCT images from LTBR study [Rittweger et al., 2005]		x		x
In vivo pQCT images from BBR study [Rittweger et al., 2010]				x
In vivo pQCT images from Valdoltra study [Rittweger et al., 2009]				x
In vivo pQCT images from ULLS study [Rittweger et al., 2006]				x

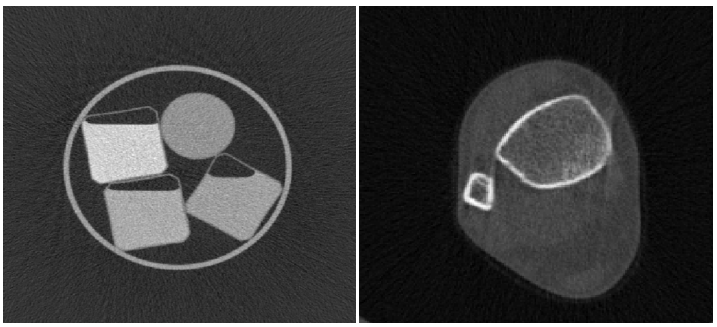


Figure 2.: Example of pQCT images of phantoms (left) and the distal tibia (right).

4.2. Preprocessing – denoising

The preprocessing of pQCT image data was performed by using non-equidistant re-quantization of the gray-scale of the image data followed by a statistical correction algorithm using *a priori* knowledge about the image data [I]. The main idea was to reduce the number of gray-levels of an image to make subsequent segmentation easier. However, due to the use of gray-level reduction, intensity corresponding to the different tissue can be falsely assigned during the transformation process. Therefore, to correct any inaccuracies caused by false assignment, the redundant wavelet transform and Markov random fields were used. Markov random fields incorporate spatial knowledge about the neighbourhood of the evaluated pixels in image data and can correct assignment according to the highest probable state. Details about Markov random fields can be found in the article [I] as well as in a flow chart of the developed method (I: Fig 1). In this study, fields with linear (LinW) and Gaussian distribution (GausW) of weight coefficients were investigated [I]. An example of the performance of the developed image preprocessing can be seen in Figure 3.

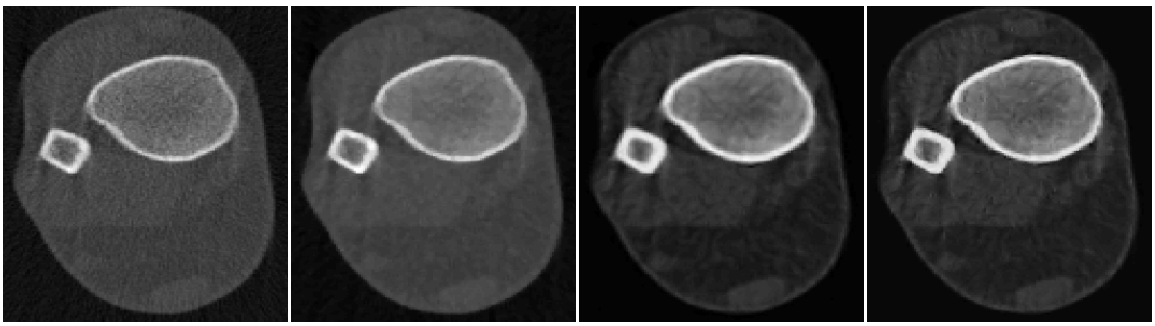


Figure 3.: Example of the different image preprocessing applied on pQCT images in this study. From left to right: the original pQCT image of the distal tibia; the resulting image after median 3x3 filtering; the resulting image after preprocessing by the developed algorithm with linear weights; the resulting image after preprocessing by the developed algorithm with Gaussian weight distribution.

4.3. Segmentation

Manual segmentation

Manual segmentation of raw unprocessed pQCT images was performed by 3 experts. The mean from these blinded independent analyses was considered to be the gold standard. The manual segmentation was implemented by in-house software developed for the segmentation and visualization of radiological images (Department of Electronics and Communications Engineering, Tampere University of Technology).

Density threshold-based segmentation

Two segmentations were performed based on standard threshold procedures used by the Stratec analysis software [Stratec, 2004a]. First, the threshold values were set to 200 mg/cm3 for the outer bone threshold and 661 mg/cm3 for the inner threshold, according to the optimal settings for cortical bone detection proposed by Kontulainen et al., [2007]. Second, the thresholds corresponding to levels automatically chosen by the contour and peel modes 2 were used. Both of these Stratec analysis protocols use 3x3 median filtering as a preprocessing method and are called C1/P2 and C2/P2, respectively.

Level set-based segmentation

In addition to simple threshold-based methods, a sophisticated segmentation method based on distance regularized level set evolution (DRLSE) developed recently by Li et al., [2010] was performed. When compared with conventional level set formulations, the DRLSE algorithm has several advantages: it allows the elimination of reinitialization; implementation is simple; large time steps significantly speed up curve evolution while ensuring numerical accuracy; and it is computationally efficient.

Proposed Algorithm for the Automatic Segmentation of Cortical Bone

The segmentation of cortical bone in pQCT image data was performed by the developed outer boundary detection and shrinking (OBS) algorithm [III]. OBS algorithm comprises a simple delineation procedure of the outer boundary of cortical bone and the subsequent shrinking procedure of the cortical pixels until the inner cortical boundary is found. This algorithm was further refined by incorporating morphological inner cortical bone boundary correction that deals with the possible “bulgy” contour caused by the blurred or lost edges of cortical bone in the image data. Moreover, two fixed density thresholds (set to 180 and 661 mg/cm^3) were used to correct and enhance cortical bone detection. The enhanced OBS algorithm was recently introduced by Cervinka et al. [2014] and its description and results are included in the thesis for completeness.

The morphological correction is triggered by values exceeding the assumed maximal cortical bone thickness (6 pixels) in detected cortical bone in a pQCT image of the distal

tibia. The algorithm detects places where “leaking” of the detected cortical shell occurred due to blurred edges resulting from insufficient resolution level, partial volume effect or movement artefact. Based upon *a priori* knowledge, thin (in sub-pixel measure) cortical bone could be assumed in place of these “leaks”. Using pixels detected in this way and taken as seed points, the morphological erosion operation is repeatedly performed until the cortical thickness in the detected place reaches a 1 pixel thin layer.

In addition to morphological refinement, two fixed threshold levels are used for the detection of falsely detected cortical bone. First, the refinement uses a lower threshold of 180 mg/cm3 to indicate places where ¼ of the pixel is filled with cortical bone. In these regions, low resolution and partial volume effects artificially increase cortical thickness (although the correct thickness is a sub-pixel measure), but decrease local cortical BMD. Therefore, pixels within the detected cortical shell that have BMD below the lower threshold are detected and the cortical shell thickness in these places is then set to a 1 px thin layer. Second, the upper threshold improves accuracy of cortical bone detection by incorporating all pixels with a high probability of containing cortical bone, (bone density level of 661 mg/cm3 or higher). Thus, the application of the upper threshold allows incorporation of high BMD pixels that were not detected by the original OBS method [III], correcting improper deletions from prior steps. The upper threshold value was selected based on strong agreement for threshold-based findings with histomorphometry results [Kontulainen et al. 2007]. The major benefits of the new method lies in its fully automatic nature (no parameters need to be adjusted) and the use of only simple segmentation algorithms. The flow chart of the enhanced OBS method can be seen in Figure 4. An example of the performance of OBS and enhanced OBS methods is depicted on Figure 5.

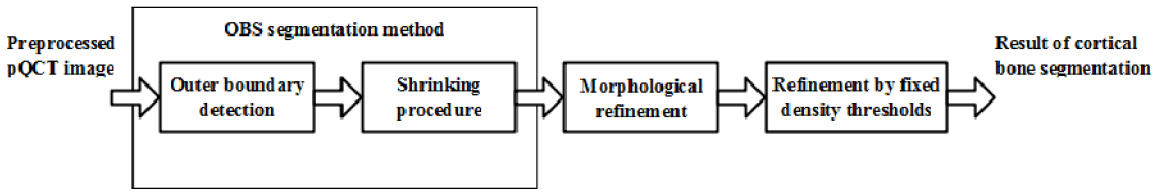


Figure 4.: Flow chart of the enhanced OBS method for cortical bone detection.

4.4. Statistical analysis

In this study, mean values, standard deviations (SD) of bone traits (inter-subject) and the SD of the differences between the repeated measurements (SD_{meas} ~ *in vivo* precision, intra-subject) are given as descriptive statistics. As another measure of the reliability of the measurements (~ *in vivo* precision), the reliability coefficient was determined as $R = 100(1 - SD^2_{meas}/SD^2)\%$. The R-value represents the error-free proportion of the inter-subject variability observed in the given bone trait [Fleiss, 1986]. To compare the performances of different segmentation approaches with the results of the manual

segmentation (which was considered a gold standard measure), the mean error as $100(X_{method} - X_{manual})/X_{manual}$ and its SD was calculated. The X denotes the geometric trait of interest. The statistical significance of the observed bone traits was evaluated with 95% confidence intervals (95%CI). For reference, if the 95%CI overlap the zero line, this implies there is no statistical difference between the manual approach and the specific automated method. In addition, as relevant measures of segmentation accuracy, the Pearson correlation coefficients and Dice similarity coefficients were determined. The Dice similarity coefficient was calculated as follows:

$$Dice\ coef. = \frac{2p}{2p+d+q} \quad (1)$$

where p is the number of pixels where the cortical bone was detected in both results from manual segmentation and the method of interest while d and q indicate the number of pixels that were detected in either the results from manual segmentation or in the method of interest, respectively. For reference, a Dice similarity coefficient is a spatial overlap index and a reproducibility validation metric where 0 indicates no spatial overlap between to segmentation results while 1 indicates complete overlap. In general, Dice coefficient above 0.75 indicates a good agreement [Zijdenbos et al., 1994].

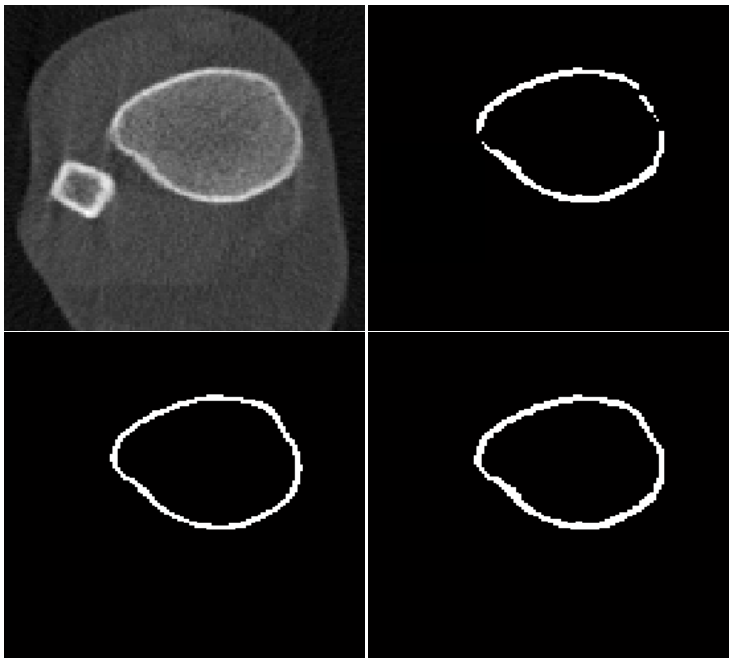


Figure 5.: Example of the performance of OBS and the improved OBS method on an image corrupted by movement artefact. Original pQCT image of the distal tibia (the upper left panel), the result of cortical bone segmentation by simple thresholding procedure with the threshold level set to 661 mg/cm^3 (the upper right panel), result of cortical bone segmentation performed by the OBS method (the lower left panel), and result of cortical bone segmentation procedure performed by the enhanced OBS method (the lower right panel).

5. Results

5.1. Performance of pQCT image analysis techniques

In publication [I], the effect of noise removal approaches applied to pQCT image data was studied in terms of the reliability of total and direction (anatomy) specific bone macro-structural traits analysis. In this work, two methods for preprocessing pQCT images were introduced and their performance tested and compared to 3x3 and 5x5 median filtering on repeated scans of four different homogeneous liquid phantoms (tap water and 3 solutions with known concentration of K_2HPO_4 – 50, 100, 250 mg/cm^3) and repeated in vivo scans of distal tibiae from 25 healthy subjects. It was shown on the phantom data that the newly introduced preprocessing improved the SNR in pQCT images by ~15 dB compared with the ~9 dB improvement of median filters. It was also shown that, when compared to raw images, preprocessing improved the reliability of the analysis of CoA by approximately 30%. Although results obtained using the LinW approach were the most consistent, no method of preference could be determined. Moreover, the preprocessing did not improve trabecular density (TrD) assessment at any anatomic region, as the reliability of TrD assessment was already high with raw images (~99%).

In publication [III], an algorithm for the automatic segmentation (OBS) of CoA was presented. The algorithm was tested on data from repeated *in vivo* pQCT scans of the distal tibiae from 25 healthy subjects. The performance of the OBS method was tested on raw pQCT images, on images after the application of median filtering and on images preprocessed using the previously developed approaches [I]. For validation, results obtained from manual segmentation were used as a gold standard. In addition, the OBS performance was compared with results obtained from commonly used threshold-based analysis and from the more sophisticated DRLSE approach [Li et al., 2010]. It was shown that the OBS method decreased the 95%CI variation by ~50% in comparison with

threshold-based analysis. In addition, the OBS method, in conjunction with the LinW preprocessing approach, agreed well with the results obtained from manual segmentation (~2% mean difference for all analysed bone geometric traits) [I]. DRLSE exhibited similar reduction of 95%CI variation to the OBS method, but DRLSE overestimated all bone geometric traits by a mean of ~15%.

Similarly, the enhanced OBS algorithm was tested on data repeated in vivo pQCT scans of the distal tibiae from 25 healthy subjects [Cervinka et al., 2014]. The results of the enhanced OBS algorithm were compared with the results of manual segmentation, the original OBS method and DRLSE. Differences in the detection of the distal tibia CoA, Pearson correlation coefficients and Dice similarity coefficients from all three segmentation methods are shown in Table 6. In general, results of all three methods were strongly positively correlated with results of manual segmentation of CoA. However, DRLSE based segmentation resulted in a nearly 15% overestimation of the CoA in comparison with 2% and 4% for the original OBS method and the enhanced OBS method, respectively. When compared with manual segmentation, the difference between the CoA and correlation results obtained from the original OBS method and the enhanced OBS method were marginal. Nevertheless, the Dice similarity coefficient showed an improvement in agreement of CoA segmented by the enhanced OBS method in comparison with manual segmentation. This improvement in the segmentation performance can be clearly seen on various examples presented in Figure 6. However, as shown by the example in the third line of Figure 6, the enhanced OBS method can fail to correctly segment CoA when severe artefacts are presented in image data.

Table 6: Descriptive data of CoA, Correlation, Dice coefficient (mean, SD) as obtained from different segmentation methods [Cervinka et al., 2014]

	CoA		
Manual segmentation	150.8 (19.2)		
Observed traits	OBS method	Enhanced OBS	DRLSE
CoA	153.0 (14.7)	156.5 (15.1)	172.6 (15.3)
Pearson correlation †	0.88	0.87	0.76
Dice coefficient †	0.83 (0.04)	0.85 (0.04)	0.81 (0.04)
Range of Dice coefficient	0.71 - 0.9	0.77 - 0.93	0.71 - 0.88

†Pearson correlation and Dice coefficients were determined between the cortical cross-sectional bone area (CoA) obtained from the manual segmentation and CoA obtained from particular segmentation methods (OBS, Enhanced OBS and DRLSE).

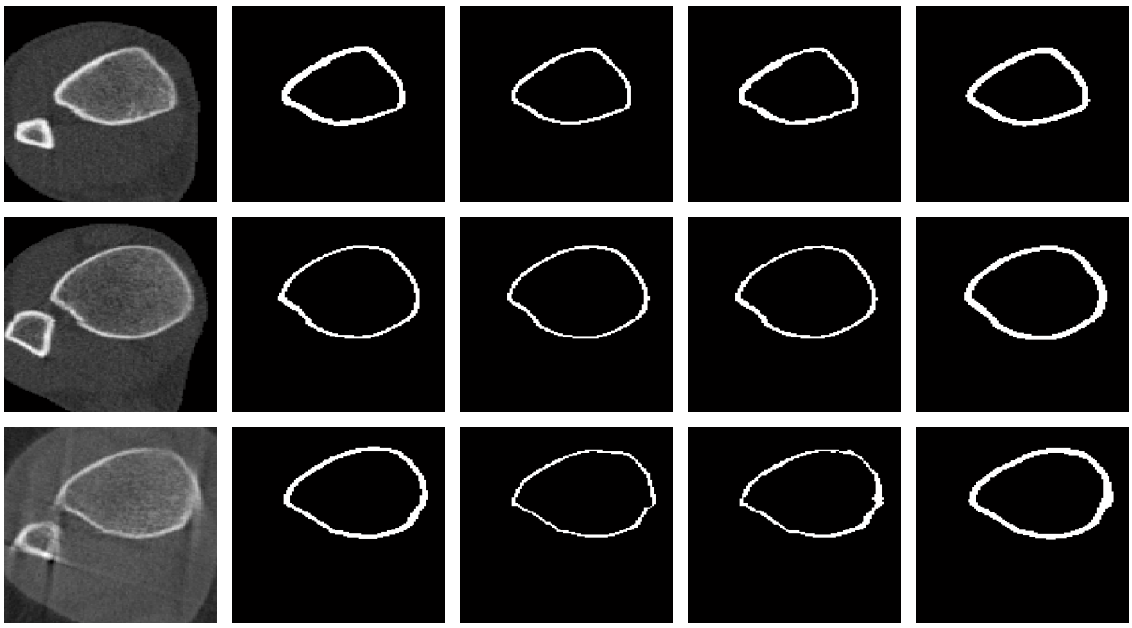


Figure 6.: Examples of the segmentation performance of particular approaches on different data sets, from left to right: Original pQCT images of the distal tibia; the results of manual segmentation; the results of the original OBS method; the results of the enhanced OBS method, and the results of the DRLSE algorithm [Cervinka et al., 2014].

5.2. Applicability of the developed methods on clinical longitudinal data

In publication [II], the reanalysis of the data from a long-term bed rest study (LTBR) was performed using the above-developed techniques, allowing more accurate sector specific bone assessment. The main goal of the study was to evaluate following hypotheses: 1) that bone loss appears mostly in bone sectors that typically undergo substantial loading during daily life (e.g. locomotion) with minor loss in other bone sectors; 2) that countermeasures used during the study can modulate sector-specific bone loss and recovery. The largest bone loss appeared to be in the trabecular bone compartment with clear modulation of bone loss depending on the countermeasure used (~3-8%). The sector-specific bone loss mostly occurred at the medial sector for countermeasure groups (~6-8%) and at the anterior sector for the control group (~10%). However, no consistent sector-specific modulation was observed during the recovery phase of the study and all bed rest-induced bone losses were practically restored.

In publication [IV], the reanalysis of the data from 4 disuse studies studies (3 bed-rest and one unilateral limb suspension study) of different durations and designs was performed using the above-developed techniques that allowed a robust reanalysis of the same subject-specific cortical, subcortical and trabecular bone cross-sectional areas throughout the longitudinal disuse studies. The primary objective was to investigate the hypothesis that a density-threshold independent analysis can reveal changes in skeletal adaptation that have

remained masked with conventional thresholding approaches. The mean absolute bone loss was systematically greater in the trabecular compartment than in the cortical or subcortical compartment. Results indicated a trend for accelerated trabecular bone loss after ~60 days of disuse but also showed a greater inter-subject variance in trabecular bone loss during the disuse period. Furthermore, bone loss seemed to continue during the first ~15-30 days after the termination of the disuse period in all bed rest studies. During this short period, bone loss became more pronounced in the cortical compartment than in the trabecular or subcortical compartment in all studies except for the longest study (of 90 days duration). In the longest study, greater bone loss was observed in the trabecular compartment rather than in the cortical or subcortical compartment (respectively: ~10 mg/mm, ~8 mg/mm, ~6 mg/mm). During the recovery phase of the studies, all bed disuse-induced bone losses were practically restored.

6. Discussion

The developed image preprocessing and segmentation algorithms provide promising practical pQCT image analysis tools for bone research applications. During the methodological development, emphasis was placed on the use and combination of simple and reliable algorithms so that their influence on fine details in clinical image data would be comprehensible. These algorithms have been applied in clinical studies including analyses of longitudinal disuse studies and, in addition to the present study, the impact of different models of exercise loading on tibial bone marrow density [Rantalainen et al., 2013].

The image-preprocessing algorithm is comprised of a basic gray-level transformation and the use of a spatial knowledge about the neighbourhood of the evaluated pixels through Markov random fields. This process required prior knowledge about noise and structures that were present in the pQCT image data. Nevertheless, if applied to pQCT image data, the current denoising process is fully automatic and was able to enhance the analysis of cortical area. An improvement in the reliability of repeated measurements (~*in vivo* precision) was consistent with reliability coefficients being nearly 90% or better, even for the direction-specific analysis [I].

The OBS segmentation method is based on bone delineation and a shrinking procedure (both use information from the first derivation of the image data) and a morphological and threshold-based correction. As the OBS comprises only simple image processing procedures, the method is fast and does not require operator involvement [III]. In conjunction with the developed preprocessing, OBS method enhanced the cortical bone analysis, halved the variation in geometrical traits in comparison with threshold-based analyses and consistently showed an excellent agreement with the results of manual segmentation (~2% overestimation) for CoA, CoTh, $CSMI_{max}$ [III].

6.1. Preprocessing – Denoising

The preprocessing of pQCT images is difficult due to (i) the lack of knowledge about the ground truth of the acquired image, (ii) the large variety of noise sources that can affect the quality of any x-ray-based image [Boas & Fleischmann, 2012], and (iii) the need to preserve fine details in the analysed images. Therefore, the need for an accurate analysis of pQCT image data limits the use of sophisticated preprocessing methods, as the influence on the data of such methods is still not well known.

The performance of developed preprocessing method was demonstrated on phantom and *in vivo* pQCT images [I]. The results of this method were compared with the results from the raw images and standard median filters that are the only currently used method in pQCT image preprocessing. In comparison with the analysis of raw images, all preprocessing methods were able to markedly improve the analysis of the cortical bone area (reliability of measurements increased by nearly 30%). However, all methods failed to enhance the trabecular density analysis and, in the end, no preference could be determined, as the between-method differences remained marginal. Nevertheless, with LinW setup, the developed method showed more consistent performance for the anatomically direction-specific analysis and was able to increase the dynamic range of the density values.

The failure in the reduction of the variance in trabecular density could be caused by (i) a limited true dynamic range of pQCT imaging that only allows the separation of six distinct gray levels (~ non-overlapping density ranges) or (ii) a number of various noise sources with a non-random nature that are difficult to suppress and need to be taken into account. These sources include: a) the high-pass filtration and the back projection of Poisson and Gaussian noise in the projections, b) the back projection reconstruction algorithm, c) a variation in the translational and rotational motion of the X-ray source and detector assembly during the scan, d) a non-uniform intensity of the X-ray beam, and d) specific detector characteristics (e.g., the detective quantum efficiency or pixel-to-pixel sensitivity differences). All of these factors may modulate the noise in the pQCT image separately or interactively, contributing to the large variance in scan and analysis parameters [Veitch et al., 2004; Ward et al., 2005; Ashe et al., 2006; Kontulainen et al., 2007].

Even though preprocessing proved beneficial in sufficiently enhancing the cortical bone analysis, the use of statistical-based preprocessing methods may not be considered appropriate. Questions may be asked as to whether any other sophisticated method, such as a diffusion-based method, would show better performance. However, due to the very low SNR (~ 20-30dB) in the pQCT images [I] and the need to preserve fine details in the image data, the selection of preprocessing based on the redundant wavelet transform and Markov random fields was a natural one. Clearly, the redundant wavelet transform decomposes the image data into separate frequency bands. While the noise appears primarily in the higher frequency bands, the important features of image data can be found in all frequency bands. Therefore, combined with geometric constraints (Markov random fields), this multiresolution approach reduces noise in the image data while allowing effective

detection and preservation of important details (e.g. corners, edges, border lines) [Dhawan et al., 2008].

Nevertheless, one may ask whether the use of any preprocessing method is meaningful, as it does not seem to improve the conventional trabecular density analysis. Obviously, depending on the type and specific parameters of the method used, image denoising suppresses noise while causing greater or lesser substantial loss of fine image details. However, the results of the trabecular density analysis of different preprocessing methods (I: Table 2) suggest that, at least at the group level, loss of fine detail may not be a major concern. In contrast, probably due to the previously mentioned limited dynamic range in pQCT image data, preprocessing revealed some apparent coarse structural patterns within the trabecular bone. These patterns reflect the spatial distribution of different trabecular density regions within the bone cross-section and may reveal association with true structural features in bones (I: Fig 5).

6.2. OBS Segmentation

Currently, density-based thresholds are commonly used for pQCT image segmentation. However, this approach has its limitations, and bone analysis accuracy greatly depends on the quality of the image data. Previously, some studies have assessed optimal density threshold settings with a comparison to phantoms [Ward et al., 2005] or histomorphometry at the tibial diaphysis [Veitch et al., 2004; Kontulainen et al., 2007]. However, the use of these settings is rather difficult at the distal epiphysis where a thin-walled (~2-4 mm) cortical shell, the coarse spatial resolution of pQCT images, and the presence of the partial volume effect limits their usability for accurate cortical bone detection [Veitch et al., 2004; Hangartner, 2007].

The novel OBS segmentation method was developed especially for rapid evaluation of thin-walled cortical bone in pQCT images of the distal tibia. The OBS method overcomes problems associated with low cortical thickness by using the first derivation of the image data. The additional benefit of the OBS method is that it evaluates cortical bone geometry automatically and does not require any operator involvement that could compromise the accuracy and the repeatability of the obtained results [III]. The performance of the novel OBS method was compared against a more sophisticated DRLSE method [C. Li et al., 2010; III] and density threshold-based Stratec C1/P2 and C2/P2 analyses [Sievänen et al., 1998; Kontulainen et al., 2007] with manual segmentation set as the gold standard. Results suggest that the OBS method outperformed not only the commonly used density threshold-based analyses but also the more sophisticated DRLSE method for accurate delineation of cortical bone (III: Figs. 4 and 5). These approaches resulted in either systematic under- or overestimations of all the observed geometrical traits of the cortical bone. However, despite the larger overestimation of the cortical geometrical traits by the DRLSE method, the generally high correlation between OBS and DRLSE method results (> 0.88), with the exception of CoTh assessment (~ 0.62), indicates that the results of both methods may be

comparable with a proper adjustment. The comparability of the OBS and DRLSE methods is also supported by a similar reduction (nearly 50%) in variance in all geometric traits compared with the results from simple threshold-based analyses. Nevertheless, in contrast to the OBS method, the DRLSE method requires a complex setting of user-defined parameters that may reduce its usability and decrease the reproducibility of the results.

Some may raise concerns about the reliability of any automated method for image segmentation, especially when its performance is evaluated in comparison with data obtained from manual segmentation that is considered to be a gold standard. Obviously, the accuracy of the results of manual segmentation is highly questionable and depends on the skills of the particular rater and his/her previous experience in analysing particular image data. A skilled rater can see not only the information that is truly present in the image but can also distinguish lines and borders that are hardly recognizable by any automated method. This skill, however, differs substantially between particular raters and studies that observed the inter- and intra-rater variability in the segmentation of the same objects demonstrated ~10% variability between ratings [Heinonen, 1999]. As an objective assessment, histomorphometry is obviously not an option for an *in vivo* study. Therefore, for the validation of the segmentation results and the reduction of possible inter-individual variances, the mean of 3 independent manual segmentations could be considered sufficient.

Another concern about the performance of the OBS method may arise due to the somewhat lower correlation between the results of the mean CoTh from the OBS method and manual segmentation. This may indicate that the assessment of the mean CoTh of relatively thin-walled bones with pQCT imaging may not be robust and the validity of the results obtained from the OBS method may remain questionable. However, the enhanced OBS method partially cleared these doubts. As was demonstrated, the Dice similarity coefficient (perfect similarity = 1), another measure of segmentation accuracy that takes into account the spatial similarity, indicated an excellent agreement between CoA results from the enhanced OBS method and manual segmentation (the range was 0.77-0.93 with a mean value of 0.85). Nevertheless, results of the present enhanced OBS method should be compared with *in vivo* HR-pQCT results to clearly determine the validity and accuracy for estimations of macro-structural traits.

6.3. Clinical Aspects

Bone fragility, or conversely its strength, is largely determined by bone structural properties (e.g. cross-sectional geometry, cortical thickness, and trabecular architecture) [Järvinen et al., 2005]. Accordingly, a proper bone structural analysis should help to identify individuals who are susceptible to fragility fractures. The potential of pQCT imaging for the clinical evaluation of whole bone strength and other various research applications is well known [Engelke et al., 2008; Zemel et al., 2008]. However, the need for better accuracy of characterized bone macro-structural traits in pQCT images has been recognized [Ashe et al., 2006; Kontulainen et al., 2007; Hangartner & Short, 2007], as the

performance and reliability of this approach is compromised by several factors such as movement artefacts and a high variance in scan and analysis parameters.

The current state of the art, threshold based approach, used in pQCT image analysis limits the accuracy of assessment of these traits [Veitch et al., 2004; Ward et al., 2005; Hangartner, 2007; Kontulainen et al., 2007], and recent technical developments are leading more towards higher resolution modalities [Elmoutaouakkil et al., 2002; Buie et al., 2007; Lublinsky et al., 2007; Burghardt et al., 2010; Valentinitsch et al., 2012] and finite element analysis (FEA) [Mueller et al., 2011; Vilayphiou et al., 2011; Rizzoli et al., 2012; Nishiyama et al., 2013]. Nevertheless, it is still not known whether the three times more expensive HR-pQCT systems with isotropic voxel size of 82 µm and derived FEA models can essentially improve the assessment of whole bone strength compared to the data yielded by standard QCT scanners [Engelke et al., 2013]. As the trabecular thickness ranges from 50 µm to 150 µm, the micro-structural properties directly measured by HR-pQCT are rather overestimated in comparison to the absolute values of properties acquired by histomorphometry [Boutroy et al., 2011]. However, there are many pQCT studies that would benefit from enhanced bone assessment. In addition, standard pQCT scanners are still widely used among bone researchers. Therefore, the proposed preprocessing and segmentation method is aimed at improving their daily practice. It should be noted that it is also possible that the proposed methods could be applied for the cortical analysis of clinical QCT data with similar voxel sizes (<1 mm) at clinically relevant sites. Nevertheless, further studies in this direction are needed.

As anticipated, the proposed preprocessing method reduced the noise present in pQCT images and improved image dynamic range. This has been confirmed in a recent study by Rantalainen et al. [2013] where the novel preprocessing method helped to reveal small but significant differences in bone marrow density in young female athletes with contrasting loading histories and bone strengths. These differences remained masked when the data were assessed with conventional threshold-based analysis only.

In addition, the proposed preprocessing in conjunction with the OBS segmentation method improved the *in vivo* precision of pQCT cortical bone analysis, and reduced variance in geometric traits. Thus, the preprocessing with the OBS method can improve detection of subtle and regional sector-specific changes in the cortical bone geometry. The regional sector-specific change may reveal details about the mechanism of regional bone adaptations, as suggested by Lai et al. [2005]. In fact, reanalysis of sector-specific differences in the long term bed rest study [II] improved detection of subtle differences and revealed skeletal responses to various interventions that cannot be inferred from the average analysis of the whole bone cross-section. Furthermore, application of a robust analysis method to longitudinal data can enhanced detection of subtle cortical bone changes within and between intervention groups [II][IV]. Moreover, as recently pinpointed by Evans et al. [2012], analysis-related improvements in sensitivity may reveal early,

anatomically specific bone changes, informing the design of patient-specific interventions to improve whole bone strength.

The OBS method is unable to correctly determine cortical bone density as it detects all pixels that belong to the cortical bone even if their density value was compromised by the partial volume effect [III]. However, the OBS method accurately identifies the cortical bone geometry. Consequently, the accurately detected cortical bone geometry should improve the assessment of the bone strength indices [Kontulainen et al., 2008] and allow a simpler comparison of between-study results [IV].

The importance of the cortical bone geometry as the major determinant of bone strength has been recently pinpointed [Pistoia et al., 2003; Seeman, 2003; Holzer et al., 2009; Melton et al., 2010; Roux et al., 2010]. Nevertheless, trabecular density is still considered to be the major trait of interest in clinical osteoporosis research [Engelke et al., 2008]. This is mainly due to higher metabolic activity of trabecular bone in comparison with cortical bone; thus greater changes in trabecular density in response to aging, medical treatment or physiological conditions are expected [Engelke et al., 2008]. Therefore, the use of the novel preprocessing method in clinical bone research might be restricted by its inability to improve the *in vivo* precision of this trait [I]. However, the apparent structural patterns revealed within the trabecular bone (I: Fig 5) may correspond to true trabecular architecture. Thus, as suggested by Findlay [2012], analysis of these patterns may open a new way to examine trabecular bone, especially in longitudinal studies. This analysis approach, however, naturally requires further elaboration.

6.4. Future Work

The developed image processing methods for the enhancement of pQCT image analysis are promising tools for future clinical bone research [I – IV]. However, for future work, HR-pQCT images should be compared with standard pQCT images to determine the effect of true noise on apparent structural traits and to determine the validity of the presented image processing approaches in revealing true structural features in bones. If their repeatability and accuracy is proven, the developed image processing algorithms can be applied for the reanalysis of existing pQCT image data of the distal tibia and the radius (similar bone structure) and enhance existing knowledge about distal tibia bone traits.

Establishing the connection between data from high- and low-resolution scanners can also reveal whether the use of the HR-pQCT scanner provides any relevant information in terms of bone strength that cannot be determined by the standard pQCT scanner. In other words, revealing whether the use of the recently developed OBS method applied to pQCT images [III] can provide similar information on cortical bone as HR-pQCT would allow more accurate bone analysis based on more widely available devices compared to more expensive high resolution devices. Consequently, the connection would clarify whether the use of HR-pQCT is a necessary step towards the more accurate prediction of whole bone strength.

Nevertheless, currently developed OBS method shows consistent over- and under-estimation of cortical area for smaller and larger bone cross-sections, respectively (III: Fig 5). Therefore, future studies should further revise the OBS method so that this over- and under-estimation of detected cortical area can be corrected.

Further, the current OBS method is unable to correctly determine cortical BMD as the method includes all pixel values that belong to detected cortical area. This results in reduction of obtained cortical BMD values due to inclusion pixel values affected by partial volume effect. However, the accurately identified cortical bone geometry may be used to improve assessment of adaptive changes in longitudinal studies (as suggested in [IV]) and further to improve an assessment of the bone strength indices as SSI (i.e., improve estimates of torsional or bending strength). Hence, future studies should involve cadaver data, histomorphometry and actual mechanical tests to reveal (i) the true accuracy of detected cortical area and (ii) whether the use of the current OBS method allows better prediction of bone strength than existing metrics.

In addition, future studies should also address whether the apparent structural patterns of BMD distributions observed in trabecular bone in pQCT images (I: Fig 5) truly reflect actual structural features in bones and spatial distribution of different trabecular density regions. Consequently, these patterns could reveal interesting associations between bone loading and density variation not only within the bone cross-section but also within the anatomic sectors. Therefore, further investigation in this and above mentioned directions could introduce new analysis protocols for coarse regional analyses of pQCT images. These protocols may comprise sector analysis of cortical bone, subcortical bone and distribution of 6 distinct density ranges in trabecular bone, and replace conventional mean BMD (BMC) measurements.

For future work, further investigations should address all these open questions while the present work that introduced the current state-of-the-art image analysis algorithms developed to enhance the assessment of pQCT data, may be considered to be the first step.

7. Conclusion

In this study, new methods were developed for preprocessing and segmentation of cortical bone in pQCT images from the distal tibia. The new methods were designed to be fast, simple and easy to understand approaches, for integration into analysis software and routine clinical research use. These novel methods were also applied to clinical and experimental data to demonstrate their benefits for more accurate analysis of bone traits used for whole bone strength estimation.

The first aim of this study was to develop an image preprocessing method that would reliably enhance pQCT image analysis and reduce noise in the image data. The statistical-based method [I] is capable of reducing the noise level in pQCT images and offers essential improvements in the analysis of the cortical bone area. Although the method failed to enhance the trabecular density analysis, it revealed apparent coarse structural patterns in the trabecular area that could lead towards a new way to analyse data from pQCT-based studies. In addition, the developed method increased the dynamic range of pQCT image intensities and enabled the detection of subtle and sector-specific changes that would have remained concealed with conventional analysis. The developed method was tested on phantoms and repeated with *in vivo* pQCT images. Further studies are, however, needed to assess its true utility.

The second aim was to develop an algorithm for cortical bone segmentation in pQCT images that enabled automatic, robust, fast and reproducible cortical detection. The OBS algorithm, together with the developed preprocessing [III], is capable of fulfilling all basic requirements and provides a promising practical tool for the fast and reliable detection of the cortical bone geometry. The developed method was tested on various *in vivo* pQCT images and demonstrated its successful use and potential to enhance bone analyses and revealed specific skeletal processes in the cortical bone geometry that cannot be acquired

using conventional threshold-based analysis. This, in a broader perspective, may improve the estimation of whole bone strength and fracture risk prediction. In addition, the OBS algorithm does not require operator involvement and performs the cortical bone segmentation automatically.

The performance of the developed methods meets the given aims and all requirements and forms a simple and applicable pQCT image analysis tool for clinical bone research applications [I][III]. Therefore, this study can be considered as a step towards the final objective which is to find the most appropriate imaging and analysis methods for predicting bone fragility and to reveal relevant changes in bone structure that affect whole bone strength.

References

Abrantes AJ., Marques JS.: A class of constrained clustering algorithms for object boundary extraction. *IEEE Transactions on Image Processing*. 1996; 5(11): 1507–1521.

Adams R., Bischof L.: Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1994; 16(6): 641–647.

Afonso MV., Bioucas-Dias JM., Figueiredo MAT.: Fast Image Recovery Using Variable Splitting and Constrained Optimization. *IEEE Transactions on Image Processing*. 2010; 19(9): 2345–2356.

Ali S., Madabhushi A.: An Integrated Region-, Boundary-, Shape-Based Active Contour for Multiple Object Overlap Resolution in Histological Imagery. *IEEE Transactions on Medical Imaging*. 2012; 31(7): 1448–1460.

Arbelaez P., Maire M., Fowlkes C., Malik J.: Contour Detection and Hierarchical Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2011; 33(5): 898–916.

Arce GR.: Multistage order statistic filters for image sequence processing. *IEEE Transactions on Signal Processing*. 1991; 39(5): 1146–1163.

Arce GR., Foster RE.: Detail-preserving ranked-order based filters for image processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*. 1989; 37(1): 83–98.

Arce GR., McLoughlin M.: Theoretical analysis of the max/Median filter. *IEEE Transactions on Acoustics, Speech and Signal Processing*. 1987; 35(1): 60–69.

Ardran GM.: Bone destruction not demonstrable by radiography. *The British journal of radiology*. 1951; 24(278): 107–109.

Ashe MC., Khan KM., Kontulainen SA., Guy P., Liu D., Beck TJ., McKay HA.: Accuracy of pQCT for evaluating the aged human radius: an ashing, histomorphometry and failure load investigation. *Osteoporosis International*. 2006; 17(8): 1241–1251.

Ballerini L., Bocchi L.: Multiple Genetic Snakes for Bone Segmentation. In: S. Cagnoni, C. G. Johnson, J. J. R. Cardalda, E. Marchiori, D. W. Corne, J.-A. Meyer, J. Gottlieb, M. Middendorf, A. Guillot, G. R. Raidl, & E. Hart eds. *Applications of Evolutionary Computing*. Lecture Notes in Computer Science. Springer Berlin Heidelberg. 346–356. Available at: http://link.springer.com/chapter/10.1007/3-540-36605-9_32.

Batenburg KJ., Sijbers J.: Optimal Threshold Selection for Tomogram Segmentation by Projection Distance Minimization. *IEEE Transactions on Medical Imaging*. 2009; 28(5): 676–686.

Bedi CS., Goyal H.: Qualitative and Quantitative Evaluation of Image Denoising Techniques. *International Journal of Computer Applications*. 2010; 8(14): 31–34.

Behiels G., Maes F., Vandermeulen D., Suetens P.: Evaluation of image features and search strategies for segmentation of bone structures in radiographs using Active Shape Models. *Medical Image Analysis*. 2002; 6(1): 47–62.

Beucher S.: The Watershed Transformation Applied To Image Segmentation. *Scanning Microscopy Supplement*. 1992; 6: 299–314.

Beutel J., Kundel HL., Van Metter RL. Handbook of Medical Imaging: Physics and psychophysics. SPIE Press, Bellingham, Washington, USA, 2000.

Bezdek JC., Hall LO., Clarke LP.: Review of MR image segmentation techniques using pattern recognition. *Medical Physics*. 1993; 20(4): 1033–1048.

Black DM., Bouxsein ML., Marshall LM., Cummings SR., Lang TF., Cauley JA., Ensrud KE., Nielson CM., Orwoll ES., Osteoporotic Fractures in Men (MrOS) Research Group.: Proximal femoral structure and the prediction of hip fracture in men: a large prospective study using QCT. *Journal of Bone and Mineral Research*. 2008; 23(8): 1326–1333.

Black MJ., Sapiro G., Marimont DH., Heeger D.: Robust anisotropic diffusion. *IEEE Transactions on Image Processing*. 1998; 7(3): 421–432.

Boas FE., Fleischmann D.: CT artifacts: causes and reduction techniques. *Imaging in Medicine*. 2012; 4(2): 229–240.

Bolotin HH.: The significant effects of bone structure on inherent patient-specific DXA in vivo bone mineral density measurement inaccuracies. *Medical Physics*. 2004; 31(4): 774–788.

Bolotin HH., Sievänen H.: Inaccuracies inherent in dual-energy X-ray absorptiometry in vivo bone mineral density can seriously mislead diagnostic/prognostic interpretations of patient-specific bone fragility. *Journal of Bone and Mineral Research*. 2001; 16(5): 799–805.

Bouchet A., Pastore J., Ballarin V.: Segmentation of Medical Images using Fuzzy Mathematical Morphology. *Journal of Computer Science and Technology*. 2007; 7(3): 256–262.

Boutroy S., Bouxsein ML., Munoz F., Delmas PD.: In vivo assessment of trabecular bone microarchitecture by high-resolution peripheral quantitative computed tomography. *The Journal of Clinical Endocrinology and Metabolism*. 2005; 90(12): 6508–6515.

Boutroy S., Vilayphiou N., Roux J-P., Delmas PD., Blain H., Chapurlat RD., Chavassieux P.: Comparison of 2D and 3D bone microarchitecture evaluation at the femoral neck, among postmenopausal women with hip fracture or hip osteoarthritis. *Bone*. 2011; 49(5): 1055–1061.

Braillon PM.: Quantitative computed tomography precision and accuracy for long-term follow-up of bone mineral density measurements: a five year in vitro assessment. *Journal of Clinical Densitometry*. 2002; 5(3): 259–266

Brownrigg DRK.: The Weighted Median Filter. *Communications of the Association for Computing Machinery*. 1984; 27(8): 807–818.

Buie HR., Campbell GM., Klinck RJ., MacNeil JA., Boyd SK.: Automatic segmentation of cortical and trabecular compartments based on a dual threshold technique for in vivo micro-CT bone analysis. *Bone*. 2007; 41(4): 505–515.

Burghardt AJ., Buie HR., Laib A., Majumdar S., Boyd SK.: Reproducibility of direct quantitative measures of cortical bone microarchitecture of the distal radius and tibia by HR-pQCT. *Bone*. 2010; 47(3): 519–528.

Cai W., Chen S., Zhang D.: Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. *Pattern Recognition*. 2007; 40(3): 825–838.

Carpenter RD., Beaupré GS., Lang TF., Orwoll ES., Carter DR.: New QCT analysis approach shows the importance of fall orientation on femoral neck strength. *Journal of Bone and Mineral Research*. 2005; 20(9): 1533–1542.

Catte F., Lions P-L., Morel J-M., Coll T.: Image selective smoothing and edge detection by nonlinear diffusion. SIAM Journal on Numerical Analysis. 1992; 29(1): 182–193.

Cervinka T., Hyttinen J., Sievänen H. Accurate Cortical Bone detection in Peripheral Quantitative Computed Tomography Images. XIII. Mediterranean conference on Medical and Biological engineering and computing. IFMBE Proceedings 2014; 41: 289-292.

Chan TF., Vese LA.: Active contours without edges. *IEEE Transactions on Image Processing*. 2001; 10(2): 266–277.

Chang D-X., Zhang X-D., Zheng C-W.: A genetic algorithm with gene rearrangement for K-means clustering. *Pattern Recognition*. 2009; 42(7): 1210–1222.

Chang SG., Yu B., Vetterli M.: Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing*. 2000; 9(9): 1532–1546.

Chen Y., Garcia EK., Gupta MR., Rahimi A., Cazzanti L.: Similarity-based Classification: Concepts and Algorithms. *Journal of Machine Learning Research*. 2009; 10: 747–776.

Chen Y-T.: A level set method based on the Bayesian risk for medical image segmentation. *Pattern Recognition*. 2010; 43(11): 3699–3711.

Cheng HD., Jiang XH., Sun Y., Wang J.: Color image segmentation: advances and prospects. *Pattern Recognition*. 2001; 34(12): 2259–2281.

Christensen L., Iqbal S., Macarios D., Badamgarav E., Harley C.: Cost of fractures commonly associated with osteoporosis in a managed-care population. *Journal of Medical Economics*. 2010; 13(2): 302–313.

Cohen LD., Cohen I.: Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1993; 15(11): 1131–1147.

Cooley JW., Lewis P., Welch P.: The finite Fourier transform. *IEEE Transactions on Audio and Electroacoustics*. 1969; 17(2): 77–85.

Cremers D., Rousson M., Deriche R.: A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape. *International Journal of Computer Vision*. 2007; 72(2): 195–215.

Cremers D., Tischhäuser F., Weickert J., Schnörr C.: Diffusion Snakes: Introducing Statistical Shape Knowledge into the Mumford-Shah Functional. *International Journal of Computer Vision*. 2002; 50(3): 295–313.

Davatzikos CA., Prince JL.: An active contour model for mapping the cortex. *IEEE Transactions on Medical Imaging*. 1995; 14(1): 65–80.

Delmas PD., Seeman E.: Changes in bone mineral density explain little of the reduction in vertebral or nonvertebral fracture risk with anti-resorptive therapy. *Bone*. 2004; 34(4): 599–604.

Dhawan AP., Huang HK., Kim D-S. Principles and Advanced Methods in Medical Imaging and Image Analysis. World Scientific Publishing Co. Pte. Ltd., Singapore, 2008.

Diop EHS., Burdin V.: Bi-planar image segmentation based on variational geometrical active contours with shape priors. *Medical Image Analysis*. 2013; 17(2): 165–181.

Donoho DL.: De-noising by soft-thresholding. *IEEE Transactions on Information Theory*. 1995; 41(3): 613–627.

Elder JH., Zucker SW.: Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1998; 20(7): 699–716.

Elmoutaouakkil A., Peyrin F., Elkafi J., Laval-Jeantet AM.: Segmentation of cancellous bone from high-resolution computed tomography images: influence on trabecular bone measurements. *IEEE Transactions on Medical Imaging*. 2002; 21(4): 354–362.

Eng H-L., Ma K-K.: Noise adaptive soft-switching median filter. *IEEE Transactions on Image Processing*. 2001; 10(2): 242–251.

Engelke K., Adams JE., Armbrecht G., Augat P., Bogado CE., Bouxsein ML., Felsenberg D., Ito M., Prevrhal S., Hans DB., Lewiecki EM.: Clinical use of quantitative computed tomography and peripheral quantitative computed tomography in the management of osteoporosis in adults: the 2007 ISCD Official Positions. *Journal of Clinical Densitometry*. 2008; 11(1): 123–162.

Engelke K., Libanati C., Fuerst T., Zysset P., Genant HK.: Advanced CT based In vivo Methods for the Assessment of Bone Density, Structure, and Strength. *Current Osteoporosis Reports*. 2013; 11(3): 246–255.

Engelke K., Libanati C., Liu Y., Wang H., Austin M., Fuerst T., Stampa B., Timm W., Genant HK.: Quantitative computed tomography (QCT) of the forearm using general purpose spiral whole-body CT scanners: accuracy, precision and comparison with dual-energy X-ray absorptiometry (DXA). *Bone*. 2009; 45(1): 110–118.

Engelke K., Mastmeyer A., Bousson V., Fuerst T., Laredo J-D., Kalender WA.: Reanalysis precision of 3D quantitative computed tomography (QCT) of the spine. *Bone*. 2009; 44(4): 566–572.

Evans RK., Negus CH., Centi AJ., Spiering BA., Kraemer WJ., Nindl BC.: Peripheral QCT sector analysis reveals early exercise-induced increases in tibial bone mineral density. *Journal of Musculoskeletal & Neuronal Interactions*. 2012; 12(3): 155–164.

Fan G., Xia X-G.: A joint multicontext and multiscale approach to Bayesian image segmentation. *IEEE Transactions on Geoscience and Remote Sensing*. 2001; 39(12): 2680–2688.

Fan J., Yau DKY., Elmagarmid AK., Aref WG.: Automatic image segmentation by integrating color-edge extraction and seeded region growing. *IEEE Transactions on Image Processing*. 2001; 10(10): 1454–1466.

Fan J., Zeng G., Body M., Hacid M-S.: Seeded region growing: an extensive and comparative study. *Pattern Recognition Letters*. 2005; 26(8): 1139–1156.

Findlay CM. Image analysis tool for the characterisation of bone turnover in the appendicular skeleton. *Doctoral thesis*. University of Glasgow, Glasgow, 2012. Available at: http://theses.gla.ac.uk/3657/.

Fitch J., Coyle EJ., Gallagher, N.C. J.: Median filtering by threshold decomposition. *IEEE Transactions on Acoustics, Speech and Signal Processing*. 1984; 32(6): 1183–1188.

Fleiss JL.: Reliability of measurement. In: J.L. Fleiss eds. *The design and analysis of clinical experiments*. John Wiley & Sons, Inc., New York, USA, 1986. p. 1–32.

Fukunaga K., Hayes RR.: The reduced Parzen classifier. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1989; 11(4): 423–425.

Gao H., Wan-Chi Siu., Hou C-H.: Improved techniques for automatic image segmentation. *IEEE Transactions on Circuits and Systems for Video Technology*. 2001; 11(12): 1273–1280.

Gärdsell P., Johnell O., Nilsson BE., Gullberg B.: Predicting various fragility fractures in women by forearm bone densitometry: a follow-up study. *Calcified tissue international*. 1993; 52(5): 348–353.

Genant HK., Engelke K., Prevrhal S.: Advanced CT bone imaging in osteoporosis. *Rheumatology (Oxford, England)*. 2008; 47 Suppl 4: iv9–16.

Gevers T.: Adaptive image segmentation by combining photometric invariant region and edge information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24(6): 848–852.

Ghazel M., Freeman GH., Vrscay ER.: Fractal-wavelet image denoising revisited. *IEEE Transactions on Image Processing*. 2006; 15(9): 2669–2675.

Gilboa G., Sochen N., Zeevi YY.: Image enhancement and denoising by complex diffusion processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2004; 26(8): 1020–1036.

Glüer CC., Faulkner KG., Estilo MJ., Engelke K., Rosin J., Genant HK.: Quality assurance for bone densitometry research studies: concept and impact. *Osteoporosis International*. 1993; 3(5): 227–235.

Goin JE.: Classification Bias of the k-Nearest Neighbor Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1984; PAMI-6(3): 379–381.

Gravel P., Beaudoin G., De Guise JA.: A method for modeling noise in medical images. *IEEE Transactions on Medical Imaging*. 2004; 23(10): 1221–1232.

Griffith JF., Genant HK.: New advances in imaging osteoporosis and its complications. *Endocrine*. 2012; 42(1): 39–51.

Griffith JF., Genant HK.: New imaging modalities in bone. *Current Rheumatology Reports*. 2011; 13(3): 241–250.

Gudbjartsson H., Patz S.: The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*. 1995; 34(6): 910–914.

Guo R., Pandit SM.: Automatic threshold selection based on histogram modes and a discriminant criterion. *Machine Vision and Applications*. 1998; 10(5-6): 331–338.

Gupta S., Chauhan RC., Saxena SC.: Locally adaptive wavelet domain Bayesian processor for denoising medical ultrasound images using Speckle modelling based on Rayleigh distribution. *IEE Proceedings - Vision, Image and Signal Processing*. 2005; 152(1): 129–135.

Hangartner TN.: Thresholding technique for accurate analysis of density and geometry in QCT, pQCT and microCT images. *Journal of Musculoskeletal & Neuronal Interactions*. 2007; 7(1): 9–16.

Hangartner TN., Short DF.: Accurate quantification of width and density of bone structures by computed tomography. *Medical Physics*. 2007; 34(10): 3777–3784.

Hanson KM.: Noise and contrast discrimination in computed tomography. In: T. H. Newton & D. G. Potts eds. Radiology of the Skull and Brain, *Technical Aspects of Computed Tomography*. C.V. Mosby, St. Louis, 1981.

Haris K., Efstratiadis SN., Maglaveras N., Katsaggelos AK.: Hybrid image segmentation using watersheds and fast region merging. *IEEE Transactions on Image Processing*. 1998; 7(12): 1684–1699.

Hassanien AE., Abraham A., Peters JF., Schaefer G., Henry C.: Rough Sets and Near Sets in Medical Imaging: A Review. *IEEE Transactions on Information Technology in Biomedicine*. 2009; 13(6): 955–968.

Heinonen T. Applications of Magnetic Resonance Image Segmentation in Neurology. *Doctoral thesis*. Tampere University of Technology, Tampere, 1999.

Hind K., Oldroyd B., Prajapati A., Rhodes L.: In vivo precision of dual-energy X-ray absorptiometry-derived hip structural analysis in adults. *Journal of Clinical Densitometry*. 2012; 15(3): 302–307.

Hipp JA., Jansujwicz A., Simmons CA., Snyder BD.: Trabecular bone morphology from micro-magnetic resonance imaging. *Journal of Bone and Mineral Research*. 1996; 11(2): 286–297.

Holzer G., von Skrbensky G., Holzer LA., Pichl W.: Hip fractures and the contribution of cortical versus trabecular bone to femoral neck strength. *Journal of Bone and Mineral Research*. 2009; 24(3): 468–474.

Huang C-C.: A novel gray-based reduced NN classification method. *Pattern Recognition*. 2006; 39(11): 1979–1986.

Huang H-C., Chuang Y-Y., Chen C-S.: Multiple Kernel Fuzzy Clustering. *IEEE Transactions on Fuzzy Systems*. 2012; 20(1): 120–134.

Ito M.: Recent progress in bone imaging for osteoporosis research. *Journal of Bone and Mineral Metabolism*. 2011; 29(2): 131–140.

Ito M., Hayashi K., Ishida Y., Uetani M., Yamada M., Ohki M., Nakamura T.: Discrimination of spinal fracture with various bone mineral measurements. *Calcified Tissue International*. 1997; 60(1): 11–15.

Jain AK., Duin RPW., Mao J.: Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000; 22(1): 4–37.

Järvinen TL., Sievänen H., Jokihaara J., Einhorn TA.: Revival of bone strength: the bottom line. *Journal of Bone and Mineral Research*. 2005; 20(5): 717–720.

Jawahar CV., Biswas PK., Ray AK.: Investigations on fuzzy thresholding based on fuzzy clustering. *Pattern Recognition*. 1997; 30(10): 1605–1613.

Jee W.: The skeletal tissues. In: L. Weiss ed. Histology: *cell and tissue biology*. Elsevier Science Ltd, New York. 1983, 206–254.

Kak AC., Slaney M. Principles of computerized tomographic imaging. IEEE Press, New York, 1988.

Kang Y., Engelke K., Kalender WA.: A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data. *IEEE Transactions on Medical Imaging*. 2003; 22(5): 586–598.

Kanis JA.: Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. WHO Study Group. *Osteoporosis International*. 1994; 4(6): 368–381.

Kanis JA., Borgstrom F., De Laet C., Johansson H., Johnell O., Jonsson B., Oden A., Zethraeus N., Pfleger B., Khaltaev N.: Assessment of fracture risk. *Osteoporosis International*. 2005; 16(6): 581–589.

Kanis JA., Johnell O., De Laet C., Johansson H., Oden A., Delmas P., Eisman J., Fujiwara S., Garnero P., Kroger H., McCloskey EV., Mellstrom D., Melton LJ., Pols H., Reeve J., Silman A., Tenenhouse A.: A meta-analysis of previous fracture and subsequent fracture risk. *Bone*. 2004; 35(2): 375–382.

Kanungo T., Mount DM., Netanyahu NS., Piatko CD., Silverman R., Wu AY.: An efficient k-means clustering algorithm: analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24(7): 881–892.

Karoui I., Fablet R., Boucher J-M., Augustin J.: Variational Region-Based Segmentation Using Multiple Texture Statistics. *IEEE Transactions on Image Processing*. 2010; 19(12): 3146–3156.

Kass M., Witkin A., Terzopoulos D.: Snakes: Active contour models. *International Journal of Computer Vision*. 1988; 1(4): 321–331.

Kazakia GJ., Majumdar S.: New imaging technologies in the diagnosis of osteoporosis. *Reviews in Endocrine & Metabolic Disorders*. 2006; 7(1-2): 67–74.

Kimia BB., Tannenbaum AR., Zucker SW.: Shapes, Shocks, and Deformations I: The Components of Two-Dimensional Shape and the Reaction-Diffusion Space. *International Journal of Computer Vision*. 1994; 15(3): 189–224.

Klinder T., Ostermann J., Ehm M., Franz A., Kneser R., Lorenz C.: Automated model-based vertebra detection, identification, and segmentation in CT images. *Medical Image Analysis*. 2009; 13(3): 471–482.

Kontulainen S., Liu D., Manske S., Jamieson M., Sievänen H., McKay H.: Analyzing cortical bone cross-sectional geometry by peripheral QCT: comparison with bone histomorphometry. *Journal of Clinical Densitometry*. 2007; 10(1): 86–92.

Kontulainen SA., Johnston JD., Liu D., Leung C., Oxland TR., McKay HA.: Strength indices from pQCT imaging predict up to 85% of variance in bone failure properties at tibial epiphysis and diaphysis. *Journal of Musculoskeletal & Neuronal Interactions*. 2008; 8(4): 401–409.

Krug R., Banerjee S., Han ET., Newitt DC., Link TM., Majumdar S.: Feasibility of in vivo structural analysis of high-resolution magnetic resonance images of the proximal femur. *Osteoporosis International*. 2005; 16(11): 1307–1314.

Krug R., Burghardt AJ., Majumdar S., Link TM.: High-resolution imaging techniques for the assessment of osteoporosis. *Radiologic Clinics of North America*. 2010; 48(3): 601–621.

Krug R., Carballido-Gamio J., Burghardt AJ., Kazakia G., Hyun BH., Jobke B., Banerjee S., Huber M., Link TM., Majumdar S.: Assessment of trabecular bone structure comparing magnetic resonance imaging at 3 Tesla with high-resolution peripheral quantitative computed tomography ex vivo and in vivo. *Osteoporosis International*. 2008; 19(5): 653–661.

Kwok SH., Constantinides AG., Wan-Chi Siu.: An efficient recursive shortest spanning tree algorithm using linking properties. *IEEE Transactions on Circuits and Systems for Video Technology*. 2004; 14(6): 852–863.

Lai YM., Qin L., Hung VWY., Chan KM.: Regional differences in cortical bone mineral density in the weight-bearing long bone shaft--a pQCT study. *Bone*. 2005; 36(3): 465–471.

Lang TF.: Quantitative computed tomography. *Radiologic Clinics of North America*. 2010; 48(3): 589–600.

Lang TF., Leblanc AD., Evans HJ., Lu Y.: Adaptation of the proximal femur to skeletal reloading after long-duration spaceflight. *Journal of Bone and Mineral Research*. 2006; 21(8): 1224–1230.

Laskey MA., de Bono S., Zhu D., Shaw CN., Laskey PJ., Ward KA., Prentice A.: Evidence for enhanced characterization of cortical bone using novel pQCT shape software. *Journal of Clinical Densitometry*. 2010; 13(3): 247–255.

Lei T., Udupa JK.: Performance evaluation of finite normal mixture model-based image segmentation techniques. *IEEE Transactions on Image Processing*. 2003; 12(10): 1153–1169.

Leung CK., Lam FK.: Performance analysis for a class of iterative image thresholding algorithms. *Pattern Recognition*. 1996; 29(9): 1523–1530.

Li C., Xu C., Gui C., Fox MD.: Distance Regularized Level Set Evolution and Its Application to Image Segmentation. *IEEE Transactions on Image Processing*. 2010; 19(12): 3243–3254.

Li EK., Zhu TY., Tam L-S., Hung VW., Griffith JF., Li TK., Li M., Wong KC., Leung PC., Kwok AW., Qin L.: Bone microarchitecture assessment by high-resolution peripheral quantitative computed tomography in patients with systemic lupus erythematosus taking corticosteroids. *The Journal of Rheumatology*. 2010; 37(7): 1473–1479.

Li L., Gong R., Chen W.: Gray level image thresholding based on fisher linear projection of two-dimensional histogram. *Pattern Recognition*. 1997; 30(5): 743–749.

Li Z., Chines AA., Meredith MP.: Statistical validation of surrogate endpoints: is bone density a valid surrogate for fracture? *Journal of Musculoskeletal & Neuronal Interactions*. 2004; 4(1): 64–74.

Liao P., Chen T., Chung P.: A fast algorithm for multilevel thresholding. *Journal of Information Science and Engineering*. 2001; 17(5): 713–727.

Lin J-H., Sellke TM., Coyle EJ.: Adaptive stack filtering under the mean absolute error criterion. *IEEE Transactions on Acoustics, Speech and Signal Processing*. 1990; 38(6): 938–954.

Lin L., Garcia-Lorenzo D., Li C., Jiang T., Barillot C.: Adaptive pixon represented segmentation (APRS) for 3D MR brain images based on mean shift and Markov random fields. *Pattern Recognition Letters*. 2011; 32(7): 1036–1043.

Link TM.: The Founder's Lecture 2009: advances in imaging of osteoporosis and osteoarthritis. *Skeletal Radiology*. 2010; 39(10): 943–955.

Liu D., Jiang Z., Feng H.: A novel fuzzy classification entropy approach to image thresholding. *Pattern Recognition Letters*. 2006; 27(16): 1968–1975.

Liu XS., Zhang XH., Sekhon KK., Adams MF., McMahon DJ., Bilezikian JP., Shane E., Guo XE.: High-resolution peripheral quantitative computed tomography can assess microstructural and mechanical properties of human distal tibial bone. *Journal of Bone and Mineral Research*. 2010; 25(4): 746–756.

Liu Z., Pan Q., Dezert J.: A new belief-based K-nearest neighbor classification method. *Pattern Recognition*. 2013; 46(3): 834–844.

Lochmüller E-M., Bürklein D., Kuhn V., Glaser C., Müller R., Glüer CC., Eckstein F.: Mechanical strength of the thoracolumbar spine in the elderly: prediction from in situ dual-energy X-ray absorptiometry, quantitative computed tomography (QCT), upper and lower limb peripheral QCT, and quantitative ultrasound. *Bone*. 2002; 31(1): 77–84.

Lochmüller E-M., Lill CA., Kuhn V., Schneider E., Eckstein F.: Radius bone strength in bending, compression, and falling and its correlation with clinical densitometry at multiple sites. *Journal of Bone and Mineral Research*. 2002; 17(9): 1629–1638.

Lochmüller E-M., Müller R., Kuhn V., Lill CA., Eckstein F.: Can novel clinical densitometric techniques replace or improve DXA in predicting bone strength in osteoporosis at the hip and other skeletal sites? *Journal of Bone and Mineral Research*. 2003; 18(5): 906–912.

Lublinsky S., Ozcivici E., Judex S.: An automated algorithm to detect the trabecular-cortical bone interface in micro-computed tomographic images. *Calcified Tissue International*. 2007; 81(4): 285–293.

Luisier F., Blu T.: SURE-LET Multichannel Image Denoising: Interscale Orthonormal Wavelet Thresholding. *IEEE Transactions on Image Processing*. 2008; 17(4): 482–492.

Luisier F., Blu T., Unser M.: A New SURE Approach to Image Denoising: Interscale Orthonormal Wavelet Thresholding. *IEEE Transactions on Image Processing*. 2007; 16(3): 593–606.

Luo W.: An efficient detail-preserving approach for removing impulse noise in images. *IEEE Signal Processing Letters*. 2006; 13(7): 413–416.

Ma W-Y., Manjunath BS.: EdgeFlow: a technique for boundary detection and image segmentation. *IEEE Transactions on Image Processing*. 2000; 9(8): 1375–1388.

MacNeil JA., Boyd SK.: Accuracy of high-resolution peripheral quantitative computed tomography for measurement of bone quality. *Medical Engineering & Physics*. 2007a; 29(10): 1096–1105.

MacNeil JA., Boyd SK.: Bone strength at the distal radius can be estimated from high-resolution peripheral quantitative computed tomography and the finite element method. *Bone*. 2008; 42(6): 1203–1213.

MacNeil JA., Boyd SK.: Load distribution and the predictive power of morphological indices in the distal radius and tibia by high resolution peripheral quantitative computed tomography. *Bone*. 2007b; 41(1): 129–137.

Majumdar S.: Magnetic resonance imaging of trabecular bone structure. *Topics in Magnetic Resonance Imaging: TMRI*. 2002; 13(5): 323–334.

Malfait M., Roose D.: Wavelet-based image denoising using a Markov random field a priori model. *IEEE Transactions on Image Processing*. 1997; 6(4): 549–565.

Mallat SG.: A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1989; 11(7): 674–693.

Marieb EN. Human anatomy & physiology. S. Beauparlant ed. Pearson Education, Inc., San Francisco, 2004.

Martini FH., Timmons MJ., Tallitsch RB. Human Anatomy. L. Berriam ed. Pearson Education, Inc., San Francisco, 2006.

Mastmeyer A., Engelke K., Fuchs C., Kalender WA.: A hierarchical 3D segmentation method and the definition of vertebral body coordinate systems for QCT of the lumbar spine. *Medical Image Analysis*. 2006; 10(4): 560–577.

McInerney T.: SketchSnakes: sketch-line initialized Snakes for efficient interactive medical image segmentation. *Computerized Medical Imaging and Graphics*. 2008; 32(5): 331–352.

McInerney T., Terzopoulos D.: Deformable models in medical image analysis: a survey. *Medical Image Analysis*. 1996; 1(2): 91–108.

McKay HA., Sievänen H., Petit MA., MacKelvie KJ., Forkheim KM., Whittall KP., Forester BB., MacDonald H.: Application of magnetic resonance imaging to evaluation of femoral neck structure in growing girls. *Journal of Clinical Densitometry*. 2004; 7(2): 161–168.

Melton LJ 3rd., Riggs BL., Keaveny TM., Achenbach SJ., Hoffmann PF., Camp JJ., Rouleau PA., Bouxsein ML., Amin S., Atkinson EJ., Robb RA., Khosla S.: Structural determinants of vertebral fracture risk. *Journal of Bone and Mineral Research*. 2007; 22(12): 1885–1892.

Melton LJ 3rd., Riggs BL., Keaveny TM., Achenbach SJ., Kopperdahl D., Camp JJ., Rouleau PA., Amin S., Atkinson EJ., Robb RA., Therneau TM., Khosla S.: Relation of vertebral deformities to bone density, structure, and strength. *Journal of Bone and Mineral Research*. 2010; 25(9): 1922–1930.

Mitra S., Banka H., Pedrycz W.: Rough-Fuzzy Collaborative Clustering. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*. 2006; 36(4): 795–805.

Monteil J., Beghdadi A.: A new interpretation and improvement of the nonlinear anisotropic diffusion for image enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1999; 21(9): 940–946.

Moser G., Serpico SB.: Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery. *IEEE Transactions on Geoscience and Remote Sensing*. 2006; 44(10): 2972–2982.

Mueller TL., Christen D., Sandercott S., Boyd SK., van Rietbergen B., Eckstein F., Lochmüller E-M., Müller R., van Lenthe GH.: Computational finite element bone mechanics accurately predicts mechanical competence in the human radius of an elderly population. *Bone*. 2011; 48(6): 1232–1238.

Muller ME., Webber CE., Bouxsein ML.: Predicting the failure load of the distal radius. *Osteoporosis International*. 2003; 14(4): 345–352.

Müller R.: Hierarchical microimaging of bone structure and function. *Nature reviews. Rheumatology*. 2009; 5(7): 373–381.

Nguyen HT., Worring M., Van den Boomgaard R.: Watersnakes: energy-driven watershed segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2003; 25(3): 330–342.

Nieminen A., Heinonen P., Neuvo Y.: A New Class of Detail-Preserving Filters for Image Processing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1987; PAMI-9(1): 74–90.

Nikander R., Sievänen H., Uusi-Rasi K., Heinonen A., Kannus P.: Loading modalities and bone structures at nonweight-bearing upper extremity and weight-bearing lower extremity: a pQCT study of adult female athletes. *Bone*. 2006; 39(4): 886–894.

Nishiyama KK., Macdonald HM., Hanley DA., Boyd SK.: Women with previous fragility fractures can be classified based on bone microarchitecture and finite element analysis measured with HR-pQCT. *Osteoporosis International*. 2013; 24(5): 1733–1740.

Nodes T., Gallagher, N.C. J.: Two-dimensional root structures and convergence properties of the separable median filter. *IEEE Transactions on Acoustics, Speech and Signal Processing*. 1983; 31(6): 1350–1365.

Osher S., Paragios N. Geometric Level Set Methods in Imaging, Vision, and Graphics. Springer-Verlag, New York, NY, 2003.

Osher S., Sethian JA.: Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*. 1988; 79(1): 12–49.

Ou G., Murphey YL.: Multi-class pattern classification using neural networks. *Pattern Recognition*. 2007; 40(1): 4–18.

Ouyang J., Patel N., Sethi I.: Induction of multiclass multifeature split decision trees from distributed data. *Pattern Recognition*. 2009; 42(9): 1786–1794.

Peng B., Zhang L., Zhang D.: A survey of graph theoretical approaches to image segmentation. *Pattern Recognition*. 2013; 46(3): 1020–1038.

Perona P., Malik J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1990; 12(7): 629–639.

Pesaresi M., Benediktsson JA.: A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*. 2001; 39(2): 309–320.

De Pierro AR., Yamagishi MEB.: Fast EM-like methods for maximum ‘a posteriori’ estimates in emission tomography. *IEEE Transactions on Medical Imaging*. 2001; 20(4): 280–288.

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

Pizurica A., Philips W.: Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising. *IEEE Transactions on Image Processing*. 2006; 15(3): 654–665.

Portilla J., Strela V., Wainwright MJ., Simoncelli EP.: Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*. 2003; 12(11): 1338–1351.

Pratt WK.: Digital image processing. John Wiley and Sons, Inc., New York, 1978.

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

Prevrhal S., Fox JC., Shepherd JA., Genant HK.: Accuracy of CT-based thickness measurement of thin structures: modeling of limited spatial resolution in all three dimensions. *Medical Physics*. 2003; 30(1): 1–8.

Pulkkinen P., Jämsä T., Lochmüller E-M., Kuhn V., Nieminen MT., Eckstein F.: Experimental hip fracture load can be predicted from plain radiography by combined analysis of trabecular bone structure and bone geometry. *Osteoporosis International*. 2008; 19(4): 547–558.

Pulkkinen P., Partanen J., Jalovaara P., Nieminen MT., Jämsä T.: Combination of radiograph-based trabecular and geometrical parameters can discriminate cervical hip fractures from controls in individuals with BMD in non-osteoporotic range. *Bone*. 2011; 49(2): 290–294.

Qi J.: Noise propagation in iterative reconstruction algorithms with line searches. *IEEE Transactions on Nuclear Science*. 2005; 52(1): 57–62.

Qin AK., Clausi DA.: Multivariate Image Segmentation Using Semantic Region Growing With Adaptive Edge Penalty. *IEEE Transactions on Image Processing*. 2010; 19(8): 2157–2170.

Rachidi M., Marchadier A., Gadois C., Lespessailles E., Chappard C., Benhamou CL.: Laws' masks descriptors applied to bone texture analysis: an innovative and discriminant tool in osteoporosis. *Skeletal Radiology*. 2008; 37(6): 541–548.

Ramamurthi K., Ahmad O., Engelke K., Taylor RH., Zhu K., Gustafsson S., Prince RL., Wilson KE.: An in vivo comparison of hip structure analysis (HSA) with measurements obtained by QCT. *Osteoporosis International*. 2012; 23(2): 543–551.

Rangayyan RM.: Biomedical Image Analysis. M.R. Neuman ed. CRC Press, Boca Raton, Florida, USA, 2005.

Rantalainen T., Nikander R., Heinonen A., Cervinka T., Sievänen H., Daly RM.: Differential effects of exercise on tibial shaft marrow density in young female athletes. *The Journal of Clinical Endocrinology and Metabolism*. 2013; 98(5): 2037–2044.

Rantalainen T., Nikander R., Heinonen A., Daly RM., Sievanen H.: An open source approach for regional cortical bone mineral density analysis. *Journal of Musculoskeletal & Neuronal Interactions*. 2011; 11(3): 243–248.

Riggs BL., Melton Iii LJ 3rd., Robb RA., Camp JJ., Atkinson EJ., Peterson JM., Rouleau PA., McCollough CH., Bouxsein ML., Khosla S.: Population-based study of age and sex differences in bone volumetric density, size, geometry, and structure at different skeletal sites. *Journal of Bone and Mineral Research*. 2004; 19(12): 1945–1954.

Riggs BL., Melton LJ 3rd., Robb RA., Camp JJ., Atkinson EJ., Oberg AL., Rouleau PA., McCollough CH., Khosla S., Bouxsein ML.: Population-based analysis of the relationship of whole bone strength indices and fall-related loads to age- and sex-specific patterns of hip and wrist fractures. *Journal of Bone and Mineral Research*. 2006; 21(2): 315–323.

Rittweger J., Beller G., Armbrecht G., Mulder E., Buehring B., Gast U., Dimeo F., Schubert H., de Haan A., Stegeman DF., Schiessl H., Felsenberg D.: Prevention of bone loss during 56 days of strict bed rest by side-alternating resistive vibration exercise. *Bone*. 2010; 46(1): 137–147.

Rittweger J., Frost HM., Schiessl H., Ohshima H., Alkner B., Tesch P., Felsenberg D.: Muscle atrophy and bone loss after 90 days' bed rest and the effects of flywheel resistive exercise and pamidronate: results from the LTBR study. *Bone*. 2005; 36(6): 1019–1029.

Rittweger J., Michaelis I., Giehl M., Wüsecke P., Felsenberg D.: Adjusting for the partial volume effect in cortical bone analyses of pQCT images. *Journal of Musculoskeletal & Neuronal Interactions*. 2004; 4(4): 436–441.

Rittweger J., Simunic B., Bilancio G., De Santo NG., Cirillo M., Biolo G., Pisot R., Eiken O., Mekjavic IB., Narici M.: Bone loss in the lower leg during 35 days of bed rest is predominantly from the cortical compartment. *Bone*. 2009; 44(4): 612–618.

Rittweger J., Winwood K., Seynnes O., de Boer M., Wilks D., Lea R., Rennie M., Narici M.: Bone loss from the human distal tibia epiphysis during 24 days of unilateral lower limb suspension. *Journal of Physiology*. 2007; 577(Pt 1): 331–337.

Rizzoli R., Chapurlat RD., Laroche J-M., Krieg MA., Thomas T., Frieling I., Boutroy S., Laib A., Bock O., Felsenberg D.: Effects of strontium ranelate and alendronate on bone microstructure in women with osteoporosis. Results of a 2-year study. *Osteoporosis International*. 2012; 23(1): 305–315.

Roux J-P., Wegrzyn J., Arlot ME., Guyen O., Delmas PD., Chapurlat R., Bouxsein ML.: Contribution of trabecular and cortical components to biomechanical behavior of human vertebrae: an ex vivo study. *Journal of Bone and Mineral Research*. 2010; 25(2): 356–361.

Saha PK., Udupa JK.: Scale-based diffusive image filtering preserving boundary sharpness and fine structures. *IEEE Transactions on Medical Imaging*. 2001; 20(11): 1140–1155.

Sahoo PK., Soltani S., Wong AKC.: A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*. 1988; 41(2): 233–260.

Salinas HM., Fernandez DC.: Comparison of PDE-Based Nonlinear Diffusion Approaches for Image Enhancement and Denoising in Optical Coherence Tomography. *IEEE Transactions on Medical Imaging*. 2007; 26(6): 761–771.

Sanches JM., Nascimento JC., Marques JS.: Medical Image Noise Reduction Using the Sylvester-Lyapunov Equation. *IEEE Transactions on Image Processing*. 2008; 17(9): 1522–1539.

Sanfeliu A., Alquézar R., Andrade J., Climent J., Serratosa F., Vergés J.: Graph-based representations and techniques for image processing and image analysis. *Pattern Recognition*. 2002; 35(3): 639–650.

Sauvola J., Pietikäinen M.: Adaptive document image binarization. *Pattern Recognition*. 2000; 33(2): 225–236.

Schalkoff RJ. Digital image processing and computer vision. John Wiley & Sons, Inc. New York, NY, 1989.

Seeley DG., Browner WS., Nevitt MC., Genant HK., Scott JC., Cummings SR.: Which fractures are associated with low appendicular bone mass in elderly women? The Study of Osteoporotic Fractures Research Group. *Annals of Internal Medicine*. 1991; 115(11): 837–842.

Seeman E.: Periosteal bone formation--a neglected determinant of bone strength. *The New England Journal of Medicine*. 2003; 349(4): 320–323.

Seeman E., Delmas PD., Hanley DA., Sellmeyer D., Cheung AM., Shane E., Kearns A., Thomas T., Boyd SK., Boutroy S., Bogado C., Majumdar S., Fan M., Libanati C., Zanchetta J.: Microarchitectural deterioration of cortical and trabecular bone: differing effects of denosumab and alendronate. *Journal of Bone and Mineral Research*. 2010; 25(8): 1886–1894.

Sendur L., Selesnick IW.: Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Transactions on Signal Processing*. 2002; 50(11): 2744–2756.

Senthilkumaran N., Rajesh R.: Edge detection techniques for image segmentation-A survey of soft computing approaches. *International Journal of Recent Trends in Engineering*. 2009; 1(2): 250–254.

Sezgin M., Sankur B.: Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*. 2004; 13(1): 146–168.

Sharma N., Aggarwal LM.: Automated medical image segmentation techniques. *Journal of Medical Physics*. 2010; 35(1): 3–14.

Shedd-Wise KM., Alekel DL., Hofmann H., Hanson KB., Schiferl DJ., Hanson LN., Van Loan MD.: The soy isoflavones for reducing bone loss study: 3-yr effects on pQCT bone mineral density and strength measures in postmenopausal women. *Journal of Clinical Densitometry*. 2011; 14(1): 47–57.

Shen T., Li H., Huang X.: Active Volume Models for Medical Image Segmentation. *IEEE Transactions on Medical Imaging*. 2011; 30(3): 774–791.

Sievänen H.: A physical model for dual-energy X-ray absorptiometry--derived bone mineral density. *Investigative Radiology*. 2000; 35(5): 325–330.

Sievänen H.: Bone densitometry and true BMD accuracy for predicting fractures: what are the alternatives? *International Journal of Clinical Rheumatology*. 2010a; 5(3): 371–385.

Sievänen H.: Immobilization and bone structure in humans. *Archives of Biochemistry and Biophysics*. 2010b; 503(1): 146–152.

Sievänen H., Kannus P., Järvinen TLN.: Bone quality: an empty term. *PLoS medicine*. 2007; 4(3): e27.

Sievänen H., Koskue V., Rauhio A., Kannus P., Heinonen A., Vuori I.: Peripheral quantitative computed tomography in human long bones: evaluation of in vitro and in vivo precision. *Journal of Bone and Mineral Research*. 1998; 13(5): 871–882.

Sievänen H., Oja P., Vuori I.: Precision of dual-energy x-ray absorptiometry in determining bone mineral density and content of various skeletal sites. *Journal of Nuclear Medicine*. 1992; 33(6): 1137–1142.

Silva TC., Zhao L.: Network-Based High Level Data Classification. *IEEE Transactions on Neural Networks and Learning Systems*. 2012; 23(6): 954–970.

Sonka M., Fitzpatrick JM. Handbook of Medical Imaging: Medical image processing and analysis. SPIE Press, Bellingham, Washington, USA, 2000.

Souza A., Udupa JK., Madabhushi A.: Image filtering via generalized scale. *Medical Image Analysis*. 2008; 12(2): 87–98.

Sowmya B., Rani BS.: Colour image segmentation using fuzzy clustering techniques and competitive neural network. *Applied Soft Computing*. 2011; 11(3): 3170–3178.

Stewart RD., Fermin I., Opper M.: Region growing with pulse-coupled neural networks: an alternative to seeded region growing. *IEEE Transactions on Neural Networks*. 2002; 13(6): 1557–1562.

Stone KL., Seeley DG., Lui L-Y., Cauley JA., Ensrud K., Browner WS., Nevitt MC., Cummings SR., Osteoporotic Fractures Research Group.: BMD at multiple sites and risk of fracture of multiple types: long-term results from the Study of Osteoporotic Fractures. *Journal of Bone and Mineral Research*. 2003; 18(11): 1947–1954.

Stratec.: XCT 2000 Manual Software Version 5.50, 2004a.

Stratec.: XCT Research Series Manual Software. Version 5.50, 2004b.

Suri JS., Liu K., Singh S., Laxminarayan SN., Zeng X., Reden L.: Shape recovery algorithms using level sets in 2-D/3-D medical imagery: a state-of-the-art review. *IEEE Transactions on Information Technology in Biomedicine*. 2002; 6(1): 8–28.

Szabo KA., Webber CE., Adachi JD., Tozer R., Gordon C., Papaioannou A.: Cortical and trabecular bone at the radius and tibia in postmenopausal breast cancer patients: a Peripheral Quantitative Computed Tomography (pQCT) study. *Bone*. 2011; 48(2): 218–224.

Toh KKV., Isa NAM.: Cluster-based adaptive fuzzy switching median filter for universal impulse noise reduction. *IEEE Transactions on Consumer Electronics*. 2010; 56(4): 2560–2568.

Treece GM., Gee AH., Mayhew PM., Poole KES.: High resolution cortical bone thickness measurement from clinical CT data. *Medical Image Analysis*. 2010; 14(3): 276–290.

Triguero I., Derrac J., Garcia S., Herrera F.: A Taxonomy and Experimental Study on Prototype Generation for Nearest Neighbor Classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*. 2012; 42(1): 86–100.

Tsai D-M.: A fast thresholding selection procedure for multimodal and unimodal histograms. *Pattern Recognition Letters*. 1995; 16(6): 653–666.

Tuia D., Volpi M., Copa L., Kanevski M., Munoz-Mari J.: A Survey of Active Learning Algorithms for Supervised Remote Sensing Image Classification. *IEEE Journal of Selected Topics in Signal Processing*. 2011; 5(3): 606–617.

Tukey JW.: Exploratory data analysis. Reading. Mass.: Addison-Wesley Pub. Co., Boston, MA, 1977.

Udupa JK., LeBlanc VR., Zhuge Y., Imielinska C., Schmidt H., Currie LM., Hirsch BE., Woodburn J.: A framework for evaluating image segmentation algorithms. *Computerized Medical Imaging and Graphics*. 2006; 30(2): 75–87.

Udupa JK., Saha PK., Lotufo RA.: Disclaimer: 'Relative fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation'. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24(11): 1485–1500.

Valentinitsch A., Patsch JM., Deutschmann J., Schueller-Weidekamm C., Resch H., Kainberger F., Langs G.: Automated threshold-independent cortex segmentation by 3D-texture analysis of HR-pQCT scans. *Bone*. 2012; 51(3): 480–487.

Vasilic B., Wehrli FW.: A novel local thresholding algorithm for trabecular bone volume fraction mapping in the limited spatial resolution regime of in vivo MRI. *IEEE Transactions on Medical Imaging*. 2005; 24(12): 1574–1585.

Veitch SW., Findlay SC., Ingle BM., Ibbotson CJ., Barrington A., Hamer AJ., Eastell R.: Accuracy and precision of peripheral quantitative computed tomography measurements at the tibial metaphysis. *Journal of Clinical Densitometry*. 2004; 7(2): 209–217.

Vilayphiou N., Boutroy S., Szulc P., van Rietbergen B., Munoz F., Delmas PD., Chapurlat R.: Finite element analysis performed on radius and tibia HR-pQCT images and fragility fractures at all sites in men. *Journal of Bone and Mineral Research*. 2011; 26(5): 965–973.

Villain N., Goussard Y., Idier J., Allain M.: Three-dimensional edge-preserving image enhancement for computed tomography. *IEEE Transactions on Medical Imaging*. 2003; 22(10): 1275–1287.

Vincent L., Beucher S. The morphological approach to segmentation: An introduction. Internal Rep. C-08/89/MM, School of Mines, Paris, France, 1989. Available at: http://www.vincent-net.com/luc/papers/.

Vincent L., Soille P.: Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1991; 13(6): 583–598.

Ward KA., Adams JE., Hangartner TN.: Recommendations for thresholds for cortical bone geometry and density measurement by peripheral quantitative computed tomography. *Calcified Tissue International*. 2005; 77(5): 275–280.

Weickert J.: Coherence-Enhancing Diffusion Filtering. *International Journal of Computer Vision*. 1999; 31(2-3): 111–127.

Williams MB., Krupinski EA., Strauss KJ., Breeden WK 3rd., Rzeszotarski MS., Applegate K., Wyatt M., Bjork S., Seibert JA.: Digital radiography image quality: image acquisition. *Journal of the American College of Radiology: JACR*. 2007; 4(6): 371–388.

Withey DJ., Koles ZJ.: Three Generations of Medical Image Segmentation: Methods and Available Software. *International Journal of Bioelectromagnetism*. 2007; 9(2): 67–68.

Wu J., Chung ACS.: A Segmentation Model Using Compound Markov Random Fields Based on a Boundary Model. *IEEE Transactions on Image Processing*. 2007; 16(1): 241–252.

Xu C., Prince JL.: Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*. 1998; 7(3): 359–369.

Xu R., Wunsch II D.: Survey of clustering algorithms. *IEEE Transactions on Neural Networks*. 2005; 16(3): 645–678.

Yang L., Maric I., McCloskey EV., Eastell R.: Shape, structural properties, and cortical stability along the femoral neck: a study using clinical QCT. *Journal of Clinical Densitometry*. 2008; 11(3): 373–382.

Yang Y., Zheng C., Lin P.: Fuzzy clustering with spatial constraints for image thresholding. *Optica Applicata*. 2005; 35(4): 943–954.

Yao W., Abolmaesumi P., Greenspan M., Ellis RE.: An estimation/correction algorithm for detecting bone edges in CT images. *IEEE Transactions on Medical Imaging*. 2005; 24(8): 997–1010.

Yen J-C., Chang F-J., Chang S.: A new criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing*. 1995; 4(3): 370–378.

Yin P-Y.: Multilevel minimum cross entropy threshold selection based on particle swarm optimization. *Applied Mathematics and Computation*. 2007; 184(2): 503–513.

Yoo TS.: Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis. A K Peters, Ltd., Wellesley, Mass., USA, 2004.

Yu W., Glüer CC., Grampp S., Jergas M., Fuerst T., Wu CY., Lu Y., Fan B., Genant HK.: Spinal bone mineral assessment in postmenopausal women: a comparison between dual X-ray absorptiometry and quantitative computed tomography. *Osteoporosis International*. 1995; 5(6): 433–439.

Yu Z., Au OC., Zou R., Yu W., Tian J.: An adaptive unsupervised approach toward pixel clustering and color image segmentation. *Pattern Recognition*. 2010; 43(5): 1889–1906.

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

Zemel B., Bass S., Binkley T., Ducher G., Macdonald H., McKay H., Moyer-Mileur L., Shepherd J., Specker B., Ward K., Hans D.: Peripheral quantitative computed tomography in children and adolescents: the 2007 ISCD Pediatric Official Positions. *Journal of Clinical Densitometry*. 2008; 11(1): 59–74.

Zhang J., Yan C-H., Chui C-K., Ong S-H.: Fast segmentation of bone in CT images using 3D adaptive thresholding. *Computers in Biology and Medicine*. 2010; 40(2): 231–236.

Zhong J., Ning R.: Image denoising based on wavelets and multifractals for singularity detection. *IEEE Transactions on Image Processing*. 2005; 14(10): 1435–1447.

Zijdenbos AP., Dawant BM., Margolin RA., Palmer AC.: Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Trans Med Imaging*. 1994; 13(4): 716–724.

Zoroofi RA., Sato Y., Sasama T., Nishii T., Sugano N., Yonenobu K., Yoshikawa H., Ochi T., Tamura S.: Automated segmentation of acetabulum and femoral head from 3-d CT images. *IEEE Transactions on Information Technology in Biomedicine*. 2003; 7(4): 329–343.

Publication I:

Cervinka T., Hyttinen J., Sievanen H.
Enhanced bone structural analysis through pQCT image preprocessing.
Medical Engineering & Physics, 2010; 32(4): 398-406.

Enhanced bone structural analysis through pQCT image preprocessing

T. Cervinka

Department of Biomedical Engineering, Tampere University of Technology, Tampere, Finland

J. Hyttinen

Department of Biomedical Engineering, Tampere University of Technology, Tampere, Finland

H. Sievanen

Department of Biomedical Engineering, Tampere University of Technology, Tampere, Finland; Bone Research Group, UKK Institute, Tampere, Finland; Pirkanmaa Hospital District, Science Center, Tampere, Finland

Finn-medi 1, 4. Floor, Biokatu 6, 33520 Tampere, Tel.: +358408490165, Fax.: +358331164013, Email: tomas.cervinka@tut.fi

ABSTRACT

Several factors, including preprocessing of the image, can affect the reliability of pQCT-measured bone traits, such as cortical area and trabecular density. Using repeated scans of four different liquid phantoms and repeated in vivo scans of distal tibiae from 25 subjects, the performance of two novel preprocessing methods, based on the down-sampling of grayscale intensity histogram and the statistical approximation of image data, was compared to 3×3 and 5×5 median filtering. According to phantom measurements, the signal to noise ratio in the raw pQCT images (XCT 3000) was low (~ 20 dB) which posed a challenge for preprocessing. Conerning the cortical analysis, the reliability coefficient (R) was 67% for the raw image and increased to 94–97% after preprocessing without apparent preference for any method. Concerning the trabecular density, the R values were already high (~99%) in the raw images leaving virtually no room for improvement. However, some coarse structural patterns could be seen in the preprocessed images in contrast to a disperse distribution of density levels in the raw image. In conclusion, preprocessing cannot suppress the high noise level to the extent that the analysis of mean trabecular density is essentially improved, whereas preprocessing can enhance cortical bone analysis and also facilitate coarse structural analyses of the trabecular region.

Key words: pQCT; cortical bone; trabecular bone; Bayesian estimation; image preprocessing

I. INTRODUCTION

Fragility fractures are common and their incidence increases substantially with age. Dual energy X-ray absorptiometry (DXA) is the present standard of diagnosis of osteoporosis and the DXA-measured bone mineral density (BMD) is the most common clinical measure to assess bone fragility and fracture risk.[1] Paradoxically, a majority of the so-called osteoporotic fractures occur among persons who are not osteoporotic according to DXA assessment, i.e., their BMD is not 2.5 standard deviations lower than the reference value.[2] Further, the reliability of BMD can be severely compromised by inherent inaccuracies and ambiguities in interpretation.[3-5]

Since a great number of non-bone factors contribute to fragility fractures, it is obvious that not even a perfect *in vivo* assessment of bone fragility could predict with high sensitivity and specificity whether a given person will sustain a fracture.[6] However, given the high socio-economic relevance of fragility fractures, there is a need for a clinical method (or methods) that can detect specific traits of bone fragility more efficiently than the conventional DXA-measured BMD.

Bone fragility, or conversely, its strength is largely determined by structural properties (e.g., cross-sectional geometry, cortical thickness, and trabecular architecture).[7] Accordingly, a proper structural analysis of bones might help identify individuals susceptible to fragility fractures. In this respect, clinical quantitative computed tomography (QCT) systems allow reasonable in vivo assessment of bone geometry and apparent cortical and trabecular densities at clinically relevant proximal femur and lumbar spine sites,[8-12] but these systems lack sufficient spatial resolution to separate between individual trabeculae making them unable to image specific structural traits of trabecular architecture (eg, trabecular thickness, separation, orientation, connectivity). Evidently, architectural features account for bone fragility.[13,14] Also, the radiation dose of the QCT investigation is large compared to that of standard DXA investigation,[15] which may pose an ethical issue and limit the widespread use of QCT in clinical bone research. Notwithstanding the facts that quantitative ultrasound (QUS) does not require ionizing radiation to

provide independent aggregate information on bone structure and material properties (~bone quality) and that the QUS measurements are associated with fracture risk similar to DXA,[16] the ability of QUS to image the bone structure in a tangible way remains quite limited at present. In the end, it is the actual bone structure that ultimately counts and even a coarse structural analysis may be helpful.[7] Obviously the low-dose peripheral QCT (pQCT) cannot be applied to the axial skeleton or hip region, but this method provides similar information on coarse structural traits for appendicular bones as does QCT for clinically relevant proximal femur and lumbar vertebral sites.[17] While the direct site-specific assessments describe best the strength of the given proximal femur or vertebral site, peripheral pQCT measurements can provide useful information in this respect.[18,19] Therefore, peripheral bone sites, such as the distal tibia showing a relatively thin cortex and trabecular bone within the cortical shell, may offer a good model for developing new algorithms and refining existing algorithms for image processing and analysis. These approaches, if found feasible and reliable, may be applied to coarse structural assessment of different bones, including the clinically relevant bones accessible with clinical QCT.

The primary objective of the present study was to evaluate whether the preprocessing of clinical pQCT images could improve the analysis of coarse bone structure in comparison to analysis of raw, unprocessed image, and if so, to what extent. We hypothesized that the preprocessing would enhance the quality of pQCT images so that the relevant traits of coarse bone structure could be more precisely detected. The other objective was to investigate the signal to noise ratio (SNR) and concomitant grayscale resolution of pQCT images. The information on noise was considered essential for devising which preprocessing method would be most appropriate for enhancing the quality of the pQCT image. Besides using appropriate phantoms, repeated *in vivo* pQCT images of the distal tibia were employed due to *a priori* knowledge that the paired images should basically provide identical results on bone structural traits. These analyses allowed us a straightforward comparison between different preprocessing methods. The new preprocessing methods described in this study were based on the statistical approximation of data

within the pQCT image. Statistical approaches have shown promising results in other fields such as pattern recognition and image processing.[20-22]

II. MATERIALS AND METHODS

II.A. The pQCT system and in vivo data

The pQCT scanner (XCT 3000, Stratec Medizintechnik GmbH, Pforzheim, Germany) used in this study represents one model of the XCT brand (960, 2000, 3000) commonly employed in clinical bone research. In short, this device produces a narrow fan beam using a collimated and filtered X-ray source (focal spot size 0.25 × 0.25 mm^2, tube voltage 60 kVp, current 0.3 mA, filter 6 mm of Al + 0.5 mm of Cu). The source collimator (dimensions 1.5 mm × 8.5 mm) was located 44 mm above the focal spot. The detector consisted of 12 cadmium telluride detectors in a row. The detector collimator had a separate aperture (dimensions 4 mm × 0.8 mm) for each detector, and the distance from the source collimator was 343 mm.

Repeated pQCT images of the distal tibia (5% of the estimated tibial length from the distal endplate) obtained from 25 healthy volunteers were used to test the performance of preprocessing methods. The age range of the subjects was 35–66 yrs; weight range was 50.4–89.1 kg; and height range was 159–179 cm. They had no history of osteoporotic fracture, but were otherwise unselected. The pQCT scans were obtained from a precision study which carried out as a part of quality assurance procedure of our bone densitometry unit. Informed consent was obtained from the subjects and the in-house review board approved the study protocol. All scans were performed the same day with repositioning between the scans, according to our standard procedure.[17] The pixel-size of the pQCT image was 0.5 mm × 0.5 mm, the slice thickness was 2.5 mm, and the translational scan speed was 30 mm/s.

Of note, in order to reduce the undue influence of excess movement artefacts on the bone analysis, differences larger than 25 mm^2 in the total cross-sectional area between the repeated scans were excluded; data from one subject was excluded.

II.B. Phantom pQCT data

For the noise assessment of pQCT imaging, four phantoms made of small plastic containers filled with one of three known concentrations (c) of homogeneous K_2HPO_4 solution (50 mg/cm^3, 100 mg/cm^3, 250 mg/cm^3) or tap water, were scanned 12 times consecutively using the same scan parameters as in the *in vivo* scans.[17] According to the previously determined relationship between the pQCT-measured densities (equivalent concentration of bone mineral in mg/cm^3) and known K_2HPO_4 concentrations,[17] the densities of the phantoms were 98.4 mg/cm^3, 146.5 mg/cm^3, and 291 mg/cm^3 , respectively. Of note, the phantom densities cover well the whole range of pQCT-measured trabecular densities in young athletic and elderly women.[23,24] In addition to being used in evaluating the noise level of pQCT images, the phantom data were used for calibrating the mean pixel values of preprocessed images against bone (trabecular) densities.

II.C. Preprocessing of pQCT data

The raw, unprocessed pQCT image data (result of the pQCT scanner back projection algorithm) was used as a reference to which the preprocessed images were compared. The standard image preprocessing methods, median filtering with a 3×3 window (employed by the Stratec analysis software in preprocessing) and a 5×5 window, were applied to the raw pQCT data.

In addition to the standard two median filters, two novel preprocessing methods were introduced. These methods were composed in two steps. First, a gray–level transformation of each image was carried out by implementing a common, iteratively employed, piecewise linear conversion function. Second, the transformed data was corrected for inaccuracies caused by the re-quantization procedure by using the Bayesian approach and Markov random fields with 3×3 linear or 5×5 Gaussian neighborhoods and redundant wavelet transform (Fig. 1).

Gray-level transformation

The gray-level transformation reduces the number of used grayscales in order to make subsequent image processing easier and faster. This transformation can be regarded as a non-equidistant re-quantization of the gray scale. The example of conversion function after the first iteration is depicted in Fig. 2.

Redundant wavelet transform

A redundant wavelet transform with bio-orthogonal filters was used to decompose the image after the above described gray-level transformation. The wavelet transform provides decomposition of a function on a particular basis of specific functions, called wavelets. A one-dimensional function *f* is decomposed as

$$f = \sum_{j,l} w_{j,l} \cdot \psi_{j,l} \qquad (1)$$

, which defines the transform of *f* to its wavelet coefficients $w_{j,l}$. The wavelets are defined in terms of a mother wavelet ψ(*t*) as $\psi_{j,l} = \sqrt{2^j}\,\psi\left(2^j t - 1\right)$. The basis wavelets differ by their position *l* and level *j*. The factor 2^j is the characteristic scale of all wavelets at level *j*. Redundant wavelet transform is used to achieve separated frequency bands of the preprocessed image with the same number of coefficients in each frequency band. Separation into the frequency bands provides information on the wavelet coefficients that represent important features in the image. These coefficients can be found at all frequency bands, while noise appears mostly at the higher frequency bands. Detailed information about wavelet transforms can be found for example in (Aldroubi A and Unser M 1996).[25]

Requantization

Through the wavelet transform separated frequency bands of the processed image are obtained. A priori knowledge, that wavelet coefficients with similar values are usually locally concentrated in the preprocessed images across all decomposition levels, is taken into account. A conditional probability can be defined as follows; if a wavelet coefficient (obtained after the decomposition of transformed image) has a certain value, it is more likely that the value at the same location before reconstruction is the same than that a different one.

Thus using the modified Bayes' rule, a posteriori probability of wavelet coefficients can be derived:

$$P(\boldsymbol{A}|\boldsymbol{B}) \propto P(\boldsymbol{B}|\boldsymbol{A}) \cdot P(\boldsymbol{A}) \qquad (2)$$

The first term in the right side is the conditional probability, the second term is the *a priori* probability. Both probabilities are usually modelled as Gibbs probability functions. This approach has an advantage that all variables can be described directly with an image model called a Markov random field. The relation between Markov random field and Gibbs probability functions is expressed in the Hammersly-Clifford theorem.[26]

In our approach the states of the Markov random field are pixels in neighbourhoods 3-by-3 (LinW) or 5-by-5 (GaussW). The a priori probability expresses that pixels in each neighbouring state vectors have the same label value, are more probable than those with different values and its calculation is based on a comparison of the central state with its neighbours. Thus the posterior probability of the value of the pixel depends on the neigbourhood as well.

Because of the different sizes and types of the Markov random fields (basic steps of both approaches were similar), we termed the preprocessing methods used the Linear window (LinW) and Gaussian window (GaussW). A variation of the present approach has been recently described by Xie et al.[27]

II.D. Noise analysis

The noise in the pQCT images of homogeneous K_2HPO_4 phantoms was characterized in the following manner. For each phantom density level and above described preprocessing method, the SNR (in dB) was determined from the data obtained from the repeated phantom scans as follows:

$$SNR = 20 \log\left(\frac{\sum mean_i}{12} \Big/ \sqrt{\frac{\sum SD_i^2}{12}} \right), \quad i = 1..12 \qquad (3)$$

, where $mean_i$ is the mean of all pixel values in the i^{th} image and SD_i is the standard deviation (SD) of separate pixel values in the i^{th} image. Pixels from the whole cross-section of the phantom (> 5000 pixels) were included in the noise analysis, the plastic wall excluded. The SNR values then used for evaluating the performance of different preprocessing methods in noise reduction. Finally, applying the above-mentioned preprocessing methods to the pQCT images of the phantoms, the maximum number of density levels that could be exclusively separated from adjacent levels was determined.

II.E. pQCT image analysis

All pQCT images, both the raw image and four preprocessed images were segmented into cortical and trabecular bone areas by using a simple and fast algorithm described by Seits et al.[28] This method was further refined to improve the tracing of outer border of cortical bone by using the maximum values after the first derivation of the image data. Then the region inside the traced contour was shrunk over the valley lying between the outer contour towards the next maximum values of the first derivation which would match with the inner border of the cortical bone. The shrunk region corresponded to the trabecular bone and subtraction of the original and shrunk bone region represented the cortical bone. Then, the conventional pQCT traits, cortical cross-sectional area (CoA) and trabecular density (TrD, expressed as the equivalent concentration of bone mineral in mg/cm^3), were determined as coarse descriptors of bone structure. In addition to analyzing the total bone cross-section, the cross-section was divided into four

anatomic sectors (lateral, posterior, medial, anterior) so that direction-specific features of the cortical and trabecular bone could be assessed (Fig. 3). The CoA was expressed either as a percentage of total bone cross-sectional area or as a percentage of total bone area of the given anatomic sector, as appropriate.

For the classification of trabecular pixels into exclusive (non-overlapping) density ranges, the k-nearest neighbor (k-NN) classifier technique was used.[29] Briefly, this method is based on assigning each trabecular pixel to the class of the majority of its k-neighbors;[30] that is, assuming that the number of voting neighbors is $k = k_1 + k_2 + \ldots + k_n$ (where k_i is the number of pixels from class i in the k-sample neighborhood of the test sample), the test sample is assigned to class m if $k_m = \max (k_i ; i = 1, 2, \ldots, n)$. The number of classes i was set according the *a priori* information gained from image analysis of the phantoms. Since the phantom images were acquired and preprocessed similarly to the *in vivo* pQCT scans, this classification approach was applied to *in vivo* data of the distal tibia in an attempt to assess structural patterns within the trabecular region.

II.F. Statistical analysis

Mean value, standard deviation (SD) of bone traits, and SD of differences between the repeated measurements (SD_{meas} ~in vivo precision) are given as descriptive statistics for each preprocessing method. It should be noted that the CoA values obtained from different methods were made comparable by presenting them as a percentage proportion of the total bone area. As another measure of reliability of the measurement (~in vivo precision), the reliability coefficient ($R = 100(1 - SD^2_{meas}/SD^2)\%$) was determined. The R value represents the error-free proportion of the inter-subject variability observed in the given bone trait.[31]

III. RESULTS

Table I shows the influence of preprocessing on the assessment of CoA. Clearly, the preprocessing of the image data improved the reliability of the cortical analysis (for the total cortical area, from R = 67% of the raw image to 94–97%), while no substantial preference was observed between different preprocessing methods. For the raw images, it is worth noting that the precision was poorer at the antero-medial regions of the distal tibia ($R < 50\%$) and much greater ($R \geq 75\%$) at the postero-lateral regions. Among the preprocessing methods, no such direction dependence was seen; however, the R-value was consistently higher with the LinW method compared to the 3×3 median filter, which is the current standard preprocessing method in pQCT image analysis.

Table II shows the influence of preprocessing on the assessment of TrD. In contrast to enhanced cortical analysis, the preprocessing did not improve the precision of the TrD assessment at any anatomic region of the distal tibia. Precision obtained with the raw images was so high (all R values ~99% or higher) that there was no room for any improvement. Noteworthy, the highest mean trabecular density was observed at the posterior region and the lowest at the anterior region.

Table III shows the mean pixel values and SD of separate pixel values within the phantom images after preprocessing, and the effect of preprocessing on the SNR in the pQCT images. For both median filters, the noise suppression was independent of density; whereas the suppression improved with density for the LinW and GaussW methods. The SNR ranged from 21(22) dB for tap water to 25(26) dB for the highest density phantom for GaussW and LinW (in brackets), respectively. Regarding the background (~air) noise, LinW and GaussW methods improved the SNR by ~15 db compared to the ~9 dB improvement for median filters. Fig. 4 indicates that the pQCT image contains a relatively large amount of random noise in addition to some nonrandom noise.

It is also worth noting that the dynamic range was greater for the LinW and GaussW than the median windows or raw image; the max/min ratios were 2.5 and 1.6, respectively. Smaller absolute mean pixel values obtained from LinW and GaussW are due to the gray-level transformation.

As can be inferred from the mean and SD values in Table III, only six distinct (equidistant) density ranges corresponding to tap water (50.2 mg/cm^3), 98.4 mg/cm^3, 146.5 mg/cm^3, 194.7 mg/cm^3, 242.8 mg/cm^3 and 291 mg/cm^3 can be exclusively separated (i.e., the density ranges were not allowed to overlap their neighbors). Four levels (three lowest and the highest one) corresponded to the densities of respective phantoms, while two intermediate levels were interpolated using the regression equation between measured density levels and phantom concentrations[17] and observed SD values in the densities of neighborhood phantoms.

Fig. 5 illustrates the results from the k-NN classifier analysis using a pair of repeated scans as an example. The colored density regions correspond to density ranges that could be unequivocally separated from each other. Some structural patterns could be seen in the preprocessed images, whereas a disperse distribution of density levels appeared in the raw images. After preprocessing of the image, the density increased consistently from the anterior to the posterior region, which was not evident in the raw image.

IV. DISCUSSION

While the pQCT imaging has potential in clinical evaluation of bone fragility and various research applications;[32,33] there are several factors that can compromise its performance and reliability. Namely, pQCT images are subject to movement artefacts that are largely attributable to the slow scan speed and long imaging time,[17] to the great variance in scan and analysis parameters,[34-37] and also to the type of scanner used, even within the same brand.[38] In this study, special emphasis was placed on evaluating the noise present in pQCT images. We elaborated this rarely discussed topic in order to compare and devise appropriate method(s) that could improve the in vivo precision of pQCT image analysis, and thus,

enhance the evaluation of commonly used bone traits, including the cortical area and trabecular density. Should the precision be improved, smaller changes in bone traits could be detected at the given sample size of study subjects, or smaller studies could be undertaken.

As anticipated, the preprocessing of pQCT image data improved the analysis of cortical bone, but to our surprise, not that of trabecular density. It appeared that the true dynamic range of pQCT imaging is poor; allowing separation of six distinct gray levels (~non-overlapping density ranges) only. However, this limited resolution may open a new way to examine the trabecular bone, particularly the apparent structural patterns within (~spatial distribution of bone mass within the bone cross-section), as illustrated in our preliminary analysis (Fig. 5). This approach naturally requires further elaboration.

Cortical bone has a vital role in maintaining bone rigidity.[40-45] However, reliable analysis of cortical bone with pQCT, particularly at the thin-walled long bone ends, is challenging because of the relatively low spatial resolution.[17,34,39] Several studies have evaluated cortical bone analysis from the viewpoint of precision, influence of the scan or analysis parameters,[17,34-36,46] but to our knowledge, the influence of the preprocessing on the pQCT image quality has not yet been evaluated. The present study showed that the preprocessing can markedly enhance the analysis of the cortical area. The improvement in reliability, judged from the R values, was approximately 27–30%. However, the differences between the preprocessing methods remained marginal and no preference could be determined. In contrast, for the direction-specific analysis, the most consistent results were obtained by LinW preprocessing, the R values were close to 90% or better.

Trabecular density is considered the major trait of interest in clinical osteoporosis research,[32] mainly because of the greater expected changes in this trait in response to medical treatment, physiological conditions, or aging. Nevertheless, the influence of preprocessing on the reliability of pQCT results has not yet been explored. Our hypothesis was that the preprocessing would render the results of trabecular density analysis more consistent, but this was not the case. Neither the novel LinW or GaussW

approaches, nor the conventional median filters were able to reduce variance in trabecular density between repeated scans. The explanation is apparent: the R values for the raw images were close to 100%, leaving no room for improvement. These negligible returns from both simple and sophisticated preprocessing of the trabecular region are probably due to the high noise level, and to some extent, also to its non-random nature. This kind of noise cannot be efficiently managed by preprocessing methods based on universal setting approaches that cannot take into account specific characteristics of noise.

In addition to observing low SNR (~20 – 30 dB) in the pQCT image, we attempted to determine specific characteristics of noise by analyzing repeated scans of homogeneous K_2HPO_4 liquid phantoms. In addition to random variation, which could be reduced, the noise appeared to comprise some features that could not be suppressed by preprocessing (Fig. 4). Apparent sources of the latter type of noise include the high-pass filtration and back projection of Poisson and Gaussian noise in the projections, the back-projection reconstruction algorithm, variation in the translational and rotational motion of the X-ray source and detector assembly during the scan, non-uniform intensity of the X-ray beam and detector characteristics, as well as the movement of the patient during the scan. All of these factors may modulate the noise in the pQCT image separately or interactively.

While the benefits of preprocessing in enhancing the cortical analysis were evident compared to the results from the raw images, one may ask whether the preprocessing of pQCT images in terms of trabecular density is worthwhile. Obviously, filtering of the image, depending on the type and specific parameters of the method, results in greater or less substantial loss of true information while suppressing noise. However, as can be judged from the similar mean trabecular density and standard deviation values obtained after different preprocessing methods (Table II), the loss of essential information, at least at the group level, does not seem to be a real concern. In contrast, preprocessing may reveal some coarse structural features (those reflecting spatial distribution of different trabecular density regions within the bone cross-section), not only within the whole cross-section, but also within the anatomic sectors. It is

noted that these apparent structural patterns were not distinct in the raw image (Fig. 5). In particular, the region-specific analysis may reveal some interesting associations between bone loading and density variation as the recent analyses based on small regions of interest by Lai et al.[47-49] suggested. Further, regarding the benefits of preprocessing, the dynamic range of the density values might be increased using sophisticated methods (LinW, GaussW). The LinW method also appeared to provide the most consistent cortical results for all anatomical sectors of the distal tibia. Further studies are needed to refine these approaches and assess their true value.

There are some issues that need further discussion. First, the benefits of sophisticated preprocessing methods need to be evaluated in terms of required processing power and time. Second, the apparent utility and validity of the obtained results need further corroboration. While being basically an important issue, the processing time is considered irrelevant since the structural analysis is generally not done in real time but after the scanning the patient scan. Obviously, simpler techniques, like median filtering, are faster by an order than the more sophisticated approaches assessed in this study, but the difference in processing time may be reduced by optimizing the code of the latter methods. Moreover, given the substantial processing power of modern PCs, the processing time cannot be considered an issue. Obviously, further studies are needed to show the utility of proposed preprocessing approaches in assessing patient-specific bone strength and fracture risk. As regards the validity of the preprocessed results, it is recalled that the spatial distribution of bone and non-bone regions of the trabecular architecture is mainly smaller than the resolution of the standard pQCT image. It may be thus possible that the noise reduction algorithms may introduce errors in the subsequent interpretation of trabecular structure rather than reveal true structural features, even at the coarse level. If so, the raw pQCT images would be a better representation of the trabecular architecture than the preprocessed images. In this respect, it would be interesting to compare images obtained from high resolution QCT (HR-QCT) to those obtained from low-resolution pQCT to determine the effect of true noise on the apparent structural traits and to determine the validity of the present preprocessing approaches in revealing true structural features in bones.

In conclusion, this study showed that the signal to noise ratio in the pQCT image can be rather low, and the noise, particularly that present at the trabecular region, cannot be suppressed with preprocessing to the extent that the analysis of trabecular density is essentially improved. However, preprocessing may enhance cortical bone analysis and also facilitate coarse structural analyses of the trabecular bone.

ACKNOWLEDGEMENTS

We would like to thank Markus Hannula for creating the testing models used for developing our method. We also acknowledge the Finnish Cultural Foundation and Research School of Tampere University of Technology for financial support of the doctoral studies for TC. This study was also supported by the Competitive Research Funding of the Tampere University Hospital (Grant 9K092).

REFERENCES

[1]Kanis JA, Borgstrom F, De Laet C, Johansson H, Johnell O, Jonsson B, Oden A, Zethraeus N, Pfleger B and Khaltaev N. Assessment of fracture risk. Osteopor. Int. 2005; **16**:581-589.

[2]Stone KL, Seeley DG, Lui LY, Cauley JA, Ensrud K, Browner WS, Nevitt MC and Cummings SR. BMD at multiple sites and risk of fracture of multiple types: Long-term results from the study of osteoporotic fractures. J. Bone Miner. Res. 2003; **18**:1947-1954.

[3]Bolotin HH and Sievanen H. Inaccuracies inherent in dual-energy X-ray absorptiometry in in vivo bone mineral density can seriously mislead diagnostic/prognostic interpretations of patient-specific fragility. J. Bone Miner. Res. 2001; **16**:799-805.

[4]Sievanen H. A physical model for dual-energy X-ray absorptiometry-derived bone mineral density. Invest. Radiol. 2000; **35**:325-330.

[5]Bolotin HH. The significant effects of bone structure on inherent patient-specific DXA in vivo bone mineral density measurement inaccuracies. Med. Phys. 2004; **31**:774-788.

[6]Järvinen TL, Sievänen H, Khan K, Heinonen A and Kannus P. Shifting the focus in fracture prevention from osteoporosis to falls. BMJ 2008; **336**:124-126.

[7]Jarvinen T, Sievanen H, Jokihaara J and Einhorn TA. Revival of bone strength: the bottom line. J. Bone Miner. Res. 2005; **20**:717-737.

[8]Riggs BL, Melton III LJ, Robb RA, Camp JJ, Atkinson EJ, Peterson JM, Rouleau PA, McCollough CH, Bouxsein ML and Khosla S. Population-based study of age and sex differences in bone volumetric density, size, geometry, and structure at different skeletal sites. J. Bone Miner. Res. 2004; **19**:1945-1954.

[9]Riggs BL, Melton III LJ, Robb RA, Camp JJ, Atkinson EJ, Oberg AL, Rouleau PA, McCollough CH, Khosla S and Bouxsein ML. Population-based analysis of the relationship of whole bone strength indices and fall-related loads to age- and sex-specific patterns of hip and wrist fractures. J. Bone Miner. Res. 2006; **21**:315-323.

[10] Carpenter DR, Beaupre GS, Lang TF, Orwoll ES and Carter DR. New QCT analysis approach shows the importance of fall orientation on femoral neck strength. J. Bone Miner. Res. 2005; **20**:1533-1542.

[11]Yang L, Masic I, McCloskey EV and Eastell R. Shape, structural properties, and cortical stability along the femoral neck: A study using clinical QCT. J. Clin. Densitom. 2008; **11**:373-382.

[12] Black DM, Bouxsein ML, Marshall LM, Cummings SR, Lang TF, Cauley JA, Ensrud KE, Nielson CM and Orwoll ES. Proximal Femoral Structure and the Prediction of Hip Fracture in Men: A Large Prospective Study Using QCT. J. Bone Miner. Res. 2008; **23**:1326-1333.

[13]Hulmea PA, Boydb SK and Ferguson SJ. Regional variation in vertebral bone morphology and its contribution to vertebral fracture strength. Bone. 2007; **41**:946-957.

[14]Verhulp E, van Rietbergen B, Huiskes R. Load distribution in the healthy and osteoporotic human proximal femur during a fall to the side. Bone. 2008; **42**:30-35.

[15]Njeh CF, Fuerst T, Hans D, Blake GM and Genant HK. Radiation exposure in bone mineral density assessment. Appl. Radiation Isot. 1999; **50**:215-236.

[16]Guglielmi G, de Terlizzi F. Quantitative ultrasond in the assessment of osteoporosis. Eur J Radiol. 2009; **71**:425-431.

[17]Sievänen H, Koskue V, Rauhio A, Kannus P, Heinonen A and Vuori I. Peripheral quantitative computed tomography in human long bones: Evaluation of in vitro and in vivo precision. J. Bone Miner. Res. 1998; **13**:871-882.

[18]Lochmüller EM, Müller R, Kuhn V, Lill CA and Eckstein F. Can Novel Clinical Densitometric Techniques Replace or Improve DXA in Predicting Bone Strength in Osteoporosis at the Hip and Other Skeletal Sites? J. Bone Miner. Res. 2003; **18**:906-912.

[19] Lochmüller EM, Bürklein D, Kuhn V, Glaser C, Müller R, Glüer CC, Eckstein F. Mechanical strength of the thoracolumbar spine in the elderly: prediction from in situ dual-energy X-ray absorptiometry, quantitative computed tomography (QCT), upper and lower limb peripheral QCT, and quantitative ultrasound. Bone. 2002; **31**:77-84.

[20]Zhang Y, Brady M and Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imag. 2001; **20**:45-57.

[21]Li J, Najmi A and Gray RM. Image classification by a two dimensional hidden Markov model," IEEE Trans. Signal Proc. 2000; **48**:517-533.

[22]Lai J, Ford JJ, O'Shea P and Walker R. Hidden Markov Model Filter Banks for Dim Target Detection from Image Sequences. presented at the DICTA 2008, Canberra, Australia, December 1-3, 2008, URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4700037&isnumber=4699978.

[23]Nikander R, Sievänen H, Uusi-Rasi K, Heinonen A, Kannus P. Loading modalities and bone structures at nonweight-bearing upper extremity and weight-bearing lower extremity: a pQCT study of adult female athletes. Bone. 2006; **39**:886-894.

[24]Uusi-Rasi K, Sievänen H, Pasanen M, Oja P, Vuori I. Associations of calcium intake and physical activity with bone density and size in premenopausal and postmenopausal women: a peripheral quantitative computed tomography study. J. Bone Miner. Res. 2002; **17**:544-552.

[25]Aldroubi A and Unser M. Wavelet transform: Theory and implementation. In *Wavelets in Medicine and Biology*, edited by Aldroubi A, Unser M (CRC Press LLC, Boca Raton, USA, 1996), pp. 3-73.

[26]Malfait M, Roose D. Wavelet based image denoising II: Wavelet based image denoising using a Markov Random Field a priori model, Technical Report TW 228, Katholieke Universiteit, Leuven, Belgium, 1995

[27]Xie H, Pierce LE and Ulaby FT. SAR speckle reduction using wavelet denoising and Markov random field modeling. IEEE Trans. Geosci. Remote Sens. 2002; **40**:2196-2212.

[28]Seits P and Ruegsegger P. Fast contour detection algorithm for high precision quantitative CT. IEEE Trans. Med. Imag. 1983; **2**:136-141.

[29]Fix E and Hodges JL. Discriminatory analysis, non-parametric discrimination. Technical report, USAF School of Aviation Medicine, Randolf Field, Tex. Project 21-49-004, Rept. 4, Contract AF41(128)-31 (1951).

[30]Kadah YM, Farag AA, Zurada JM, Badawi AM and Youssef A-BM. Classification Algorithms for Quantitative Tissue Characterization of Diffuse Liver Disease from Ultrasound Images Medical Imaging. IEEE Trans. Med. Imag. 1996; **15**:466-478.

[31] Fleiss JL. Reliability of measurement. In: *The Design and Analysis of Clinical Experiments*, edited by Fleiss JL (John Wiley & Sons, New York, USA, 1986), pp. 1-32.

[32]Engelke K, Adams JE, Armbrecht G, Augat P, Bogado CE, Bouxsein ML, Felsenberg D, Ito M, Prevrhal S, Hans DB and Lewiecki EM. Clinical Use of Quantitative Computed Tomography and Peripheral Quantitative Computed Tomography in the Management of Osteoporosis in Adults: The 2007 ISCD Official Positions. J. Clin. Densitom. 2008; **11**:123-162.

[33]Zemel B, Bass S, Binkley T, Ducher G, Macdonald H, McKay H, Moyer-Mileur L, Shepherd J, Specker B, Ward K and Hans D. Peripheral Quantitative Computed Tomography in Children and Adolescents: The 2007 ISCD Pediatric Official Positions. J. Clin. Densitom. 2008; **11**:59-74.

[34]Veitch SW, Findlay SC, Ingle BM, Ibbotson CJ, Barrington A, Hamer AJ and Eastel R. Accuracy and Precision of Peripheral Quantitative Computed Tomography Measurements at the Tibial Metaphysis. J. Clin. Densitom. 2004; **7**:209-217.

[35]Ward KA, Adams JE and Hangartner TN. Recommendations for Thresholds for Cortical Bone Geometry and Density Measurement by Peripheral Quantitative Computed Tomography. Calcif. Tissue Int. 2005; **77**:275-280.

[36]Kontulainen S, Liu D, Manske S, Jamieson M, Sievänen H and McKay H. Analyzing Cortical Bone Cross-Sectional Geometry by Peripheral QCT: Comparison With Bone Histomorphometry. J. Clin. Densitom. 2007; **10**:86-92.

[37]Ashe MC, Liu-Ambrose T, Khan KM, White N and McKay HA. Optimizing results from pQCT: reliability of operator-dependent pQCT variables in cadavers and humans with low bone mass. J. Clin. Densitom. 2005; **8**:335-340.

[38]Burrows M, Cooper DM, Liu D and McKay HA. Bone and Muscle Parameters of the Tibia: Agreement Between the XCT 2000 and XCT 3000 Instruments. J. Clin. Densitom. 2009; **12**:186-194.

[39]Hangartner TN and Gilsanz V. Evaluation of cortical bone by computed tomography. J. Bone Miner. Res. 1996; **11**:1518-1525.

[40]Augat P, Reeb H and Claes LE. Prediction of fracture load at different skeletal sites by geometric properties of the cortical shell. J. Bone Miner. Res. 1996; **11**:1356-1363.

[41]Bell KL, Loveridge N, Power J, Garrahan N, Stanton M, Lunt M, Meggitt BF and Reeve J. Structure of the Femoral Neck in Hip Fracture: Cortical Bone Loss in the Inferoanterior to Superoposterior Axis. J. Bone Miner. Res. 1999; **14**:111-119.

[42]Crabtree N, Loveridge N, Parker M, Rushton N, Power J, Bell KL, Beck TJ and Reeve J. Intracapsular Hip Fracture and the Region-Specific Loss of Cortical Bone: Analysis by Peripheral Quantitative Computed Tomography. J. Bone Miner. Res. 2001; **16**:1318-1328.

[43]Pistoia W, van Rietbergen B and Ruegsegger P. Mechanical consequences of different scenarios for simulated bone atrophy and recovery in the distal radius. Bone 2003; **33**:937-945(9).

[44]Mayhew PM, Thomas CD, Clement JG, Loveridge N, Beck TJ, Bonfield W, Burgoyne CJ and Reeve J. Relation between age, femoral neck cortical stability, and hip fracture risk. The Lancet 2005; **336**:129-135.

[45]Nikander R, Kannus P, Dastidar P, Hannula M, Harrison L, Cervinka T, Narra NG, Aktour R, Arola T, Eskola H, Soimakallio S, Heinonen A, Hyttinen J and Sievänen H. Targeted exercises against hip fragility. Osteoporos. Int. 2009; in press.

[46]Groll O, Lochmüller EM, Bachmeier M, Willnecker J and Eckstein F. Precision and intersite correlation of bone densitometry at the radius, tibia and femur with peripheral quantitative CT. Skeletal Radiol. 1999; **28**:696-702.

[47]Lai YM, Qin L, Hung VWY and Chan KM. Regional differences in cortical bone mineral density in the weight-bearing long bone shaft—a pQCT study. Bone 2005; **36**:465-471.

[48]Lai YM, Qin L, Yeung HY, Lee KKH and Chan KM. Regional differences in trabecular BMD and micro-architecture of weight-bearing bone under habitual gait loading – a pQCT and microCT study in human cadavers. Bone 2005; **37**:274-282.

[49]Lai YM, Qin L, Hung VWY, Choy WY, Chan ST, Chan LWC and Chan KM. Trabecular Bone Status in Ultradistal Tibia Under Habitual Gait Loading: A pQCT Study in Postmenopausal Women. J. Clin. Densitom. 2006; **9**:175-183.

LIST OF FIGURES

Fig. 5. Repeated pQCT scans after being processed by different methods: a) the raw, unprocessed image of the first scan (left-most panel), its result after k-NN analysis, the original image of the repeated scan and its results after k-NN analysis (right-most panel); b) preprocessed image by linear window; c) preprocessed image by Gaussian window; d) preprocessed image after 3×3 median filtering; and e) preprocessed image after 5×5 median filtering. ROI – region of interest (densities investigated from phantom images and cortical bone).

Table I. Descriptive data and precision of the proportional cortical area (in %) at the total bone cross-section and four anatomic regions of the distal tibia after different pre-processing methods.

Pre-processing method		Cortical area	Posterior region	Medial region	Anterior region	Lateral region
Raw image	Mean (SD)	15.5 (1.6)	15.8 (1.7)	15.7 (1.7)	15.6 (2.0)	15.0 (1.7)
	SD_{meas}	0.9	0.7	1.3	1.5	0.9
	R	67	84	47	48	75
3x3 median window	Mean (SD)	15.6 (1.8)	15.8 (1.7)	16.0 (2.1)	15.7 (2.1)	15.2 (2.0)
	SD_{meas}	0.5	0.7	0.8	0.6	0.8
	R-value	94	85	85	91	83
5x5 median window	Mean (SD)	16.4 (1.6)	16.4 (1.8)	17.1 (2.0)	16.7 (2.0)	15.3 (1.6)
	SD_{meas}	0.4	0.7	0.5	0.6	0.7
	R-value	94	84	93	92	81
Linear window	Mean (SD)	17.6 (1.8)	18.0 (1.7)	17.7 (1.8)	17.9 (2.3)	17.0 (1.8)
	SD_{meas}	0.3	0.5	0.6	0.5	0.5
	R-value	97	93	88	95	91
Gauss window	Mean (SD)	16.5 (1.7)	16.9 (1.7)	16.6 (1.6)	16.5 (2.4)	16.0 (1.8)
	SD_{meas}	0.3	0.6	0.6	0.7	0.5
	R-value	97	87	85	92	91

SD_{meas} is the standard deviation of the differences between the repeated scans.

R is the reliability coefficient ($R = 100(1 - SD^2_{meas}/SD^2)\%$)

Table II. Descriptive data and precision of the trabecular density (in mg/cm^3)* in total and at the four anatomic sectors of the distal tibia after different preprocessing methods.

Preprocessing method		Total area	Posterior region	Medial region	Anterior region	Lateral region
Raw image	Mean (SD)	282 (42)	319 (42)	281 (45)	260 (49)	270 (52)
	SD_{meas}	2	5	4	5	5
	R (in %)	~100	99	99	99	99
3×3 median window	Mean (SD)	283 (42)	320 (42)	281 (45)	261 (49)	271 (52)
	SD_{meas}	3	5	5	5	5
	R-value	~100	99	99	99	99
5×5 median window	Mean (SD)	279 (42)	316 (41)	278 (44)	254 (49)	267 (52)
	SD_{meas}	3	5	5	4	5
	R-value	~100	99	99	99	99
Linear window	Mean (SD)	281 (41)	319 (41)	280 (46)	259 (48)	269 (51)
	SD_{meas}	2	5	4	3	5
	R-value	~100	99	99	~100	99
Gauss window	Mean (SD)	279 (41)	316 (41)	279 (46)	255 (48)	265 (51)
	SD_{meas}	2	5	4	4	5
	R-value	~100	99	99	99	99

SD_{meas} is the standard deviation of the differences between the repeated scans.

R is the reliability coefficient ($R=100(1 - SD^2_{meas}/SD^2)\%$).

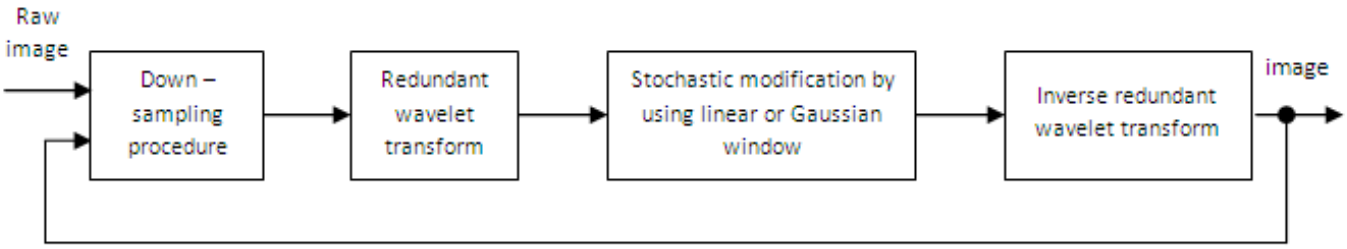
* obtained from the mean pixel values after applying the calibration equation between the measured mean pixel values and known densities of the K_2HPO_4 phantoms: density = 1.577mean_pixel_value + 357.13 for 3×3, 5×5 median, density = 6.452mean_pixel_value + 125.82 for the LinW and GaussW. Both relationships between the density values and mean pixel values were linear ($r \approx 1$).

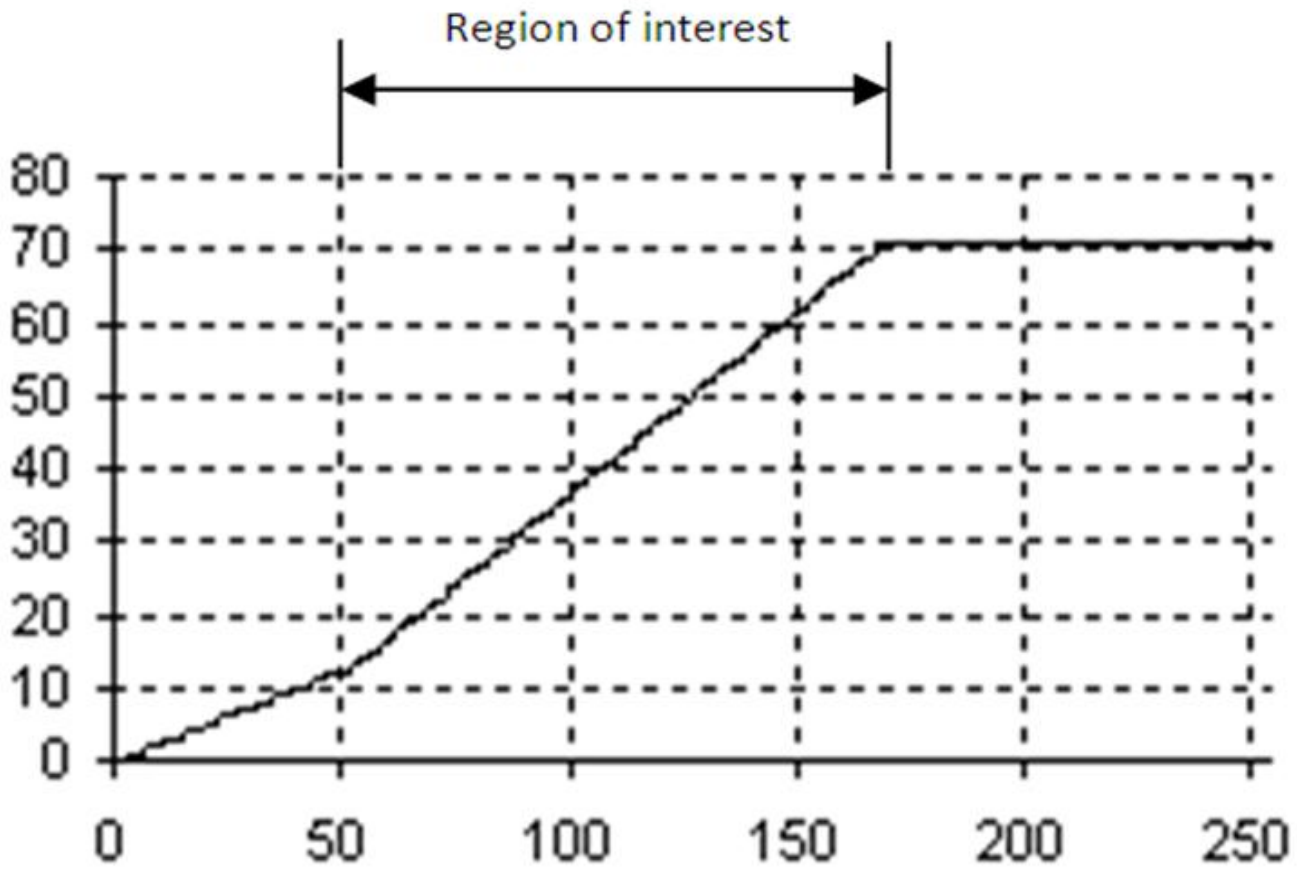
Table III. Mean pixel values and SDs of separate pixel values (in brackets) within the pQCT image and SNR at different density levels of the homogeneous phantoms after different preprocessing methods.

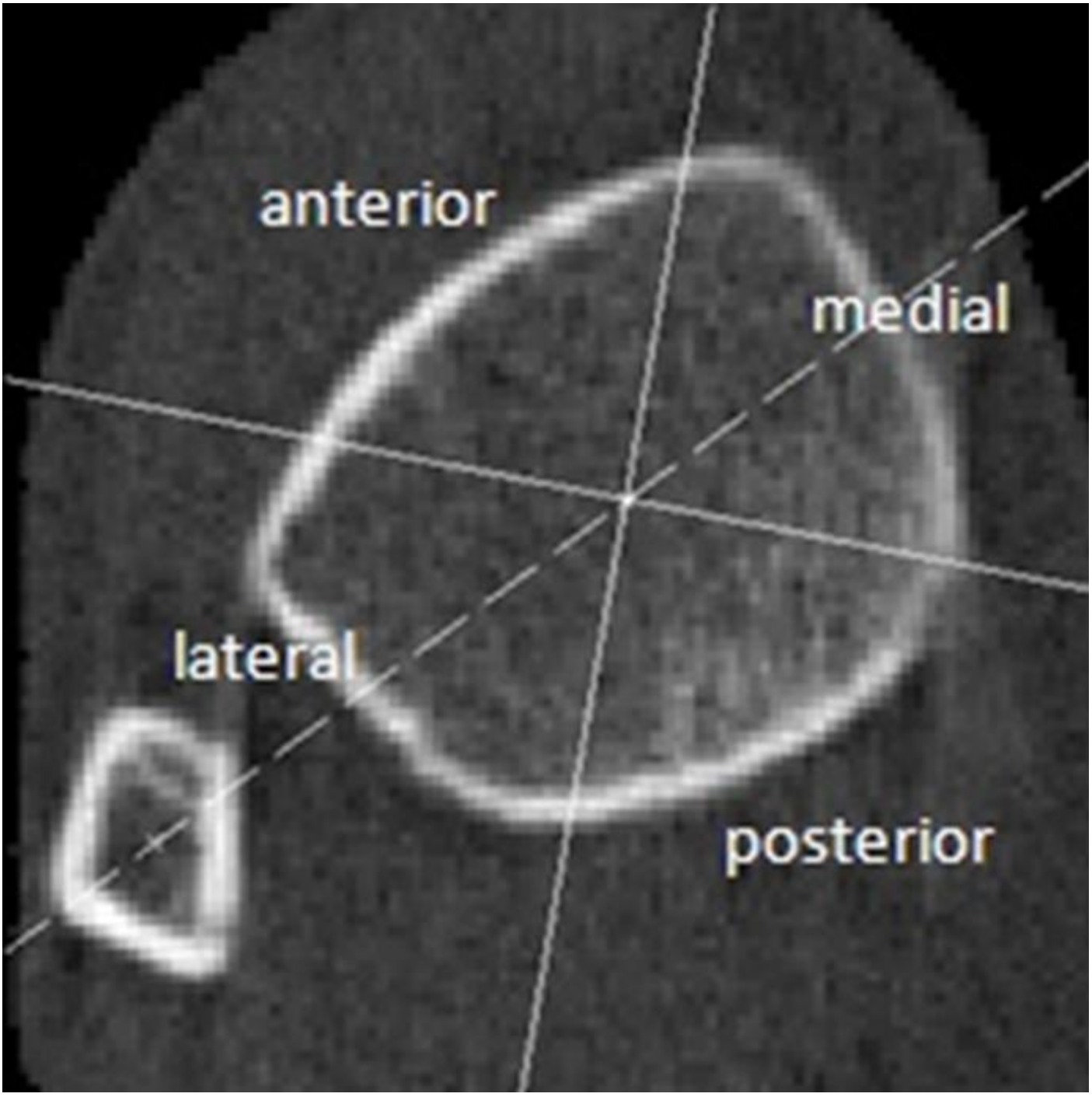
Phantom (concentration)		Background*	Tap water	98.4 (mg/cm^3)	146.5 (mg/cm^3)	291 (mg/cm^3)
Preprocessing method						
Raw image	Mean (SD)	12 (21)	259 (28)	292 (29)	323 (30)	417 (35)
	SNR†	-4.6	19	20	21	21
3×3 median window	Mean (SD)	12 (10.3)	259 (14.5)	291 (14.5)	322 (14.6)	416 (20.9)
	SNR†	1.2	25	26	27	26
5×5 median window	Mean (SD)	12 (7.2)	258 (11.2)	290 (11.2)	321 (11.1)	415 (18.0)
	SNR†	4.5	27	28	29	27
Linear window	Mean (SD)	1.4 (1.0)	36.9 (3.0)	45.0 (3.0)	52.9 (3.1)	76.2 (4.0)
	SNR†	9.4	22	23.5	25	26
Gauss window	Mean (SD)	1.4 (1.0)	37.0 (3.3)	45.0 (3.3)	52.9 (3.4)	76.4 (4.4)
	SNR†	9.4	21	22.5	24	25

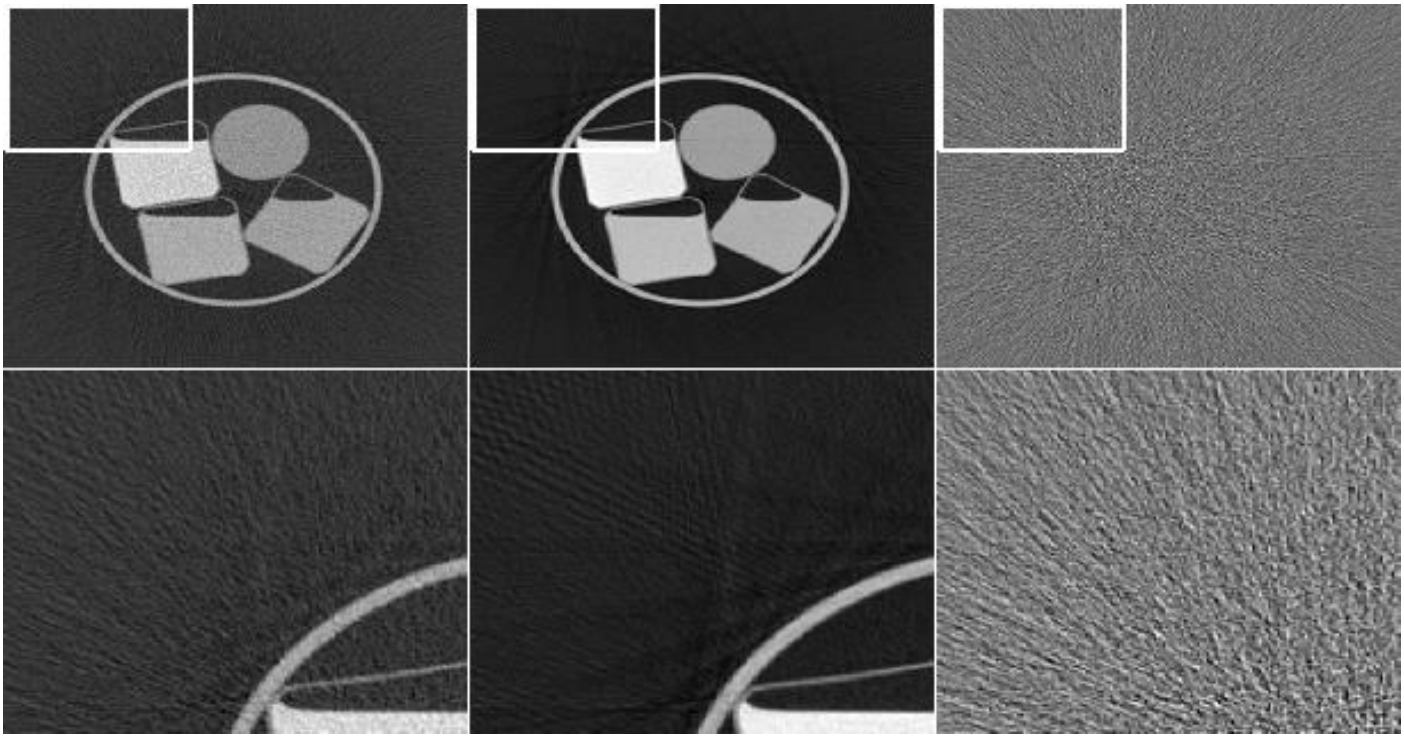
*Background denotes the air region of the pQCT image.

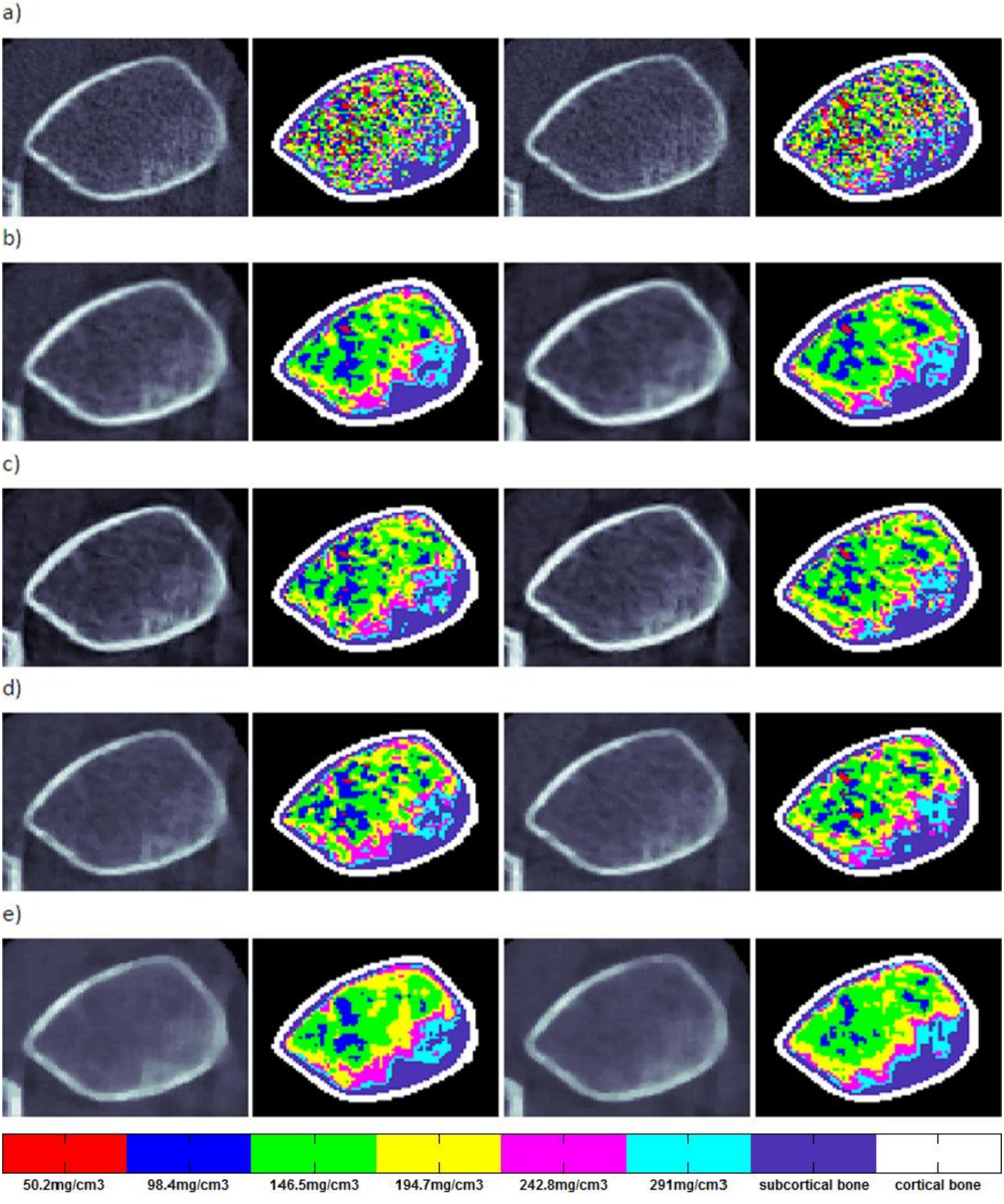
†the derivation is given in the Section II.D (Equation 3)


Raw
image
Down –
sampling
procedure
Redundant
wavelet
transform
Stochastic modification by
using linear or Gaussian
window
Inverse redundant
wavelet transform
image


Region of interest
80
70
60
50
40
30
20
10
0
0
50
100
150
200
250


anterior
medial
lateral
posterior




a)
b)
c)
d)
e)
50.2mg/cm3
98.4mg/cm3
146.5mg/cm3
194.7mg/cm3
242.8mg/cm3
291mg/cm3
subcortical bone
cortical bone

Publication II:

Cervinka T., Rittweger J., Hyttinen J., Felsenberg D., Sievanen H.
Anatomical sector analysis of load-bearing tibial bone structure during 90-day bed rest and 1-year recovery.
Clinical Physiology and Functional Imaging, 2011; 31(4): 249-257.

Clin Physiol Funct Imaging (2011) **31,** pp249–257 doi: 10.1111/j.1475-097X.2011.01009.x

Anatomical sector analysis of load-bearing tibial bone structure during 90-day bed rest and 1-year recovery

Tomas Cervinka[1], Jörn Rittweger[2,3], Jari Hyttinen[1], Dieter Felsenberg[4] and Harri Sievänen[5,1]

[1]Department of Biomedical Engineering, Tampere University of Technology, Tampere, Finland, [2]Institute for Biomedical Research into Human Movement and Health, Manchester Metropolian University, Manchester, UK, [3]Institute of Aerospace Medicine, German Aerospace Center, Linder Hoehe, Cologne, Germany, [4]Charité Universitätsmedizin Berlin, Center for Muscle and Bone Research, Berlin, Germany and [5]Bone Research Group, UKK Institute, Tampere, Finland

Summary

The aim of this study was to investigate whether the bone response to long bed rest–related immobility and during subsequent recovery differed at anatomically different sectors of tibial epiphysis and diaphysis. For this study, peripheral quantitative tomographic (pQCT) scans obtained from a previous 90-day 'Long Term Bed Rest' intervention were preprocessed with a new method based on statistical approach and re-analysed sector-wise. The pQCT was performed on 25 young healthy males twice before the bed rest, after the bed rest and after 1-year follow-up. All men underwent a strict bed rest intervention, and in addition, seven of them received pamidronate treatment and nine did flywheel exercises as countermeasures against disuse-related bone loss. Clearly, 3–9% sector-specific losses in trabecular density were observed at the tibial epiphysis on average. Similarly, cortical density decreased in a sector-specific way being the largest at the anterior sector of tibial diaphysis. During recovery, the bed rest–induced bone losses were practically restored and no consistent sector-specific modulation was observed in any subgroup. It is concluded that the sector-specific analysis of bone cross-sections has potential to reveal skeletal responses to various interventions that cannot be inferred from the average analysis of the whole bone cross-section. This approach is considered also useful for evaluating the bone responses from the biomechanical point of view.

Correspondence
Tomas Cervinka, Finn-medi 1, 4. Floor, Biokatu 6, 33520 Tampere, Finland
E-mail: tomas.cervinka@tut.fi

Accepted for publication
Received 29 October 2010; accepted 3 January 2011

Key words
bone loss; cortical bone; disuse; image preprocessing; trabecular bone

Introduction

Bone structure can deteriorate in response to several factors such as disuse, immobilization, ageing, diseases and hormonal disturbances. Because the lower limb skeleton is primarily locomotive organ and capable to adapt to varying loading conditions (Frost, 2003; Ruff *et al.*, 2006), disuse irrespective of its primary cause provides a useful model to investigate responses of bone structure to reduced loading. Experimental bed rest with −6° head down tilt is an established ground-based model to simulate the physiological effects of spaceflight (Pavy-Le Traon *et al.*, 2007). Earlier bed rest studies have mostly relied on dual-energy X-ray absorptiometry (DXA) (LeBlanc *et al.*, 1990; Zerwekh *et al.*, 1998; Shackelford *et al.*, 2004; Armbrecht *et al.*, 2010) but with regard to bone structure these studies are limited by the inherent inability of DXA to yield tangible information on actual cross-sectional bone geometry, let alone separating the measured bone into trabecular and cortical compartments (Sievänen, 2000). Peripheral quantitative computed tomography (pQCT) offers a reasonable option to assess bone geometry and density without evident limitations of planar DXA (Sievänen *et al.*, 1998; Sievänen, 2000). The few studies which have used pQCT or QCT technology indicated that bone loss is most prominent at endocortical bone regions both after bed rest (Rittweger *et al.*, 2005, 2009, 2010) and space flight (Lang *et al.*, 2006).

It is obvious that different bone sectors experience specific loading environment in terms of biomechanics of the given site. A good example of sector-specific bone adaptation comes from patients with spinal cord injury, in whom electrical stimulation of the soleus muscle led to site-specific bone accrual in the posterior aspect of the distal tibia (Dudley-Javoroski & Shields, 2008). Apparently, normal locomotive muscle contractions mostly affect the anterior and posterior sectors of distal tibia via the ground reaction forces imposed on the feet in the front and the pulling forces mediated by the Achilles tendon in the back. Therefore, we hypothesized that the most substantial bone losses would occur at those sectors where the lack of locomotive loading is most evident because of immobility during bed rest and accordingly a faster recovery at the same sectors when the

locomotive loading was resumed. To our knowledge, disuse-related changes in anatomical sectors of bone structure have not been studied before.

To elaborate our hypothesis, we re-analysed the pQCT data from the 'Long Term Bed Rest' (LTBR) study that was carried out in 2001 and 2002 in Toulouse/France. The LTBR results indicated that administered countermeasures, pamidronate and flywheel exercise could only partly prevent the mean loss in tibial bone mineral content during bed rest (Rittweger *et al.*, 2005). However, after the resumption of locomotion, the bone loss fully recovered at the tibial shaft and also to a large extent at the distal epiphysis (Rittweger & Felsenberg, 2009). In this study, we assessed specifically whether (i) the bone responses to bed rest were different at anatomically different sectors of the distal tibia and tibial shaft in different subgroups and (ii) whether the sector-specific responses were different during the recovery. The latter research question can give valuable information as to whether the recovery is achieved not only in quantitative terms but also with anatomical specificity in line with anticipated biomechanics. We also expected different responses between the groups treated with countermeasures and the control group. While the flywheel exercise could halve the bone loss in the LTBR study (Watanabe *et al.*, 2004; Rittweger *et al.*, 2005), it is unclear whether this benefit is evenly distributed within the bone cross-section. Likewise, it is not known whether bisphosphonate treatment could modulate the sector-specific maintenance of bone mass during bed rest, while the general benefits of such treatments on DXA-measured bone mineral density have been consistently observed (Chappard *et al.*, 1989; Grigoriev *et al.*, 1992; Ruml *et al.*, 1995; LeBlanc *et al.*, 2002; Watanabe *et al.*, 2004; Rittweger *et al.*, 2005).

Material and methods

The details of the LTBR study as well as the countermeasures and scanning procedures have been described in detail elsewhere (Rittweger *et al.*, 2005). In short, 25 healthy young male volunteers were recruited to undergo 90 days of strict bed rest with −6° head down tilt. These men were randomly assigned into three groups: nine to the control group (Ctrl, bed rest only), seven received a single infusion of 60 mg pamidronate i.v. prior to bed rest (Pam) and the other nine participants practiced resistive flywheel exercises every 2–3 days (FW). The age range of the subjects was 23–41 years; weight range 60–80 kg and height range 167–185 cm. The study was approved by the local Ethics Committee and all participants gave their written informed consent before they were included into the study.

Peripheral quantitative computed tomography (XCT 2000; Stratec Medizintechnik GmbH, Pforzheim, Germany) was performed at distal 4% (epiphysis) and at 66% (diaphysis) sites of the tibia (Rittweger *et al.*, 2005). For the present study, pQCT data measured 2 weeks before the bed rest (BDC-14), 1 week before the bed rest (BDC-7), 2 weeks after cessation of the bed rest (R + 14) and 1 year after the bed rest (R + 360) were used. The reason to select R + 14 rather than the penultimate day of bed rest (HDT89) was that the bone changes continued and were largest around R + 14 (Rittweger *et al.*, 2010). Three subjects (one from Pam group and two from FW group) were excluded because they did not attend the R360 follow-up measurement. Also, one subject from FW group was moved into Ctrl group for second half of the study because of knee pain that was partly related to the flywheel training (Rittweger *et al.*, 2007). Repositioning for the pQCT scans included in this study was considered acceptable because the differences in total cross-sectional bone areas were within 25 mm^2 (∼1%) in all consecutive four scans of distal tibia. The mean total area and mean intra-individual standard deviation (in parentheses) between BDC-14 and BDC-7 were 2122 (10) mm^2 for epiphysis and 662 (0·1) mm^2 for diaphysis.

Peripheral quantitative computed tomography image analysis

A new preprocessing method for pQCT data, proposed recently by Cervinka *et al.* (2010), was applied to the raw pQCT data that were produced by the back projection algorithm of the pQCT system without preprocessing. The new method was chosen because it permitted more consistent detection of outer and inner cortical boundaries compared to common image-preprocessing techniques, such as median filtering. In brief, preprocessing of the image data was composed in two steps. First, a gray-level transformation was carried out by implementing a common, iteratively employed, piecewise linear conversion function. This transformation represents a non-equidistant re-quantization of the gray scale. Second, the transformed data were corrected for inaccuracies caused by the re-quantization procedure by using the Bayesian approach and Markov random fields with 3 × 3 linear neighbourhoods and redundant wavelet transform. The method is described in detail elsewhere (Cervinka *et al.*, 2010).

After preprocessing of the pQCT images, a modification of simple and fast edge tracing algorithm (described in Seits & Ruegsegger, 1983) was used for segmenting the bone cross-section into cortical and trabecular bone regions (Cervinka *et al.*, 2010). After detecting the outer cortical boundary according to the maximum values of the first derivative of the preprocessed image data, the region inside the traced bone contour was radially shrunk until the inner cortical boundary was found. The next local maximal points along the radii were considered to coincide with the inner boundary of the cortical shell (Cervinka *et al.*, 2010). The shrunken region corresponded to the trabecular bone area and subtraction of the original and shrunk bone regions provided the cortical bone area. The procedure for cortical contour detection is illustrated in Fig. 1.

Then, the conventional pQCT traits, cortical cross-sectional area (CoA) and trabecular density (TrD) for epiphysis, and CoA and cortical density (CoD) for diaphysis, were determined as descriptors of bone structure. Besides analysing the total bone cross-section, the cross-section was divided into four anatomical

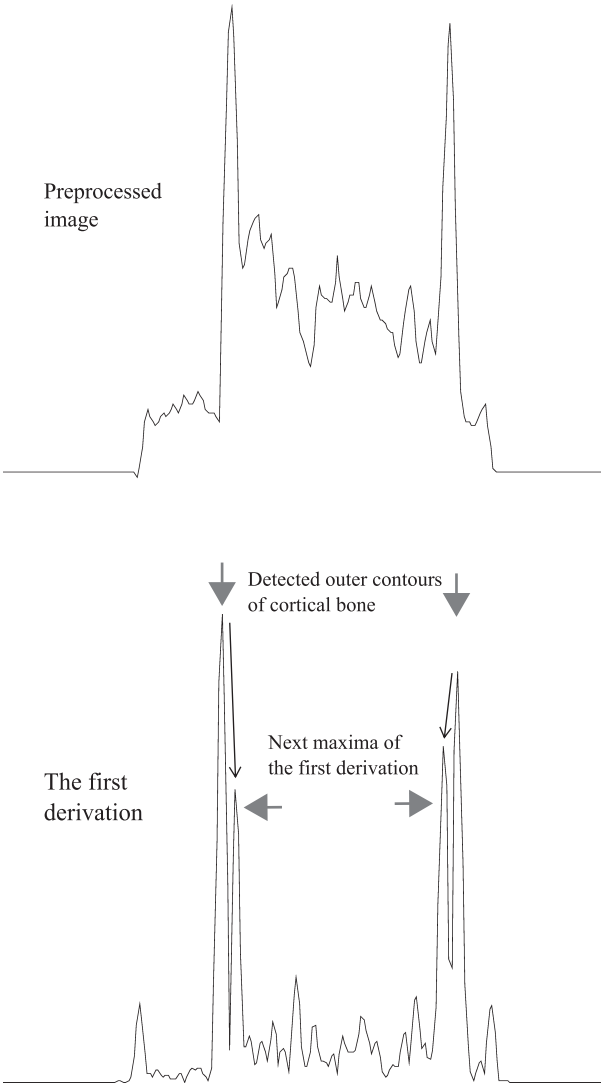


Figure 1 Cortical bone contour detection. The upper curve shows one radial X-ray attenuation profile along a radial line through the preprocessed bone image and the lower curve represents its first derivate. The greater attenuation of cortical bone is distinct. The maximum of the first derivative denotes the outer cortical boundary and the next local maximum denotes the inner cortical boundary. This algorithm is repeated across the whole bone cross-section and the area remaining between the maxima and next local maxima of the first derivative represented the cortical area.

sectors (lateral, posterior, medial and anterior) so that direction-specific features of the cortical and trabecular bone could be assessed (Fig. 2). The CoA was expressed as a percentage of total bone cross-sectional area or as a percentage of total bone area within the given anatomical sector, as appropriate. The TrD and CoD values represented the volumetric bone mineral apparent density of the given bone region. The conversion from gray scale values to bone mineral density was based on previously determined relationship between the pQCT-measured densities, known K_2HPO_4 concentrations and gray scale values (Sievänen *et al.*, 1998; Cervinka *et al.*, 2010).

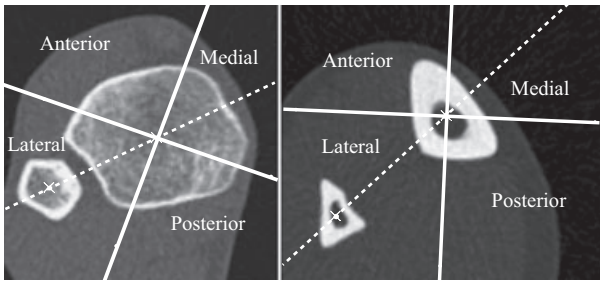


Figure 2 Definition of the anatomical sectors of interest for the distal tibia (4%) and tibia diaphysis (66%). First, a line was drawn through the centres-of-mass of the tibia and fibula. Then, two lines were drawn through the centre-of-mass of the tibia at angles of 45° and −45°. The sectors between these lines denote the four anatomical (posterior, medial, anterior and lateral) regions of interest.

Statistical analysis

Group mean values and 95% confidence intervals (95% CI) of bone traits are given as descriptive statistics. Changes in bone traits after bed rest (R + 14) and after recovery periods (R + 360) were expressed as percentage changes from baseline value, which was defined as the mean of BDC-14 and BDC-7 data measured before bed rest. These two pQCT assessments prior to bed rest provided also relevant information on short-term precision of bone traits and facilitated appropriate interpretation of changes in bone traits. Given the small number of subjects in study groups, the 95% CIs of percentage changes were considered adequate for indicating within-group changes during bed rest and subsequent recovery period or substantial between-group differences at different time points. In addition to mean changes over time, potential group differences in the variance of responses to bed rest and accompanying counter-measures at R + 14 and R + 360 were evaluated with the F-test.

Results

Epiphysis

Changes in CoA and TrD at the distal tibia during bed rest and 1-year recovery are shown in Fig. 3. Mean changes in CoA generally remained within ∼±2% in all groups without apparent trends or indication of statistical significance. However, the Pam group showed lower variance in the anterior and posterior changes in CoA during the bed rest compared to FW and Ctrl groups, respectively, ($P<0{\cdot}05$ for both) but higher variance in medial and anterior changes in CoA during recovery compared to Ctrl group ($P<0{\cdot}05$ for both sectors).

Mean trabecular loss during bed rest was systematically greater (∼3–8%) than the cortical loss, and it was most pronounced in Ctrl group and smallest in FW group. Besides the general decline within the whole bone cross-section, some anatomical sectors appeared to be more affected. During bed rest, the greatest (∼6–8%) mean losses in trabecular density occurred at the medial sector in Pam and FW groups, whereas in

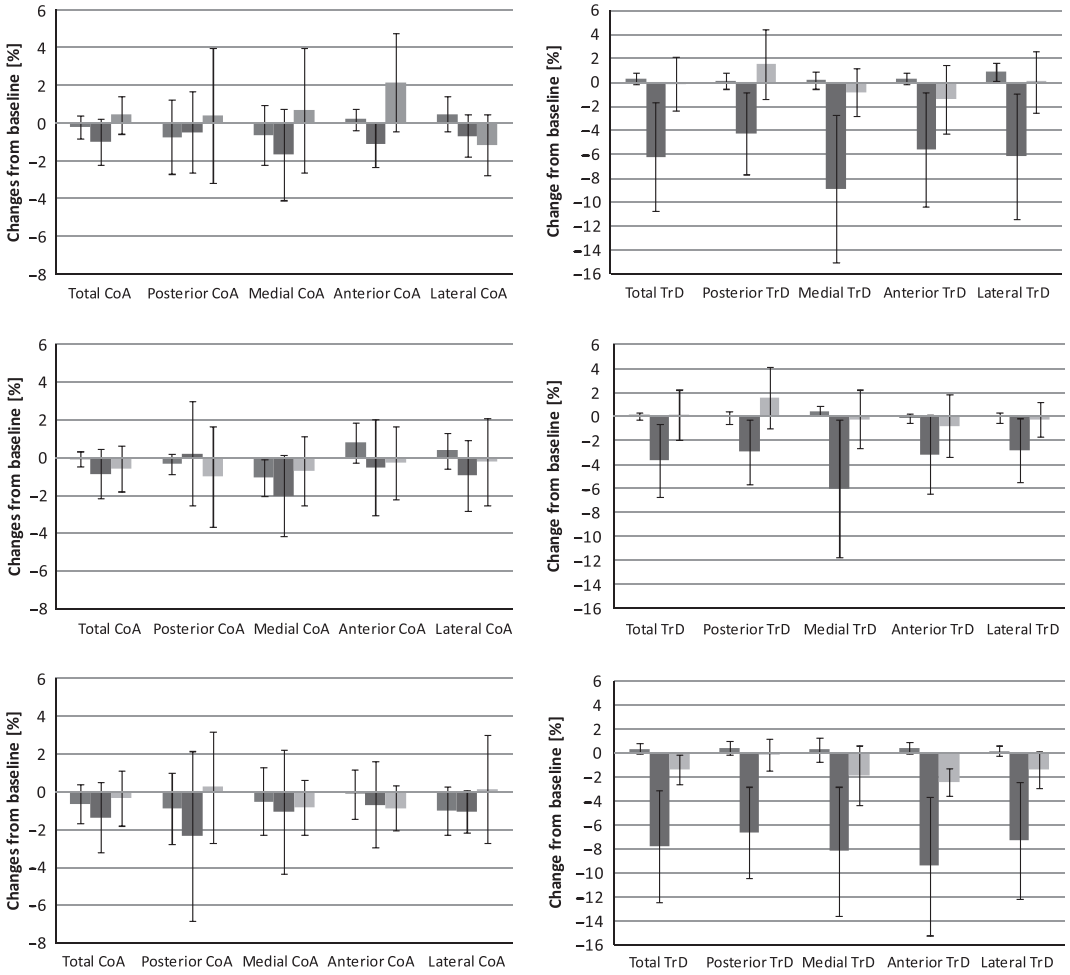


Figure 3 Distal tibia (4%): mean percentage changes (95% CI) of CoA (left panel) and TrD (right panel) in Pam (upper row), FW (middle row) and Ctrl (lower row) groups at baseline (mean of BDC-14 and BDC-7; left bars), from baseline to 2 weeks after 90 days bed rest (R + 14; middle bars) and from baseline to the end of 1-year recovery period (R + 360; right bars) in total and at four anatomical sectors of distal tibia.

Ctrl group, the greatest mean loss (∼10%) took place at the anterior sector. Compared to FW group, Ctrl group showed higher variance in density changes at anterior and lateral sectors during bed rest. During recovery, virtually all mean density levels returned back to baseline, the quite consistent ∼2% anterior loss in Ctrl group excluded. The variance at the anterior density in Ctrl group was lower compared to Pam group ($P<0{\cdot}05$). Individual changes in TrD during the bed rest and recovery periods are shown in Fig. 4.

Diaphysis

Mean changes in CoA and CoD at the tibial diaphysis during bed rest and 1-year recovery are shown in Fig. 5. In general, mean changes in CoA during the whole study period were marginal (mainly within ±1%) without apparent trends or indication of statistical significance. However, the variance in individual responses appeared to be great and some between-group differences were indicated. Compared to Ctrl group, changes in total CoA were less variant in Pam and FW groups during bed rest ($P<0{\cdot}05$ for both). In FW group, in turn, the variance in bed rest–related changes in medial CoA was higher compared to those in Ctrl and Pam groups ($P<0{\cdot}05$ for both). During recovery, changes in lateral CoA were more variant in Pam group than in Ctrl group.

The mean losses in CoD remained marginal (mostly clearly within ±1%) in Pam and FW groups but were somewhat larger at posterior (∼1·5%) and anterior (∼2%) cortical regions in Ctrl group. During bed rest, changes in medial and anterior CoD were more variant in FW group than in Pam or Ctrl groups and in Pam group, respectively ($P<0{\cdot}05$). In addition, changes in total, posterior and anterior CoD were more variant in Ctrl group compared to Pam group ($P<0{\cdot}05$). During recovery, there was virtually no systematic sector-wise effect on R + 360 and all mean density levels returned back to baseline. The variances were also similar in all groups. Individual changes in CoD during the bed rest and recovery are shown in Fig. 6.

Discussion

The main purpose of this study was to investigate whether the bone responses to bed rest and subsequent recovery differ

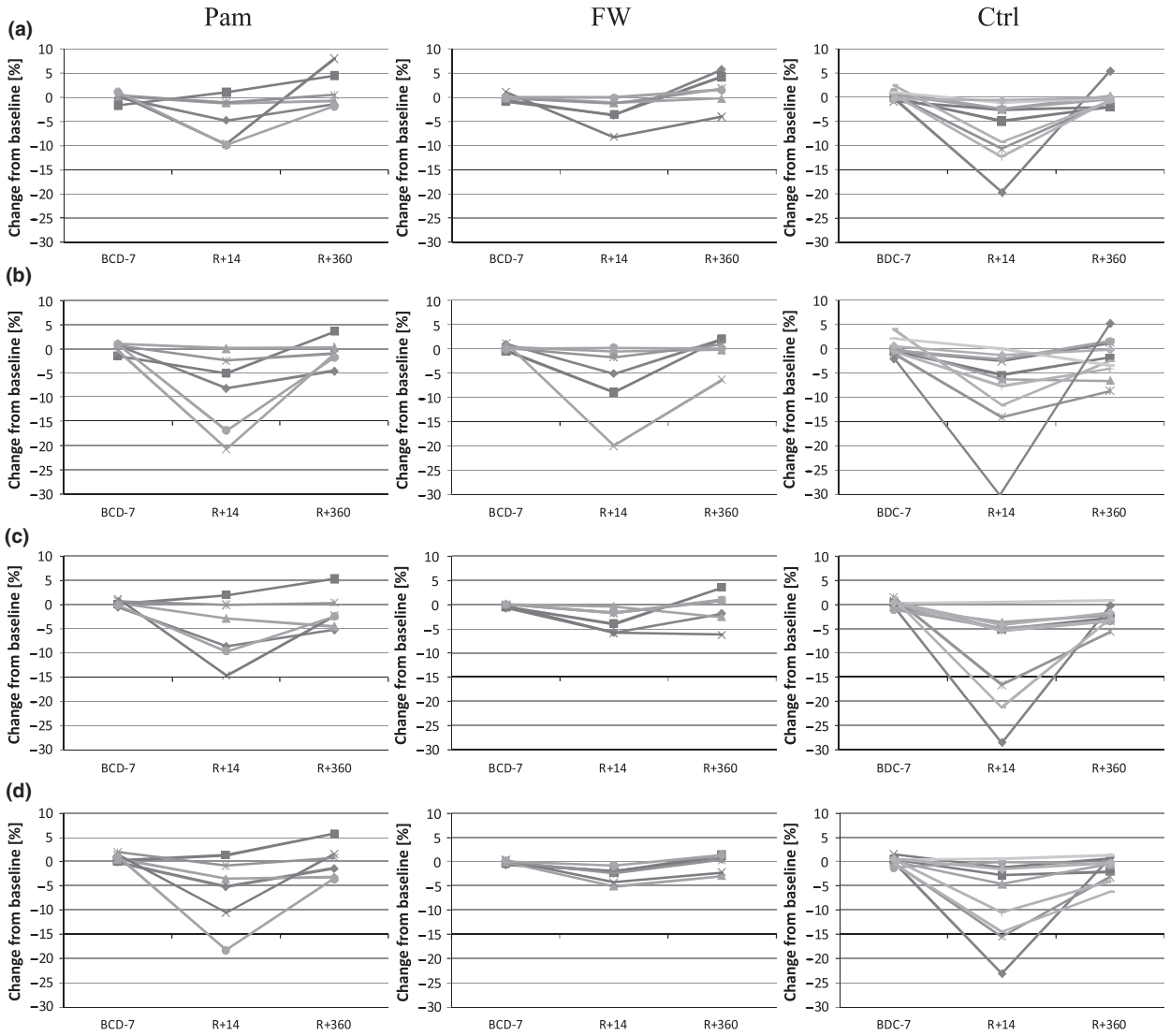


Figure 4 Individual percentage changes in TrD at distal tibia (4%) in Pam (left column), FW (middle column) and Ctrl (right column) at four anatomical sectors; posterior (a), medial (b), anterior (c) and lateral (d) at base line (mean of BDC-14 and BDC-7), R + 14 and R + 360.

between anatomically different sectors of the tibial epiphysis and diaphysis and whether they differ between different subgroups. Sector-specific modulation was evident and the bone losses were most pronounced in the trabecular compartment (reaching ~9% on average) of the distal tibia. This is in line with the results from long-term immobilization caused by spinal cord injury (Sievänen, 2010) but in apparent contrast to recent findings from bed rest studies (Rittweger & Felsenberg, 2009; Rittweger *et al.*, 2009, 2010; Armbrecht *et al.*, 2010). The latter have suggested that the most prominent bone loss occurred at the endocortical surface of cortical bone. During 1-year recovery, no consistent sector-specific modulation was observed in any subgroup and the bed rest–induced bone losses were practically restored, albeit with a small deficit was seen in trabecular density at the anterior sector of the control group. The general recovery of bone mass is perfectly in line with recent results of Rittweger & Felsenberg (2009).

Obviously, the structure and geometry of lower limb bones have developed under the regular influence of mechanical loading comprising largely the ground impacts and accompanying muscle forces during bipedal locomotion in different physical activities. Therefore, bed rest induces a drastic change to the loading environment of lower limb skeleton. Reportedly, significant (2–6%) mean losses occur in bone in response to bed rest and the magnitude of mean loss is related to the total duration of intervention (Rittweger & Felsenberg, 2009; Rittweger *et al.*, 2009, 2010; Armbrecht *et al.*, 2010; Sievänen, 2010). In the present study, the emphasis was uniquely placed on the sector-specific evaluation of cortical and trabecular bone responses to bed rest and consequent recovery in three different groups, one without countermeasure and two with countermeasures against bed rest–related bone loss.

Sector-specific analysis of bone structure has been previously applied in assessing the association of physical activity or

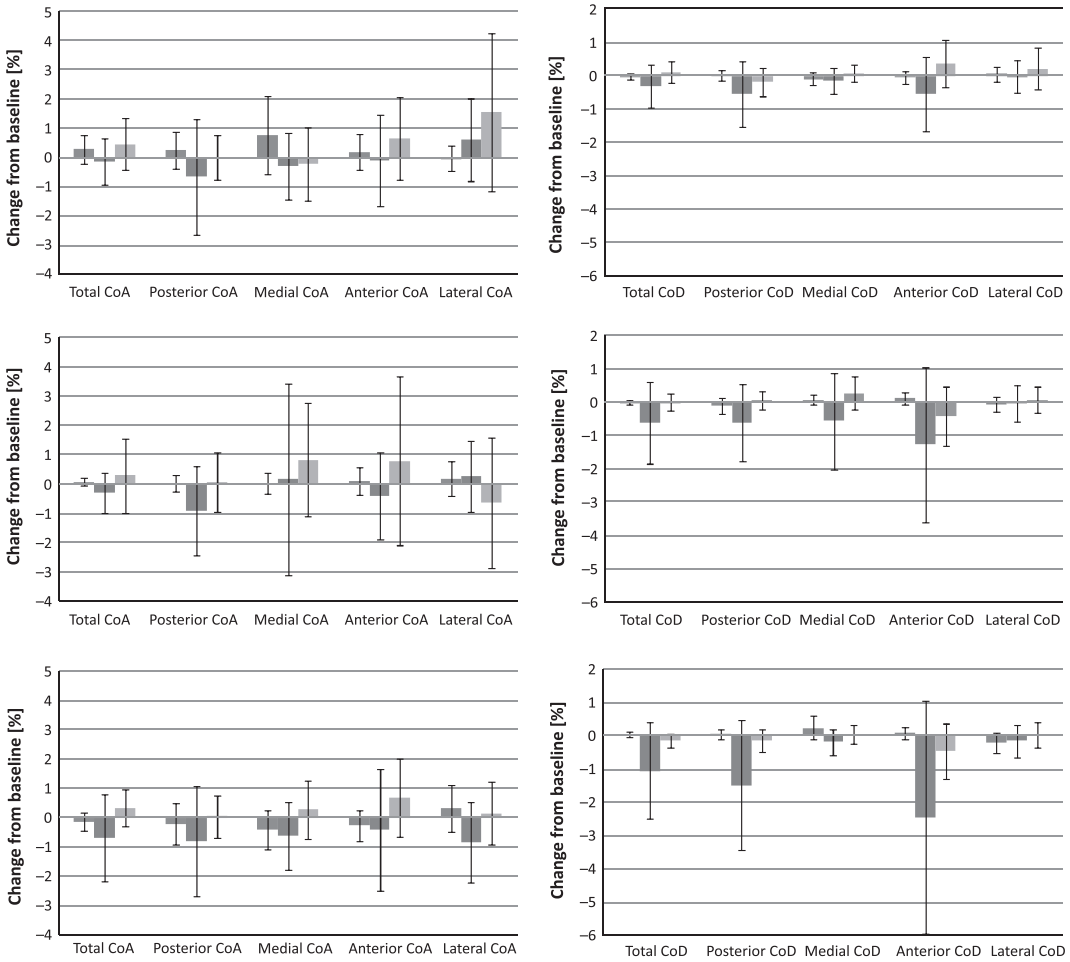


Figure 5 Tibia diaphysis (66%): mean percentage changes (95% CI) of CoA (left panel) and CoD (right panel) in Pam (upper row), FW (middle row) and Ctrl (lower row) groups at baseline (mean of BDC-14 and BDC-7; left bars), from baseline to 2 weeks after 90 days bed rest (R + 14; middle bars) and from baseline to the end of 1 year recovery period (R + 360; right bars) in total and at four anatomical sectors of the tibia shaft.

exercise loading with tibial geometry (Ma *et al.*, 2009; Macdonald *et al.*, 2009; Shaw & Stock, 2009; Rantalainen *et al.*, 2010). With regard to the present bed rest data, this approach allowed a good opportunity to test, at least tentatively, three following hypotheses all pertinent to bone physiology. First, a general bone loss should occur as a consequence of the lack of locomotive loading; second, the flywheel exercise should reduce this bone loss especially at sectors that are loaded most by pedalling (i.e. mainly along the antero-posterior axis because of muscle forces delivered via Achilles tendon); third, the pamidronate treatment should show a general preventive effect because of its apparently systemic influence on bone metabolism. Quite surprisingly, all three hypotheses were virtually confirmed for the cortical density only (Fig. 5), whereas the responses in trabecular density were not so distinct between groups (Fig. 3), let alone the marginal responses in cortical area (Figs 3 and 5).

As expected, the variance in individual skeletal responses was impressively great during the bed rest phase. In the light of such substantial inter-individual variability, the inter-subject variation during recovery has to be described as astoundingly small. It therefore seems that bone adaptation is more accurately governed when bone is accrued than when it is lost (Figs 4 and 6). Increased bone resorption, in particular, has been shown in a recent bed rest study on bone biomarkers (Armbrecht *et al.*, 2010).

Long bones, such as tibia, allow efficient movement of the body and provide stiffness against muscle contractions. The apparent goal of skeletal adaptation is to keep the loading-induced deformations (strains) within a specific physiological range (Frost, 2003), and when the regular loading is essentially reduced, the bone loses some of its rigidity. The contribution of cortical bone (geometry in conjunction with material elasticity) to the whole bone rigidity and strength is evident (Currey, 2001). Because the elasticity of cortical bone as a material is proportional to the third power of cortical density (Martin, 1991), the reduction in cortical density may be the first adaptive process to increase the flexibility of the given bone to meet the new loading environment – provided that the bone normal metabolism is not affected by antiresorptive medication. Indeed,

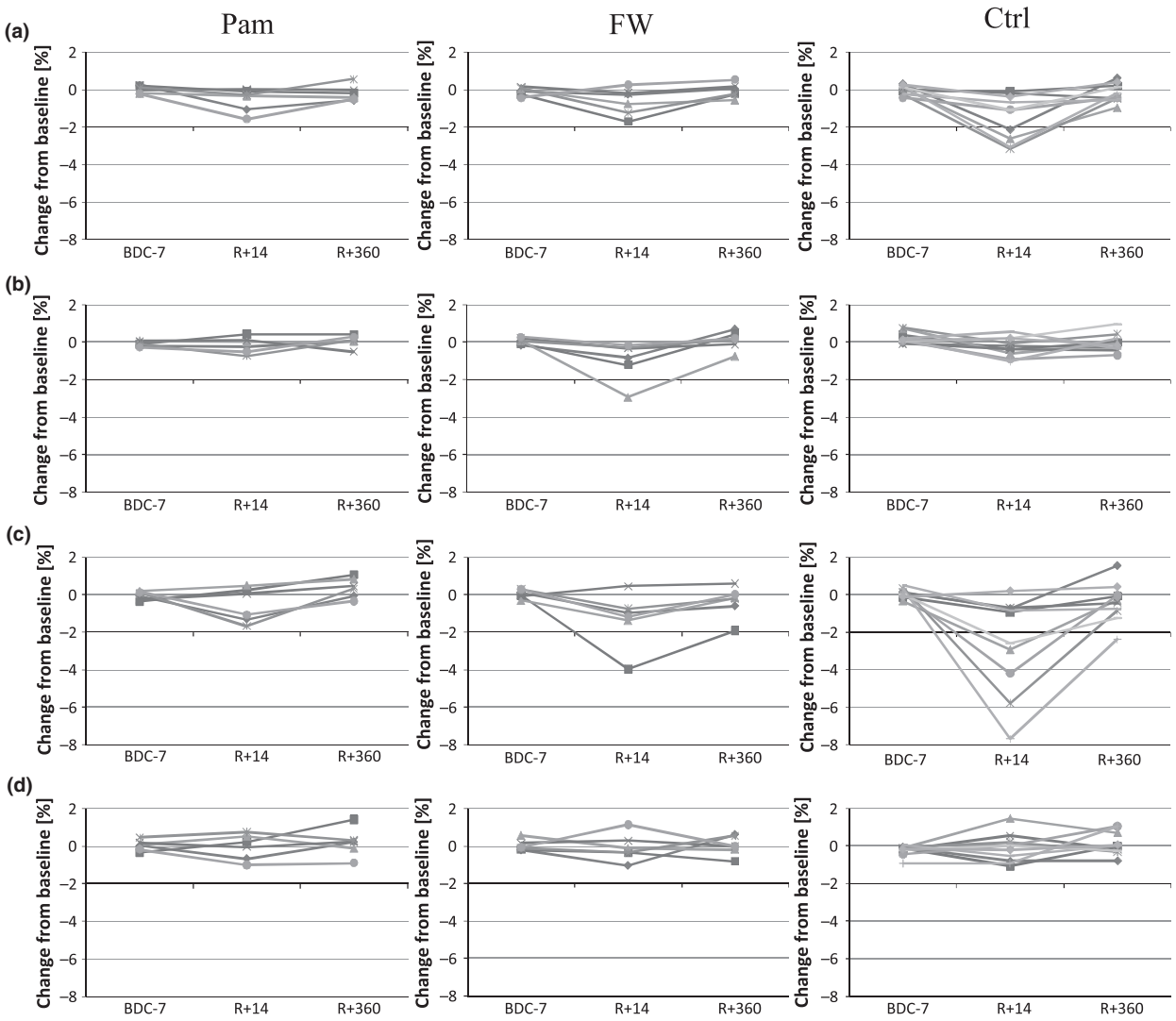


Figure 6 Individual percentage changes in CoD at tibia diaphysis (66%) in Pam (left column), FW (middle column) and Ctrl (right column) at four anatomical sectors: posterior (a), medial (b), anterior (c) and lateral (d) at base line (mean of BDC-14 and BDC-7), R + 14 and R + 360.

cortical density is shown to decline by 7–8% during the first year after spinal cord injury (Frey-Rindova *et al.*, 2000). Later, this density deficit may reduce back to some percents only while substantial (>30%) cortical thinning ensues (Eser *et al.*, 2004). This initial reduction in cortical density among paralysed persons is basically concordant what was seen in the present study. Apparently, the 3-month duration of the present bed rest intervention was too short to result in substantial reduction in cortical area, at least to the extent that could be reliably captured by low resolution pQCT.

Trabecular loss was much more substantial in magnitude and rather uniform without a distinct indication of direction-specific loss. The largest (∼9%) bone loss in the anterior sector and relative insensitivity of the medial loss to any countermeasure may be noted, however. In general, substantial trabecular bone loss is a well-known consequence of bone disuse (Sievänen, 2010). However, the specific nature of disuse, e.g. in bed rest studies, the body position and the total duration (0° and 35 days in (Rittweger *et al.*, 2009), 0° and 56 days in (Armbrecht *et al.*, 2010), and −6° head tilt down and 90 days in LTBR study (Rittweger *et al.*, 2005)) may account for the results.

Somewhat surprisingly, the most effective countermeasure against bed rest–induced trabecular loss appeared to be the flywheel exercise, not antiresorptive medication. This observation could stipulate the speculation whether incident mechanical loading is necessary for such a medication to be effective. Among ambulatory postmenopausal women, however, there is no indication for interaction between bisphosphonate treatment and exercise intervention; both interventions showed anticipated effects independently without interaction (Uusi-Rasi *et al.*, 2003). The time course of the bone losses in the present pamidronate group suggests that there was good efficacy in the first 4 weeks and that further injections might have led to better maintenance of bone mass (Rittweger *et al.*, 2005).

The strengths of the present study are the randomized prospective design with three interventional arms and a long follow-up period: the knowledge on the baseline variance which helped to put the observed skeletal responses in proper context; the sector-specific analysis of bone structure based on apparent ankle biomechanics; and the application of a new image processing method known to reduce noise in pQCT images (Cervinka *et al.*, 2010). It is known that the use of different thresholds and analysis methods can affect the results of pQCT studies (Kontulainen *et al.*, 2007) that may also account for the above-mentioned discrepant results observed in the present study compared to other bed rest studies (Rittweger & Felsenberg, 2009; Rittweger *et al.*, 2009, 2010; Armbrecht *et al.*, 2010).

The bed rest studies undoubtedly provide a useful model for bone loss (Pavy-Le *et al.*, 2007), but some inherent limitations of these demanding studies need to be recognized. First, the group sizes are deemed quite limited because of obvious challenges in financial, technical and practical execution of the intervention. Second, besides the small number of subjects in each study group, subject background characteristics within and between the groups can be heterogeneous, which may increase the variance in responses, confound their interpretation and further compromise the power of the study to detect between-group differences with statistical confidence. Therefore, many results obtained from bed rest studies remain descriptive and indicative only rather than conclusive in the statistical sense. Third, the bed rest intervention was too short to result in measurable changes in cortical area, but on the other hand, 3 months being bound to bed is close to maximum among healthy volunteers. Fourth, during the recovery, the assessment of biomechanical environment was based on coarse qualitative description because of lack of specific quantitative data on bone loading (the type, amount and intensity of physical activity of particular subjects were not known). Notwithstanding these limitations, observations from bed rest interventions because of their specific and well-controlled design are of utmost importance in unravelling not only the skeletal responses to disuse but also the effects of accompanying countermeasures on disuse-induced bone loss.

Keeping the inherent statistical limitations of bed rest interventions in mind, it is concluded that the sector-specific analysis of bone cross-sections has potential to reveal skeletal responses to various interventions that cannot be inferred from the average analysis of the whole bone cross-section. This approach is considered particularly useful for evaluating the responses from the biomechanical point of view and thus strengthening the interpretation of bone data from this relevant aspect of bone physiology.

Acknowledgments

We acknowledge the Finnish Cultural Foundation and Research School of Tampere University of Technology for financial support of the doctoral studies for TC. This study was also supported by the Competitive Research Funding of the Tampere University Hospital (Grant 9K121).

The Long Term Bed Rest Study was carried out in 2001 and 2002 at MEDES, Toulouse/France and was organized by ESA and NASDA. Our participation in the study was supported by DLR (50 WB 0156). Last but not least, we are grateful to the participants of the LTBR study – without their selfless contribution; this work would not have been possible.

References

Armbrecht G, Belavý DL, Gast U, Bongrazio M, Touby F, Beller G, Roth HJ, Perschel FH, Rittweger J, Felsenberg D. Resistive vibration exercise attenuates bone and muscle atrophy in 56 days of bed rest: biochemical markers of bone metabolism. *Osteoporos Int* (2010); **21**: 597–607.

Cervinka T, Hyttinen J, Sievanen H. Enhanced bone structural analysis through pQCT image preprocessing. *Med Eng Phys* (2010); **32**: 398–406.

Chappard D, Alexandre C, Palle S, Vico L, Morukov BV, Rodionova SS, Minaire P, Riffat G. Effects of a bisphosphonate (1-hydroxy ethylidene-1,1 bisphosphonic acid) on osteoclast number during prolonged bed rest in healthy humans. *Metabolism* (1989); **38**: 822–825.

Currey JD. Bone strength: what are we trying to measure? *Calcif Tissue Int* (2001); **68**: 205–210.

Dudley-Javoroski S, Shields RK. Asymmetric bone adaptations to soleus mechanical loading after spinal cord injury. *J Musculoskelet Neuronal Interact* (2008); **8**: 227–238.

Eser P, Frotzler A, Zehnder Y, Wick L, Knecht H, Denoth J, Schiessl H. Relationship between the duration of paralysis and bone structure: a pQCT study of spinal cord injured individuals. *Bone* (2004); **34**: 869–880.

Frey-Rindova P, de Bruin ED, Stüssi E, Dambacher MA, Dietz V. Bone mineral density in upper and lower extremities during 12 months after spinal cord injury measured by peripheral quantitative computed tomography. *Spinal Cord* (2000); **38**: 26–32.

Frost HM. Bone's Mechanostat: a 2003 update. *Anat Rec* (2003); **274**: 1081–1101.

Grigoriev AI, Morukov BV, Oganov VS, Rakhmanov AS, Buravkova LB. Effect of exercise and bisphosphonate on mineral balance and bone density during 360 day antiorthostatic hypokinesia. *J Bone Miner Res* (1992); **7**: 449–455.

Kontulainen S, Liu D, Manske S, Jamieson M, Sievänen H, McKay H. Analyzing cortical bone cross-sectional geometry by peripheral QCT: comparison with bone Histomorphometry. *J Clin Densitom* (2007); **10**: 86–92.

Lang TF, Leblanc AD, Evans HJ, Lu Y. Adaptation of the proximal femur to skeletal reloading after long-duration spaceflight. *J Bone Miner Res* (2006); **21**: 1224–1230.

LeBlanc AD, Schneider VS, Evans HJ, Engelbretson DA, Krebs JM. Bone mineral loss and recovery after 17 weeks of bed rest. *J Bone Miner Res* (1990); **5**: 843–850.

LeBlanc AD, Driscol TB, Shackelford LC, Evans HJ, Rianon NJ, Smith SM, Feeback DL, Lai D. Alendronate as an effective countermeasure to disuse induced bone loss. *J Musculoskelet Neuronal Interact* (2002); **2**: 335–343.

Ma H, Leskinen T, Alen M, Cheng S, Sipilä S, Heinonen A, Kaprio J, Suominen H, Kujala UM. Long-term leisure time physical activity and properties of bone: a twin study. *J Bone Miner Res* (2009); **24**: 1427–1433.

Macdonald HM, Cooper DM, McKay HA. Anterior-posterior bending strength at the tibial shaft increases with physical activity in boys: evidence for non-uniform geometric adaptation. *Osteoporos Int* (2009); **20**: 61–70.

Martin RB. Determinants of the mechanical properties of bones. *J Biomech* (1991); **24**: 79–88.

Pavy-Le Traon A, Heer M, Narici MV, Rittweger J, Vernikos J. From space to earth: advances in human physiology from 20 years of bed rest studies (1986–2006). *Eur J Appl Physiol* (2007); **101**: 143–194.

Rantalainen T, Nikander R, Heinonen A, Suominen H, Sievänen H. Direction-specific diaphyseal geometry and mineral mass distribution of tibia and fibula: a pQCT study of female athletes representing different exercise loading types. *Calcif Tissue Int* (2010); **86**: 447–454.

Rittweger J, Felsenberg D. Recovery of muscle atrophy and bone loss from 90 days bed rest: results from a one-year follow-up. *Bone* (2009); **44**: 214–224.

Rittweger J, Frost HM, Schiessl H, Ohshima H, Alkner B, Tesch P, Felsenberg D. Muscle atrophy and bone loss after 90 days' bed rest and the effects of flywheel resistive exercise and pamidronate: results from the LTBR study. *Bone* (2005); **36**: 1019–1029.

Rittweger J, Felsenberg D, Maganaris C, Ferretti JL. Vertical jump performance after 90 days bed rest with and without flywheel resistive exercise, including a 180 days follow-up. *Eur J Appl Physiol* (2007); **100**: 427–436.

Rittweger J, Simunic B, Bilancio G, De Santo NG, Cirillo M, Biolo G, Pisot R, Eiken O, Mekjavic IB, Narici M. Bone loss in the lower leg during 35 days of bed rest is predominantly from the cortical compartment. *Bone* (2009); **44**: 612–618.

Rittweger J, Beller G, Armbrecht G, Mulder E, Buehring B, Gast U, Dimeo F, Schubert H, de Haan A, Stegeman DF, Schiessl H, Felsenberg D. Prevention of bone loss during 56 days of strict bed rest by side-alternating resistive vibration exercise. *Bone* (2010); **46**: 137–147.

Ruff C, Holt B, Trinkaus E. Who's afraid of the big bad Wolff?: ''Wolff's law'' and bone functional adaptation. *Am J Phys Anthropol* (2006); **129**: 484–498.

Ruml LA, Dubois SK, Roberts ML, Pak CY. Prevention of hypercalciuria and stone-forming propensity during prolonged bed-rest by alendronate. *J Bone Miner Res* (1995); **10**: 655–662.

Seits P, Ruegsegger P. Fast contour detection algorithm for high precision quantitative CT. *IEEE Trans Med Imag* (1983); **2**: 136–141.

Shackelford LC, LeBlanc AD, Driscoll TB, Evans HJ, Rianon NJ, Smith SM, Spector E, Feeback DL, Lai D. Resistance exercise as a countermeasure to disuse-induced bone loss. *J Appl Physiol* (2004); **97**: 119–129.

Shaw CN, Stock JT. Intensity, repetitiveness, and directionality of habitual adolescent mobility patterns influence the tibial diaphysis morphology of athletes. *Am J Phys Anthropol* (2009); **140**: 149–159.

Sievänen H. A physical model for dual-energy X-ray absorptiometry – derived bone mineral density. *Invest Radiol* (2000); **35**: 325–330.

Sievänen H. Immobilization and bone structure in humans. *Arch Biochem Biophys* (2010); **503**: 146–152.

Sievänen H, Koskue V, Rauhio A, Kannus P, Heinonen A, Vuori I. Peripheral quantitative computed tomography in human long bones: evaluation of ***in vitro*** and *in vivo* precision. *J Bone Miner Res* (1998); **13**: 871–882.

Uusi-Rasi K, Kannus P, Cheng S, Sievänen H, Pasanen M, Heinonen A, Nenonen A, Halleen J, Fuerst T, Genant H, Vuori I. Effect of alendronate and exercise on bone and physical performance of postmenopausal women: a randomized controlled trial. *Bone* (2003); **33**: 132–143.

Watanabe Y, Ohshima H, Mizuno K, Sekiguchi C, Fukunaga M, Kohri K, Rittweger J, Felsenberg D, Matsumoto T, Nakamura T. Intravenous pamidronate prevents femoral bone loss and renal stone formation during 90-day bed rest. *J Bone Miner Res* (2004); **19**: 1771–1778.

Zerwekh JE, Ruml LA, Gottschalk F, Pak CY. The effects of twelve weeks of bed rest on bone histology, biochemical markers of bone turnover, and calcium homeostasis in eleven normal subjects. *J Bone Miner Res* (1998); **13**: 1594–1601.

Publication III:

Cervinka T., Hyttinen J., Sievanen H.
Threshold-Free Automatic Detection of Cortical Bone Geometry by Peripheral Quantitative Computed Tomography.
Journal of Clinical Densitometry, 2012; 15(4): 413-421.

Threshold Free Automatic Detection of Cortical Bone Geometry by Peripheral Quantitative Computed Tomography

Abstract

An accurate assessment of bone strength is an important goal in clinical bone research. For appropriate information on bone strength, precise segmentation of actual cross-sectional bone geometry is needed. In this paper, we introduce an automatic, simple and fast approach for reliable segmentation of cortical bone cross-sectional area based on the outer boundary detection and subsequent shrinking procedure (OBS). Using repeated in vivo pQCT images of distal tibia from 25 subjects we compared new segmentation results with those obtained from commonly applied simple density thresholds and from a recent advanced analysis based on distance regularized level set evolution (DRLSE). Manual segmentation of cortical bone done by three independent evaluators was considered a gold standard. The new approach showed nearly 50% less variation in error compared to threshold based analysis in conjunction with a recently introduced statistical preprocessing method, and agreed well with results obtained from manual segmentation. The DRLSE segmentation resulted consistently in ~15% mean overestimation of all geometrical traits with a similar variation of data as obtained from the OBS method. In conclusion, the OBS method improved assessment of all observed measures of cortical geometry and can enhance the cortical bone analysis of pQCT images in clinical research studies.

Key words: pQCT, Cortical bone, Image preprocessing, Segmentation, distance regularized level set evolution, Bone strength

Introduction

Fragility fractures are a common health problem in aging populations. The present diagnostic assessment of bone fragility and fracture risk rests largely on dual-energy X-ray absorptiometry (DXA) measured bone mineral density (BMD) (1). However, the BMD approach is limited not only by the inherent inaccuracy of DXA caused by the violation of the two-component (homogeneous soft tissue and bone) assumption in individual patients (2) but also by the inability of planar DXA to yield reliable information on actual bone geometry and structure which is necessary for proper assessment of bone strength (3). Peripheral quantitative computed tomography (pQCT) allows a reasonable option to assess bone cross-sectional geometry and separate it into trabecular and cortical compartments (4). Although the commonly used pQCT systems lack sufficient spatial resolution to capture specific structural traits compared to present high resolution pQCT (HR-pQCT) systems, they are still more available and widely used in many recent studies (5-12). While pQCT cannot be applied to clinically relevant vertebral and proximal femur sites due to obvious technical limitations, it provides similar information on apparent structural traits from appendicular bones as does the clinical quantitative CT (QCT) from proximal femur and lumbar spine (13-16).

Typically the separation of cortical bone from trabecular bone and soft tissues in pQCT image is based on application of density thresholds (17-19). This practice comes not only from technical simplicity of analysis algorithms but also from the need for reproducible and tangible results in clinical settings. Many sophisticated segmentation algorithms, e.g., level sets based segmentations (20-22), need relatively large operator involvement.

Obviously, both accurate and precise description of the cortical bone compartment is essential for appropriate assessment of bone strength (14,23-25). This is, however, not always guaranteed by simple threshold based analyses. It is well known that the use of different threshold values can substantially modulate pQCT results (17-19), mainly because of the partial volume effect, relatively low signal to noise ratio in pQCT images, and presence of movement artifacts leading to blurred or discontinued cortical edges.

To account for the above described limitations of simple thresholding, we developed a new fast algorithm for cortical bone segmentation following the basis ideas of fast contour detection algorithm for QCT images described earlier by Seitz and Ruegsegger (26). Specifically, we addressed the following four research questions: first, what is the in vivo accuracy of the refined algorithm in analyzing raw pQCT images using the manual segmentation of cortical bone as the gold standard; second, whether could the preprocessing of pQCT images enhance the cortical analysis compared to the analysis based on raw images; third, to what extent can the new segmentation algorithm improve the analysis of real pQCT images in comparison to simple, density threshold based analyses and to a recently propoced, more advanced analysis based on a new variational level set formulation (22)?

Material and Methods

The pQCT system and in vivo data

The pQCT scanner (XCT 3000, Stratec Medizintechnik GmbH, Pforzheim, Germany) used in this study represents one model of the widely used XCT brand (models 960, 2000, 3000) in clinical bone research. In this study, the mean of results from two

repeated pQCT scans of the distal tibia (5% of the estimated tibial length from the distal endplate) from 25 volunteers was used in the analysis to reduce random variation so that the differences between different approaches could be better determined. The age range of the subjects was 35–66 yrs; weight range was 50–89 kg; and height range was 159–179 cm. The repeated pQCT scans were obtained from a precision study which was carried out as a part of quality assurance procedure of our bone densitometry unit. Informed consent was obtained from the subjects and the in-house review board approved the study protocol. Both scans were performed on the same day with repositioning according to our standard procedure (4). The pixel-size of the pQCT image was 0.5 mm × 0.5 mm, the slice thickness was 2.5 mm, and the translational scan speed was 30 mm/s.

Preprocessing procedures

For preprocessing of the raw pQCT image data, standard image preprocessing methods, median filtering with a 3x3 window and a 5x5 window, were applied. In addition to these two median filters, two preprocessing methods introduced recently to reduce noise in pQCT images were used (27). Details of these methods have been described elsewhere (27). In short, the methods were composed of two steps. First, a gray level transformation of image intensity histogram was reduced by implementing a common, iteratively employed, piecewise linear conversion function. Second, the transformed data was corrected by using Bayesian approach and Markov random fields with 3x3 linear or 5x5 Gaussian neighborhoods and redundant wavelet transform. Because of the different sizes and types of the Markov random fields (basic steps of both approaches were identical), the preprocessing methods were coined the Linear window (LinW) and Gaussian window (GaussW).

Manual segmentation

The manual segmentation of cortical bone in raw unprocessed pQCT images was performed by three experts in segmenting radiological images. The mean of the cortical cross-sectional area (CoA) from these three blinded, independent analyses was considered the gold standard of CoA. The manual segmentation was implemented by in-house software developed for segmentation and visualization of radiological images (Department of Biomedical Engineering, Tampere University of Technology). Manual segmentation of a single pQCT image took 3 to 6 minutes.

Density threshold based segmentation

Two segmentations based on standard threshold procedures employed by the Stratec analysis software were performed. First, the threshold values were set to 200 mg/cm^3 for the outer bone threshold and to 661 mg/cm^3 for the inner threshold according to optimal settings for cortical bone detection proposed by Kontulainen et al. (19). Second, the thresholds corresponding to levels automatically chosen by the contour and peel modes 2 were used in line with our standard procedures (4). Both of these Stratec analysis protocols use 3x3 median filtering as a preprocessing method and are called C1/P2 and C2/P2, respectively.

Level set based segmentation

In addition to simple threshold based methods, a sophisticated segmentation method based on distance regularized level set evolution (DRLSE) developed recently by Li et al. (22) was performed. Compared to conventional level set formulations, the DRLSE algorithm has several advantages. It allows: elimination of re-initialization; use of large time steps to significantly speed up curve evolution, while ensuring numerical accuracy;

and a computationally efficient simple implementation. In the present study, the DRLSE method was applied both on raw pQCT images and those processed with standard 3x3 median filtering so that the influence of additional preprocessing on segmentation results could be assessed. The parameters of the DRLSE method were experimentally set to the following values so that as common setting of parameters as possible could be used for analyzing the whole set of pQCT images. For the inner contour evolution, time step was 5.0, lambda 40.0, alpha -11.0, epsilon 2, and sigma 1.5. For the outer contour evolution time step was 5.0, lambda 5.0, alpha 3.5, epsilon 2, and sigma 1.5. However, for some images, the alpha parameter had to be changed within the limits -11.0 − -5.0 and 3.0 − 4.0 for inner and outer contours, respectively. With the above mentioned settings, the DRLSE segmentation of CoA of a single slice pQCT image with space resolution 375x375 pixels took about 20 - 25s including the contour initialization.

Proposed algorithm for automatic segmentation of cortical bone

The new method comprises the delineation procedure of the outer boundary of cortical bone and the subsequent shrinking procedure of cortical pixels until the inner cortical boundary is found. The present method is called the outer boundary detection and shrinking algorithm and coined as the OBS-method. The OBS-method is described in detail below.

Detection of outer bone contour

The outer contour of the cortical bone in raw and four preprocessed pQCT images was detected by using a simple and fast algorithm described by Seits and Ruegsegger (26). This algorithm was chosen for its good performance in detecting the outer contour of cortical bone without additional preprocessing of the raw images and for yielding highly

reproducible data. In brief, the initialization of the algorithm was based on exhaustive search in the successive lines of the image for the first (seed) pixel representing bone tissue (i.e., the pixel value was greater or equal threshold value T). The algorithm further assumed an initial set of contour elements P_{n-1}, …, P_{n-k} and the traced contour being situated on the left side of initial element. Selection of the next contour element P_n was based on searching the nearest neighborhood connected to element P_{n-1} (8 possible paths) and testing possible successor elements C_1, …, C_8 against a given threshold value T. The first element which fulfilled the threshold condition was chosen as the new contour element P_n. Seits and Ruegsegger stated that they achieved satisfactory results with only two predecessor elements $P_{n-1,}$ P_{n-2} and three possible successor elements C_1, C_2, C_3 situated in line with direction of predecessor element. Hence, our contour detection was also based on this approach.

However, in some cases (e.g., blurred or discontinued edges caused by movement artifacts), the setting of the correct threshold value T can be critical, making the algorithm unable to follow the correct contour without operator intervention (Figure 1). Therefore to improve the contour detection, we refined the method further by using additional information from the first derivation of the image data and by adding a supreme condition to the threshold condition for assessment of the new contour element P_n. Hence, the threshold value T was only used to coarsely distinguish between bone and soft tissues, and to search initial contour elements (the T value was experimentally set to 200 mg/cm^3). The new condition compared the values of the first derivation of image data corresponding to the new candidate elements C_1, C_2, and $C_{3,}$ and the new candidate element with the largest value was taken as the new contour element P_n. Then the difference between the largest and the second largest value was determined. If this

difference, representing possible candidate elements of traced contour (e.g. C_2 and C_3), was smaller than a pre-defined decision value D (set to 40mg/cm^4 based on knowledge of a profile of the first derivation of pQCT image), it remained uncertain which path should be followed. Therefore, the algorithm continued searching the next candidate states ($C_{2.1,}$ $C_{2.2}$ and $C_{3.1}$, $C_{3.2}$) in possible path directions of the traced contour. Values of the first derivation of image data representing the next candidate states ($C_{2.1,}$ $C_{2.2}$ and $C_{3.1}$, $C_{3.2}$) were then compared with each other and the largest value determined which one of the possible candidates (C_2 and C_3) was chosen as the new contour element P_n. The example of pattern of the candidate and following candidate states for two possible configurations is depicted in Figure 2.

Shrinking procedure

After detecting the outer cortical contour, our method detected the inner cortical boundary by radial shrinking of the region inside the outer contour (pixel by pixel) until the inner cortical boundary was found. For every pixel at the outer contour, the shrinking procedure aimed at searching the second peak of the first derivation of the image data along a line starting from the given pixel to the center of mass of the bone region. Until the next local maximum of the first derivation was found, the preceding pixels along the line were removed. These maximal points were considered to coincide with the inner boundary of the cortical shell (27). The shrunken region corresponded to the trabecular bone area (TrA) and subtraction of the original and the shrunken bone regions provided the CoA. In order to prevent undue extension of the inner cortical boundary, e.g. in case of partially disrupted edges, the maximum radial change in cortical thickness was experimentally set to one mm (two pixels) along 2 mm distance

(four pixels) of the bone contour. The procedure for the inner cortical contour detection is illustrated in Figure 3.

The time of CoA analysis from a single slice pQCT image comprising 375x375 pixels was about 3-4s for all automatic segmentation approaches depending on used preprocessing technique. The unsupervised preprocessing and cortical segmentation were implemented in MATLAB ver. R2011a and performed on a laptop computer (Lenovo X61) with Intel Core 2 Duo CPU (2.1 GHz) and 4GB of RAM.

Statistical analysis

As descriptive statistics, the means and standard deviation (SD) are given. Besides the segmented CoA, cortical thickness (CoTh) as obtained from a commonly used ring model of the bone, and maximal (I_{max}) and minimal (I_{min}) areal moments of inertia of cortical cross-sections[1] were determined as relevant measures of cortical geometry. Accordingly, we compared the CoA, CoTh, I_{max} and I_{min} results obtained from C1/P2 and C2/P2 segmentations, DRLSE based segmentation and the OBS-method after applying different preprocessing approaches to results obtained from the manual segmentation. We calculated the mean error as $100*(X_{method} - X_{manual})/X_{manual}$ and its standard deviation, where X denotes the geometric trait of interest. Statistical significance was evaluated by 95% confidence intervals (95% CIs) of the mean error. Pearson correlation coefficients were determined between cortical traits obtained from the manual segmentation and those obtained from different segmentation and preprocessing methods. As the delineation of cortical area was the primary target of

[1] The cortical cross-sections represented either the actual segmented cortical areas, or for the density threshold based methods C1/P2 and C2/P2, cortical pixels exceeding the inner (higher) bone thresholds.

segmentation, the Bland-Altman plots illustrating the differences (or agreement) between the manual segmentation and the method of interest are shown only for CoA results.

Results

Differences in distal tibia CoA, CoTh, I_{max} and I_{min} from various segmentation and preprocessing methods in comparison to manual segmentation are shown in Table 1 and Figure 4. In general, substantial between-method differences did manifest as statistically significant under- or overestimations. Only the results of the OBS-method in conjunction with the LinW preprocessing agreed well on average with results obtained from the manual segmentation, the slight 2% overestimation of Imin excluded. Without preprocessing of the image data, the OBS method resulted in 4-11% mean underestimation of all geometrical traits. Preprocessing with the GaussW method resulted consistently in 4 % mean underestimation, the accurate assessment of Imin excluded. Median filtering with 3x3 and 5x5 windows showed opposite effects leading to 5-11% mean underestimation and to 3-10% mean overestimation of geometrical traits. The C1/P2 threshold based segmentation resulted consistently in underestimated geometrical traits whereas the C2/P2 method using lower thresholds led to clear overestimation. It is also worth noting that these threshold based methods showed almost twice greater variation (SD) of measured traits. The sophisticated DRLSE segmentation resulted consistently in ~15% mean overestimation of all geometrical traits with a similar variation of data as obtained from the OBS method irrespective of image preprocessing. Median filtering had a negligible (increasing) effect on the results in contrast to more distinct effect of preprocessing on the OBS method.

The Bland-Altman plots illustrating the mean agreement of the CoA assessment between different segmentation methods and the manual segmentation are shown in Figure 5. Correlation between the cortical traits obtained from the manual segmentation with those obtained from different segmentation and preprocessing methods is shown in Table 2. As expected, all correlations were significant ($r>0.39$, $p<0.05$) most of them indicating very strong associations (r about 0.9 or greater). Strikingly, the correlations were lower for density threshold based C1/P2 and C2/P2 segmentations, excluding the correlations obtained from CoTh values.

Discussion and Conclusion

In this work a novel automatic segmentation algorithm, coined the OBS method, for evaluation of cortical bone geometry from pQCT images was presented. The segmentation comprised a fast and simple contour detection algorithm for the outer cortical boundary and a subsequent shrinking procedure for the inner cortical boundary both employing the information from the first derivation of the image data. Further, the aim of present study was to determine whether the OBS method could improve the in vivo assessment of tibial cortical geometry compared to commonly used simple density threshold based analyses (4,19) and also to a more sophisticated image processing method based on the DRLSE method (22). Among the several methods evaluated in the present study, only the results of the OBS method used in conjunction with LinW preprocessing showed consistently an excellent agreement on average with the results of manual segmentation whereas the other approaches mostly resulted in either systematic under- or overestimations of geometrical traits of cortical bone.

The importance of bone geometry and that of cortical bone in particular as major determinants of bone strength has been recently pinpointed (14,23-25,28). Obviously, overestimation or underestimation of cortical geometry can confound the estimation of bone strength in clinical research. Thus, to reliably estimate bone strength, accurate assessment of relevant cortical bone traits is indispensable. Previously, some studies have assessed optimal density threshold settings with comparison to phantoms (17) or histomorphometry at the tibial diaphysis (19,29). However, to our knowledge, no study has so far compared several automated analysis methods with results obtained by manual segmentation of cortical bone.

The present study showed that the novel OBS segmentation of cortical bone could outperform not only the commonly used threshold based analyses but also a sophisticated level set approach for image segmentation and delineated the cortical bone more accurately (Figures 4 and 5). Nevertheless, it is noted that the correlation between the OBS and DRLSE methods was generally very high (even up to 0.96 for Imax and 0.97 for Imin; data not shown) indicating that the result from these methods may be made comparable through proper adjustment. Further, both the new OBS method and the sophisticated DRLSE methods were associated with almost halved variation in geometrical traits compared to the results from density threshold based analyses. However, it is recalled that the OBS method was faster and moreover, the user input is more straightforward. In DRLSE method the complex setting of user defined parameters may reduce its usability and degrease the reproducibility of the results.

The preprocessing in itself did not significantly affect the variability in results, but the most consistent findings with the manual segmentation were obtained by LinW preprocessing (Figure 5). Be it noted that the smaller variance of results together with

better in vivo precision (27) can make the analysis of cortical bone more sensitive to detect between-group differences or within-group changes.

Notwithstanding the promising results, there are some issues that warrant further discussion and studies. Using the manual segmentation based on visual assessment as the gold standard may rise some concerns about its reliability. Obviously, the human eye can see not only the information which truly is present in the image but also distinguish lines and borders which cannot be recognized by any automatic algorithm and therefore reduce the influence of partial volume effect. However, this skill rests on previous experience in analyzing image data and can differ substantially between raters. Apparently an objective assessment such as histomorphometry and high resolution digitization is not an option for an in vivo study. Therefore, we considered the mean of three independent manual segmentations of cortical bone as the gold standard, which reduces possible between-individual variance in determination of cortical area. In the present study, the relative standard deviation of differences between raters in manual segmentations of CoA was 4.4% (< 10mm^2). This is comparable to the in vivo precision of cortical area measurement (4).

The coarse spatial resolution of pQCT images obviously limits the performance of all image processing methods, but especially this concerns the threshold based analyses when the thickness of cortical bone is less than 4 mm (29) as is the case with distal tibia (and all epiphyseal sites of long bones). The latter influence is clearly seen in Bland-Altman plots (Figure 5) which indicate the dependence of errors on the size of CoA (ie, the smaller CoA the more underestimation) in the threshold based analyses. The thinness of cortical bone may also account for somewhat conflicting correlations indicating seemingly better performance for simple density threshold based

segmentation methods. Apparently, the validity of the estimation of mean cortical thickness of relatively thin-walled bones with coarse pQCT imaging remains questionable.

The strength of present study pertains to the assessment of several different segmentation approaches and preprocessing methods that were applied to the distal tibia, the most common site of interest in pQCT studies. As the bone structure of distal radius, the other common site in pQCT studies, is similar to distal tibia, the present results are likely applicable to distal radius as well. Furthermore, it is also possible that the OBS method could be applied for cortical analysis of clinical QCT images with similarly coarse resolution (< 1 mm) and also for HR-pQCT with much better spatial resolution (< 0.1 mm) with only small modifications. However, further studies are needed to show this.

In conclusion, this study showed that the new OBS algorithm performed reasonably well and it offers a promising practical tool to enhance the assessment of cortical bone geometry in pQCT images. As a broader perspective, a more reliable cortical analysis may facilitate estimation of bone strength and related prediction of fracture risk.

References

1. Kanis JA, Borgstrom F, De Laet C, et al. 2005 Assessment of fracture risk. Osteopor Int. 16:581-589.
2. Bolotin HH, Sievänen H. 2001 Inaccuracies inherent in dual-energy X-ray absorptiometry in vivo bone mineral density can seriously mislead diagnostic/prognostic interpretations of patient-specific bone fragility. J Bone Miner Res. 16:799-805.
3. Sievänen H, Kannus P, Järvinen TL. 2007 Bone quality: an empty term. PLoS Med. 4:e27.
4. Sievänen H, Koskue V, Rauhio A, Kannus P, Heinonen A, Vuori I. 1998 Peripheral quantitative computed tomography in human long bones: Evaluation of in vitro and in vivo precision. J Bone Miner Res. 13:871-882.
5. Lochmüller EM, Bürklein D, Kuhn V, et al. 2002 Mechanical strength of the thoracolumbar spine in the elderly: prediction from in situ dual-energy X-ray absorptiometry, quantitative computed tomography (QCT), upper and lower limb peripheral QCT, and quantitative ultrasound. Bone 31:77-84.
6. Rittweger J, Frost HM, Schiessl H, et al. 2005 Muscle atrophy and bone loss after 90 days' bed rest and the effects of flywheel resistive exercise and pamidronate: results from the LTBR study. Bone 36:1019-1029.
7. Nikander R, Sievänen H, Uusi-Rasi K, Heinonen A, Kannus P. 2006 Loading modalities and bone structures at nonweight-bearing upper extremity and weight-bearing lower extremity: a pQCT study of adult female athletes. Bone 39:886-894.

8. Rittweger J, Simunic B, Bilancio G, et al. 2009 Bone loss in the lower leg during 35 days of bed rest is predominantly from the cortical compartment. Bone 44:612-618.
9. Dowthwaite JN, Hickman RM, Kanaley JA, Ploutz-Snyder RJ, Spadaro JA, Scerpella TA. 2009 Distal radius strength: a comparison of DXA-derived vs pQCT-measured parameters in adolescent females. J Clin Densitom. 12:42-53.
10. Rittweger J, Beller G, Armbrecht G, et al. 2010 Prevention of bone loss during 56 days of strict bed rest by side-alternating resistive vibration exercise. Bone 46:137-147.
11. Szabo KA, Webber CE, Adachi JD, Tozer R, Gordon C, Papaioannou A. 2011 Cortical and trabecular bone at the radius and tibia in postmenopausal breast cancer patients: a Peripheral Quantitative Computed Tomography (pQCT) study. Bone 48:218-224.
12. Shedd-Wise KM, Alekel DL, Hofmann H, et al. 2011 The soy isoflavones for reducing bone loss study: 3-yr effects on pQCT bone mineral density and strength measures in postmenopausal women. J Clin Densitom. 14:47-57.
13. Lang TF, Leblanc AD, Evans HJ, Lu Y. 2006 Adaptation of the proximal femur to skeletal reloading after long-duration spaceflight. J Bone Miner Res. 21:1224-1230.
14. Melton 3rd LJ, Riggs BL, Keaveny TM, et al. 2007 Structural determinants of vertebral fracture risk. J Bone Miner Res. 22:1885-1892.
15. Riggs BL, Melton 3rd LJ, Robb RA, et al. 2008 A population-based assessment of rates of bone loss at multiple skeletal sites: evidence for substantial trabecular bone loss in young adult women and men. J Bone Miner Res. 23:205-214.

16. Melton 3rd LJ, Riggs BL, Keaveny TM, et al. 2010 Relation of vertebral deformities to bone density, structure, and strength. J Bone Miner Res. 25:1922-1930.
17. Ward KA, Adams JE, Hangartner TN. 2005 Recommendations for thresholds for cortical bone geometry and density measurement by peripheral quantitative computed tomography. Calcif Tissue Int. 77:275-280.
18. Ashe MC, Khan KM, Kontulainen SA, et al. 2006 Accuracy of pQCT for evaluating the aged human radius: an ashing, histomorphometry and failure load investigation. Osteoporos Int. 17:1241-1251.
19. Kontulainen S, Liu D, Manske S, Jamieson M, Sievänen H, McKay H. 2007 Analyzing cortical bone cross-sectional geometry by periphearal QCT: comparison with bone Histomorphometry. J Clin Densitom. 10:86-92.
20. Osher S, Paragios N, ed. 2003 Geometric Level Set Methods in Imaging, Vision, and Graphics. Springer-Verlag, New York.
21. Sethian J, ed. 2003 Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science. Cambridge University Press, Cambridge, UK.
22. Li C, Xu C, Gui C, Fox MD. 2010 Distance Regularized Level Set Evolution and Its Application to Image Segmentation. IEEE Trans Image Proc. 19:3243-3254.
23. Pistoia W, van Rietbergen B, Rüegsegger P. 2003 Mechanical consequences of different scenarios for simulated bone atrophy and recovery in the distal radius. Bone 33:937-945.

24. Roux JP, Wegrzyn J, Arlot ME, et al. 2010 Contribution of trabecular and cortical components to biomechanical behavior of human vertebrae: an ex vivo study. J Bone Miner Res. 25: 356-361.
25. Holzer G, von Skrbensky G, Holzer LA, Pichl W. 2009 Hip fractures and the contribution of cortical versus trabecular bone to femoral neck strength. J Bone Miner Res. 24:468-474.
26. Seits P, Ruegsegger P. 1983 Fast contour detection algorithm for high precision quantitative CT. IEEE Trans Med Imag. 2:136-141.
27. Cervinka T, Hyttinen J, Sievanen H. 2010 Enhanced bone structural analysis through pQCT image preprocessing. Med Eng Phys. 32:398-406.
28. Seeman E. 2003 Periosteal bone formation a neglected determinant of bone strength. N Engl J Med. 349:320-323.
29. Veitch SW, Findlay SC, Ingle BM, et al. 2004 Accuracy and Precision of Peripheral Quantitative Computed Tomography Measurements at the Tibial Metaphysis. J Clin Densitom. 7:209-217.
30. Chan TF, Vese LA. 2001 Active Contours Without Edges. IEEE Trans Im Proc. 10:266-277.

Figure captions:

Figure 1: Examples of distorted edges (left) and movement artifacts (right) in raw pQCT images, and development of detected contours for particular threshold setting T corresponding to density values (in mg/cm^3) from top to bottom: 590, 540, 460, 380 and 590, 460, 350, 270, respectively.

Figure 2: An example of newly proposed pattern of tracing algorithm for two out of eight eligible pixel arrangements with two predecessor elements $P_{n-1,}$ P_{n-2}, three successor elements $C_1 - C_3$ and four possible following candidate states $C_{1.1} - C_{3.2}$.

Figure 3: A raw pQCT image (left), radial X-ray attenuation profile along a radial line through the raw bone image (specified by cut line) with small distortion of the cortical bone on right side (middle) its first derivate (right). The maximum of the first derivative denotes the outer cortical boundary and the next local maximum clearly denotes the inner cortical boundary. This algorithm is repeated across the whole bone cross-section by rotating the line pixel by pixel along the outer contour.

Figure 4: Mean percentage differences in CoA, CoTh, Imax and Imin (95% CIs) between results from different analysis methods and the manual segmentation.

Figure 5: Bland-Altman plots for each CoA analysis approach with depicted regression lines for all cases where correlation was significant ($p<0.05$).

Table 1: Descriptive data of CoA, CoTh, I_{max} and I_{min} (mean, SD) as obtained from different segmentation and preprocessing methods

	CoA [mm^2]	CoTh [mm]	Imax [mm^4]	Imin [mm^4]
Manual segmentation	150.8 (19.2)	1.5 (0.1)	22684 (6809)	17198 (4975)

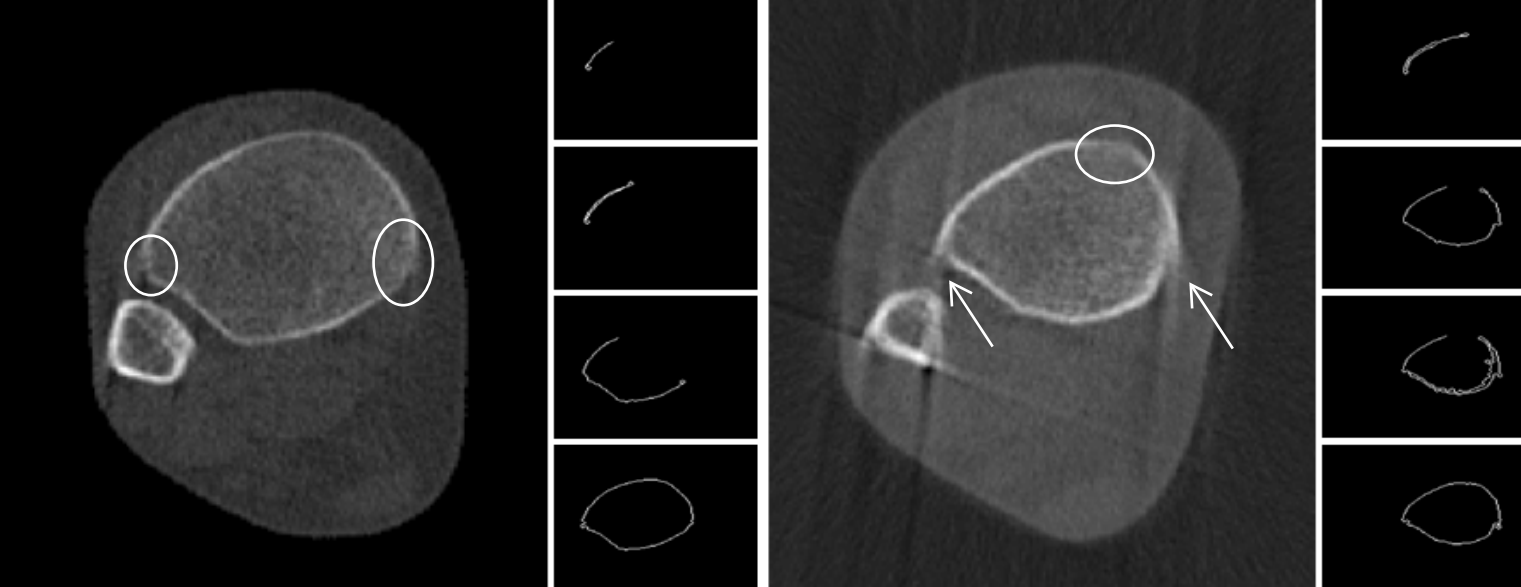
	Raw images	LinW	GaussW	Median 3x3	Median 5x5	C1/P2	C2/P2	DRLSE	med3x3 - DRLSE
CoA [mm^2]	132.7 (13.5)	153.0 (14.7)	142.9 (13.9)	133.2 (16.8)	160.2 (15.3)	137.0 (24.1)	175.3 (35.4)	172.6 (15.3)	176.4 (15.5)
CoTh [mm]	1.5 (0.1)	1.5 (0.1)	1.5 (0.1)	1.4 (0.1)	1.6 (0.1)	1.4 (0.3)	1.9 (0.4)	1.8 (0.1)	1.8 (0.1)
Imax [mm^4]	19776 (5101)	21961 (5580)	21355 (5408)	20055 (5295)	22970 (5591)	10224[a] (3990)	25261[b] (8080)	24886 (6306)	25319 (6528)
Imin [mm^4]	15973 (4392)	17593 (4838)	17322 (4618)	16097 (4531)	18701 (4902)	7152[a] (3613)	19084[b] (6063)	19169 (5095)	19623 (5192)

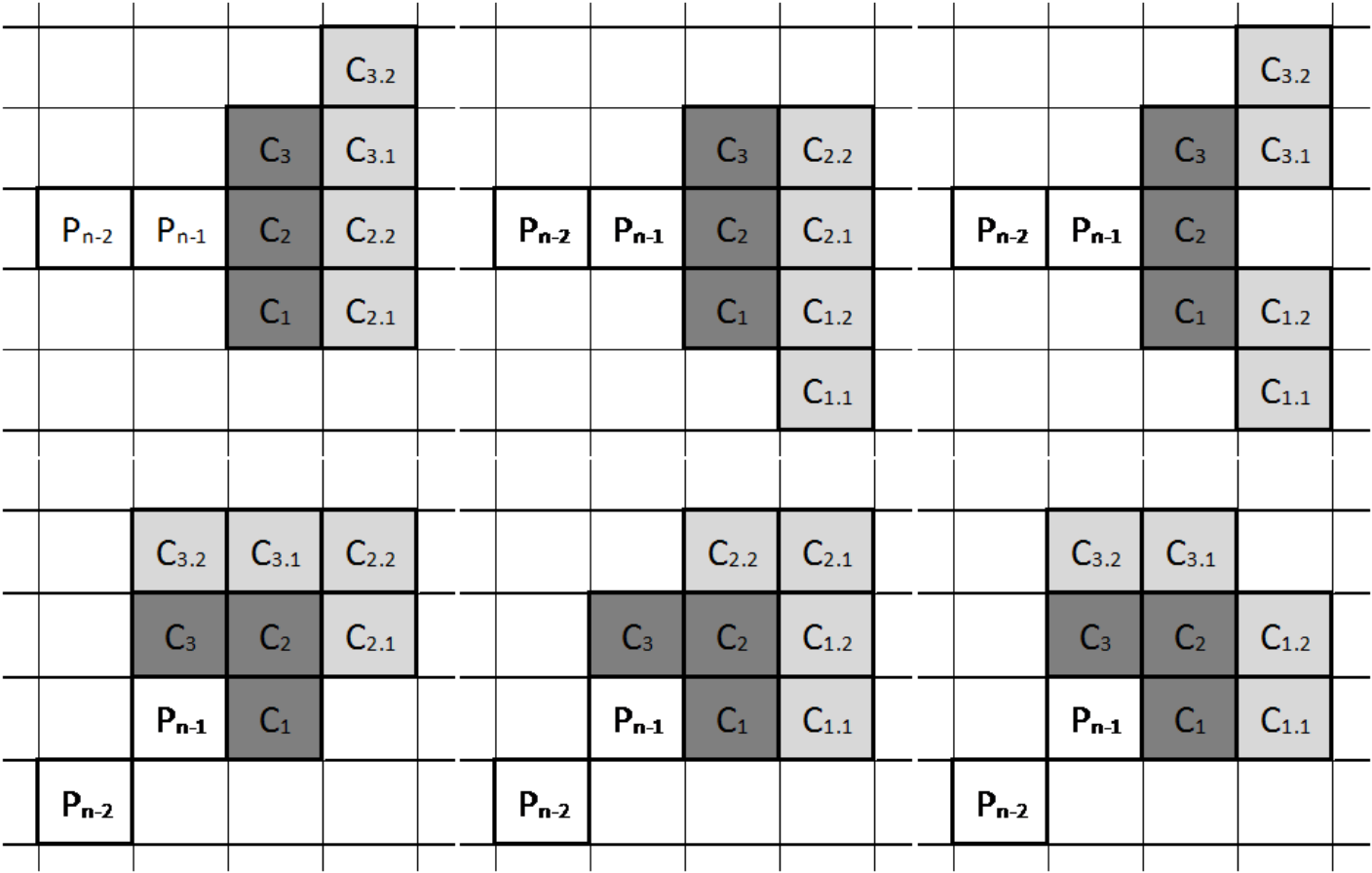
[a] based on cortical pixels exceeding the density threshold of 661 mg/cm^3

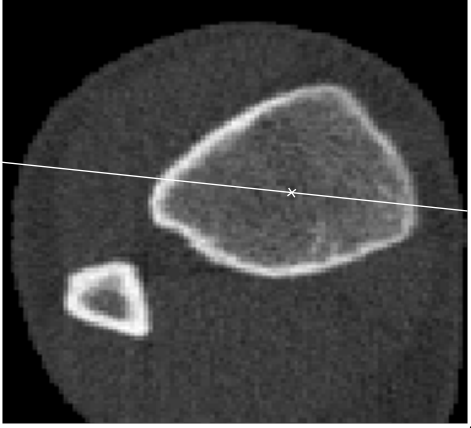
[b] based on cortical pixels exceeding the density threshold of 405 mg/cm^3

Table 2: Correlation between CoA, CoTh, Imax and Imin values obtained from manual segmentation and different segmentation methods

	Raw images	LinW	GaussW	Median 3x3	Median 5x5	C1/P2	C2/P2	DRLSE	med3x3 - DRLSE
CoA	0.71	0.88	0.81	0.79	0.84	0.62	0.61	0.76	0.79
CoTh	0.32	0.65	0.58	0.56	0.29	0.72	0.70	0.46	0.57
Imax	0.90	0.97	0.94	0.90	0.91	0.58	0.80	0.95	0.95
Imin	0.94	0.99	0.96	0.94	0.94	0.42	0.81	0.96	0.96




C3.2
C3
C3.1
Pn-2
Pn-1
C2
C2.2
C1
C2.1
C3.2
C3.1
C2.2
C3
C2
C2.1
C1
C1.2
Pn-1
C1
Pn-2
C2.2
C2.1
C3
C2
C1.2
Pn-1
C1
C1.1
Pn-2
C3.2
C3.1
C3
C2
C1.2
Pn-1
C1
C1.1
Pn-2

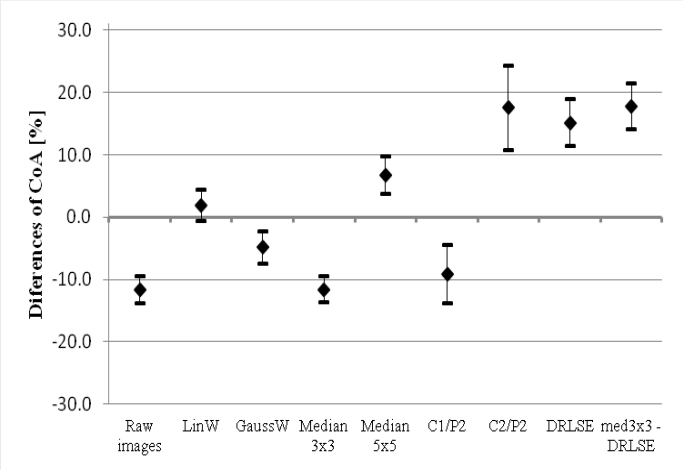


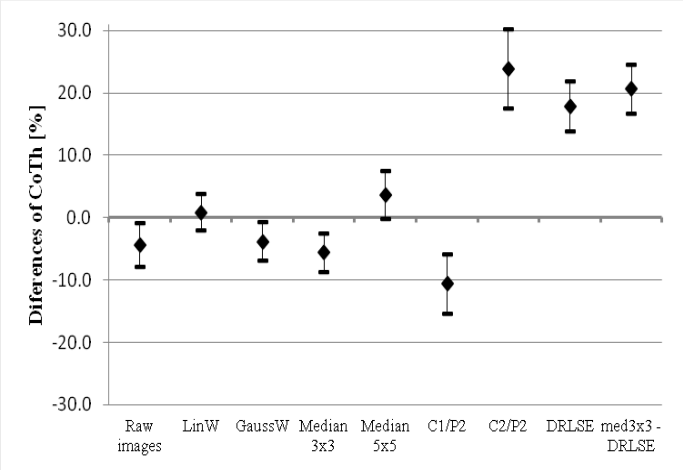
X-ray attenuation profile

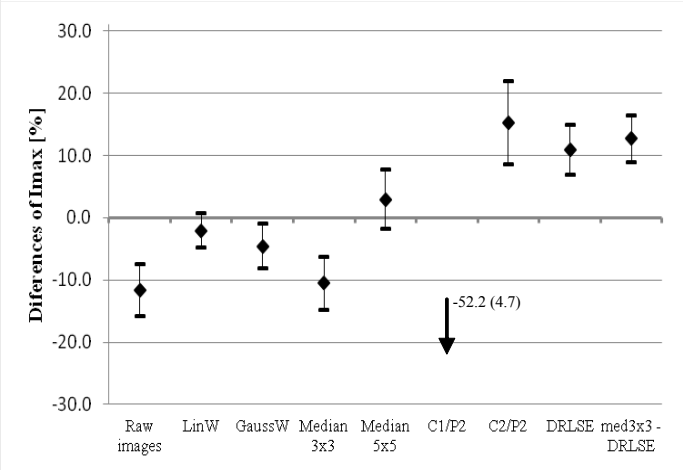
The first derivation

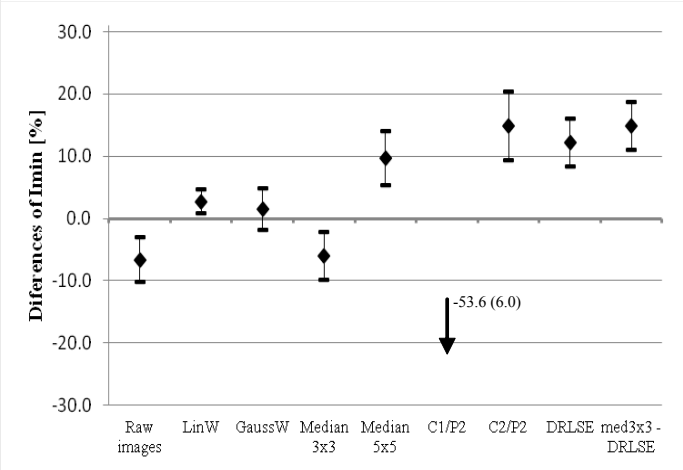
Detected outer contours of cortical bone

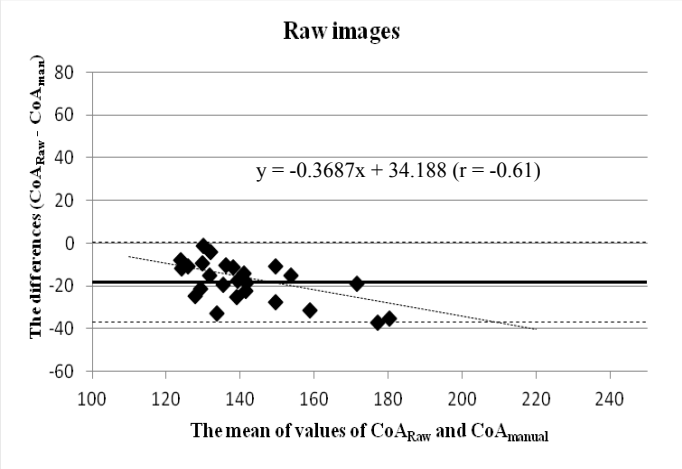
Next maxima of the first derivation

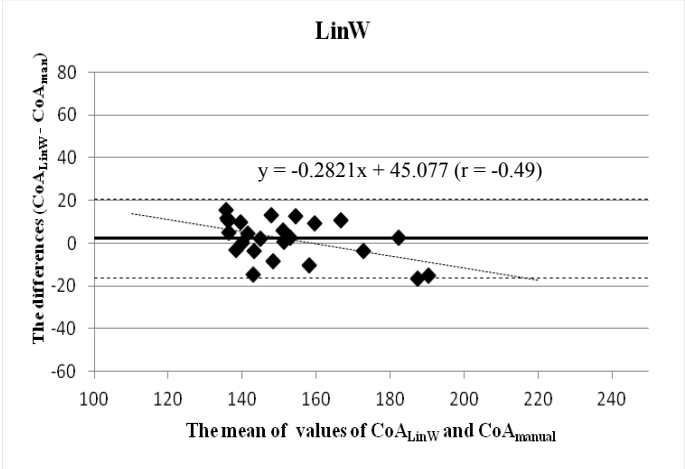

Diferences of CoA [%]
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
Raw images
LinW
GaussW
Median 3x3
Median 5x5
C1/P2
C2/P2
DRLSE
med3x3 - DRLSE

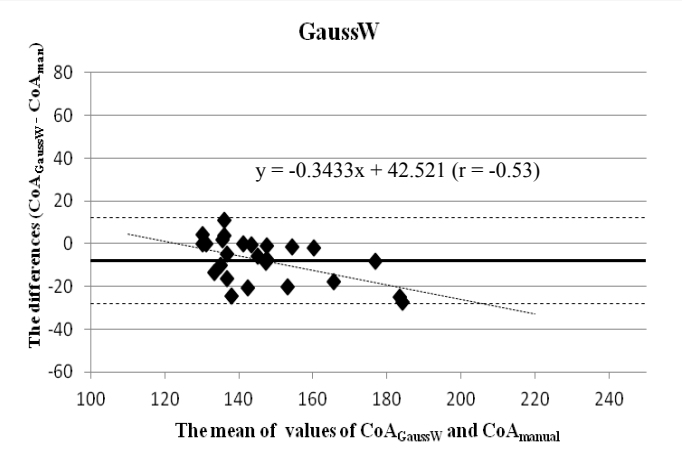

Diferences of CoTh [%]
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
Raw images
LinW
GaussW
Median 3x3
Median 5x5
C1/P2
C2/P2
DRLSE
med3x3 - DRLSE

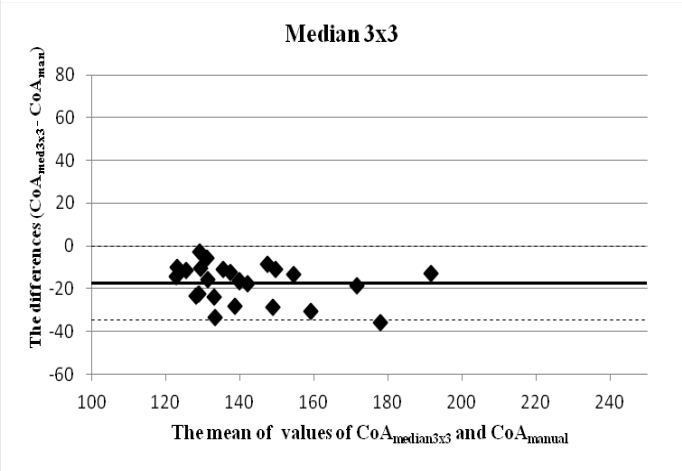

Diferences of Imax [%]
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
-52.2 (4.7)
Raw images
LinW
GaussW
Median 3x3
Median 5x5
C1/P2
C2/P2
DRLSE
med3x3 - DRLSE

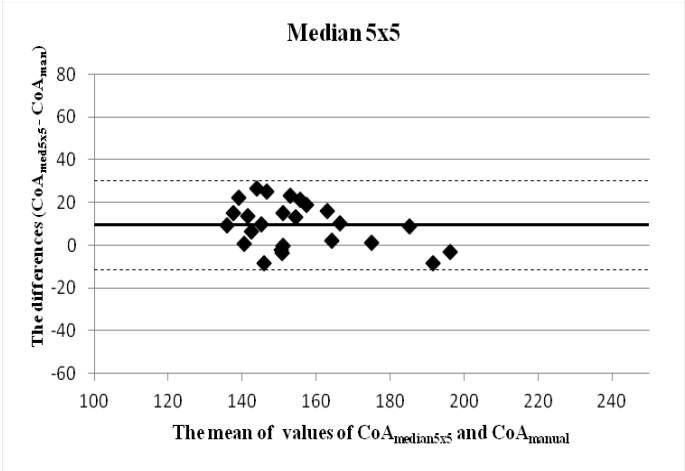

Diferences of Imin [%]
30.0
20.0
10.0
0.0
-10.0
-20.0
-30.0
-53.6 (6.0)
Raw images
LinW
GaussW
Median 3x3
Median 5x5
C1/P2
C2/P2
DRLSE
med3x3 - DRLSE

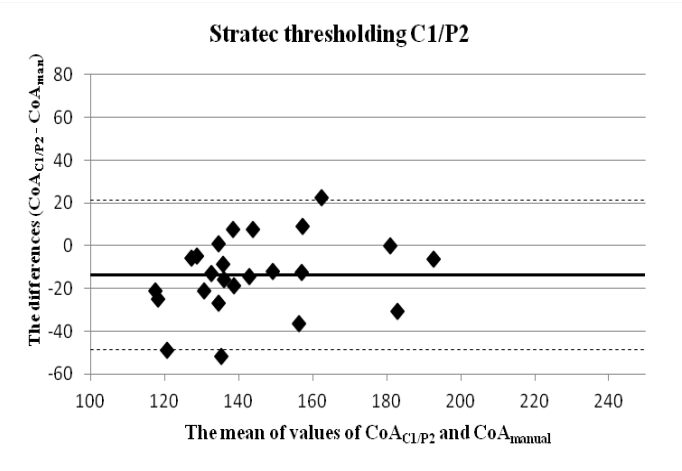

Raw images
y = -0.3687x + 34.188 (r = -0.61)
The differences (CoA_Raw - CoA_man)
The mean of values of CoA_Raw and CoA_manual

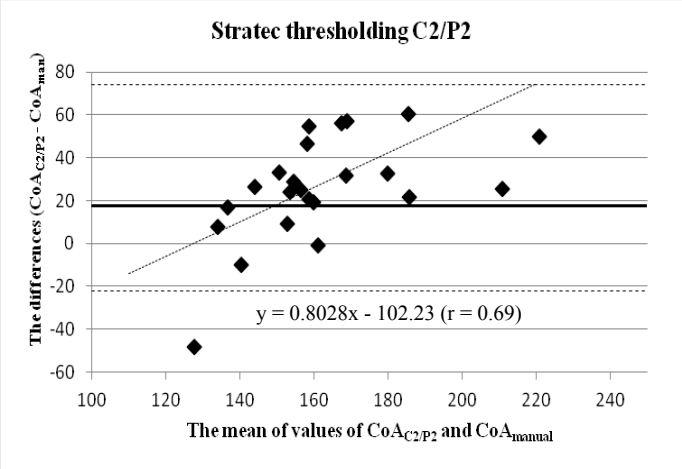

LinW
y = -0.2821x + 45.077 (r = -0.49)
The differences (CoA_LinW - CoA_man)
The mean of values of CoA_LinW and CoA_manual

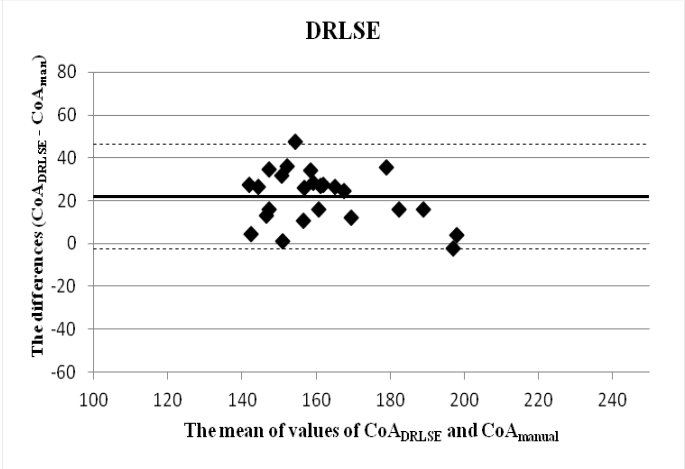

GaussW
y = -0.3433x + 42.521 (r = -0.53)
The differences (CoA_GaussW - CoA_man)
The mean of values of CoA_GaussW and CoA_manual

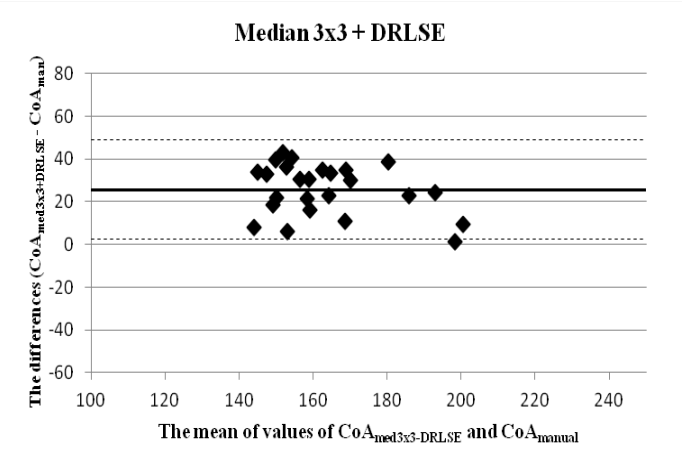

Median 3x3
The differences (CoA_med3x3 - CoA_man)
The mean of values of CoA_median3x3 and CoA_manual


Median 5x5
The differences (CoA_med5x5 - CoA_man)
The mean of values of CoA_median5x5 and CoA_manual


Stratec thresholding C1/P2
The differences (CoA_C1/P2 - CoA_man)
The mean of values of CoA_C1/P2 and CoA_manual


Stratec thresholding C2/P2
y = 0.8028x - 102.23 (r = 0.69)
The differences (CoA_C2/P2 - CoA_man)
The mean of values of CoA_C2/P2 and CoA_manual


DRLSE
The differences (CoA_DRLSE - CoA_man)
The mean of values of CoA_DRLSE and CoA_manual


Median 3x3 + DRLSE
The differences (CoA_med3x3+DRLSE - CoA_man)
The mean of values of CoA_med3x3-DRLSE and CoA_manual

Publication IV:

Cervinka T., Sievanen H., Hyttinen J., Rittweger J.
Bone loss patterns in cortical, subcortical and trabecular compartments during simulated microgravity.
Journal of Applied Physiology (1985), 2014; 117(1): 80-88.

J Appl Physiol 117: 80–88, 2014.
First published May 8, 2014; doi:10.1152/japplphysiol.00021.2014.

Bone loss patterns in cortical, subcortical, and trabecular compartments during simulated microgravity

Tomas Cervinka,[1,2] Harri Sievänen,[3] Jari Hyttinen,[1,2] and Jörn Rittweger[4,5]

[1]*Department of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland;* [2]*Institute of Bioscience and Medical Technology, Tampere, Finland;* [3]*Bone Research Group, UKK Institute, Tampere, Finland;* [4]*Institute of Aerospace Medicine, German Aerospace Center, Cologne, Germany; and* [5]*Institute for Biomedical Research into Human Movement and Health, Manchester Metropolitan University, Manchester, United Kingdom*

Submitted 13 January 2014; accepted in final form 5 May 2014

Cervinka T, Sievänen H, Hyttinen J, Rittweger J. Bone loss patterns in cortical, subcortical, and trabecular compartments during simulated microgravity. *J Appl Physiol* 117: 80–88, 2014. First published May 8, 2014; doi:10.1152/japplphysiol.00021.2014.—Disuse studies provide a useful model for bone adaptation. A direct comparison of these studies is, however, complicated by the different settings used for bone analysis. Through pooling and reanalysis of bone data from previous disuse studies, we determined bone loss and recovery in cortical, subcortical, and trabecular compartments and evaluated whether the study design modulated skeletal adaptation. Peripheral quantitative tomographic (pQCT) images from control groups of four disuse studies with a duration of 24, 35, 56, and 90 days were reanalyzed using a robust threshold-free segmentation algorithm. The pQCT data were available from 27 young healthy men at baseline, and at specified intervals over disuse and reambulation phases. The mean maximum absolute bone loss (mean ± 95% CI) was 6.1 ± 4.5 mg/mm in cortical, 2.4 ± 1.6 mg/mm in subcortical, and 9.8 ± 9.1 mg/mm in trabecular compartments, after 90 days of bed rest. The percentage changes in all bone compartments were, however, similar. During the first few weeks after onset of reambulation, the bone loss rate was systematically greater in the cortical than in the trabecular compartment ($P < 0.002$), and this was observed in all studies except for the longest study. We conclude that disuse-induced bone losses follow similar patterns irrespective of study design, and the largest mean absolute bone loss occurs in the cortical compartment, but apparently only during the first 60 days. With longer study duration, trabecular loss may become more prominent.

bed rest; immobilization; osteoporosis; pQCT; image processing

BONE LOSS, RESULTING IN REDUCED whole bone strength, is triggered by several factors such as aging, hormonal disturbances, disease, disuse, or microgravity during space flight. Therefore, the study of processes in bone tissue as responses to various conditions or treatments that could prevent bone deterioration is one of the main goals of clinical bone research.

Over the past two decades, bed rest studies have been used to investigate skeletal responses to reduced loading, disuse, or simulated microgravity (15). The first bed rest studies used planar dual-energy X-ray absorptiometry (DXA) imaging (1, 14, 24, 28, 30, 35, 36), but this method is known to be compromised by inherent inaccuracy due to the violation of the two-component assumption (homogeneous soft tissue and bone) and by the inability to yield reliable information on actual bone geometry and structure (6, 26). Recently, several studies have used peripheral quantitative computed tomography (pQCT) (4, 18, 21–23, 29, 33) or its high-resolution version (3). In contrast to DXA, pQCT provides a reliable assessment of bone cross-sectional geometry and allows separation into trabecular and cortical compartments (25).

The pQCT-based bed rest studies primarily used manufacturer-proposed thresholding procedures for bone analysis, and results have therefore not been reported in the same way. At the distal tibia site, some studies reported changes in total bone mineral content (BMC) (21) or cortical and trabecular BMC (22, 29) from baseline data collection until the onset of reambulation. Other studies reported changes in cortical and trabecular BMC (18, 23) or total and trabecular bone mineral density (BMD) (4, 33) at various time points during bed rest and subsequent reambulation. As a result, the direct between-study comparison of results is complicated. Simple threshold-based approaches can also result in a considerable variation in the cross-sectional areas of bone compartments (2, 11, 12) and further compromise bone analysis in a longitudinal setting. Therefore, we hypothesized that a robust analysis of the same subject-specific cortical, subcortical, and trabecular bone cross-sectional areas of all consecutive pQCT images throughout the entire longitudinal study would reveal compartment-specific changes in skeletal responses that have remained masked in conventional thresholding approaches. Moreover, the pooling of the largest set of available data obtained from four out of seven disuse studies performed with pQCT so far, and reanalyzing the image data in a similar robust way, is expected to provide more accurate insight into bone adaptation patterns with higher statistical power.

The studies from which the pQCT data were obtained were the 90-day Long Term Bed Rest (LTBR) study, carried out in Toulouse, France, in 2001 and 2002 (21); the 56-day first Berlin Bed Rest (BBR) study, carried out in Berlin, Germany, in 2003 and 2004 (18); the 24-day Unilateral Lower Limb Suspension (ULLS) study, carried out in Alsager, U.K., in 2005 (23); and the 35-day Valdoltra study, carried out in Ankaran, Slovenia, in 2007 (22).

The primary objective of the present study was to carry out a commensurate analysis of pQCT-measured cortical, subcortical, and trabecular bone mineral content (BMCc, BMCsub, BMCt, respectively) at the distal tibia by applying a robust, recently developed, threshold-free segmentation algorithm (8) on the same bone areas using data pooled from the above-mentioned four prospective disuse trials each with a different duration and design (18, 21–23). More specifically, we assessed *1*) whether this analysis could confirm the notion that the majority of bone loss occurs in the cortical compartment (1, 18, 20, 22) as opposed to the trabecular compartment, which

Address for reprint requests and other correspondence: T. Cervinka, Finnmedi 1, 4th Floor, Biokatu 6, 33520 Tampere, Finland (e-mail: tomas.cervinka@tut.fi).

Table 1. *Characteristics of the four disuse studies*

Study	Place	Year	Duration, days	Position	Legs immobilized	Countermeasures
LTBR	Toulouse, France	2001/2002	90	−6°	2	Flywheel exercise, pamidronate
BBR	Berlin, Germany	2003/2004	56	0°	2	Resistive vibration exercise
Valdoltra	Ankaran, Slovenia	2007	35	0°	2	
ULLS	Alsager, UK	2005	24	+90°	1	

LTBR, Long Term Bed Rest; BBR, Berlin Bed Rest; ULLS, Unilateral Lower Limb Suspension.

seems to be affected in patients with spinal cord injury or in astronauts during long space flights (27, 32); and *2*) whether the different study designs modulated the bone loss and recovery patterns.

MATERIALS AND METHODS

Study procedures. In this study, the pQCT data were reanalyzed from the control groups only (i.e., the subjects who underwent disuse only) of the LTBR study ($n = 9$), the first BBR study ($n = 10$), the Valdoltra study ($n = 10$), and the ULLS study ($n = 8$). The protocols of these four studies are described elsewhere in detail (18, 21–23) and summarized in Table 1. The ethical committee at each facility approved the study protocols, and each participant (all were male) in the study gave written informed consent before the study began (18, 21–23).

In short, LTBR study participants were randomly divided into three groups in which the influence of a one-time treatment by pamidronate prior to bed rest and a flying wheel exercise during bed rest, both used as counter measures to bone loss, was investigated and compared with results from the control group (strict bed rest only). BBR study participants were randomly divided into control and resistive vibration exercise groups. All Valdoltra study participants underwent strict bed rest only. This study was designed primarily to investigate the differences in changes in cortical and trabecular bone compartments and to determine whether cortical bone loss exceeded trabecular bone loss during immobilization. All ULLS study participants underwent a right leg suspension. This was achieved with a 7.5-cm soled shoe on the left foot, and participant stability was achieved with bilateral crutches. For the present study, the left leg data of control participants in the LTBR and BBR studies and the right leg data of all participants of the Valdoltra and ULLS studies were used. The anthropometric characteristics of these groups are shown in Table 2.

Bone measurements. In all four studies, pQCT (XCT 2000; Stratec Medizintechnik, Pforzheim, Germany) was performed in the same way according to a previously published imaging protocol (19). In short, pQCT images of the distal tibia of the left leg (LTBR study), the right leg (Valdoltra and ULLS studies), and both legs (BBR study) were obtained at the 4% distal site of the tibia. The pixel size of the pQCT image was 0.4 × 0.4 mm in the LTBR study and 0.5 × 0.5 mm in the BBR, Valdoltra, and ULLS studies. The translational scan speed was 30 mm/s, and the slice thickness was 2.5 mm in all studies.

In the LTBR study, measurements were obtained 14 and 7 days prior to the start of bed rest (BR−14 and BR−7 were considered as baseline measurements), and on days 28, 44, 57, 69, and 89 during the bed rest phase (BR+28, BR+44, BR+57, BR+69, and BR+89, respectively), and on days 14, 90, 180, and 360 after reambulation (R+14, R+90, R+180, and R+360, respectively). In the BBR study, measurements were obtained 3 days prior and 2 days after the start of the bed rest phase (BR−3 and BR+2 baseline measurements); on days 17, 31, 45, and 55 during the bed rest phase (BR+17, BR+31, BR+45, and BR+55, respectively); and on days 4, 14, 28, 90, 180, and 360 after reambulation (R+4, R+14, R+28, R+90, R+180, and R+360, respectively). In the Valdoltra study, measurements were obtained 2 days prior to bed rest and in the morning of the first day of bed rest phase (BR−2 and BR+1 baseline measurements), on the last day of bed rest phase (BR+35), and 14 days after reambulation (R+14). In the ULLS study, measurements were obtained 9 days and 1 day prior to unilateral suspension (ULS−9 and ULS−1 baseline measurements); on days 7, 14, and 21 during the unilateral suspension (ULS+7, ULS+14, and ULS+21, respectively); and on days 4, 9, 35, and 90 after reambulation (R+4, R+9, R+35, and R+90, respectively).

Data processing and analysis. For each participant, the raw data from the first baseline pQCT image was used to determine an individual reference template of the cortical area (CoA) that was applied to all follow-up images of the given participant to determine changes in cortical, subcortical, and trabecular bone mass within the same cross-sectional area through the entire disuse and reambulation periods. In short, an accurate threshold-free segmentation algorithm recently developed by Cervinka et al. (8) was used to determine the CoA of the baseline image. Next, the reference CoA template was created by adding a 1-pixel-thick layer periosteally and a 2-pixel-thick layer endocortically to the detected CoA to compensate for any small differences in bone alignment between the baseline image and follow-up images. Then, the reference CoA template was anatomically aligned with every follow-up image by employing a fast, rigid image registration implemented in Slicer 3D version 3.6 (available at http://www.slicer.org) (16). Thereafter, two peeling thresholds, 180 mg/cm^3 (outer border) and 480 mg/cm^3 (inner border), were applied to the CoA to exclude soft tissues and bone tissue pixels that apparently did not represent actual cortical bone. After this peeling procedure, the subcortical area (SubA) was defined as the immediate 4-pixel-wide region inside the CoA, and the trabecular area (TrA) was defined as the area inside the SubA. A flow chart of the image processing procedures is depicted in Fig. 1 and an example of BMD distribution within the detected CoA is shown in Fig. 2.

Table 2. *Anthropometric characteristics of the study subjects*

Study	Group	Number of subjects	Age, years	Height, cm	Weight, kg	BMI, kg/m^2
LTBR	Control	8	31.3 (3.3)^{U,V}	173.4 (3.2)^{B}	73.0 (3.8)^{b}	24.3 (0.9)
BBR	Control	10	33.4 (6.6)^{U,V}	184.5 (7.0)^{L}	79.9 (8.0)^{l}	23.4 (1.3)
Valdoltra	Control	10	23.9 (2.5)^{B,L,U}	179.7 (7.7)	77.3 (11.0)	23.9 (2.8)
ULLS	Control	8	19.0 (0.8)^{B,L,V}	179.3 (5.0)	72.0 (9.1)	22.4 (2.6)
P for variance		—	0.005	0.13	0.32	0.36
P for means		—	<0.001	<0.001	0.03	0.42

Columns for age, height, weight, and body mass index (BMI) are means and (variances). The bottom row contains *P* value differences across studies; the row above it contains *P* values for homogeneity of variance. Superscript letters *L*, *B*, *V*, and *U* denote differences from the LTBR, BBR, Valdoltra, and ULLS studies, respectively, with $P < 0.05$ for small letters, $P < 0.01$ for capital letters, and $P < 0.001$ for bold italic capital letters. See Table 1 for abbreviations.

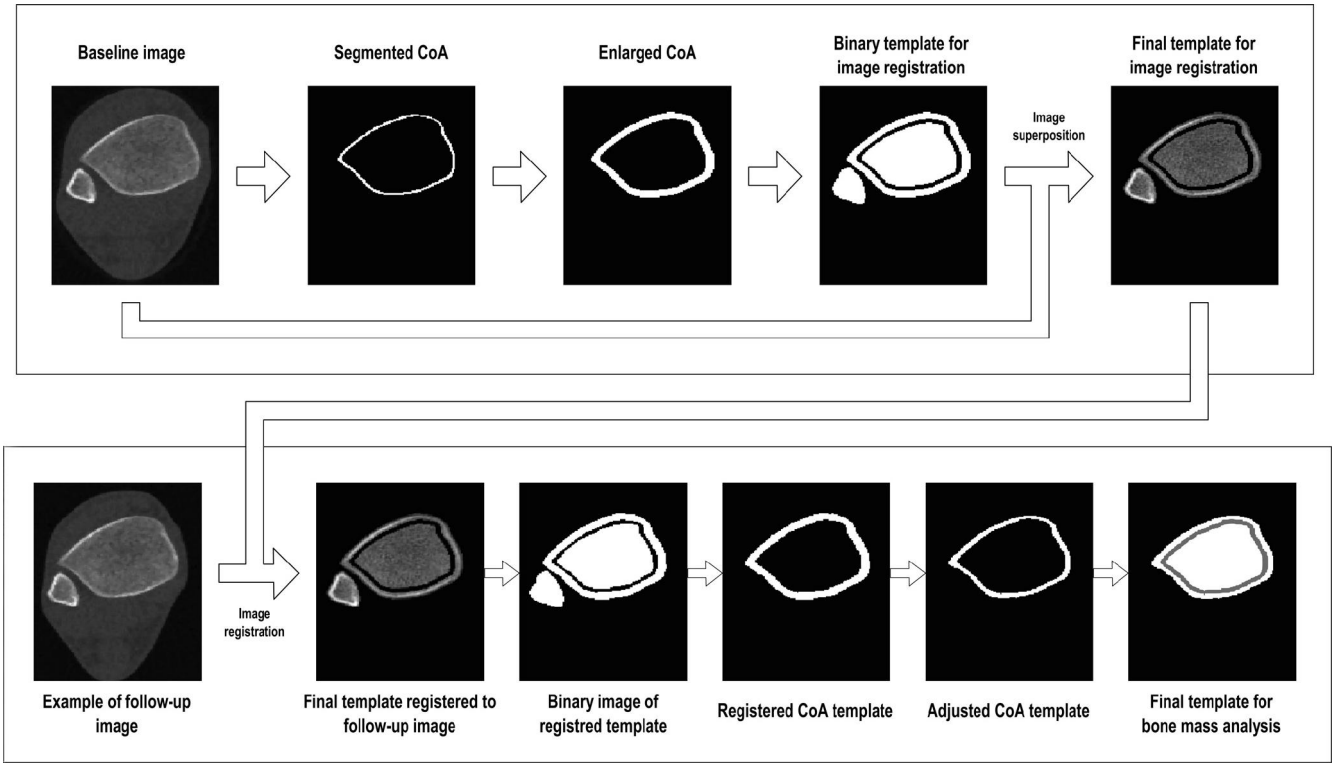


Fig. 1. Example of image processing procedure for cortical area (CoA), subcortical area (SubA), and trabecular area (TrA) determination. First, the CoA, in baseline image, was determined by analysis on the basis of a statistical approach and novel segmentation method developed previously by Cervinka et al. (7, 8). Second, the CoA template obtained from baseline data was created on the basis of segmented CoA enlarged by 2 pixels along the periosteal border and by 2 pixels along the endosteal border. This enlargement was carried out to compensate for small, expected variations in the total bone cross-sectional area that arose from positioning errors between the reference (the first baseline) image and all follow-up images. Third, the baseline CoA template was fitted to all follow-up images by using a fast, rigid registration algorithm implemented in Slicer 3D ver. 3.6. Fourth, registered CoA templates were adjusted using the superposition of the original data (processed by median filtering with a 3 × 3 window) into the templates. This adjustment was performed to compensate the previous enlargement procedure by removing pixels along the endosteal border that clearly contain nonbone tissues (peeling with threshold 180 mg/mm^3) and along the periosteal border contour that contains trabecular bone (peeling 1-pixel-thick contour with retaining pixels whose density was above threshold set to 480 mg/mm^3). The adjusted CoA templates were considered as CoAs for the following cortical bone mass analysis. Fifth, the SubA was defined as a 4-pixel-wide border zone of area enclosed by the CoAs and TrA, as the area enclosed by SubA.

Data sets from two participants in the LTBR study and one participant in the ULLS study had to be excluded because of a failed image registration process and an inability to perform image registration within acceptable limits (the cortical compartments of the follow-up data did not fit into the reference CoA template due to apparent bone misalignment). Because of similar image registration problems with data from the left leg in three BBR study participants and because longitudinal data sets from the right leg were available,

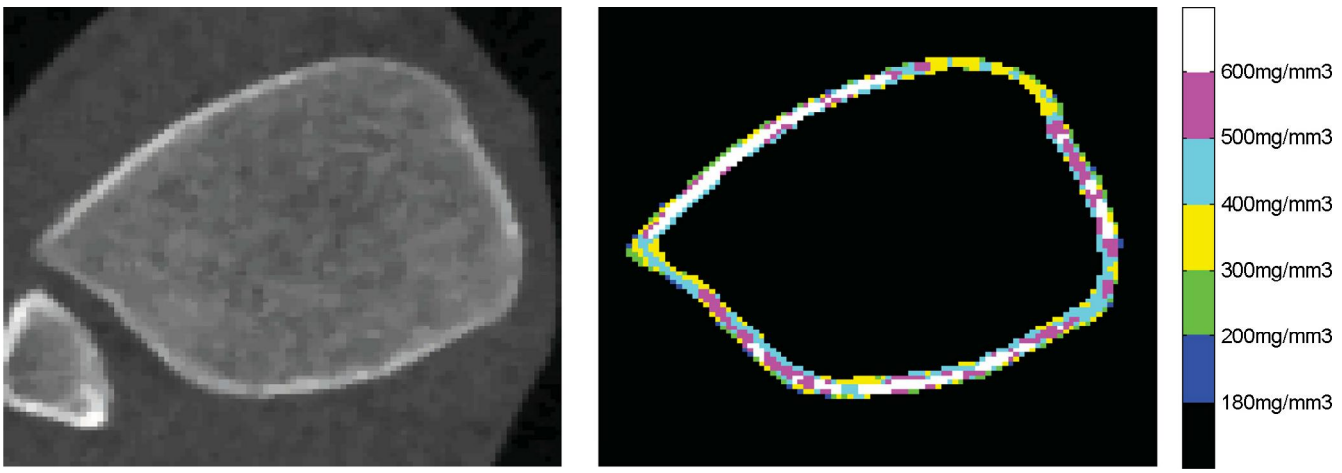


Fig. 2. Example of detected CoA: original peripheral quantitative computed tomographic image (*left*), and corresponding detected CoA with marked bone mineral density distribution (*right*).

Table 3. *Baseline bone mineral content values for cortical, subcortical, and trabecular compartments of all disuse studies*

Study	LTBR	BBR	Valdoltra	ULLS
BMCc	132.3 (20.3)	141.7 (23.8)	146.3 (29.2)	134.0 (18.4)
BMCsub	56.7 (6.7)	66.4 (8.1)	71.5 (8.9)	71.0 (9.1)
BMCt	221.0 (28.8)	205.9 (24.9)	218.8 (31.1)	226.1 (41.8)

Values are means and (SD) in mg/mm. BMCc, cortical bone mineral content; BMCsub, subcortical bone mineral content; BMCt, trabecular bone mineral content.

the three data sets from the left leg were replaced with right leg data. Furthermore, to avoid any inaccuracy caused by substantial bone misalignment during the follow-up, data from the follow-up images were also excluded when the CoA differed more than 10% from the baseline CoA (5 images from the LTBR study, 3 images from the BBR study, 3 images from the Valdoltra study, and 8 images from the ULLS study). The proportion of excluded images out of all the 348 images was 5.5%, which was considered satisfactory.

Because of the generally limited number of participants in the disuse studies and for the sake of retaining statistical power, we imputed values for the missing or excluded data points. The imputed values were estimated by linear interpolation using data from two neighboring time points. Note that imputation was not performed for the last follow-up image if it was missing (i.e., two images in the BBR study, and one image in the ULLS study).

Statistical analyses. Anthropometric data are given as means and standard deviation (SD). ANOVA was used to assess whether any significant differences existed across the studies in terms of age, height, weight, or body mass index (BMI). The level for statistical significance was set to $\alpha = 0.05$. Further, mean values and 95% confidence intervals (95% CI) of BMC are given as descriptive statistics, if not otherwise indicated. Given the small number of participants in the disuse studies, no formal statistical testing was performed, but the 95% CI was considered adequate to indicate within-group changes and to compare between-group changes during the disuse and subsequent reambulatory periods. Changes in BMC during the disuse and reambulation periods were expressed as mean absolute (in mg/mm) and relative (in %) changes from baseline (Table 3). The baseline value was specifically defined as the mean of the two baseline data points described above for each study (see section *Bone measurements*). These two data points also provided relevant information on the short-term error of BMC measurements and facilitated an appropriate interpretation of the changes in BMC (Table 4). Using the two baseline measurements, the short-term error (in %) was calculated as follows:

$$Er_{sh} = \frac{100}{N} \sum \frac{\left| \bar{x}_{i,BL} - x_{i,BL1} \right|}{\bar{x}_{i,BL}} \quad (1)$$

where $\bar{x}_{i,BL}$ is the mean baseline value of i-th subject, $\bar{x}_{i,BL1}$ is the value of the first baseline measurement of i-th subject, and N is the total number of participants.

A two-sample t-test with assumption of unequal variances was performed to assess whether the bone loss at the end of immobilization (EIBL), the maximal postreambulatory bone loss ($PRABL_{max}$, present at day R+14 for bed-rest studies and R+35 for the ULLS study) in relation to EIBL, as well as their rates [defined as the slope of the tangent in the above-mentioned time points (bone loss/day)] were higher in the cortical than in the trabecular compartment. Multiple-regression analysis was performed to assess the influence of the disuse duration and baseline BMC values on EIBL, PRABL, and their rates.

Furthermore, to characterize the time courses of the mean absolute changes in BMC during disuse, we used simple quadratic and linear models and an exponential model proposed by Fyhrie and Schaffler (10). The formulation of these models is the following:

$$\text{quadratic model}: \quad BMC_l(t) = at^2 + bt \quad (2)$$

$$\text{linear model}: \quad BMC_l(t) = bt + c \quad (3)$$

$$\text{exponential model}: \quad BMC_l(t) = 0.7BMC_{baseline}(1 - e^{-bt}) \quad (4)$$

where BMC_l is the absolute change from the baseline at time t, a and b are time constants describing the rate of change in BMC, and c is an intercept coefficient of the linear model. It was assumed that the biological minimum is 30% of the baseline BMC (10). Therefore, $0.7BMC_{baseline}$ represents the theoretical maximal bone loss. Of note, pooled data from all three bed rest studies was used for modeling (the ULLS study was excluded because of its different study design), with the exception of the baseline value in the linear model. The goodness of fit of different models in explaining changes in BMC was compared by an F-test.

The time course of recovery in BMC during reambulation from minimum BMC values was analyzed with the modified exponential model for the LTBR and BBR studies only because the Valdoltra study did not include data after 14 days of reambulation. In *Equation 4*, the $0.7BMC_{baseline}$ value was replaced by the mean $PRABL_{max}$ of the given study. The time course of the reambulatory changes in BMC was analyzed only for the LTBR and BBR studies because the data sets from these studies comprised sufficiently long 1-yr follow-up data. Because these studies had different durations of disuse, the follow-up data had to be analyzed separately. The level of significance was set at 0.05 for all the above-mentioned tests.

RESULTS

Anthropometric data. Table 2 displays participants' descriptive data in different disuse studies. Participants in the ULLS study were the youngest, followed by the Valdoltra participants, and the participants from the LTBR and BBR studies. Height was also different across the studies ($P < 0.001$), with BBR subjects being taller than LTBR subjects ($P = 0.003$). BBR subjects were also heavier than LTBR subjects ($P = 0.032$), but BMI was comparable across the studies ($P = 0.42$).

Bone loss patterns during disuse. Absolute and percentage changes in BMCc, BMCsub, and BMCt at the distal tibia during disuse are shown in Fig. 3 and Fig. 4, respectively. For bed rest studies, mean BMC BMC changes across the disuse period were generally consistent, with the exception of the non–bed rest ULLS study. The mean absolute bone loss was systematically greater in BMCt than in BMCc or BMCsub, but also showed much more variation during disuse (Fig. 3 and Table 5). The mean absolute EIBL rates indicated a trend toward accelerated trabecular bone loss after ~60 days of disuse (of note, according to the LTBR data only). No such trend was observed in the cortical compartment (see Fig. 3 and Table 5). In contrast, the mean relative EIBLs were similar in all bone compartments (Table 5).

The multiple-regression analysis confirmed a positive relationship between EIBL and duration of disuse for both the cortical compartment in absolute ($R^2 = 0.42$; $P = 0.024$) and relative ($R^2 = 0.40$; $P = 0.034$) values, and the trabecular compartment in absolute ($R^2 = 0.38$; $P = 0.045$) and relative

Table 4. *Short-term error*

Study	BMCc	BMCsub	BMCt
LTBR	0.7%	1.3%	0.4%
BBR	0.8%	0.9%	0.6%
Valdoltra	1.5%	1.1%	0.7%
ULLS	1.8%	0.9%	0.8%

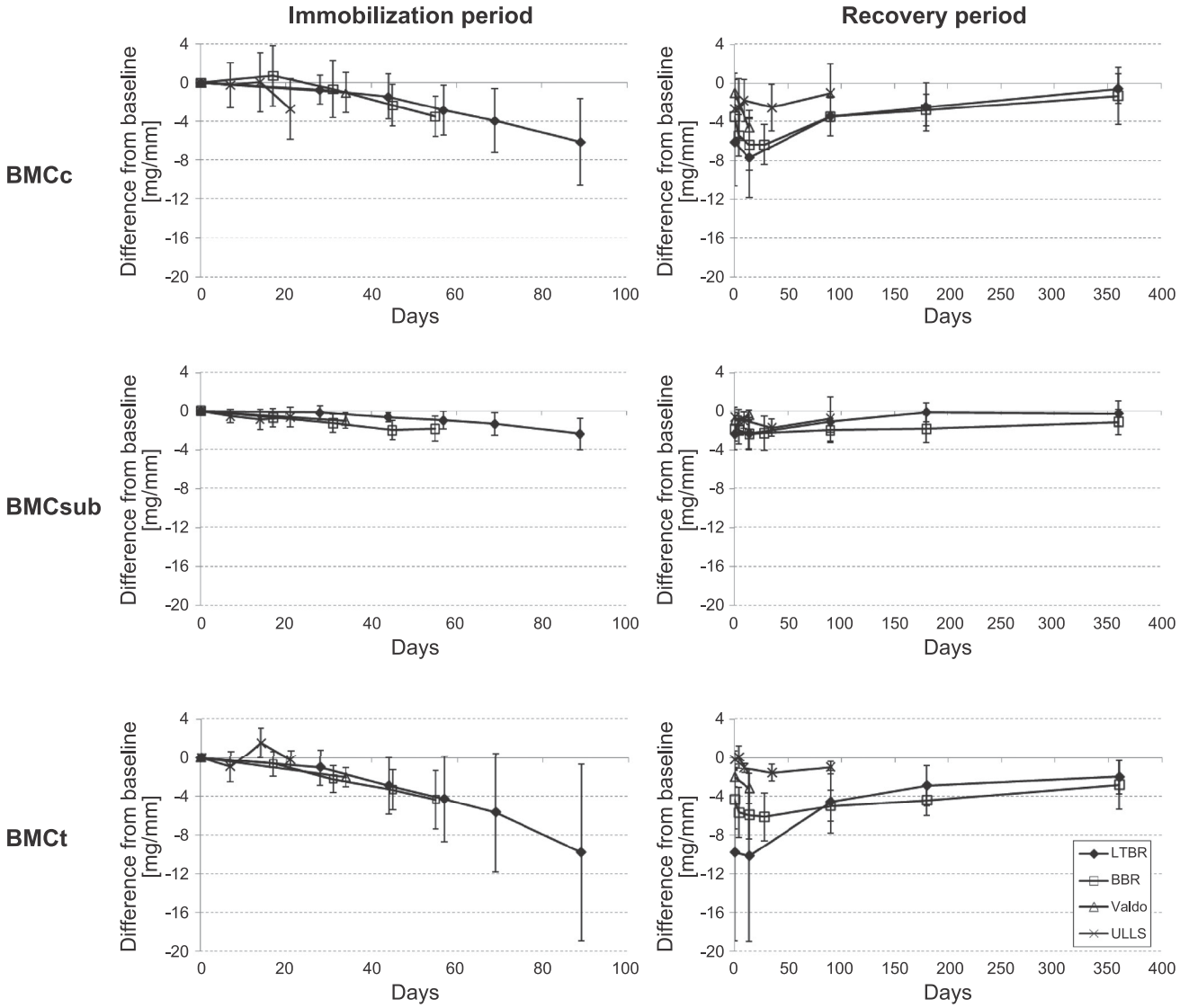


Fig. 3. Mean absolute bone losses (95% CI) during disuse and reambulatory periods for the Long Term Bed Rest (LTBR), Berlin Bed Rest (BBR), Valdoltra, and Unilateral Lower Limb Suspension (ULLS) studies. From *top* to *bottom*: cortical, subcortical, and trabecular compartments. The baseline bone mineral content values for each compartment and studies are shown in Table 3. BMCc, cortical bone mineral content; BMCsub, subcortical bone mineral content; BMCt, trabecular bone mineral content.

(R^2 = 0.37; P = 0.046) values. No association, however, was found between EIBL and baseline BMC values ($P > 0.4$).

The goodness of fit of the exponential model (*Equation 4*) during the disuse period, given in R^2 values, was 0.80, 0.61, and 0.87 (b values were −0.00054, −0.00056, and −0.00055) for cortical, subcortical, and trabecular compartments, respectively. The simple quadratic and linear models (*Equation 2* and *Equation 3*) did not improve prediction of bone loss in the subcortical compartment (R^2 = 0.62 and 0.61; parameters, a = 0.00006, b = 0.0293, and b = −0.0258, and c = 0.6112; $P >$ 0.12), but these models were better descriptors of bone loss in the cortical (R^2 = 0.94, a = −0.0007, b = −0.0071; R^2 = 0.97, b = −0.092, c = 2.14) and trabecular compartments (R^2 = 0.99, a = −0.0009, b = −0.0264; R^2 = 0.95, b = −0.12, c = 2.11), and clearly outperformed, within the given range of data, the exponential model ($P < 0.005$). The linear model better described changes in the cortical compartment (P = 0.003), whereas the quadratic model better described changes in the trabecular compartment (P = 0.004).

Bone recovery patterns during reambulation. Absolute and relative recovery of BMCc, BMCsub, and BMCt at the distal tibia during the reambulation period is shown in Figs. 3 and 4, respectively. In general, mean BMC in all bone compartments returned virtually to baseline during reambulation, excluding the consistent ~2 mg/mm trabecular loss in the longest LTBR and BBR studies. It is notable that the bone loss seemed to continue during the first approximately 15–30 days after termination of disuse in all bed rest studies. At the onset of reambulation, a systematically greater PRABL rate was observed in BMCc than in BMCt (P = 0.002 in terms of absolute changes, and $P < 0.0001$ in terms of relative changes), the non–bed rest ULLS study excluded, whereas the PRABL rate was marginal in BMCsub. Of interest, as shown in Table 5, the highest PRABL rate in the cortical compartment was in the shortest

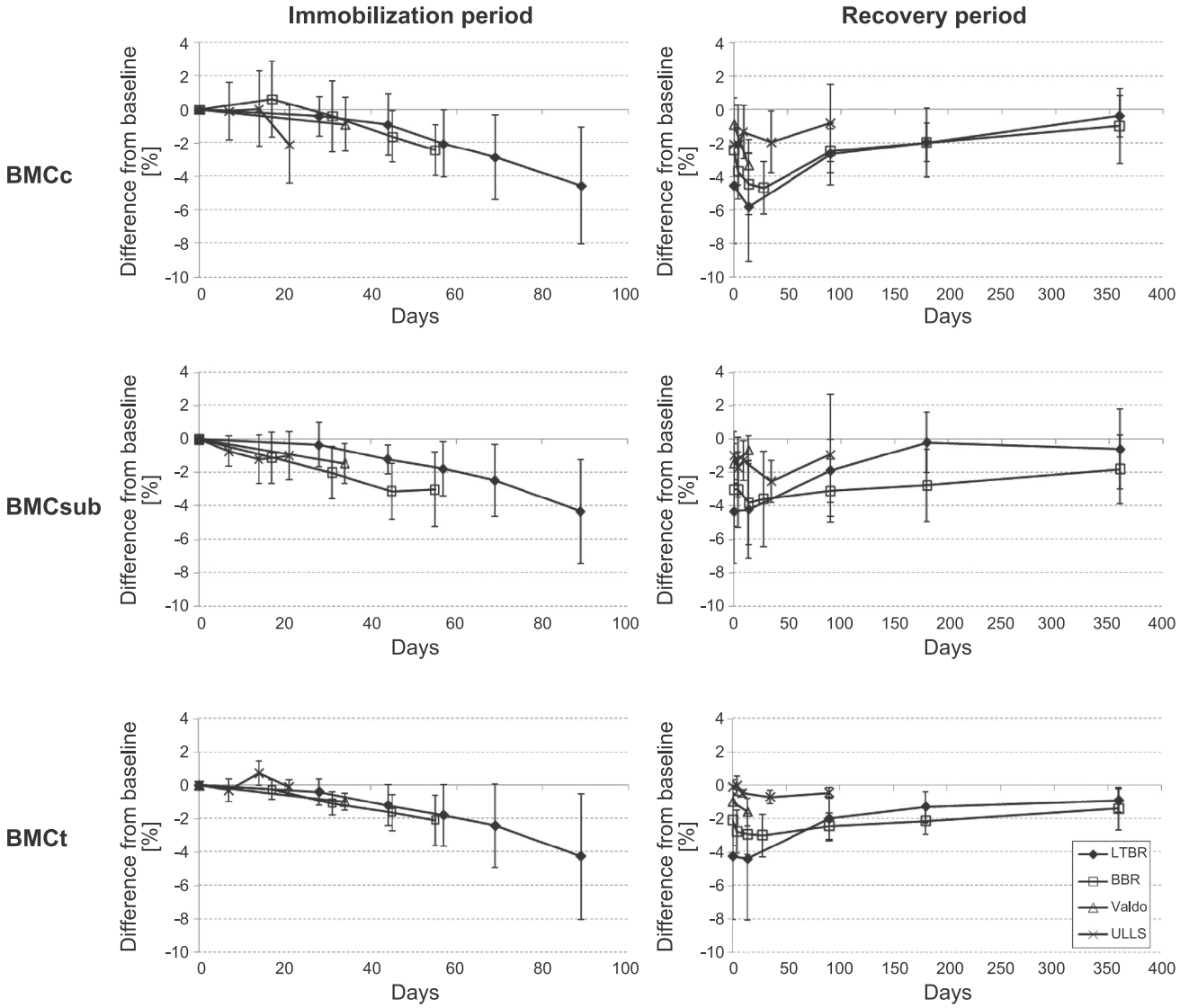


Fig. 4. Mean relative bone losses (95% CI) during disuse and reambulatory periods for the LTBR, BBR, Valdoltra, and ULLS studies. From *top* to *bottom*: cortical, subcortical, and trabecular compartments. The baseline bone mineral content values for each compartment and studies are shown in Table 3.

Valdoltra study, and was significantly higher than in the longest LTBR study ($P = 0.034$). Differences in PRABL rates in the trabecular compartment between these studies were not significant ($P > 0.29$). The mean absolute $PRABL_{max}$ seemed to be larger in the cortical than in the trabecular compartment for all studies except for the longest LTBR study (Table 5); however, none of the absolute differences between the cortical and trabecular compartments reached statistical significance ($P = 0.46$). In contrast, the relative mean $PRABL_{max}$ was systematically greater ($P = 0.048$) in the cortical compartment than in the trabecular compartment.

The multiple-regression analysis gave support to the view that the PRABL rate in the cortical compartment would be negatively related to the duration of disuse ($P < 0.05$) and EIBL rate ($P < 0.0007$), and positively related to $PRABL_{max}$ ($P < 0.028$) both in absolute ($R^2 = 0.70$; $P = 0.0006$) and relative ($R^2 = 0.67$; $P = 0.0016$) terms. Similarly for the trabecular PRABL rate, $PRABL_{max}$ showed positive ($P < 0.03$) and the EIBL rate showed negative ($P < 0.003$) association. These factors appeared to be good descriptors of the PRABL rate both in absolute ($R^2 = 0.59$; $P = 0.013$) and relative ($R^2 = 0.60$; $P = 0.01$) changes in trabecular BMC, whereas the duration of disuse showed no association ($P > 0.25$).

Bone gains during recovery were well explained by the exponential model (*Equation 4*) in the cortical ($R^2 = 0.93$, $b = 0.0047$; $R^2 = 0.96$, $b = 0.0073$), subcortical ($R^2 = 0.97$, $b = 0.0019$; $R^2 = 0.92$, $b = 0.0091$), and trabecular ($R^2 = 0.96$, $b = 0.0019$; $R^2 = 0.93$, $b = 0.0068$) compartments, respectively.

DISCUSSION

The main purpose of the present work was to reanalyze the pQCT data pooled from several disuse studies with a robust method to obtain a larger study cohort to investigate whether bone loss is different in cortical and trabecular bone compart-

Table 5. *Mean absolute and relative values, with 95% CI, of maximal bone losses and bone loss rates induced during the end of immobilization and after the reambulation period*

	Compartment	Study	Duration	EIBL	$PRABL_{max}$	EIBL-rate	PRABL-rate
Absolute values	Cortical	LTBR	90	−6.1 ± 4.5	−7.7 ± 4.1	−0.11 ± 0.08	−0.11 ± 0.08
		BBR	56	−3.5 ± 2.1	−6.3 ± 2.7	−0.12 ± 0.18	−0.19 ± 0.10
		Valdoltra	35	−1.0 ± 2.1	−4.6 ± 1.7	−0.03 ± 0.06	−0.24 ± 0.10
		ULLS	24	−2.7 ± 3.1	−2.5 ± 2.4	−0.39 ± 0.34	0.005 ± 0.04
	Trabecular	LTBR	90	−9.8 ± 9.1	−10.1 ± 8.9	−0.21 ± 0.16	−0.02 ± 0.06
		BBR	56	−4.4 ± 3.0	−6.0 ± 2.4	−0.10 ± 0.15	−0.11 ± 0.08
		Valdoltra	35	−2.0 ± 1.0	−3.2 ± 1.6	−0.06 ± 0.03	−0.08 ± 0.07
		ULLS	24	−0.2 ± 1.0	−1.6 ± 0.9	−0.25 ± 0.20	−0.03 ± 0.03
Relative values	Cortical	LTBR	90	−4.6 ± 3.5	−5.8 ± 3.2	−0.08 ± 0.06	−0.08 ± 0.07
		BBR	56	−2.4 ± 1.5	−4.4 ± 1.8	−0.08 ± 0.14	−0.13 ± 0.07
		Valdoltra	35	−0.9 ± 1.6	−3.3 ± 1.6	−0.03 ± 0.05	−0.16 ± 0.06
		ULLS	24	−2.1 ± 2.3	−1.9 ± 1.8	−0.12 ± 0.09	−0.01 ± 0.01
	Trabecular	LTBR	90	−4.3 ± 3.7	−4.4 ± 3.6	−0.09 ± 0.07	−0.01 ± 0.03
		BBR	56	−2.1 ± 1.5	−2.9 ± 1.2	−0.05 ± 0.07	−0.05 ± 0.04
		Valdoltra	35	−1.0 ± 0.5	−1.6 ± 0.8	−0.03 ± 0.02	−0.04 ± 0.04
		ULLS	24	−0.1 ± 0.5	−0.7 ± 0.4	−0.12 ± 0.09	−0.01 ± 0.01

Maximal bone losses are shown as mg/mm and %; bone loss rates are shown as $mg \cdot mm^{-1} \cdot day^{-1}$ and %/day. EIBL, bone loss at the end of immobilization in relation to baseline; EIBL rate, the rate of bone loss at the end of immobilization period (bone loss/day); $PRABL_{max}$, maximal postreambulatory bone loss in relation to EIBL; PRABL rate, postreambulatory bone loss rate (bone loss/day).

ments, as previous small studies had separately indicated (13, 21, 22, 32), and whether the bone responses to disuse and subsequent reambulation were modulated by the different designs of the studies.

To achieve this, the present work uniquely focused on consistent image processing that permitted analysis of the same bone cross-sectional areas throughout the entire follow-up periods so that changes in bone compartments were monitored and analyzed in a more comparable fashion. Our results indicate that during the disuse period, mean absolute bone losses were most pronounced in the trabecular compartment (Fig. 3). Of interest, mean absolute bone losses in the cortical compartment increased during a short period (2–3 wk) after reambulation, the longest LTBR study excluded. During this immediate postambulation period, the total absolute cortical bone loss in disuse studies shorter than 60 days (about 2 mo) exceeded the absolute trabecular bone losses induced during the preceding disuse period. This is in line with recent findings from immobilization studies (1, 18, 20, 22). In the longer 90-day LTBR study, however, the total mean absolute cortical bone loss did not exceed the total mean trabecular bone loss induced during disuse. This was despite faster cortical bone deterioration immediately after reambulation, and mean trabecular bone loss remained greater than cortical bone loss.

Study design. The specific nature of the disuse design and the body position in particular [0 degrees in the Valdoltra study (22) and in the BBR study (1), and −6 degrees head tilted down in the LTBR study (21)] seemed not to substantially modulate bone loss. The mean bone loss in all bone compartments in all bed rest studies were concordant during the disuse period and showed similar patterns during the reambulation period as well. Clearly, the magnitude of disuse-induced mean bone loss is related to the total duration of intervention (4, 18, 20–22, 27), as was also confirmed by the present regression analysis. It should be noted that interindividual variability seemed to increase with disuse duration.

Apparently, the seminal LTBR study had a direct impact on the design of subsequent bed rest studies. The shorter 35-day Valdoltra study was performed to corroborate the findings of the LTBR study (21) and findings from a long-duration space flight (13) suggesting that the majority of bone loss occurred primarily in the cortical compartment. Thus the duration of successive bed rest studies was limited to a maximum of 60 days. This was considered to be a sufficient duration for the prediction of future bone loss (18). The ULLS study, in turn, was considered a feasible alternative to more complicated and expensive bed rest study designs that might cause not only extensive physical effects but also psychological effects (23). However, the ULLS study, with a completely different mode of disuse and younger participants, showed different patterns in bone responses to disuse compared with bed rest studies (Figs. 3 and 4).

Specifically, the mean BMCt in the ULLS study showed an unexpected increase during the first days of disuse. This increase could be related to an increase in venous pressure that partially could have prevented the trabecular bone loss, as suggested by a rat study by Bergula et al. (5). However, a similar increase was also reported in a trabecular BMD in a short, 21-day, −6 degrees head-tilted-down bed rest study (29), but not in the bed rest studies assessed similarly in the present work. That the majority of bone loss in the ULLS study was predominantly cortical is concordant with findings from the bed rest studies (1, 18, 22). On the other hand, the short duration (24 days) and poor quality of image data compromised this analysis. Longer limb suspension studies are needed to more appropriately assess whether the bone loss patterns would agree with the bed rest studies or have a unique pattern. At present, it seems that unilateral suspension is not a feasible model for bed rest studies.

Patterns of disuse induced bone losses. In the assessed bed rest studies, trabecular bone loss showed a short delay (~17 days) at the onset of disuse. This delay is in line with a recent finding from the bed rest study of Wang et al. (33), in which they reported only marginal trabecular bone loss at the distal tibia during the first 30 days of disuse. After this initial quiescent period, trabecular bone loss increased rather linearly until the ~60th day of disuse. With longer disuse, mean trabecular bone loss seemed to gradually increase and showed

more substantial loss than occurred in the cortical bone compartment. However, the number of participants who underwent long disuse (duration >60 days) remained small because only one bed rest study lasted that long. In general, substantial trabecular bone loss is a well known consequence of bone disuse (27) or a long-duration space flight mission (32). Nevertheless, previous findings from bed rest studies have suggested that the most pronounced bone loss would be in the cortical compartment (3, 18, 21, 22). It may be that bone losses from the cortical and trabecular compartments have different time courses such that cortical loss starts almost immediately and is mostly linear over time, whereas trabecular loss is initially delayed but then becomes aggravated with longer immobilization. The ultimate question is which factors would underlie the apparently different dynamics in response to an altered loading environment?

Patterns of postreambulation bone loss. In agreement with the previously reported disuse studies (1, 4, 18, 21–23), the current analysis showed that bone deterioration continues during the first days after reambulation and culminates between the 15th and 30th days after reambulation. Furthermore, bone loss in all compartments is almost completely restored (99% on average) within 1 yr after termination of disuse. The immediate postreambulation cortical losses excluded, no study-specific modulations were observed during the later reambulation period. However, the greater postreambulatory bone-loss rate within the cortical compartment compared with the trabecular compartment seems to be related to maximal postreambulatory bone loss and negatively to the rate of bone loss at the end of immobilization. This relation may indicate a hysteresis or lag in skeletal adaption by which the hysteresis is attributable to the delays in the differentiation of the osteoblasts while the osteoclasts remain active. Nevertheless, further studies are needed to better understand the cellular-level processes behind cortical bone loss, because only a small part of the variation could be explained.

Strengths and limitations. The strengths of the present analysis are *1*) the evaluation of pooled data from four out of the seven so-far-published disuse interventions performed with pQCT, and *2*) completely harmonized bone analysis by applying the new threshold-free image processing and segmentation method known to enhance pQCT image analysis (7, 9, 17). The method allowed accurate detection of the cortical bone without the use of simple thresholds that can substantially modulate pQCT results (12, 34). A direct comparison of disuse studies was not previously possible because each study used different threshold settings for bone analysis (18, 21–23).

Some limitations warrant further discussion. Basically, even the use of an advanced image processing technique cannot overcome inherent limitations caused by the coarse spatial resolution of pQCT images, as resolution limits the performance of any image processing, the threshold-based analyses in particular (31). In addition, despite having the largest ever pooled and carefully analyzed data sets from several disuse studies, the total sample size still remained small. This is even more pronounced at later time points of the disuse or recovery period when data from only a limited number of subjects and studies were available. The small sample size limited the statistical power of the analysis to detect potential, subtle differences in time-specific skeletal responses to disuse. This might have affected the bone loss models. Thus the possibility that the linear and quadratic models would not accurately represent the actual adaptation processes in bones at the later time points cannot be entirely ruled out. Furthermore, when considering the significance of the magnitude of bone loss, one must take into account the short-term error of this analysis approach (Table 4) and precision of pQCT-measured BMC values at the distal tibia [<1.5% (25)]. Therefore, given the relatively small differences observed in BMC values and the sensitivity of pQCT, the presented results should be interpreted with caution.

Conclusion. Keeping inherent limitations of this analysis in mind, the present pooled analysis provided new and profound insights into patterns of compartment-specific bone losses during disuse and subsequent recovery. First, the specific design of the bed rest study seemed not to substantially influence bone loss and recovery patterns. Second, the largest mean absolute bone losses seemed to take place in the cortical compartment, but apparently only during the first 2 mo (<60 days). Third, as indicated by the longer LTBR study, continued disuse would result in the largest mean absolute bone loss in the trabecular compartment. Obviously, more long-duration studies are needed to confirm this. Because of the short duration and young subject ages of the ULLS study, the present analysis could not establish whether the unilateral suspension could be considered a feasible model for expensive and challenging bed rest study designs. Keeping the inherent statistical limitations of small disuse studies in mind, the present analysis showed that the use of robust image processing applied to the same subject-specific bone cross-sectional areas throughout the longitudinal studies enhanced bone analysis and revealed specific skeletal responses that would have remained concealed in conventional analysis. In addition, our analysis approach, if applied to future longitudinal studies, would allow direct interstudy comparisons. Obviously, more studies are needed to prove this concept.

ACKNOWLEDGMENTS

We thank all personnel of the participating institutions and coworkers for their advice and assistance during the disuse studies. We are grateful to the study participants. This work would not have been possible without their selfless contribution.

GRANTS

Support for this study was provided by the Finnish Cultural Foundation, the Finnish Funding Agency for Technology and Inovations (Tekes), Project 718/31/2011, and the Graduate School of Tampere University of Technology to T. Cervinka. Support was also provided by Competitive Research Funding of Tampere University Hospital Grant 9K121. Support for our participation in the LTBR study was supported by Institute Aerospace Medicine (DLR in Hamburg) Grant 50-WB 0156. The ULLS study was funded by European Space Agency MESM2 Project Grant 15097/01/NL/SH-CCN3. This study was also funded in part by the Osteoporosis and Muscle Atrophy Group of the Italian Space Agency, the Slovene Ministry of Defence, and the Gösta Fraenckel Foundation for Medical Research.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: T.C., H.S., and J.R. conception and design of research; T.C. analyzed data; T.C. prepared figures; T.C. drafted manuscript; T.C., H.S., J.H., and J.R. edited and revised manuscript; T.C., H.S., J.H., and J.R. approved final version of manuscript; H.S., J.H., and J.R. interpreted results of experiments; J.R. performed experiments.

REFERENCES

1. **Armbrecht G, Belavý DL, Gast U, Bongrazio M, Touby F, Beller G, Roth HJ, Perschel FH, Rittweger J, Felsenberg D.** Resistive vibration exercise attenuates bone and muscle atrophy in 56 days of bed rest: biochemical markers of bone metabolism. *Osteoporos Int* 21: 597–607, 2010.
2. **Ashe MC, Khan KM, Kontulainen SA, Guy P, Liu D, Beck TJ, McKay HA.** Accuracy of pQCT for evaluating the aged human radius: an ashing, histomorphometry and failure load investigation. *Osteoporos Int* 17: 1241–1251, 2006.
3. **Belavý DL, Beller G, Armbrecht G, Perschel FH, Fitzner R, Bock O, Börst H, Degner C, Gast U, Felsenberg D.** Evidence for an additional effect of whole-body vibration above resistive exercise alone in preventing bone loss during prolonged bed rest. *Osteoporos Int* 22: 1581–1591, 2011.
4. **Beller G, Belavý DL, Sun L, Armbrecht G, Alexandre C, Felsenberg D.** WISE-2005: bed-rest induced changes in bone mineral density in women during 60 days simulated microgravity. *Bone* 49: 858–866, 2011.
5. **Bergula AP, Huang W, Frangos JA.** Femoral vein ligation increases bone mass in the hindlimb suspended rat. *Bone* 24: 171–177, 1999.
6. **Bolotin HH, Sievanen H.** Inaccuracies inherent in dual-energy X-ray absorptiometry in in vivo bone mineral density can seriously mislead diagnostic/prognostic interpretations of patient-specific fragility. *J Bone Miner Res* 16: 799–805, 2001.
7. **Cervinka T, Hyttinen J, Sievänen H.** Enhanced bone structural analysis through pQCT image preprocessing. *Med Eng Phys* 32: 398–406, 2010.
8. **Cervinka T, Hyttinen J, Sievänen H.** Threshold-free automatic detection of cortical bone geometry by peripheral quantitative computed tomography. *J Clin Densitom* 15: 413–421, 2012.
9. **Cervinka T, Rittweger J, Hyttinen J, Felsenberg D, Sievänen H.** Anatomical sector analysis of load-bearing tibial bone structure during 90-day bed rest and 1-year recovery. *Clin Physiol Funct Imaging* 31: 249–257, 2011.
10. **Fyhrie DP, Schaffler MB.** The adaptation of bone apparent density to applied load. *J Biomech* 28: 135–146, 1995.
11. **Hangartner TN, Short DF.** Accurate quantification of width and density of bone structures by computed tomography. *Med Phys* 34: 3777–3784, 2007.
12. **Kontulainen S, Liu D, Manske S, Jamieson M, Sievänen H, McKay H.** Analyzing cortical bone cross-sectional geometry by peripheral QCT: comparison with bone histomorphometry. *J Clin Densitom* 10: 86–92, 2007.
13. **Lang TF, Leblanc AD, Evans HJ, Lu Y.** Adaptation of the proximal femur to skeletal reloading after long-duration spaceflight. *J Bone Miner Res* 21: 1224–1230, 2006.
14. **LeBlanc AD, Driscol TB, Shackelford LC, Evans HJ, Rianon NJ, Smith SM, Feeback DL, Lai D.** Alendronate as an effective countermeasure to disuse induced bone loss. *J Musculoskelet Neuronal Interact* 2: 335–343, 2002.
15. **Pavy-Le Traon A, Heer M, Narici MV, Rittweger J, Vernikos J.** From space to Earth: advances in human physiology from 20 years of bed rest studies (1986–2006). *Eur J Appl Physiol* 101: 143–194, 2007.
16. **Pieper S, Lorensen W, Schroeder W, Kikinis R.** The NA-MIC kit: ITK, VTK, pipelines, grids and 3D slicer as an open platform for the medical image computing community. *Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 2006, pp. 698–701.
17. **Rantalainen T, Nikander R, Heinonen A, Cervinka T, Sievänen H, Daly RM.** Differential effects of exercise on tibial shaft marrow density in young female athletes. *J Clin Endocrinol Metab* 98: 2037–2044, 2013.
18. **Rittweger J, Beller G, Armbrecht G, Mulder E, Buehring B, Gast U, Dimeo F, Schubert H, de Haan A, Stegeman DF, Schiessl H, Felsenberg D.** Prevention of bone loss during 56 days of strict bed rest by side-alternating resistive vibration exercise. *Bone* 46: 137–147, 2010.
19. **Rittweger J, Beller G, Ehrig J, Jung C, Koch U, Ramolla J, Schmidt F, Newitt D, Majumdar S, Schiessl H, Felsenberg D.** Bone-muscle strength indices for the human lower leg. *Bone* 27: 319–326, 2000.
20. **Rittweger J, Felsenberg D.** Recovery of muscle atrophy and bone loss from 90 days bed rest: results from a one-year follow-up. *Bone* 44: 214–224, 2009.
21. **Rittweger J, Frost HM, Schiessl H, Ohshima H, Alkner B, Tesch P, Felsenberg D.** Muscle atrophy and bone loss after 90 days' bed rest and the effects of flywheel resistive exercise and pamidronate: results from the LTBR study. *Bone* 36: 1019–1029, 2005.
22. **Rittweger J, Simunic B, Bilancio G, De Santo NG, Cirillo M, Biolo G, Pisot R, Eiken O, Mekjavic IB, Narici M.** Bone loss in the lower leg during 35 days of bed rest is predominantly from the cortical compartment. *Bone* 44: 612–618, 2009.
23. **Rittweger J, Winwood K, Seynnes O, de Boer M, Wilks D, Lea R, Rennie M, Narici M.** Bone loss from the human distal tibia epiphysis during 24 days of unilateral lower limb suspension. *J Physiol* 577, Pt 1: 331337, 2006.
24. **Shackelford LC, LeBlanc AD, Driscoll TB, Evans HJ, Rianon NJ, Smith SM, Spector E, Feeback DL, Lai D.** Resistance exercise as a countermeasure to disuse-induced bone loss. *J Appl Physiol* 97: 119–129, 2004.
25. **Sievänen H, Koskue V, Rauhio A, Kannus P, Heinonen A, Vuori I.** Peripheral quantitative computed tomography in human long bones: evaluation of in vitro and in vivo precision. *J Bone Miner Res* 13: 871–882, 1998.
26. **Sievänen H.** A physical model for dual-energy X-ray absorptiometry-derived bone mineral density. *Invest Radiol* 35: 325–330, 2000.
27. **Sievänen H.** Immobilization and bone structure in humans. *Arch Biochem Biophys* 503: 146–152, 2010.
28. **Smith SM, Zwart SR, Heer M, Lee SM, Baecker N, Meuche S, Macias BR, Shackelford LC, Schneider S, Hargens AR.** WISE-2005: supine treadmill exercise within lower body negative pressure and flywheel resistive exercise as a countermeasure to bed rest-induced bone loss in women during 60-day simulated microgravity. *Bone* 42: 572–581, 2008.
29. **Smith SM, Zwart SR, Heer MA, Baecker N, Evans HJ, Feiveson AH, Shackelford LC, Leblanc AD.** Effects of artificial gravity during bed rest on bone metabolism in humans. *J Appl Physiol* 107: 47–53, 2009.
30. **Uebelhart D, Bernard J, Hartmann DJ, Moro L, Roth M, Uebelhart B, Rehailia M, Mauco G, Schmitt DA, Alexandre C, Vico L.** Modifications of bone and connective tissue after orthostatic bedrest. *Osteoporos Int* 11: 59–67, 2000.
31. **Veitch SW, Findlay SC, Ingle BM, Ibbotson CJ, Barrington A, Hamer AJ, Eastell R.** Accuracy and precision of peripheral quantitative computed tomography measurements at the tibial metaphysis. *J Clin Densitom* 7: 209–217, 2004.
32. **Vico L, Collet P, Guignandon A, Lafage-Proust MH, Thomas T, Rehaillia M, Alexandre C.** Effects of long-term microgravity exposure on cancellous and cortical weight-bearing bones of cosmonauts. *Lancet* 355: 1607–1611, 2000.
33. **Wang H, Wan Y, Tam KF, Ling S, Bai Y, Deng Y, Liu Y, Zhang H, Cheung WH, Qin L, Cheng JC, Leung KS, Li Y.** Resistive vibration exercise retards bone loss in weight-bearing skeletons during 60 days bed rest. *Osteoporos Int* 23: 2169–2178, 2012.
34. **Ward KA, Adams JE, Hangartner TN.** Recommendations for thresholds for cortical bone geometry and density measurement by peripheral quantitative computed tomography. *Calcif Tissue Int* 77: 275–280, 2005.
35. **Zerwekh JE, Ruml LA, Gottschalk F, Pak CY.** The effects of twelve weeks of bed rest on bone histology, biochemical markers of bone turnover, and calcium homeostasis in eleven normal subjects. *J Bone Miner Res* 13: 1594–1601, 1998.
36. **Zwart SR, Hargens AR, Lee SM, Macias BR, Watenpaugh DE, Tse K, Smith SM.** Lower body negative pressure treadmill exercise as a countermeasure for bed rest-induced bone loss in female identical twins. *Bone* 40: 529–537, 2007.

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O.B. 527
FI-33101 Tampere, Finland

ISBN 978-952-15-3340-2
ISSN 1459-2045