Delft University of Technology

Bones don't lie

What does bone shape tell us about skeletal diseases?

Tümer, Nazli

DOI
10.4233/uuid:d9b3f849-087e-46a9-96d2-f15d1b573a50

Publication date
2018

Document Version
Final published version

Citation (APA)
Tümer, N. (2018). *Bones don't lie: What does bone shape tell us about skeletal diseases?* [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:d9b3f849-087e-46a9-96d2-f15d1b573a50

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Bones Don't Lie:
What Does Bone Shape Tell Us About Skeletal Diseases?

Bones Don't Lie:
What Does Bone Shape Tell Us About Skeletal Diseases?

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus, Prof. dr. ir. T.H.J.J. van der Hagen
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
donderdag 13 december 2018 om 10:00 uur
door

Nazlı TÜMER

Master of Science in Biomedical Engineering,
Technische Universiteit Delft, Nederland
geboren te Istanbul, Turkije

Dit proefschrift is goedgekeurd door de promotoren:
Prof.dr. A.A. Zadpoor en Prof.dr.ir. H.H. Weinans
en copromotor:
Dr.ir. G.J.M. Tuijthof

Samenstelling promotiecommissie:

Rector Magnificus	voorzittter
Prof.dr. A.A. Zadpoor	Technische Universiteit Delft
Prof.dr.ir. H.H. Weinans	Technische Universiteit Delft
Dr.ir. G.J.M. Tuijthof	Universiteit Maastricht, Universiteit van Amsterdam

Onafhankelijke leden:

Prof.dr.ir. J. Harlaar	Technische Universiteit Delft
Prof.dr. W.R. Taylor	ETH Zurich
Prof.dr. J. Dankelman	Technische Universiteit Delft
Dr. R. Agricola	Erasmus MC
Prof.dr. H.E.J. Veeger	Technische Universiteit Delft, reservelid

Copyright © 2018 by N. Tümer
Een digitale versie van dit proefschrift is beschikbaar op
http://repository.tudelft.nl/

Contents

1 General Introduction **9**
1.1 Introduction 10
1.1.1 Skeletal Diseases and Bone Shape 10
1.2 Thesis Outline 12

2 Statistical Shape and Appearance Models of Bones **15**
2.1 Introduction 16
2.2 Statistical Models of Shape and Appearance 16
2.2.1 Statistical Shape Models 17
2.2.2 Statistical Appearance Models 19
2.2.3 Combined Statistical Models 20
2.3 Search Algorithms 20
2.4 Applications 21
2.4.1 Bone Model Reconstruction 21
2.4.2 Osteoporosis 23
2.4.3 Osteoarthritis (OA) 25
2.4.4 Patient-Specific Finite Element Models of Bones 28
2.4.5 Implant Design 29
2.4.6 Surgery Planning 30
2.5 Discussion and Conclusions 31
2.5.1 Steps Involved in Building and Using SSM and SAM . . . 32
2.5.2 Accuracy and Reproducibility of the Techniques 33
2.5.3 Limitations and Challenges 35
2.5.4 Potential Applications and Future Research 36

3 Statistical Shape and Appearance Models for Fast and Automated Estimation of Proximal Femur Fracture Load Using 2D Finite Element Models **39**
3.1 Introduction 40
3.2 Materials and Methods 41
3.2.1 Training and Evaluation Data 41
3.2.2 Generation of Statistical Shape and Appearance Model . . 41
3.2.3 The Search Algorithm 42
3.2.4 Generation of Patient-Specific FE Models 42

3.3 Results 43
3.4 Discussion 45

4 Bone Shape Difference Between Control and Osteochondral Defect Groups of the Ankle Joint 51
4.1 Introduction 52
4.2 Materials and Methods 53
4.2.1 Image Acquisition 54
4.2.2 Segmentation 54
4.2.3 Transection of Tibia 55
4.2.4 Registration 55
4.2.5 Statistical Shape Model(s) 55
4.2.6 Evaluation of Statistical Shape Model(s) 56
4.2.7 Comparison of Shapes Between Control and OCD Groups 57
4.3 Results 58
4.4 Discussion 60

5 Typical Shape Differences in Talus and Calcaneus Between Subjects with Chronic Ankle Instability and Controls 63
5.1 Introduction 64
5.2 Materials and Methods 65
5.2.1 Data Collection 65
5.2.2 Segmentation of Bones from CT Scans 65
5.2.3 Registration of Bones 66
5.2.4 Extraction of Bone Shape Variations 66
5.2.5 Comparison of Bone Shapes Between the Groups 67
5.3 Results 68
5.3.1 Cohort Characteristics 68
5.3.2 General Bone Dominant Shape Variations 68
5.3.3 Description of Bone Shape Variations 68
5.3.4 Comparison of Bone Shape Variations Between Three Groups 69
5.4 Discussion 70
5.5 Conclusion 72

6 Three-Dimensional Analysis of Shape Variations and Symmetry of the Fibula, Tibia, Calcaneus, and Talus 79
6.1 Introduction 80
6.2 Materials and Methods 80
6.2.1 Data Collection 81
6.2.2 Segmentation of Bones from CT Scans 81
6.2.3 Registration of Bones 82
6.2.4 Statistical Analyses Ipsi- and Contralateral Sides as Separate Groups 83
6.2.5 Statistical Analyses Ipsi- and Contralateral Sides Pooled into One Group 84

6.3 Results . . . 84
6.3.1 Shape Patterns in Ipsi- and Contralateral Sides as Separate Groups . . . 84
6.3.2 Shape Patterns in Ipsi- and Contralateral Sides Pooled into One Group . . . 85
6.3.3 Effects of Gender on Age-Adjusted Shape Variations . . . 86
6.3.4 Intra- and Inter-Subject Shape Variations . . . 88
6.4 Discussion . . . 88
6.5 Conclusions . . . 91

7 Three-Dimensional Registration of Freehand-Tracked Ultrasound to CT Images of the Talocrural Joint 97
7.1 Introduction . . . 98
7.2 Materials and Methods . . . 99
7.2.1 Test Data Acquisition . . . 100
7.2.2 Data Preprocessing . . . 102
7.2.3 3D Bone Response Data and US to CT Registration . . . 103
7.2.4 Evaluation of the Registrations . . . 104
7.2.5 Clinical Study . . . 105
7.3 Results . . . 106
7.4 Discussion . . . 108
7.5 Conclusions . . . 111

8 A Novel Ultrasound Technique for Detection of Osteochondral Defects in the Ankle Joint: A Parametric and Feasibility Study 113
8.1 Introduction . . . 114
8.2 Materials and Methods . . . 115
8.2.1 Simulations . . . 115
8.2.2 Experiments . . . 118
8.2.3 Validation of the Reference Model . . . 118
8.2.4 Assessment of Parameters Effects . . . 119
8.3 Results . . . 119
8.4 Discussion . . . 120

9 General Discussion and Summary 131
9.1 General Discussion . . . 132
9.2 Summary . . . 136
9.3 Samenvatting . . . 138

Appendices 141

A Appendix A 143
A.1 Statistical Models of Shape and Appearance . . . 144
A.1.1 Statistical Shape Models . . . 144
A.1.2 Statistical Models of Appearance . . . 150
A.2 Search Algorithms . . . 153
A.2.1 Initialization . . . 154

A.2.2 Active Shape Models (ASM) 154
A.2.3 Active Appearance Models (AAM) 158

B Appendix B **161**
B.1 Removing the Scaling and Rotational Effects 162
B.2 Combined Appearance Model 162
B.3 Derivation of Derivative Matrix 163
B.4 Definition of Proximal Femur Thickness 164
B.5 Definition of Material Properties 165
B.6 Definition of Peak Impact Force 166

C Appendix C **167**
C.1 Determination of Subchondral Bone Surface 168
C.2 Addition of Extra Points to Obtain Closed Form of Transected Tibia 169
C.3 Unbiased Point Registration Algorithm 170
C.4 Optimization of Registration Parameters (Numerical Experiments) 171
C.5 Robustness of Statistical Model(s) 173
C.6 Shape Parameters (Mean and 95% Confidence Interval) 174
C.7 Effects of Age and Gender on Bone Shape Parameters 174

D Appendix D **177**
D.1 Statistical Analyses of Ipsi- and Contralateral Sides as Separate Groups 178
D.1.1 *D*-Statistic Used in Group Location Test 178
D.1.2 *D*-Statistic Used in Variance-Covariance Scale Test 179
D.2 Parallel Analysis 179

E Appendix E **181**
E.1 Monogenic Signal, Local Phase and Phase Symmetry Measure . . 182
E.2 Surface Point Extraction From CT Data 183

References **185**

Acknowledgement **213**

About the Author **219**

List of Publications **221**

1

General Introduction

1.1 Introduction

If someone asks you to prepare a bunch of absolutely identical homemade cookies, your answer would probably be that even if you evenly mix all the ingredients and use a mold, none of the cookies will be exactly the same. The thickness of one cookie may simply be different than that of another. If you are asked a second time to bake another batch of cookies at a different location, the answer will not change. As in the previous case, the cookies will not be exactly the same. Dissimilarity between the cookies of two batches could be even higher as compared to that within the same batch due to variations in the ingredients, their quality, and mix level. Changes in variables (e.g. higher or lower volume of ingredients) would make cookies more or less tasty. Although distortions in cookies' shapes would not alter the taste of cookies, they could adversely affect the quality of visual presentation and decrease the palatability of the cookies.
In a way, human bones are like homemade cookies. Bones of the same type may look the same, but they are never absolutely identical to each other. Their shapes and constituents vary within and between individuals due to intrinsic (individual related) and extrinsic (environment related) factors. Unlike cookies, deviations in the variables of a bone from their optima could have serious implications for individuals. Indeed, it is a relatively recent discovery that the shape of bones could make them either susceptible to skeletal diseases, protect them against certain diseases, or play a role in the onset and progression of skeletal diseases [1, 2, 3]. The unifying theme of this thesis is the relationship between bone shape and skeletal diseases. We have considered 3 types of skeletal diseases (Section 1.1.1) and have tried to find answer to questions that are directly or indirectly related to bone shape. In doing so, we have used advanced techniques such as statistical shape models (SSM), statistical shape and appearance models (SSAM), and finite element models (FEM).

1.1.1 Skeletal Diseases and Bone Shape

Osteoporosis is one of the most well-known skeletal diseases characterized by decreased bone strength predisposing the individual to an increased risk of low-energy fractures [4, 5, 6, 7]. Based on estimations, 200 million individuals worldwide and 54 million men and women in the United States [6] suffer from the disease. After the age of 50, approximately 50% of women and 20% of men sustain osteoporosis-related fractures [5]. Although osteoporotic fractures are observed in most of the bone types, hip fractures are considered most serious and are associated with high morbidity and mortality [8, 9, 10, 11]. In clinical practice, dual-energy x-ray absorptiometry (DEXA) is the routinely used technique to evaluate the risk of osteoporosis-related fractures by performing areal bone mineral density measurements (aBMD) [12, 13, 14]. Nevertheless, the use of BMD alone is not sufficient to predict all of the osteoporotic fractures [11]. For instance, in a pilot study performed by Gregory et al. [15], osteoporosis-related hip fractures could be predicted with 82% accuracy based on BMD measurements alone. Besides BMD, bone geometry, microarchitecture and loading conditions [8, 12, 16]

are among the many other risk factors that can influence the fracture occurrence. Referring to previously mentioned study, the accuracy in the prediction of hip fractures increased by 8% (i.e. up to 90% accuracy) with an inclusion of the proximal femur shape next to BMD measurements [15].
Osteoarthritis (OA) is another serious skeletal disease that decreases level of patient activity, patient quality of life, and poses a large economic burden on society [17, 18, 19, 20, 21]. It has been reckoned that OA affects 10% of men and 18% of women over 60 years of age [17] and it will be the fourth leading cause of disability worldwide by 2020 [18]. Current joint-preserving interventions to treat OA consist of lifestyle modification as well as pharmaceutical and surgical modalities [17]. Nevertheless, these options have limited efficacy. Only a few of them have been shown to stop or postpone disease progression [17, 22]. To more efficiently deal with the disease, the current focus is to diagnose, prevent and treat the disease at an early stage as possible [17]. Like in the case of any other diseases, identification of risk factors is one of the keys to discern the causes of OA and to define individuals who would benefit the most from prevention and treatment options. An extensive effort has been put to determine multivariate risk factors (e.g. age, gender, hormonal status, genetic factors, bone density, previous damage, muscle weakness, ligament laxity, obesity, sport activities) [22] contributing to the onset and progression of OA. In recent years, there is an increasing interest in bone morphology and its relation with OA. Mounting evidence points that a slight difference in joint mechanics driven by variability in the shapes of bones may play a role in the initiation and progression of OA [3, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32]. For example, abnormalities in the contact between the proximal femur and acetabulum (i.e. cam impingement caused by an aspherical femoral head) have been reported to be strongly correlated with the development of hip OA [24].
There are several other types of lesions that are often caused by an injury and could progress into OA, if left untreated. One such type of lesion is an osteochondral defect (OCD), which is mostly manifested as the disruption of the articular cartilage together with its subchondral bone [33, 34]. Another important case is joint instability caused by induced laxity of ligaments [35] that could become chronic in time. OCD or chronic instability may be observed in any joint. Ankle joint is one of the common sites for these injuries. This is partially due to the high incidence rate of ankle sprain that comprise up to 46% of all sports-related injuries [36]. Despite extensive research on the risk factors associated with these lesions and the growing evidence showing that the shape of a bone could increase the risk of sustaining various skeletal diseases, the factors with bone shape origin have been mostly overlooked in the studies of OCD and chronic instability.
Determining the risk factors relevant for the initiation and progression of a disease, and early identification of an individual who is under risk of sustaining a disease or post-treatment monitoring of patients is highly important to take preventive measures and to decide on effective patient-specific treatment strategies. Besides, bone shape may be one of the risk factors and have something to tell us about skeletal diseases, as it contributes to the bone strength and to the stability

of the joints and determines their kinematics. Considering all these important aspects, the main focus of this thesis is on the clinically important problems, which are associated with the morphology of bones, and either have been overlooked or necessitate technical improvements. In the remainder of this chapter, clinical problems that have been handled, the rationale behind these studies, and our approaches are briefly presented.

1.2 Thesis Outline

In **Chapter 2**, we provide an overview on the main concepts, methods, and applications of SSM and SAM in bone research. The main motivation behind this literature review is to make readers familiar with these models that are powerful in describing bone shape variations (i.e. SSM), density variations (i.e. SAM), or both shape and density variations (i.e. SSAM) within a specific population, as these tools are extensively used in the following chapters.
In **Chapter 3**, we focus on the prediction of osteoporosis-related proximal femur fracture load based on FE modelling, as it is a promising alternative to densitometry-based techniques (e.g. DEXA) in estimating the bone strength and the risk of fracture under specific loading conditions. In return for their power, the development of patient-specific FE models is a time-consuming and complex task, which necessitates the type of expertise that is not commonly available in clinical settings. To decrease the time and user dependency in the generation of patient-specific FE models, we present an automated platform. The presented procedure starts of by fitting a SSAM to a new unseen DEXA scan of the patient, thereby determining the shape of a patient's proximal femur as well as its bone density distribution. The automated process continues with assignment of material properties, application of loads and boundary conditions, and discretization of the proximal femur model. The final outcome of the process is an estimation of the fracture load calculated using FE analysis.
In **Chapter 4**, we investigate whether there are statistically significant shape differences between the distal tibiae and tali of patients with an OCD and those of healthy subjects. One of the ultimate reasons to perform this study is that the etiology of an OCD is still not fully understood. Despite extensive research aimed to identify the risk factors of OCD (e.g. local avascular necrosis, acute trauma, chronic microtrauma, metabolic factors, malalignment of the lower limb [37, 38], the relationship between bone shape and OCD risk in the ankle has been so far overlooked. To gain insight into three-dimensional (3D) shape variations in the bones forming the talocrural joint (i.e. the distal tibia and talus), we build 3D SSM of the talus and distal tibia based on the mixed data of the CT scans of subjects with and without an OCD. Using those SSMs and statistical tools, we quantitatively compare the bone shapes of both groups.
In **Chapter 5**, we systematically describe and compare 3D shape variations of the talus and calcaneus within a population composed of subjects with no known history of the ankle joint pathology and patients who had developed chronic ankle instability (CAI) after sustaining lateral ankle sprain. This study was con-

ducted, as risk factors with bone shape origin have been in general disregarded. During this study, we use a similar methodology as described in Chapter 4.
Chapter 6 differs from Chapters 3-5 and Chapters 7-8 with its scope. It does not directly address the association of bone shape with a specific skeletal disease (i.e. investigation of potential risk factors or early identification of a disease or long-term monitoring of a patient with a disease). Instead, we focus on the bone shape itself, and analyze 3D shape variations and (a)symmetry of the lower extremity bones (i.e. the fibula, tibia, talus and calcaneus). This study was motivated by the fact that the lower extremity bones are often assumed to be bilaterally symmetric in clinical assessments and research studies (e.g. Chapter 4). However, there is a lack of sufficient documentation on 3D shape variations and (a)symmetry of the lower extremity bones. Therefore, we evaluate whether (1) bone shape patterns in both sides of an individual are alike, (2) bone shapes vary based on gender, and (3) intra-subject bone shape variations are smaller than inter-subject variations using 3D spatially dense descriptions of the bones and advanced statistical techniques, namely 3D SSM and distance based permutational statistics.
In **Chapter 7**, we present a registration scheme for computed tomography (CT) and freehand-tracked ultrasound (US) data of the talocrural joint, which can be used for noninvasive detailed longitudinal monitoring of morphologic changes in the bone shapes. The main motivation behind this study was that US imaging is the preferred technique as compared to CT, magnetic resonance imaging (MRI) and arthroscopy in the detection of an OCD or longterm monitoring of patients with a disease [33, 39, 40, 41] due to its inherent properties, such as non-invasiveness and cost-effectiveness. Nevertheless, interpretation and comparison of US images over time is difficult without having access to anatomical data provided by other imaging modalities, such as CT. Therefore, we proposed the registration scheme based on evolutionary approaches to make interpretation and comparison of US images over time possible. To assess the robustness of the registration algorithm to different initialization positions, and the sampling of US data, two different tests are applied on a dataset collected from cadaveric ankles. Moreover, a practical case study is performed by using the presented registration scheme to evaluate whether US imaging could be an alternative to CT for pre-operative planning of arthroscopic access to anterior talar OCDs.
In **Chapter 8**, we propose a new diagnostic technique to discriminate OCDs in the ankle joint and assess its feasibility using experimentally-validated 2D finite-difference time-domain models of the ankle joint. Specifically, we tackle the limited ability of US waves in conventional US imaging to penetrate through bones so that we could benefit from the inherent non-invasiveness and cost-effectiveness of US. The main working principle of the new technique is the transmission of ultrasonic waves from one side of the ankle joint in such a way that they could propagate through the entire joint space of the ankle and recording the transmitted waves at the other side. When the recorded US signal deviates from the one associated with a healthy joint, the presence of OCDs is identified. In this chapter, within the scope of the study, the effects of the variables of the ankle joint (i.e. the width of the joint space), those of the US transducer (i.e. the positioning of the US transducer acting as transmitter), and the defect (i.e.

width, depth and location) on acoustic wave response are analyzed.
In **Chapter 9**, we provide some concluding remarks, summarize the studies presented in the preceding chapters, and present some suggestions for future research.

2

Statistical Shape and Appearance Models of Bones

Published as:
N. Sarkalkan, H. Weinans, A.A. Zadpoor, Statistical shape and appearance models of bones, Bone. 60 (2014) 129-140.

2.1 Introduction

Statistical shape models (SSM) and statistical appearance models (SAM) are two types of statistical models that respectively describe the average shape and average density distribution together with the main modes of variation of shape and density distribution within a population. SSM and SAM are generally created based on a training dataset that provides the model with a-priori knowledge about the expected shape and density distribution of bones. In many applications, SSM and SAM are fitted to new unseen sets of images. Active Shape Models (ASM) and Active Appearance Models (AAM) are probably the best known methods that are used to respectively fit SSM and SAM to a new set of two-dimensional (2D) or three-dimensional (3D) images [42, 43].
The possibility of fitting SSM and SAM to new unseen images creates many opportunities that could be used for facilitating the study, diagnosis, and treatment of skeletal diseases. For example, 3D SSM can be used in generation of 3D patient-specific bone models from 3D sparse data or from a set of 2D images. These patient-specific models are used in a variety of medical applications including in vivo bone motion tracking [44] and computer aided orthopedic surgeries [45, 46, 47, 48].
When a 3D patient-specific bone model is correctly reconstructed from a set of 2D images by using 3D SSM, the method basically provides us with an excellent bargain: the level of detail normally obtained by 3D imaging for the price of 2D imaging. The use of 2D imaging modalities means significantly reduced imaging costs, decreased exposure of patients to radiation, the possibility of repetitive follow-up imaging, and ultimately many new opportunities for improved diagnosis and treatment.
The bone research community has recently realized the potential of SSM and SAM. In this paper, we present an overview of SSM and SAM, and their applications in bone research. First, the main concepts and algorithms are presented. Most of the technical details are presented in Appendix A that is organized in parallel with the main text and provides the interested reader with the details of algorithms. Second, the applications of SSM and SAM in bone research are discussed.

2.2 Statistical Models of Shape and Appearance

Statistical models of shape and appearance are basically composed of two components: 1. the average shape or appearance and 2. the main modes of variation of shape or appearance from the average values. Based on these two components, they could describe the shape and/or density distribution of the bones within a certain population of individuals. The population may, for example, refer to healthy individuals with a shared ethnical background or to patients with a certain type of skeletal disease such as osteoporotic patients. The main idea is that the shape and density distribution of the bones of every individual within

that population could be obtained by adding the contributions of (a limited number of) the principal modes of variation to the average shape and/or density distribution. There are, however, two steps that need to be taken. First, one needs to determine the average shape and/or density distribution and the principal modes of variation within the population of which the individual is a member (Section 2.2). Second, one needs to use a fitting algorithm to determine what the contributions of the principal modes of variations are to the shape and/or density distribution of the bones of that particular individual (Section 2.3).

2.2.1 Statistical Shape Models

SSM are either developed in 2D or 3D. For 2D SSM, x-ray or dual-energy x-ray absorptiometry (DXA) images are generally used. 3D SSM are often built using computed tomography (CT) or magnetic resonance imaging (MRI) images. A number of training images need to be first collected from a sufficiently large number of bones that represent the variation of shape within the statistical population for which the SSM is going to be built.
The first step in obtaining the average shape and the principal modes of variation is to find a way for representing the shape of bones. Among the various techniques available for shape representation (see Appendix A), landmarks are used most frequently in representing the shape of bones. In this method, the coordinates of a number of bony landmarks, $\mathbf{x}_i$, describe the shape of the bone (Fig. 2.1).

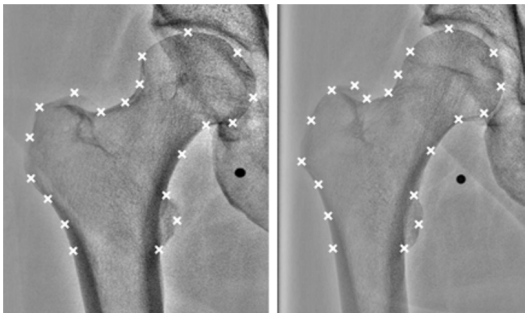


Figure 2.1: Two examples of manually positioned bony landmarks on radiographs.

The second step involves defining the bony landmarks on a number of training bone shapes. The landmarks should be defined such that they consistently refer to the same anatomical location on every instance of the bone shape. We therefore speak of a concept called correspondence, meaning that the landmarks placed on all training instances should correspond to each other. The simplest way for ensuring correspondence is manual positioning of the landmarks on the periphery of bones in training images. The persons who perform the manual positioning of landmarks should be anatomically trained so that they consistently

assign the landmarks to anatomical locations found on training images. Once a few clearly and uniquely defined landmarks are manually positioned on every training image, the other landmarks can be automatically placed in between those manually positioned landmarks by interpolation. Despite its simplicity, manual positioning becomes prohibitively difficult and time-consuming for 3D images as well as for large datasets of 2D images. That is why automated methods might be needed for establishing correspondence in large 2D and 3D datasets (see Appendix A).

Once the landmarks are positioned on the training images and the correspondence between the landmarks is established, the different bones within the training database are aligned to eliminate the variations between the training instances that are caused by factors other than variation in shape. For example, rotations and translations of training images with respect to each other are eliminated during the alignment procedure. There are certain algorithms for automated alignment that are described in Appendix A among which Generalized Procrustes Analysis (GPA) [49, 50, 51, 52, 53] is the most widely used algorithm. After alignment, the coordinates of the bony landmarks are assumed to represent only shape variations. The average shape of bones, $\bar{\mathbf{x}}$, can be easily calculated as:

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i \tag{2.1}$$

where N is the number of training instances. In order to calculate the variation of shapes from the mean shape, one could calculate the covariance matrix, $\mathbf{S}$, that measures the variation of different shapes, $\mathbf{x}_i$, from the mean shape $\bar{\mathbf{x}}$ by:

$$\mathbf{S} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \tag{2.2}$$

Principal component analysis (PCA) is then performed on the covariance matrix to calculate the most important modes of variation of the training shapes, $\boldsymbol{\phi}_i$, from the mean shape, $\bar{\mathbf{x}}$. The principal modes of variation are ordered in the descending order of the percentage of shape variation explained by that principal mode of variation. Therefore, the importance of the first few modes of variation in describing the shapes of other individuals within the population is much more than the importance of other modes (Fig. 2.2). The bones of other individuals within the population, $\mathbf{x}$, not present in the training dataset can be then described as:

$$\mathbf{x} = \bar{\mathbf{x}} + \sum_{s=1}^{c} b_s \boldsymbol{\phi}_s \tag{2.3}$$

where b_s values describe the contributions of the first c modes of shape variation to the average shape of the bone. The number of modes retained depends on the desired accuracy in description of the bone shape. In general, a larger number of retained modes results in a more accurate description of the bone shape. For a more detailed explanation of the above-mentioned procedures see Appendix A.

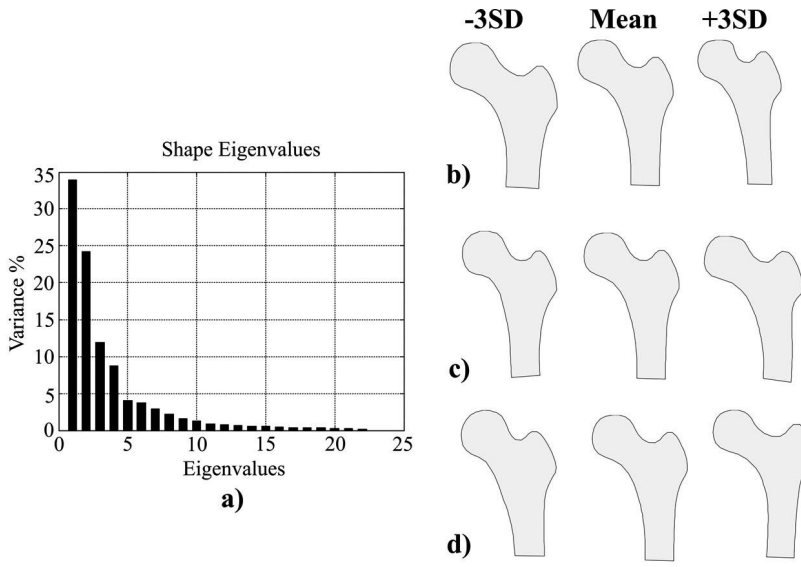


Figure 2.2: The percentage of shape variations explained by different shape modes in a 2D statistical shape model of the femur (a) together with the first three modes of shape variation (b-d).

2.2.2 Statistical Appearance Models

While SSM provide useful information regarding the shape of bones, they do not contain any information regarding the density distribution within bones. SAM, on the other hand, can describe the average density distribution within bones as well as the main modes of variations of density distribution from the mean values. Since density distributions are often obtained from the intensity of gray values within medical images, SAM are also called statistical intensity models. To generate a SAM, one needs to capture information related to the pixels/voxels (e.g. pixel intensities) within the training bone instances. It is also important to make sure that the settings and illumination properties of different imaging machines that might result in different gray values for the same bone density do not influence the SAM. The gray values are therefore normalized in a normalization procedure that tries to eliminate the effects of machine settings and global illumination properties. In a process somewhat similar to the one used in generation of SSM (see Appendix A), the average density distribution, $\bar{\mathbf{g}}$, and the main modes of variation of the density distribution from the average values, $\boldsymbol{\phi}_g$, are calculated. The density distribution of another bone, $\mathbf{g}$, can be described as:

$$\mathbf{g} = \bar{\mathbf{g}} + \boldsymbol{\phi}_g \boldsymbol{b}_g \tag{2.4}$$

where $\boldsymbol{b}_g$ is the vector that describes the contributions of the principal modes of variation in density distribution to the average density distribution of that particular bone. Similar to shape, only a limited number of the principal modes

of variations are generally retained for describing the density distribution within a particular bone (Fig. 2.3).

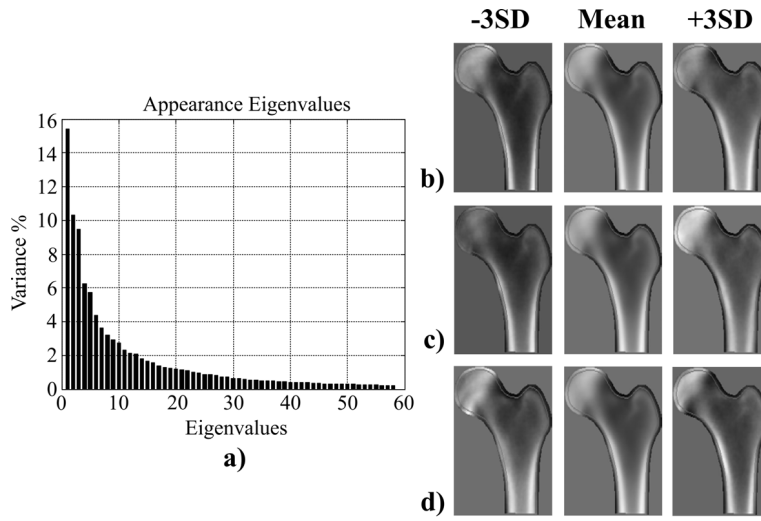


Figure 2.3: The percentage of variation in appearance explained by different modes in a 2D statistical appearance model of the femur (a), together with the first three modes of appearance variations (b-d).

2.2.3 Combined Statistical Models

In order to explain shape and density distribution simultaneously and to discard the correlations between shape and density distribution, the SSM and SAM can be combined into a so-called combined statistical shape and appearance model that describes both shape and density distribution. See Appendix A for the details of the applied procedure.

2.3 Search Algorithms

Once the statistical models of shape and appearance are created, the only remaining aspect is to fit the models to (a set of) unseen images. In different applications and contexts, the process of fitting a statistical model of shape and/or appearance to a new set of images may be called a search, matching, registration, or fitting process. Regardless of their names, all those procedures try to do one thing: finding b_s values in Eq. 2.3 or $\boldsymbol{b}_g$ vectors in Eq. 2.4 such that the statistical models of shape and appearance are fitted to the bone seen on the new images as well as possible. One can therefore conclude that search algorithms are optimization algorithms that try to minimize the difference between the shape and density distribution represented by the statistical model and the shape and density distribution of the bone seen on a new set of images by adjusting b_s values

or $\boldsymbol{b}_g$ vectors. In addition to traditional optimization algorithms, two specific local search algorithms called active shape models (ASM) and active appearance models (AAM) are often used for fitting statistical models of shape and appearance to images. While ASM only uses shape data for finding the best fit, AAM uses both shape and appearance data available in the images for finding the best fit of the model to the images. The technical details of the search algorithms can be found in Appendix A.

2.4 Applications

2.4.1 Bone Model Reconstruction

3D patient-specific models are generally derived from pre-operative CT or MRI images [54]. Nevertheless, those two imaging modalities are expensive, and CT imaging exposes patients to high doses of radiation. Therefore, 3D/3D or 3D/2D registration techniques that enable construction of 3D models from intra-operative data such as ultrasound, fluoroscopy, optical images, 3D digitized points [55] or 2D images such as x-ray are of great interest. Some of the most important methods used for registration of 3D statistical models of shape and appearance to 2D and 3D image are reviewed here. These methods can be categorized into two main groups, namely feature-based and intensity-based techniques. For an extensive review of the related techniques, the reader is referred to Markelj et al. [55].

Feature based registration methods use geometrical entities such as sets of points, edges, contours, or surfaces [56]. The main principle of feature-based registration methods is to minimize the distance between the 3D features extracted from the 3D statistical models and the corresponding 3D or 2D features obtained from images [44, 55].

One representative study that used feature-based registration for reconstructing 3D patient-specific bone models is reported in reference [46]. The first step of their method is to build a 3D SSM of the relevant anatomy from CT images. The interventional data, i.e. bone surface points, were extracted from ultrasound images of a cadaveric bone that was not used in the construction of the statistical model. To generate a 3D patient-specific model of the bone, the statistical model was matched to the bone surface points by using a well-known method for rigid registration, i.e. iterative closest point algorithm (ICP). In this method, the shape parameters are found in such a way that the distance between the corresponding features of model and those of the ultrasound images is minimized [46].

In another study, Fleute et al. suggested a method to represent the complete surface of the bone of interest from very sparse 3D point data. In the proposed method, the ICP global search algorithm is used to estimate the rigid-body transformation before applying the non-rigid local registration based on a simulated annealing technique and the downhill simplex algorithm [47]. An alternative method to reconstruct 3D patient-specific model from sparse 3D intra-operative point data is explained in reference [48]. Based on that method, a SSM is built and is registered to sparse 3D point data by updating the shape parameters such

that the least square error between the model data and the interventional data is minimized. Rajamani et al. also consider incorporation of a regularization term, namely Mahalanobis distance of the estimated model, to more robustly guess the parameters of the model [48].
Registration methods can be also used for matching 3D statistical models to 2D image data (Fig. 2.4). In another study of Fleute et al., a variation of ICP is used to match a 3D statistical model to the contours of a bone extracted from intra-operative x-ray images [57]. Zheng et al. introduced a method to build up a patient-specific 3D bone surface model from calibrated 2D x-ray images. The proposed method basically establishes correspondence between the apparent contours of the 3D SSM's bone surface and the edge points identified in 2D images. The estimated 2D point pairs are back-projected to 3D to build a set of 3D points. Thus, the 2D/3D reconstruction problem is converted to a 3D/3D reconstruction problem. Afterwards, the 3D point pairs are registered to the corresponding 3D point set generated by deforming the SSM [58].
The previously mentioned methods establish the correspondence explicitly and that may result in matching wrong point pairs. Several methods have been proposed to avoid such inaccuracies. One of the alternatives is the use of distance maps to determine the correspondence implicitly. Kurazume et al. introduced a method that estimates the position and shape parameters of 3D statistical model from two fluoroscopic images by using a distance map [59]. Hurvitz et al. proposed an intensity-based method to reduce the frequency of incorrect correspondences as well as to discard the possible edge detection problems that may be encountered in the studies that use interventional images [60]. They first build up a CT-like intensity atlas similar to AAMs. The registration of the statistical model to a few fluoroscopic images is a recursive process. It starts with generation of an instance of the statistical model based on some initially-estimated transformations, and of simulated x-ray projection images known as digitally reconstructed radiographs (DRR). Following generation of DRRs, the correspondence is established between DRRs and fluoroscopic images. As a result of correspondence, unknown transformations are re-estimated. The above-mentioned steps continue repeatedly until the convergence condition is satisfied for the transformation estimates [60].
Another intensity-based method is described by Tang et al. [61]. They use a hybrid 3D shape model that can be implemented in the point distribution model scheme proposed by Cootes et al. [62]. In the 3D shape reconstruction step, a recursive process is applied to determine the patient-specific shape by optimizing a similarity measure (gradient correlation in that study) between the DRRs of the shape model and the 2D interventional images. Moreover, as the shape model used in the study cannot provide information on image volume to generate DRRs, Tang et al. grow the surface of the model inward to simulate image volume and the thickness of the cortical bone [61].

Whitmarsh et al. proposed an intensity-based method to reconstruct both 3D shape and bone mineral density distribution (BMD) from a single 2D DXA image as interventional data (Fig. 2.4) [64, 65]. They first set up a statistical model similar to AAMs described by Cootes et al. [66]. Following this step, the combined

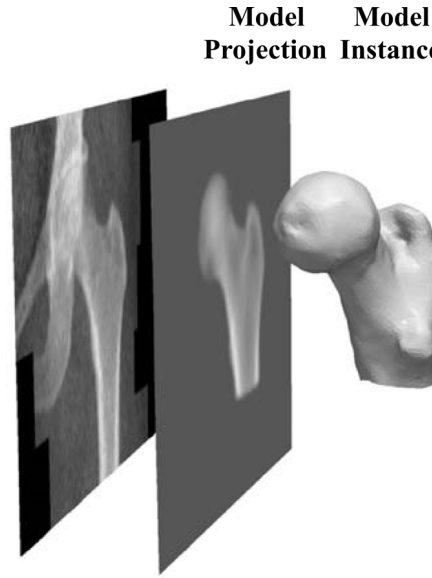


Figure 2.4: The different steps involved in matching a 3D SSM with a DXA scan [63]. First, digitally reconstructed radiographs (DRRs) are generated by projecting an instance of the 3D statistical shape and appearance model. The generated DRRs are then compared with the DXA scan. The parameters of the statistical shape and appearance model are adjusted such that the 2D projection of the model instance matches the DXA scan as well as possible.

model is registered to a DXA image. During a recursive process of registration, the model parameters are updated repeatedly such that the similarity measure between the DRRs and the DXA image is optimized [64].

2.4.2 Osteoporosis

Low bone mass caused by osteoporosis increases the risk of bone fractures. There are several techniques to measure bone mass per unit volume (volumetric bone mineral density, vBMD, $\frac{g}{cm^3}$) or per unit area (areal bone mineral density, aBMD, $\frac{g}{cm^2}$). These techniques include DXA, quantitative ultrasound (QUS), quantitative computed tomography (QCT) and digital x-ray radiogrammetry [67]. The most widely used way of assessing bone mass in clinical practice is based on aBMD measures derived from DXA scans [52, 13, 63]. However, low BMD measures alone are not sufficient to explain all osteoporotic fractures [68]. This is due to the fact that other factors such as the shape of bones contribute to their strength as well and may be potential risk factors for fractures. Therefore, several methods are proposed to evaluate the fracture risk by taking other possible risk factors into account. It has been, for example, established that the shape of the proximal femur is a risk factor for femoral neck fractures [8]. Referring to the re-

sults of several studies, Gregory et al. indicate that larger hip axis length, larger neck width, and larger neck-shaft angle are associated with an increased risk of femoral neck fractures [8]. However, these geometric properties are highly correlated with each other.

In order to determine independent shape factors that are correlated with higher risks of fracture, Gregory et al. used ASM in their study [15]. They found that the morphology of the femur quantified using ASM may be helpful for identifying the individuals who are at increased risk of proximal femoral fracture [15]. On the other hand, an alternative method to improve the prediction of hip fracture is suggested by Goodyear et al. [13]. The authors use statistical shape and appearance models to capture shape and texture information within the femoral head and neck. Goodyear et al. indicate that modes of variations, derived from the statistical model that is built from DXA training images, enhance prediction of hip fracture when they are used in combination with BMD [13]. It should be, however, noted that identifying shape modes that may be related to increased risk of fracture is of purely diagnostic value as bone shape factors cannot be modified in the advanced ages when osteoporotic fractures normally occur.

Furthermore, assessment of hip fracture using aBMD measurements derived from DXA is limited by its two-dimensionality. To overcome this limitation, vBMD can be used together with 3D structural measurements obtained from QCT for fracture risk evaluation. However, QCT is expensive and exposes patients to high radiation doses. With the aim of keeping DXA as the standard modality while improving hip fracture predictions, Whitmarsh et al. suggest to use a statistical model of both the 3D shape and BMD distribution [63]. The basic steps of the proposed method are generation of the statistical model and registration of it onto DXA images. Following the registration process, the estimated model parameters are used for hip fracture discrimination. Based on their results, the model parameters enhance the predictive ability of DXA derived aBMD [63]. Thus, the proposed method has the potential to improve hip fracture risk estimation.

Vertebral fractures are also of interest. Various semi- and fully-quantitative methods are proposed for diagnosing and describing vertebral fractures. As for quantitative methods, vertebral fractures are determined based on assessment of decrease in the anterior, middle, and posterior heights of vertebral bodies. Although quantitative methods are objective and reproducible, they have limited ability to distinguish between vertebral fracture and vertebral deformity [69]. Semi-quantitative assessment of vertebral fractures is performed by expert radiologists or trained readers. Semi-quantitative methods benefit from additional information such as typical changes in image texture when endplate fracture occurs. That kind of information is not normally used in quantitative methods even though it could be helpful for distinguishing deformity from fracture. Semi-quantitative assessment is, however, more subjective particularly for mild fractures [70]. With the aim of combining some of the advantages of quantitative techniques with those of semi-quantitative techniques, Roberts et al. defined a quantitative approach [71] in which statistical shape and appearance models were built for vertebrae at different levels, i.e. lumbar spine, lower, and up-

per thoracic spine. Following the construction of shape models, shape parameters are defined by fitting the model to each vertebra in the training set. On the other hand, appearance parameters are determined after the generation of statistical appearance models. Subsequently, linear discriminant fracture classifiers are trained by using the defined parameters. They concluded that the quantitative classifiers are more efficient compared to the standard height ratio method for detecting vertebral fractures particularly the mild ones [71]. In another study, Roberts et al. found that the quantitative classifiers of vertebral fracture require a detailed and accurate segmentation of the vertebral endplate. As manual segmentation of vertebra is a time consuming task, Roberts et al. proposed a method to semi-automatically segment the vertebra from DXA images. During the segmentation process, a variant of AAM is used. Based on the results, an appearance based classifier is found to be adequate for detecting vertebral fractures, although there is a sensitivity loss of 7% compared to manual segmentation [72]. Afterwards, the semi-automatic detection method based on AAM was applied to spinal radiographs with the purpose of evaluating the accuracy of the method in locating vertebrae on radiographs [70]. There are also some other (semi-)automatic segmentation techniques that have been developed to segment vertebrae using the ASM [73, 74, 75] and AAM algorithms [76]. It is not easy to compare the methods proposed in above-mentioned studies against each other, because the techniques are evaluated on different datasets that might have variations in their image and annotation qualities. It is, however, clear that the proposed (semi-) automatic segmentation techniques decrease the time required to locate vertebra in images and have the potential to be nearly as accurate as manual segmentation.

2.4.3 Osteoarthritis (OA)

About 85% of individuals aged 75 years or more have radiographic or clinical evidence of OA [27]. While radiological OA is often assessed using (semi-) quantitative systems such as Kellgren-Lawrence score that measure the changes in cartilage and in bone [27, 23, 29], clinical OA is evaluated based on self-reported measures such as patient pain experience [26].
It has been conventionally believed that degeneration of cartilage is the initiating factor of OA and the changes in bones occur afterwards, due to the altered biomechanics of the joint. However, it is now understood that the changes in bone may also play important roles in the onset and progression of OA [77]. Certain patterns of bone geometry (e.g. femoral head-neck-shaft angle) or shape are potential risk factors for the development of OA [27, 29]. The geometry of bones is generally quantified using predefined measures (e.g. angles) on radiological images. As these measures represent the properties of the same object, there can be significant correlations between different geometric properties. Highly correlated measures might not measure what would be relevant [26]. On the other hand, complex shapes cannot be fully described using a few geometric measures. Statistical models could help in quantifying the shape of bones in radiological images. To overcome the above-mentioned limitations, Gregory et al. [27] used

a SSM of the femoral head and neck to examine the ability of statistical models in representing the shape differences between healthy controls and subjects with hip OA. Moreover, the ability of statistical models to quantify the deformation of the femoral head caused by OA progression was assessed. They showed that their SSM could capture the shape differences between the healthy subjects and subjects developing radiological OA of the hip. Therefore, it was concluded that SSM facilitates early identification of individuals who are at increased risk of developing radiographic OA. Furthermore, they found that SSM could quantify femoral head deformations more accurately as compared to conventional scorings such as Kellgren-Lawrence [27].

Lynch et al. [29] used a similar methodology to study OA in a homogenous population (i.e. Caucasian women), as there might be significant differences between the bone features of men and women. They modeled the complete proximal femur rather than only modeling the femoral head and neck. According to their observations, three specific shape factors characterized by variations in femoral head, neck, and trochanteric regions represent the potential risks for developing radiographic OA in elderly Caucasian women [29]. Later, Barr et al. examined the potential of two SSM for predicting the need for total hip replacement (THR) apart from the traditional scoring (i.e. Kellgren-Lawrence) and other risk factors such as clinical factors including pain duration [23]. The first statistical model they used was similar to the one used by Gregory et al. and modeled the shape of femoral head and neck [27]. The second statistical model included the whole proximal femur together with the osteophytes and acetabulum. They found that both statistical models were powerful tools for estimating the risk of the progression of radiographic hip OA and the need for THR. Moreover, they observed that the second statistical model is more powerful in predicting the need for THR. They concluded that SSM enhances the predictive ability for identifying the individuals at increased risk of THR [23].

The previously explained studies mainly show that specific shape patterns correlate with radiographic OA. However, it remains unknown whether there is a relationship between shape variants and clinical OA. Agricola et al. studied the relationship between hip shape at baseline and clinical OA [3]. They used a SSM of the proximal femur and pelvis and showed that it could predict the need for THR. However, none of the shape modes were associated with clinical OA as defined by the American College of Rheumatology (ACR) criteria. Furthermore, they observed that SSM could predict two clinical criteria independently, i.e. hip pain and decreased internal rotation, although none of the shape modes were related to the clinical OA as determined by combination of the ACR criteria [3]. The study also supports the idea that SSM may be used as a radiographic marker to predict the need for THR.

Another relevant study that takes clinical OA into consideration is conducted by Waarsing et al. (Fig. 2.5) [26]. There are inconsistencies between the radiological measures of OA and the symptoms that are experienced by patients [26]. Therefore, the study investigates the potential of statistical shape and appearance models in capturing symptom-related information from DXA images of OA hip that cannot be obtained by conventional radiological measures. The

study finds several appearance modes that are related to radiological and clinical OA (Fig. 2.5). It is also observed that statistical shape and appearance models of the proximal femur could capture information relevant to clinical OA that is not provided by traditional radiological measures [26].

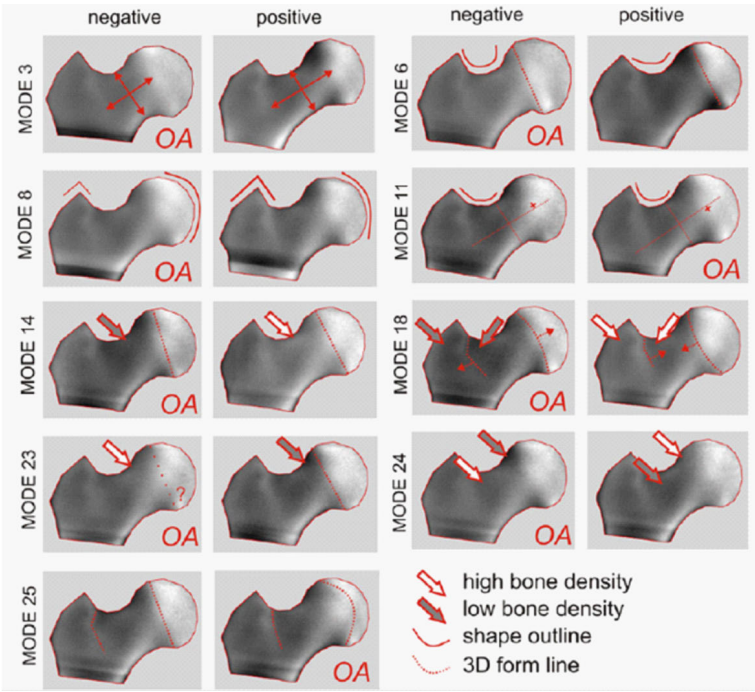


Figure 2.5: Demonstration of the appearance modes that showed significant association with OA [26]. Positive: +2 times the standard deviation of the population, Negative: -2 times the standard deviation of the population.

An another study, Waarsing et al. investigated the influence of osteoarthritis susceptibility genes on the relationship between the hip shape and osteoarthritis [78]. They developed a SSM of the femur together with the acetabulum and pelvis. The relationship between the shape modes that were found to be related to the radiographic characteristics of hip OA, as defined by Kellgren-Lawrence criterion, and the osteoarthritis susceptibility genes was examined [78]. It was observed that certain susceptibility genes may affect the association between hip morphology and OA, possibly by increasing the vulnerability of cartilage to unfavorable bone shapes [78].
Another application of SSM is described by Chan et al. [79]. Two common pediatric disorders, namely Legg-Calvé-Perthes disease (LCPD) and slipped capital femoral epiphysis (SCFE) that cause disrupted proximal femur morphology were considered. As the morphology of the hip joint is one of the risk factors for developing OA, it is important to understand possible shape variations in LCPD

and SCFE. Such morphological deformities are generally quantified on 2D radiographs [79]. However, evaluation of proximal femur deformations is limited in accuracy, when 2D radiographs are used. To gain more insight into the shape deformations of the proximal femur, Chan et al. propose to use 3D SSM to quantify shapes during normal growth and in LCPD and SCFE. They built a 3D SSM of the proximal femur using a training dataset that consisted of 24 proximal femurs (three training samples for each of the four considered asymptomatic age groups and four LCPD/SCFE groups). Shape modes were found to be distinct with respect to age and disease. It was shown that shape variants could describe the shape differences between asymptomatic and diseased femurs and could provide information regarding the morphological development of the deformities. Moreover, it was found out that the shape modes could capture shape details that cannot be easily described by conventional variables (e.g. head-neck transition) [79].
Recently, a fully automatic shape model matching system was developed by Lindner et al. [80] to derive statistical shape models from radiographs. They used the developed system for global representation of the proximal femur shape. In another study, Neogi et al. [30] examined whether 3D bone shape based on MRI images together with AAM could be used to predict the later onset of OA. They found that "bone shape at baseline, often several years before incidence, predicted later OA" [30].
Previously explained studies mainly focus on the effect of hip joint geometry on the development of OA. However, the studies that use SSM and/or SAM to analyze the relationship between the knee joint and OA are relatively limited in number. In one study by Haverkamp et al., a SSM that covered the contours of the femur, tibia, patella and the back of the medial condyle was built to investigate the association of specific shape patterns with radiographic OA [31]. The study showed that some shape modes, that were found to be related to radiographic OA, have a potential role to play as predictors of OA progression [31].
A different application of statistical models is described by Väänänen et al. [81]. Diagnosis of degenerative hip diseases using 2D femoral radiograph may not be reliable, since the geometry and density measurements derived from radiographs may be influenced by the 3D rotation of the femur. Therefore, identification of 3D rotation of the femur in 2D radiographs is helpful to increase the accuracy of diagnostic measurements. Väänänen et al. combined statistical models of shape and appearance with artificial neural networks to determine the 3D rotations of the femur in 2D radiographs [81].

2.4.4 Patient-Specific Finite Element Models of Bones

Patient-specific finite element (FE) models are currently used in many different orthopedic applications. They could be used to calculate the stress and strain distribution within different tissues including bone [82] and cartilage [83] both in physiological and pathological conditions. Moreover, one could use FE models to study bone tissue adaptation [84] in response to changes in musculoskeletal loading either due to different patterns of physical activity or because of the use

of prostheses and implants [85]. The applications of FE models in orthopedics include, for example, design of orthopedic implants [86, 87, 88, 89], bone fracture risk prediction [90, 91, 92], simulating bone tissue adaptation [93, 94, 95], orthopedic surgery planning [96], and relating musculoskeletal loads to bone density distribution [97, 98].
In some earlier studies, patient-specific finite element models were created using a relatively slow and time-consuming procedure based on manual or automated segmentation of either 3D or 2D image data obtained for individual patients [99]. The alternative approach that is based on statistical shape and appearance models can be used to automate the entire process of patient-specific model generation. A search algorithm such as ASM or AAM could be used for fitting a SSM or SAM to the image data obtained for the individual patient, thereby eliminating the need for manual segmentation. Once the model is fitted to the image data, the obtained model parameters could be used for generating patient-specific FE models. Since there are methods to fit 3D SSM or SAM to 2D images, one has the opportunity to obtain 3D FE models from 2D patient images. Moreover, the complete process of generating FE models could be automated. The obtained patient-specific geometry could be also used for adapting/scaling musculoskeletal models [100, 101] or simpler mass-spring-damper models of the musculoskeletal system [102, 103, 104]. Those musculoskeletal models could then be used for estimating the musculoskeletal loads applied on bones and automatically applying the estimated loads on the geometry of the FE models.
The use of such procedures based on SSM or SAM could greatly facilitate generation of patient-specific FE models and will facilitate clinical applications of such models. First, the time, required level of technical expertise, and cost associated with generation of patient-specific FE models will greatly reduce. Second, one could use 2D imaging modalities such as x-ray, DXA, or bi-planar x-rays instead of 3D imaging modalities such as CT for obtaining the images that are required for patient-specific FE modeling, thereby decreasing the level of ionizing radiation received by the patient. The above-mentioned approach for generating patient-specific FE models is relatively new and not many studies have so far used it. In the next section, some of the studies that use similar approaches for implant design have been reviewed. In addition to implant design, statistical FE models are useful for estimating the bone fracture risk [105, 106].

2.4.5 Implant Design

The design process of orthopedic implants includes considerable tests to verify their reliability before releasing them into the market. Nevertheless, the performance of the designed implants may not be at the expected level due to significant variations in bone geometry and bone quality among different individuals. SSM and SAM could be helpful for design and analysis of orthopedic implants. First, SSM and SAM could be used for identifying bone shape variations [107] and designing appropriately shaped and sized implants. Second, as explained in the previous section, the conventional process for generating individualized FE models of implanted joints is based on 3D images obtained using clinical CT

scanners. Nevertheless, only a limited number of such finite element models can be generated, because the process of creating individualized FE models involves several laborious and complex steps that cannot be easily automated.
Statistical models of shape and appearance could be used to generate finite element (FE) bone models by alleviating the drawbacks of the conventional model generation process. A limited number of studies have used statistical shape and appearance methods for generation of FE models [105, 106, 108, 109]. These statistical FE models can be used during the design process of new orthopedic implants to find the optimal shape and mechanical properties of implants that are favorable for the majority of the target population [108, 109] or to design patient-specific implants that best serve a specific individual or a portion of the statistical population that have certain anatomical characteristics.

2.4.6 Surgery Planning

Computer aided orthopedic surgery (CAOS) systems including the computer and robotic technology have become increasingly popular for assisting pre-operative surgical planning and intra-operative instrument navigation [45, 110].
CAOS systems, while benefitting from the recent advances in imaging technology, increase the accuracy of intra-operative navigation. One of the approaches used in CAOS systems is to register the 3D model of bony anatomy to intra-operatively obtained patient-specific data [45]. While the intra-operative data may have been obtained by x-ray imaging, surface digitization using ultrasound imaging, or digitization of bone surface and/or anatomical landmarks using a tracked pointer [45], the pre-operative 3D models are generally derived from CT or MRI scans [48, 54]. The high cost and complexity of these imaging modalities and the high level of ionizing radiation to which the patients are exposed during CT scanning has sparked increasing interest in generic 3D bone models built using statistical means [48, 54]. The use of 3D statistical models derived from pre-operative CT or MRI scans results also in a simpler workflow and reduces the time required for surgery [60]. Nonetheless, care should be taken when using statistical models for construction of 3D patient-specific bone models from interventional data. Statistical models may yield inaccurate results if the shape of the patient bone includes local abnormalities that have not been present in the training dataset or if there is too much difference between the patient bone and the bones used for training the statistical model [54].
Studies that use 3D/3D or 3D/2D registration techniques for reconstructing 3D patient-specific models can be classified based on the modeled anatomical structure. Benamur et al. modeled the vertebrae of the scoliotic spine by statistical means to reconstruct 3D patient-specific model from two conventional radiographic images. With the proposed method, it is possible to evaluate 3D deformations of the spine and to plan surgical corrections, if surgery is necessary [111]. Some other researchers have modeled the femur [45, 46, 47, 48, 58, 59, 60, 61, 112] and pelvis [45]. Models of both pelvis and femur or only the femur can be used for computer-assisted total hip replacement (THR) surgeries [45, 46] and for image-guided hip surfacing surgeries [113]. Moreover, distal femur mod-

els are shown to be beneficial for computer-aided total knee replacement (TKR) surgeries or for anterior cruciate ligament (ACL) reconstruction surgeries [47]. An important issue that should be considered for surgical guidance is the need for sufficiently accurate reconstructed models. In most studies, several experiments are conducted to evaluate the efficacy of the proposed methods. Many authors have assessed the capabilities of their methods in terms of surface reconstruction accuracy [45, 46, 48, 59, 60, 112, 113]. As various studies use different training and intra-operative datasets, it is not possible to directly compare the results of different studies with each other in terms of reconstructed model accuracy. Moreover, the applied evaluation methodologies in these studies are not always clear, rendering the comparison even more difficult [55]. In addition, the accuracy of surface reconstruction only represents the averaged error over the whole reconstructed surface. Accuracy results that are evaluated based on clinically significant morphometric parameters may be more relevant and consistent for orthopedic applications. Therefore, Schumann et al. proposed a method to evaluate the reconstruction accuracy in terms of clinically relevant morphometric parameters [114].

2.5 Discussion and Conclusions

The main concepts and applications of SSM and SAM in orthopedics and orthopedic biomechanics were reviewed in this paper. It is clear that there are many potential applications in orthopedics that could benefit from the availability of SSM and SAM. Only a relatively limited number of studies have so far used this great potential, partly because many active researchers within the biomechanics community are not familiar with statistical models of shape and appearance and how they could help them in their research. There are, however, three important issues that need to be addressed before a more widespread use of statistical models can be realized.

First, the shapes and density distributions generated by using a-priori knowledge available in statistical models need to be carefully validated against gold standards, i.e. manually obtained shape and density distribution data. The performance of search algorithms such as ASM and AAM in finding the best fits to the image data needs to be evaluated as well. Finally, the accuracy of methods in constructing 3D bone geometries from 2D and 3D image need to be carefully studied, particularly because such accuracies may differ depending on the anatomical location, type of imaging protocol (e.g. the project plane in 2D imaging, number of 2D images), and the specifications of the registration algorithm.

Second, not many statistical models of shape and appearance are currently available. Statistical models are only available for a limited number of bones such as femur, acetabulum, tibia, and scapula. Even when the models are created for a specific bone, they are often not publicly or commercially available, meaning that the access of the bone research community to statistical models is currently limited.

Third, the use of statistical models often requires describing pathological bone

shapes and density distributions for which very few statistical models are available. For example, osteoporotic bone fracture risk estimation requires availability of statistical models that are developed based on osteoporotic training datasets. It is not clear how accurately can the statistical models based on healthy bones describe the variations of shape and density distribution in pathological bones. Studying cam deformity is another example for which a statistical model that takes the pathological bone shapes into account is needed.

2.5.1 Steps Involved in Building and Using SSM and SAM

The different steps involved in creating and using SSM and SAM were described before. More extensive descriptions of the steps can be found in Appendix A. In this sub-section, we analyze some aspects of those steps.
Among the various techniques available for shape representation, landmarks are used most frequently. As landmarks are easy to understand and to implement, it is expected that the use of landmarks will continue to be popular in the future. As far as the correspondence is concerned, it is currently not clear which technique performs the best in establishing the correspondence. That is partly because standard measures of performance do not exist for assessment of the quality of the resulting statistical models. Davies et al. indicate that group-wise approaches are more suitable compared to pair-wise approaches for establishing the correspondence [115]. The main reason is that group-wise approaches consider all the training samples at the same time and optimization of the cost function is performed in a way that the resulting statistical model shows certain desirable properties such as compactness. It is, however, important to develop standard quality measures that enable comparison between different approaches and to clearly evaluate the advantages and disadvantages of the different correspondence techniques when they are used for building SSM and SAM using the same training database.
There are certain algorithms for automated alignment that are described in Appendix A among which Generalized Procrustes Analysis (GPA) [49, 50, 51, 52, 53] is the most widely used algorithm. This method is efficient and simple to implement and is therefore expected to remain popular for alignment of the training shapes.
PCA is at the heart of SSM and SAM. An important point when using PCA is the linearity assumption. There is no guarantee that this assumption is valid for all training databases. It is therefore important to check linearity before applying PCA. If there is nonlinearity in the data, the linear approximation model will be suboptimal. An alternative decomposition method is Kernel PCA. As indicated by Heimann et al., there has been little interest in Kernel PCA for landmark-based shape modeling [43]. That is partly because non-linear methods are not robust for 3D statistical shape modeling [43].
When trying to fit statistical models to image data, ASM has generally been the preferred method as compared to AAM. That is partly due to the excessive memory usage when using AAM and partly because implementation of ASM is relatively easier. ASM is, however, less powerful in detecting the global minima

when fitting statistical models to image data and may converge to a local minimum due to multiple nearby edges in the image. The ASM contour may therefore be locally detached from the correct boundary locations [116]. With the aim of increasing the robustness of the method against such outliers, several variants of ASM have been developed [43]. The proposed variations need to be systematically compared against the original ASM to clarify which one yields the most accurate results.
Considering 3D model reconstruction from 3D or 2D image data, there is currently not enough data to decide which methods is the most accurate [55]. The success of the applied algorithms seems to be largely dependent on the specifications of the considered problem. Generally speaking, feature-based methods are used for cases where edge and surface detection of the examined object can be performed quite reliably, while intensity-based methods are more common for cases where there is a possible edge detection problem [55].

2.5.2 Accuracy and Reproducibility of the Techniques

Two important topics regarding the techniques presented in this review are 1. the accuracy of previously discussed techniques in reconstruction of 3D bony structures and segmentation of bones in images and 2. the role of human input on reproducibility of the obtained results. In this sub-section, we review both topics and present the accuracy and reproducibility values reported in the literature.
When using the previously discussed techniques for reconstruction of bone shape and/or appearance, it is important to evaluate the accuracy of the reconstruction techniques. Towards that end, the output of the image registration method is normally compared to the ground truth. In the case of predicting 3D shape from 2D images, the predicted 3D shape can be compared with the shapes constructed based on CT or MRI images. Many authors have assessed the accuracy of their methods in terms of point-to-surface distances of every vertex in the predicted model to the surface of the ground truth. In non-pathological bones, the reported mean errors vary between 0.16 mm [112] to 1.68 mm [44]. There are several factors that could affect the accuracy of 2D/3D registration including image noise, image distortion, the number of 2D views and the angle between 2D views. It has been shown that increasing the number of 2D views improves the registration accuracy [52, 117]. It is interesting to note that while adding a second view significantly improves the registration accuracy, inclusion of more views has only slight impact. Furthermore, it has been found that the best registration accuracy using two views can be achieved when the views are approximately orthogonal to each other [117].
As previously mentioned, different studies use different training and test datasets. It is therefore not always possible to directly compare the results of different studies with each other in terms of the reconstructed model accuracy. Moreover, the applied evaluation methodologies in these studies are not always clear, rendering the comparison even more difficult.
The registration and segmentation accuracies of different anatomical structures may be different due to the fact that some bones have more complex shapes

than the others. Humbert et al. reported a mean registration error of 1.3 mm for the femoral shape reconstructed from a single DXA image [52]. The accuracy improved to 0.9 mm when two views (sagittal-frontal) were used [52]. Benameur et al. reported mean registration errors of 0.71 mm and 1.48 mm respectively for scoliotic lumbar vertebrae and thoracic vertebrae [111]. The authors attributed the less accurate results for thoracic vertebrae to the additional complexity caused by the presence of ribs [111]. In general, the reported mean segmentation errors of normal vertebrae vary between 0.47 mm [73] and 0.82 mm [70] for semi-automatic segmentation and between 0.73 mm [74] and 0.93 mm [118] for automatic segmentation. The mean segmentation errors of fractured vertebrae are higher and vary between 0.54 mm [73] and 2.32 mm [74] for semi-automatic segmentation and between 2.17 mm [74] and 2.27 mm [118] for automatic segmentation. As indicated by Mysling et al. [74], possible reasons for less accurate segmentation of fractured vertebrae are the difficulties to model huge variations in the shape of fractured vertebrae and undertraining of the statistical models due to availability of only a few fractured vertebrae in the training databases. It is proposed that increasing the number of fractured vertebrae in the training database could improve the segmentation accuracy of fractured vertebra. Furthermore, multiple statistical models can be built up for various types of vertebral fractures.

It is also important to study the role of human input on the reproducibility of the results obtained using statistical models of shape and appearance. For example, establishing correspondence based on manually placed landmarks is a subjective task. There may be both inter-grader and intra-grader variability in the positioning of landmarks. According to Styner et al., the variations are generally in the range of a few millimeters [119]. That is why it is important to report inter-grader and intra-grader variability when evaluating the accuracy of methods that require user input. In some texts, intra- and inter-grader variability are called repeatability and reproducibility [120]. In a study by Agricola et al., intra- and inter-grader reliabilities in assessing hip shape on radiographs using SSM were evaluated using intraclass correlation coefficient (ICC) [3]. Intra- and inter-grader reliabilities of the shape modes that were associated with total hip replacement respectively varied between 0.01 and 0.98 and between 0.43 and 0.97 [3]. In another study, Gooβen et al. assessed the quality of manually segmented models that were used as the ground truth in evaluation of segmentation accuracy of their proposed segmentation technique [121]. In their study, inter-grader variability was evaluated based on manual annotation of 20 legs by five experts. Gooβen et al. point out that this variability represents the lower bound for the measurable accuracy of any automatic segmentation method. The reproducibility is reported to be 0.49 mm for the hip, 0.43 mm for the knee, and 0.35 mm for the ankle joint. The mean segmentation error of their proposed technique is found to be 0.59, 0.47, and 0.37 mm for the hip, knee and ankle joints, respectively. User interaction may also be necessary in initialization of a search process, thereby affecting the reproducibility of the segmentation techniques. In a study by Guglielmi et al. [122] that deals with semi-automated annotation of osteoporotic vertebrae, multiple initialization on the same image were performed to

evaluate the reproducibility of the segmentation technique. They chose a total of 20 randomly selected test images and simulated multiple user initializations by perturbing the initialization points with a radial offset computed from a random Gaussian distribution. The mean reproducibility error is reported to be $2.6 \pm 1.3\%$ of vertebral width [122]. In another study, Roberts et al. assessed the operator precision in initialization of AAM search algorithms [70]. Multiple perturbed initializations were performed for each test image and the mean precision of the AAM solutions was calculated based on the mean point-to-curve error from the mean shape. The mean precision is reported to be 0.25 mm for lumbar radiographs [70]. In another study of Roberts et al., the mean precision of AAM solutions in semi-automatic determination of vertebral shape from DXA images is reported to be 0.14 mm [123].

2.5.3 Limitations and Challenges

One of the challenges in application of SSM and SAM is that individuals with different ethnic backgrounds may have different bone morphologies. SSM and SAM are not necessarily capable of describing shape and appearance variations that do not exist in their training databases. It is therefore unclear whether SSM and SAM created using training datasets of one single ethnic group could describe the shape and appearance variations of individuals from other ethnic backgrounds. In an exploratory study, Mahfouz et al. [124] used SSM to study shape differences between the knees of different ethnic groups. They found different mean shapes in African-American and East Asian populations as compared to the Caucasian population. Furthermore, differences were observed between the female and male populations of different ethnic groups. One strategy to overcome such ethnic differences in statistical modeling of shape and appearance is to create separate SSM and SAM for different ethnic groups. Creating separate SSM and SAM is probably one of the most accurate approaches for capturing the shape and appearance variations between ethnic groups but is more costly. Another approach is to create SSM and SAM based on a mixed training dataset that includes bones from different ethnic groups. It is not clear what the optimal way of creating the training database in the second approach is nor is it clear how accurately could the SSM and SAM created using the second approach describe the shape and appearance variations of individuals with different ethnic backgrounds.

Another point is that SSM and SAM may need to be modified over time to capture secular trends in bone morphology. There is currently no data in the literature as to how the secular trends can be captured using SSM and SAM. The problem of segmenting bones with different ages is, however, somewhat similar to the problem of capturing secular trends in bone morphology. In a recent study [121], it was demonstrated that ASM could capture growth-related changes in bone morphology and segment the bones of adult and child patients from digital radiographs with accuracies of 0.48 mm and 0.64 mm, respectively.

2.5.4 Potential Applications and Future Research

One can think of many possible applications for SSM and SAM. One important application is FE modeling of bones and joints. Since every patient has a certain anatomy within the population described by SSM and SAM, one could precalculate the results of FE models for different values of mode parameters and use a look-up table approach for obtaining the results of FE analysis, thereby eliminating the need for presence of high-end computational facility in every hospital and reducing computation time. It should be, however, noted that the results of FE models generated based on SSM and SAM need to be corroborated against the simulations results of the FE models generated using standard segmentation techniques. The material properties of bones can be mapped based on the empirical relationships between the appearance and bone density on the one hand and the relationships between bone density and elastic properties of bones on the other hand. Whether or not the SAM provides accurate appearance and, thus, density distribution is dependent, among other factors, on the training population and whether it can well represent the individual for whom the patient-specific FE model is being built.

The other possible area of application is grafting of large bony defects. Modern additive manufacturing (3D printing) techniques have enabled production of highly porous patient-specific bone substitutes that are considered very promising biomaterials for bone substitution [125, 126, 127]. However, one needs an accurate representation of bone shape so that the missing part of bone can be manufactured using additive manufacturing techniques. In such applications, ASM could be used for estimating the original shape of bone, for example, by fitting the SSM to the intact parts of the bone.

Many possible applications can be envisioned for SSM and SAM in diagnosis and treatment of pathologies that involve bone deformities. Examples of such pathologies are cam deformities, genu valgum, varus deformities, etc. In such application, SSM and SAM can be used for two main purposes. Firstly, one could use SSM or SAM to estimate how the undeformed anatomy of the patient might have looked like. Secondly, new SSM and/or SAM can be trained using the data from pathological population to study, among others, the main modes of deformity, the shape changes caused by the deformity, and the modifications caused by deformities in musculoskeletal loading.

The potential applications of SSM and SAM may increase as imaging modalities improve. For example, it may be possible to enrich SSM and SAM using imaging data acquired at the micro-scale such as data acquired using micro-CT regarding the shape, orientation, and distribution of individual trabeculae. That kind of information could, for instance, be used for representing anisotropy in SSM and SAM.

In summary, statistical model of shape and appearance have been already used in different areas of in bone-related research. However, more efforts in generation of statistical models of shape and appearance and testing their accuracy is needed to allow for better exploitation of their potential in orthopedics research, diagnosis, and treatment. That requires intense collaboration between engineers,

radiologists, and orthopedic surgeons.

3

Statistical Shape and Appearance Models for Fast and Automated Estimation of Proximal Femur Fracture Load Using 2D Finite Element Models

Published as:
N. Sarkalkan, J.H. Waarsing, P.K. Bos, H. Weinans, A.A. Zadpoor. Statistical shape and appearance models for fast and automated estimation of proximal femur fracture load using 2D finite element models. Journal of Biomechanics. 47(2014) 3107-3114.

3.1 Introduction

Every year, more than 8.9 million osteoporotic fractures occur around the world with one third taking place in Europe and accounting for 2 million disability-adjusted years [67]. Osteoporotic fractures whose prevalence increases with age, mostly occur at hip, spine and distal forearm [72]. Hip fractures are among the most serious and necessitate hospitalization and surgical intervention. Early identification of individuals who are at increased risk of fracture is essential, because fracture prevention measures such as medicinal treatment could then be taken [128]. In clinical settings, dual-energy x-ray absorptiometry (DXA) is the most widely used densitometry technique to measure areal bone mineral density (aBMD). Measured aBMD is used for diagnosis of osteoporosis and for fracture risk prediction [67, 63]. However, aBMD alone cannot explain all low energy fractures [68]. That is because bone mass is not the only parameter that determines bone strength. Other factors such as bone shape, the distribution of bone mass, and loading conditions that should be also considered [128, 129].
Several attempts have been made to take other risk factors such as bone shape and density distribution into consideration [130] including development of patient-specific finite element (FE) models [131]. These models are powerful tools to predict bone strength in vivo, because they include information about bone shape, bone tissue material properties and external loading conditions. Either 2D or 3D images could be used for generating patient-specific FE models [99]. While 3D computed tomography (CT) images are more accurate and informative, they are more costly to obtain and expose the patient to higher doses of ionizing radiation.
The currently used process for estimating fracture risk using manually generated patient-specific FE model includes segmentation of the proximal femur from 2D or 3D images, assignment of material properties, generation of mesh, application of loads and boundary conditions, running the simulations, and post-processing the simulation results. Such a manual process is costly, time-consuming, and laborious as compared to diagnosis based on aBMD values. This has so far delayed the clinical application of manually generated patient-specific FE models. The aim of the current study was to develop automated procedures for generation of patient-specific FE models, thus enabling fast and inexpensive fracture risk estimation. First, a statistical shape and appearance model (SSAM) of the proximal femur is built using DXA images. Then, an active appearance model (AAM) [132] is used for automated segmentation of bone from new DXA images. The extracted information is then used for automated generation of FE mesh and density-based assignment of mechanical properties. Automated procedures also assign loads and boundary conditions, run FE simulations, and post-process the simulations results to estimate fracture load. The results of the automated process are benchmarked against those of a similar manual process.

3.2 Materials and Methods

3.2.1 Training and Evaluation Data

A set of 87 anonymous DXA scans (Lunar Prodigy, GE Healthcare, USA) of the proximal femur (left or right) of patients screened for osteoporosis following a low fracture energy was randomly divided into two groups: 70 DXA scans (56 female and 14 male, 61.5 ± 12.6 years) for generating statistical models and another 17 DXA scans (12 female and 5 male, 61.8 ± 10.7 years) for independent evaluation of the process. The DXA scans were taken from patients who already had a fracture that made clinicians suspicious of osteoporosis. This population was chosen, because it best represents the population for which the methods developed in this study will be ultimately applied. The Min-Max filter method described in [133] was used for enhancing the contrast of all images.

3.2.2 Generation of Statistical Shape and Appearance Model

A statistical shape model (SSM), a statistical appearance model (SAM), and a combined SSAM were created using the training data. For every proximal femur of the training dataset, 70 landmarks were manually positioned along the contour of the femur. As the lesser trochanter was not visible on several images, the lesser trochanter region was not considered. All shapes were then aligned in a common coordinate system using the generalized Procrustes analysis. Translational effects were removed by positioning all the shapes in a way that their centroids were at the origin [134]. Scaling and rotational effects (see Appendix B) were removed using the methods described in [135, 136]. The mean shape, $\bar{\mathbf{x}}$, was then calculated by averaging the position vectors of all landmarks. Principal component analysis (PCA) [137] was performed to calculate the principal modes of variation, and their respective variances, by applying singular value decomposition (SVD) on the aligned data [43]. Subsequently, a shape instance, $\mathbf{x}$, can be described by adding contributions of most important modes of variation to the average shape, $\bar{\mathbf{x}}$:

$$\mathbf{x} = \bar{\mathbf{x}} + \boldsymbol{\phi}_s \boldsymbol{b}_s \tag{3.1}$$

where $\boldsymbol{b}_s$ is a set of shape parameters and $\boldsymbol{\phi}_s$ represents the matrix of ordered eigenvectors (i.e. principal modes of variation). The number of modes was increased until the ratio of the accumulated variance to the total variance, r, reached 0.98 [43].
The SAM was created by applying the following three steps: capture of pixel information, normalization of the data to discard the global illumination effects, and modeling the normalized texture variation [138]. A piece-wise affine warp based on the Delaunay triangulation was used to capture texture information [138]. Each training example was warped to the mean shape and the spurious texture variations that might rise due to shape differences were removed [139]. The intensity information from the shape normalized image was sampled using bilinear interpolation to form a texture vector, $\mathbf{g}$. The texture vector was then

normalized [140]. Subsequently, the texture variation was extracted by applying PCA. Any texture instance, $\mathbf{g}$, can be created by deforming the mean texture, $\bar{\mathbf{g}}$, by a linear combination of eigenvectors [138]:

$$\mathbf{g} = \bar{\mathbf{g}} + \boldsymbol{\phi}_g \boldsymbol{b}_g \tag{3.2}$$

where $\boldsymbol{b}_g$ is a set of texture parameters and $\boldsymbol{\phi}_g$ represents the matrix of ordered eigenvectors. The number of appearance modes increased until the model could explain 98% of appearance variation. Since bone shape and density distribution are related through bone tissue adaptation process [97, 98], the principal modes of shape and texture variation are not independent and need to be combined [139] (see Appendix B).

3.2.3 The Search Algorithm

AAM was used to fit the generated SSAM to new unseen DXA images. The method described by Cootes et al. [140] was used to decrease the complexity of problem. The presented approach assumes that there is a linear relationship between the parameter updates (combined model and pose parameters), $\delta \boldsymbol{p}$, and the texture residuals, $\boldsymbol{r}(\boldsymbol{p})$, over the entire search [43, 141]:

$$\delta \mathbf{p} = -\boldsymbol{R}\boldsymbol{r}(\boldsymbol{p}) \tag{3.3}$$

where $\boldsymbol{R}$ is the derivative matrix. Leave-one-out tests were conducted on all images in the training datasets, meaning that the generation of the SSAM was repeated 70 times every time leaving one image out of the training dataset to see how the AAM can fit the SSAM to that image. The point-to-curve error [138] was used for evaluating the accuracy of the fit. The new unseen images were also segmented manually to evaluate the performance of the automated segmentations. The intra-observer segmentation variability was determined by asking three observers to segment 10 images.

3.2.4 Generation of Patient-Specific FE Models

For every DXA scan in the second group (17 scans), patient-specific 2D FE models were generated both using the automated and manual procedures. The user initiated the automated procedure by moving the mean shape over the previously unseen DXA image and fixing the position of the mean shape with a mouse click. The automated process then found the best fit of the SSAM to the DXA image, generated the mesh, assigned material properties, loads, and boundary conditions, and performed post-processing (see Appendix B). The meshing tool available in Abaqus CAE was employed to discretize the geometry using 6 node modified second-order plane-stress triangle elements (element CPS6M). The thickness of the element (see Appendix B) was defined as described in [128]. A mesh convergence study was performed to determine the minimum number of elements needed for accurate FE analysis according to which the von Mises stress values and displacements of 6 selected elements (Fig. 3.1) converged within $\approx 1\%$ when c.a. 28000 elements were used.

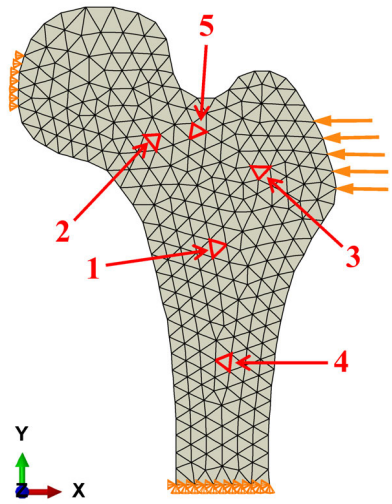


Figure 3.1: The loads and boundary conditions used in the current study together with the 5 elements chosen for convergence study.

The number of elements was therefore fixed around that number for the remainder of the study. Material properties of each element were defined by first averaging the gray values of pixels lying within the element. The mean gray values were used to calculate aBMD values that were then converted to vBMD values [128], and, subsequently, to ash density. Apparent density was defined using the ratio of ash density to apparent density [131]. The relationship proposed by Morgan et al. [142] was used to relate apparent density and Young's modulus as it has been found to be the most accurate relationship [143]. The loads and boundary conditions were defined to simulate a fall on the greater trochanter [128]. The impact force defined by body weight and height of individual was applied to the greater trochanter. The femoral head and the distal shaft were, respectively constrained in x direction and x and y directions (Fig. 3.1). The FE models were solved using a nonlinear implicit solver (Abaqus Standard 6.10). A failure criterion based on von Mises stress [130] was used. The failure load was defined as the load which causes failure of contiguous elements with an area of 25 mm^2[128].

3.3 Results

The first three modes of variation could explain $\approx$ 70% of shape variations and $\approx$ 35% of appearance and combined shape and appearance (Figs. 3.2 - 3.4). To explain 98% of variations in shape, appearance, and combined shape and appearance, respectively 16, 50, and 44 principal modes were needed (Figs. 3.2 - 3.4). Based on the leave-one-out tests of the training dataset, the mean point-to-curve error was 1.25 ± 0.65 mm (95% confidence interval) (Fig. 3.5). Segmen-

tation errors were larger close to the femoral head and greater trochanter area (Fig. 3.5). The point to curve error was 1.42 ± 0.75 mm (95% confidence interval) for segmentation of 17 DXA scans not used in the training (Fig. 3.5). The largest segmentation errors were observed in the femoral head area (Fig. 3.6). In comparison, the intra-observer variability in manual segmentation of DXA scans measured in terms of point to curve error was 1.03 ± 0.48 mm (95% confidence interval).
For the 17 DXA scans, the predicted fracture loads determined using manually and automatically generated models were respectively 3871 ± 933 N and 3804 ± 850 N. In addition, the fracture loads estimated using manually generated FE models by three operators were 4061 ± 1252 N, 4257 ± 1036 N and 3846 ± 972 N.

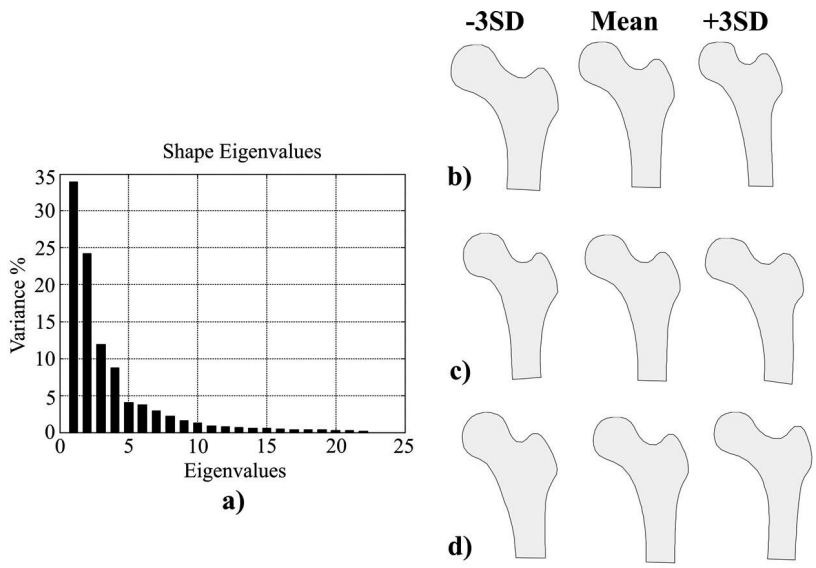


Figure 3.2: The percentage of shape variation described by different shape modes (a) as well as the first three modes of shape variation (b-d).

The fracture loads calculated using both types of models were very strongly correlated with a coefficient of determination, R, around 0.82 (Fig. 3.7) and a relative percentage difference of $8.8\% \pm 6.5\%$ (considering manual segmentation as the ground truth). Repeated measures ANOVA showed no significant differences between the fracture loads calculated using FE models manually generated by three independent observers and those calculated using the automatically generated FE models ($p > 0.05$).
The femur for which the difference between both fracture loads was closest to the mean value of the fracture load difference, i.e. 8.8%, was chosen to further examine the correspondence between manually and automatically generated FE models (Figs. 3.8 and 3.9). In general, distributions of density and von Mises

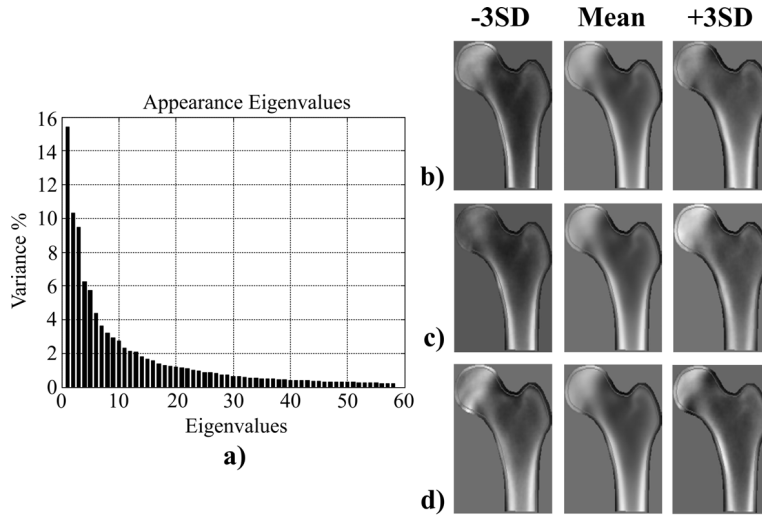


Figure 3.3: The percentage of appearance variation described by different appearance modes (a) as well as the first three modes of appearance variation (b-d).

stress were similar between manually and automatically generated FE models (Figs. 3.8 and 3.9). Considering all test scans, the relative percent differences between manually and automatically generated models (averaged over the entire femur geometry) were 1.7% $\pm$ 0.5% for density distribution and 5.7% $\pm$ 1.0% for von Mises stress.

3.4 Discussion

Currently, aBMD values that are known to lack sufficient sensitivity and specificity [144] are used for fracture risk estimation in clinical setting. The alternative approach presented here uses automatically generated patient-specific FE models based on DXA images. This type of models are shown to be promising tools for accurate fracture risk estimation of the proximal femur [128], and do not subject the patients to additional ionizing radiation or substantially increase the required cost or time. The time for the automated procedure to complete is approximately 7 min. The automated segmentation process is completed within a few seconds on a standard PC, and requires only a few seconds of user interaction for initialization. Another advantage of the proposed framework is that non-expert users could use it. Moreover, the proposed framework could be further improved, for example, by incorporating previously proposed initialization techniques [43] that could make it fully automated.
The mean point-to-curve error of the automated segmentation procedure was 1.25 $\pm$ 0.65 mm for the leave-one-out test and 1.42 $\pm$ 0.75 mm for the 17 validation DXA scans. This level of error is only slightly larger than that of the intra-

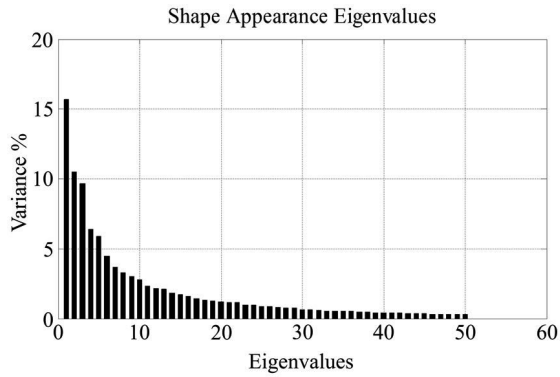


Figure 3.4: The percentage of shape and appearance variation described by different shape and appearance modes.

observer variability in manual segmentation of DXA scans, i.e. 1.03 ± 0.48 mm. The maximum point-to-curve error observed in our automated segmentations, i.e. both leave-one-out test of the training material and 17 validation test DXA scans, was 2.2 mm.

A similar study was performed very recently [80] where a similar technique was applied in osteoarthritis research. In that study, AAM was used for segmentation of the femoral shape from x-ray images. When the results of 1% worst performing images were excluded, they could achieve a mean error of 0.9 mm [80]. However, their maximum error increased to 3.3 mm when worst performing images were not excluded from the analysis [80]. In their study, [80] did not need the manual initiation of the process and that could be considered an advantage particularly when the technique is applied in analysis of large clinical databases. Moreover, they had access to a much larger database of osteoarthritic patients to generate and examine their SSAM.

It should be noted that automated segmentation of femoral shapes is more difficult when osteoporotic patients are involved. That is because the density distributions of osteoporotic patients may significantly deviate from that of the general population in the ways that are not necessarily well captured using SSAM. One novelty and advantage of the current study is that the SSAM used here is based on training images originating from osteoporotic patients. The SSAM created using those images should therefore include more information regarding the pathological density distributions observed in osteoporotic patients and may therefore be more capable in segmenting the femoral shapes of osteoporotic patients as compared to the SSAM built using x-ray images originating from general population or osteoarthritic patients. To speed up the model generation phase and to improve the segmentation accuracy, canonical correlation analysis as described by Donner et al. can be used for construction of the derivative matrix $\boldsymbol{R}$ [43, 145]. This approach was found to be more accurate in prediction of parameter updates as compared to the numeric differentiation applied in this study.

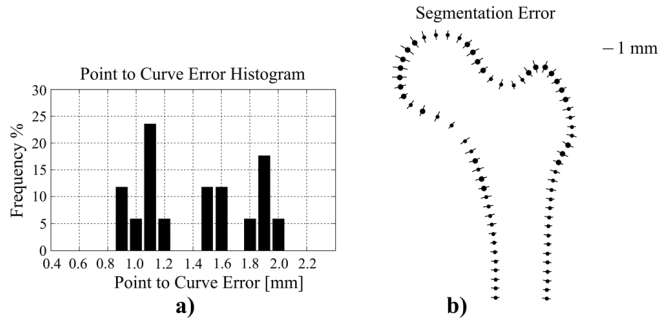


Figure 3.5: The histogram of point-to-curve error in segmentation of femoral shapes measured using the leave-one-out test of the training dataset (N=70) (a) and the corresponding spatial distribution of error (b). The size of the marker shows the mean error value while the length of the crossing line shows the standard deviation of the point-to-curve error values.

The fracture loads estimated using manually generated and automatically generated FE models (i.e. 17 FE models) were, respectively 3871 ± 933 N and 3804 ± 850 N with an absolute difference of $8.8 \pm 6.5\%$ (Fig. 3.7). According to the repeated measures ANOVA, the fracture loads predicted using the manually generated FE models by three observers were not significantly different from those predicted using automatically generated FE models. The differences between manually and automatically generated patient-specific FE models originate from the segmentation process. Since the segmentation error in the automated process is not much different than the intra-observer variability, it is expected that the estimated fracture loads using automatically generated FE models do not differ much from the fracture loads predicted using manually generated FE models.
Amin et al. found in their 3D FE study that the femoral fracture load is around 2500 N (2577.7 ± 1172.2 N) for women and 3200 N (3217.4 ± 1270.9 N) for men with prevalent osteoporotic fractures [146]. The corresponding values for the men and women with no-fracture were respectively around 4000 N (3866.9 ± 1186.6 N) for women and 4600 N (4602.0 ± 1287.4 N) for men [146]. The proximal femur fracture loads estimated in the current study are within the range of failure loads [146]. On the other hand, estimated fracture loads for both segmentation cases were higher compared to those reported by Naylor et al. [128]. The difference may originate from the differences between datasets used in both studies. In [128], only women with a mean age of 82 years (ranging between 75 and 95 years) were considered. The dataset used here includes men and women with an average age of 61.82 ± 10.7 years. As indicated by Lochmüller et al., although no significant changes in bone strength were observed in men, the failure loads decreased in females with an increasing age [147]. Furthermore, failure loads in men were higher than those of women [146, 147].
When the whole proximal femur or four specific regions of the femur, i.e. neck,

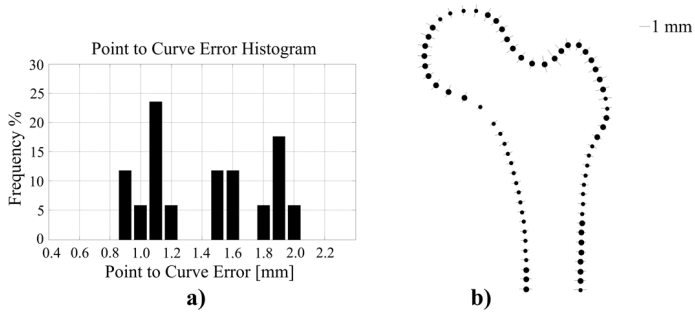


Figure 3.6: The histogram of point-to-curve error in segmentation of the 17 femoral shapes not used in the creation of the SSAM (a) and the corresponding spatial distribution of error (b). The size of the marker shows the mean error value while the length of the crossing line shows the standard deviation of the point-to-curve error values.

trochanter, shaft and head, were considered, no associations could be observed between the segmentation errors and relative difference between the failure loads. We noticed that even small errors, e.g. 1.06 mm, in segmentation process might in certain cases result in relatively large errors, e.g. 24.1%, in prediction of fracture loads. This is consistent with the findings of another study [148] that found small errors in shape reconstruction result in larger errors in von Mises stress values.

One of the limitations of this study is the use of DXA scans of patients screened for osteoporosis. As the dataset does not include DXA scans of healthy individuals, evaluation of the specificity and sensitivity of the proposed method in fracture risk estimation will be impossible. Moreover, the dataset does not incorporate non-fracture controls nor is there any comparison with experimentally determined in vitro fracture loads. The next step would be to apply the proposed technique to a large dataset including DXA scans of healthy individuals and osteoporotic patients for whom long-term follow-up data is available. Only then it would be possible to assess the utility of the proposed techniques in clinical assessment of fracture risk. It should be, however, noted that since the results of the proposed method are not much different from those of manually generated patient-specific FE models, the utility of the proposed techniques should be similar to those of already available manually generated FE models. Moreover, it would be important to validate the calculated strain distributions against full-filed strains measured using optical methods such as digital image correlation (DIC) [149, 150] during in vitro mechanical testing experiments [151, 152, 153]. 2D FE models have limitations in representing the distribution of material properties and force components in the third direction. To avoid such limitations,

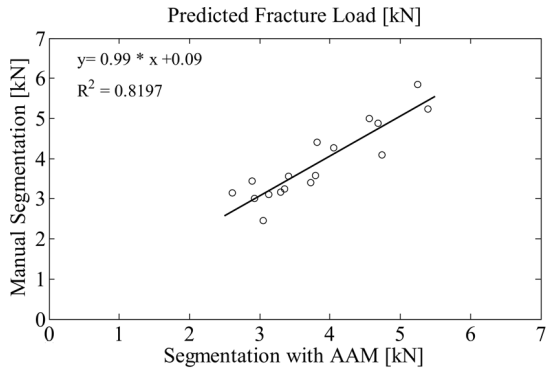


Figure 3.7: Fracture loads estimated using the manually segmented FE models vs. the fracture load predicted using automatically segmented FE models.

3D FE models could be built based on 3D images and patient-specific loads estimated using musculoskeletal [154, 155, 100] or mass-spring-damper models [103, 104, 102]. 3D FE models [93, 14] require 3D imaging modalities or use of search algorithms that match 3D SSAM with 2D images [64]. In contrast, 2D FE models do not need additional imaging as compared to current clinical practice. Various structures, density patterns and risk factors are observed for cervical and trochanteric fractures [156]. In future studies, it would be useful to investigate methods that could be used to distinguish between the different types of hip fracture. This might be helpful in development of strategies for prevention of hip fractures and for patient-specific treatment.

Due to the overlap of the femoral head with the pelvis in DXA scans, BMD is overestimated at the femoral head. As there is an increase in the predicted strength of the 2D proximal femur model, overestimation of BMD at the femoral head would influence the estimated fracture load. In future studies, it is important to take the overestimation of BMD at the femoral head into account.

In summary, an automated procedure based on SSAM and AAM was proposed in this study for generation of patient-specific FE models and estimation of proximal femur fracture load. It was shown that the mean error of the proposed one-click segmentation technique is between 1.2-1.4 mm and is only slightly larger than the mean intra-observer variability, i.e. 1.0 mm. The fracture loads predicted using the manual and automated models were strongly correlated, i.e. $R^2 = 0.82$, and showed a mean difference of 8.8%. There were no significant differences between the fracture loads estimated using manually and automatically generated models. The required time, cost, and radiation associated with the proposed technique are comparable with those of the currently used clinical methods.

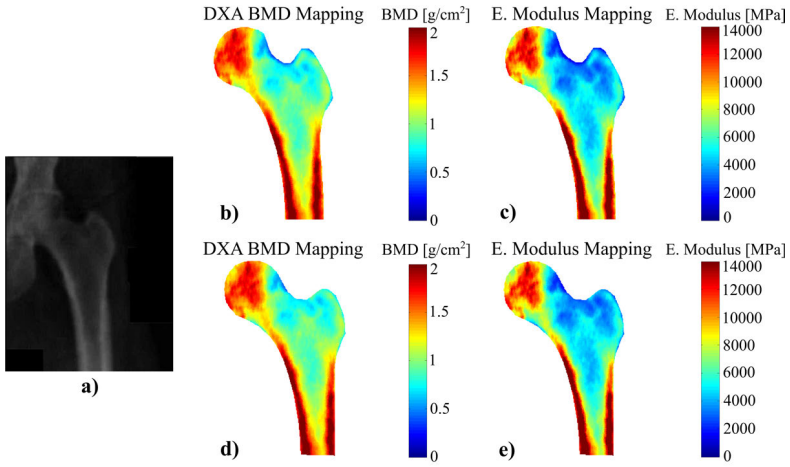


Figure 3.8: Normalized DXA image (a), BMD mapping according to manual (b) and automated segmentation (d), and mapping of the elastic modulus according to manual (c) and automated segmentation (e).

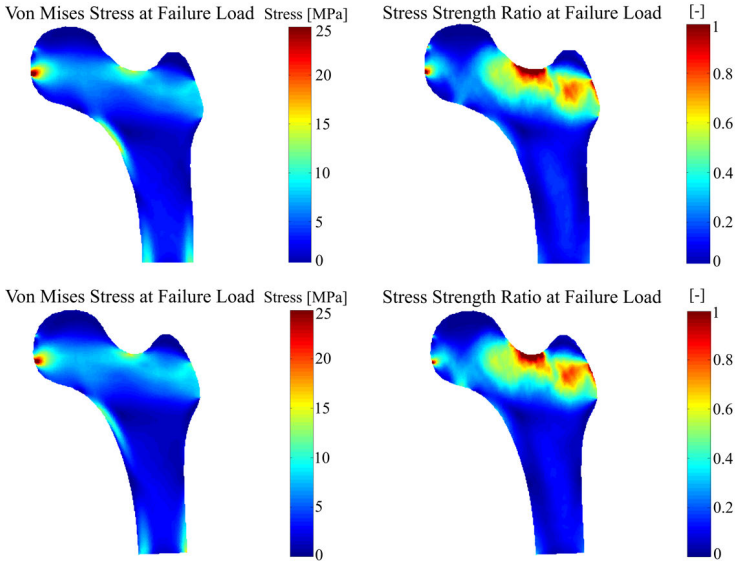


Figure 3.9: The distribution of von Mises stress at the failure load according to manual (a) and automated segmentation (c) and distribution of stress strength ratio at the failure load according to manual (b) and automated segmentation (d).

4

Bone Shape Difference Between Control and Osteochondral Defect Groups of the Ankle Joint

Published as:
N. Tümer, L. Blankevoort, M. van de Giessen, M.P. Terra, P.A. de Jong, H. Weinans, G.J. Tuijthof, A.A. Zadpoor. Bone shape difference between control and osteochondral defect groups of the ankle joint. Osteoarthritis and Cartilage. 24(2016) 2108-2115.

4.1 Introduction

The ankle (talocrural) joint is the second most common site for osteochondral defects (OCDs), exceeded in frequency only by the knee [41]. The true incidence of OCDs in general population is currently unknown [157]. In a study following US military personnel during a 10 years period (1999 to 2008), the average incidence of OCDs of the talus has been stated to be 27 per 100000 people [157]. In another study, it has been reported that 15-25% of ankle injuries, from which approximately 1 in 10000 persons per day suffer, result in OCDs [37]. In the ankle, tibia OCDs have been stated to be rare as compared to the incidence of talus OCDs. In a study consisting of 428 ankles with OCDs, the medial talar dome lesions (i.e. 269 OCDs, 62%) have been found to be more common in comparison to lateral (i.e. 143 OCDs, 34%) and central (i.e. 16 OCDs, 4%) talar dome lesions [38]. OCDs can lead to chronic ankle joint pain, decreased level of patient activity, and osteoarthritis (OA) [157]. To take preventive measures, it is essential to know risk factors that induce OCDs in the talocrural joint. Although OCDs of the ankle joint are extensively studied, its etiology is not yet fully understood [33]. Several factors including "local avascular necrosis, systemic vasculopathies, acute trauma, chronic microtrauma, endocrine or metabolic factors" [37], articular incongruency [38], malalignment of the lower limb [38] and genetic predisposition [38] have been proposed to contribute to the risk of developing an OCD. Trauma has been extensively stated to be the main cause of ankle OCDs [158]. 93% - 98% of lateral and 61% - 70% of medial dome lesions were found to be related to trauma [158]. The lateral lesions are thought to be caused by a shear mechanism, while medial lesions are results of torsional impaction and axial loading of the talocrural joint [33]. Up to now, no study has investigated possible relation between bone shape and OCD risk in the ankle.

There is a growing evidence [1, 2, 3] suggesting that a slight difference in joint mechanics driven by variability in joint morphology may be a risk factor for development and progression of joint diseases. Variations in the acetabular cup of the hip joint (i.e. acetabular dysplasia) [2], the femoral head-neck junction (i.e. cam type impingement), and the acetabular depth (i.e. pincer type impingement) have been found to be associated with OA [2]. Moreover, significantly different shapes of the distal femur and the proximal tibia have been observed between subjects with and without anterior cruciate ligament injuries [1].

Taking all findings into account, it seems plausible that relatively subtle differences in ankle morphology may be associated with the risk of sustaining an OCD. In this study, we hypothesize that the morphologies of the talus and the distal tibia are different between subjects without an OCD and patients with an OCD. Therefore, we aimed to describe the complex geometries of the ankle bones together with shape variances within a population using statistical shape modelling (SSM) technique [159, 43] and to quantitatively compare the morphology of the talus and the distal tibia between an OCD group and a control group.

4.2 Materials and Methods

Two three-dimensional (3D) SSMs were built by segmenting bones from computed tomography (CT) images of the mixed data of subjects without an OCD and patients with an OCD, each corresponding to either talus or distal tibia. Using those SSMs, bone shapes of control and OCD groups were quantitatively compared (Fig. 4.1).

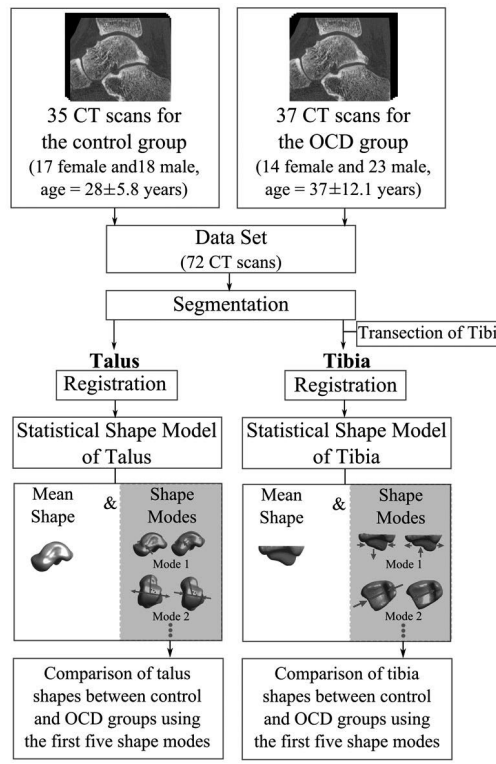


Figure 4.1: A general scheme to compare the morphology of the talus and the distal tibia between the OCD group and control group. 35 CT scans for the control group and 37 CT scans for the OCD group were collected from medical centers. Based on the segmentations of the talus and tibia from all CT scans (i.e. 72 CT scans), two SSMs (i.e. SSM of talus and SSM of tibia) have been constructed, being independent from each other. Using shape parameters for the first five modes of bone shape variation, bone shapes of the OCD and control groups were quantitatively compared.

4.2.1 Image Acquisition

Seventy two anonymized CT scans (i.e. 35 CT scans for the control group: 17 female and 18 male, age = 28 ± 5.8 years; and 37 CT scans for the OCD group: 14 female and 23 male, age = 37 ± 12.1 years; there exists a significant difference between the mean ages of two groups based on unpaired t-test, $p = 0.0006$), of both left and right ankles, were collected from the Academic Medical Center (AMC, Amsterdam, The Netherlands) and Utrecht Medical Center (UMC, Utrecht, The Netherlands). Control and OCD groups CT scans were checked by experienced radiologists to confirm absence or presence of OCDs in the talocrural joint, respectively. In 5 of 37 OCD group CT scans, bilateral lesions were present in the talocrural joint. Lateral, medial, and central talar dome OCDs were observed in 11, 29, and 2 CT scans of OCD group, respectively.
CT images collected from the AMC were acquired using Philips MX-8000 multidetector CT scanner (Philips Medical Systems, Best, The Netherlands). The acquisition parameters were: effective dose 150 mAs/slice, rotation time 0.75 s per 360°, pitch 0.875, slice thickness 0.6 mm, and ultra-high-resolution mode. Tomographic reconstructions were made with a field of view of 154 mm, a slice increment of 0.3 mm, and a matrix of 512×512 pixels. The voxel sizes were 0.3 mm $\times$ 0.3 mm $\times$ 0.3 mm. CT images collected from the UMC were acquired using a 40 detector row CT scanner (LSO PET 40-slice CT scanner, Siemens Healthcare, Erlangen, Germany). The acquisition parameters were: collimation 40×0.6 mm, tube voltage 120 kVp, effective dose 40 mAs, rotation time 0.5 s, pitch 0.8, slice thickness 1.5 mm, care DOSE4D, and Care kV dose modulation. Tomographic reconstructions were made with a field of view of 500 mm, and a matrix of 512×512 pixels. The B60f sharp reconstruction filter was used. The voxel sizes were 1.0 mm $\times$ 1.0 mm $\times$ 0.75 mm.

4.2.2 Segmentation

To build the SSMs, the unilateral talocrural joint was segmented from each CT scan. A second user re-segmented three of the joints to evaluate inter-observer variability and its effect on reproducibility of the method used in generation of SSMs. For the OCD group, the contralateral unaffected talocrural joint was segmented assuming bilateral symmetry of the ankle bones [160, 161]. In the cases with bilateral lesions (i.e. in 5 subjects), the talocrural joint that had been less affected by the lesions was selected. The main reason behind these selections was to minimize any inaccuracies that might arise in estimation of the complete shape of the bones, if some part of the bone surface was missing due to lesions. The bones were segmented using Mimics (version 14.01, Materialise, Leuven, Belgium). During the reconstruction of the 3D bone models, a smoothing factor (i.e. 0.5 with a smoothing iteration of 1) was applied. Smoothing effects for each bone were visually checked to ensure proper definition of the contour of the bones. Using the same software, triangulated bone surfaces were extracted from the segmentation results. As the dataset used in this study was a mixture of left and right-side ankles, left side ankles were mirrored in the sagittal plane.

4.2.3 Transection of Tibia

The entire volume of the tibia was not visible on all CT scans, necessitating transection of this bone in the axial plane. For each tibia, its subchondral bone surface was automatically determined using a custom-made code developed in Matlab (Matlab R2013b, The Mathworks, Inc., Natick, MA) (see Appendix C) (Fig. 4.2). The vertices of the triangles located on the subchondral bone surfaces were extracted (Fig. 4.2). As variations due to differently oriented coordinate systems of CT scans might be seen during transection of the bones and might contribute to bone shape variations, each tibia was scaled and aligned with respect to its subchondral bone surface before transection (Fig. 4.2). For alignment and scaling, an unbiased registration algorithm was used [162]. Aligned and scaled bones were transected at a level that was defined by the bone for which the smallest volume was visible on the CT-scans (Fig. 4.2). Transected tibia surfaces were closed by automatic addition of points (Fig. 4.2) (see Appendix C). Subsequently, they were triangulated using a custom-made Matlab code developed based on the Crust algorithm [163] (Fig. 4.2).

4.2.4 Registration

Position, orientation, and size differences among all bones of the same type were minimized by aligning and scaling them using an unbiased registration [162] (see Appendix C). This ensures that the remaining differences could be attributed to shape variations only. During the registration process, each bone was represented with 2000 points, which were randomly chosen from the entire set of surface points in a way that sub-selected points were uniformly distributed over the bone surface. The total number of points used to describe each bone shape during the registration process (i.e. 2000 points) and few other important parameters in the registration algorithm required to be set by the user were determined in a similar way as described by van de Giessen [162] (see Appendix C).

4.2.5 Statistical Shape Model(s)

The SSMs were constructed based on the vertices of the triangulated surfaces of the bones [164, 165]. Corresponding points (i.e. 26842 and 16246 points for talus and tibia models, respectively) on all registered bone surfaces were automatically determined [166]. Principal component analysis (PCA) was applied to the covariance matrix of the data vectors (i.e. sizes of data vectors were $26842 \times 3 \times 72$ and $16246 \times 3 \times 72$ respectively for talus and tibia), which include the 3D coordinates of the corresponding points for each bone sample. This gave the eigenvectors (i.e. modes of variation) and eigenvalues (i.e. variance) for each bone type. This implies that new or existing bone shapes can be generated with a weighted summation of the eigenvectors and the mean bone shape:

$$\mathbf{x} = \bar{\mathbf{x}} + \boldsymbol{\phi} \boldsymbol{b} \tag{4.1}$$

in which $\boldsymbol{x}$ and $\bar{\mathbf{x}}$ are vectors respectively describing the shape of any (new) bone and the mean bone shape. $\boldsymbol{\phi}$ is a matrix describing the modes of shape variation

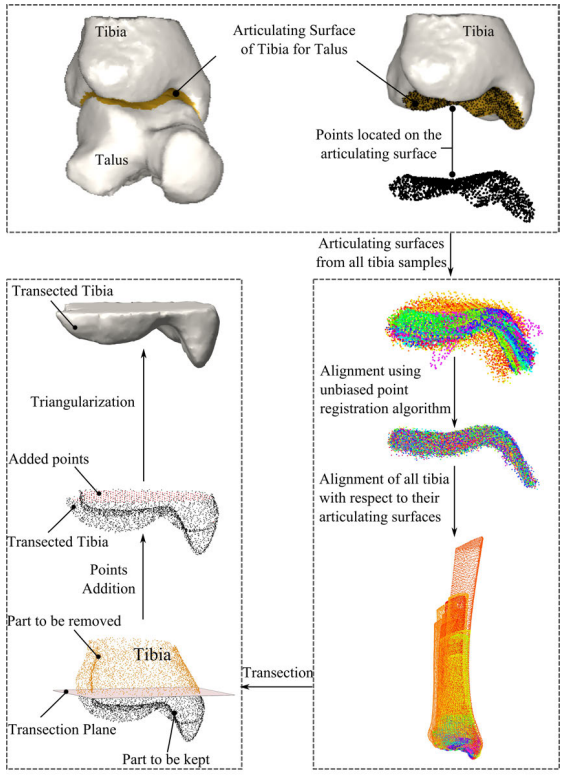


Figure 4.2: Transection procedure of tibia. Subchondral surface of the tibia in contact with the talus was automatically determined. Subchondral surface points on the tibia were automatically extracted. Each tibia was scaled and aligned using solely its subchondral surface. Aligned and scaled bones were transected axially. The open surfaces of transected bones were closed with automatic addition of points and were subsequently triangulated.

(i.e. eigenvectors) and $\boldsymbol{b}$ is a vector specifying the shape parameters. It is this $\boldsymbol{b}$ vector that is different between the distinct bones of a given population. Since the mean bone shape, $\bar{\mathbf{x}}$, and the main modes of shape variation, $\boldsymbol{\phi}$, are the same for all bone samples within the same population, the entries of vector $\boldsymbol{b}$ are the only parameters that are different from one bone sample to another (e.g. from one tibia to another tibia).

4.2.6 Evaluation of Statistical Shape Model(s)

To check whether the number of bone shapes (i.e. 72 samples) included in the study was sufficient for an accurate description of all possible shape variations with a similar accuracy as the voxel size in the CT scans, a numerical experiment

described by van de Giessen [164] was performed independently for the talus and distal tibia SSMs.
Throughout this numerical experiment, a SSM was generated using N randomly selected training shapes. Subsequently, the SSM was fit to a test bone, which was not included in the training set. The total number of modes used during the fit was $N-1$. To quantitatively evaluate the quality of the fit, the mean distance between the surface points estimated by the model and those of the test bone was calculated. This experiment was repeated 10 times with different and randomly selected training shapes for $N = 5, 10, \ldots, 35$ shapes.
The reliability of inter-user segmentation was assessed by calculating the mean surface distances between the surfaces of three samples obtained following the segmentation and re-segmentation steps. To evaluate reproducibility of the method, SSM of talus and SSM of tibia were rebuilt based on the dataset (i.e. 72 bones for both SSMs) consisting of three samples segmented twice by a different user. The surfaces of three samples modelled from re-built SSMs in each first five modes of shape variation (i.e. $\boldsymbol{x} = \bar{\mathbf{x}} \pm \sum_{s=i}^{i} b_s \boldsymbol{\Phi_s}, i = 1, 2, \ldots, 5$) were compared to those obtained from original SSMs in terms of the mean surface distances.
To assess the robustness of SSMs, a bootstrap analysis was performed as described by Pedoia [1] (see Appendix C).

4.2.7 Comparison of Shapes Between Control and OCD Groups

In this study, the SSMs were built based on all 72 training shapes, which are mixture of the control and OCD groups (Fig. 4.1). Therefore, the mean shape and the modes of variation were identical for each training shape regardless of its group (OCD vs. control). Only the entries of the vector $\boldsymbol{b}$ (Eq. 4.1), representing the location of the shape in the multi-dimensional distribution, were different for each training shape. The differences in bone shape between the control and OCD groups can therefore be assessed by comparing the entries of the vector $\boldsymbol{b}$ (i.e. shape parameters).
Presence of a normal distribution of shape parameters was assessed with a Kolmogorov-Smirnoff test. The shape parameters for the first five modes of bone shape variation were analyzed using SPSS (Version 22, Chicago, IL) with an ANOVA test to determine the presence of significant differences between the bone shapes of the OCD and control groups. Modes of variation higher than five were not considered in the analysis, as they were each representing less than 5% of the total variance observed [164]. The results of the reliability and robustness experiments were also supporting the choice made. The significance level ($p = 0.05$) was adjusted for multiple comparisons using the Bonferroni correction, resulting in a significance threshold of $p = 0.005$ (i.e. 0.05 / (2 groups $\times$ 5 modes)). To test whether age and gender were predictors of shape parameters, a linear regression analysis was performed.

4.3 Results

As the number of training shapes increased, the SSMs of the talus and tibia could more accurately describe the shape of bones that were not included in the training sets (Fig. 4.3). For both the talus and tibia, the mean estimation errors were below 1 mm (i.e. maximum voxel size of the CT scans) when more than 5 training shapes were used (Fig. 4.3). The first five modes of variation of the talus and tibia respectively described 49% and 40% of the total variance seen in the studied population (Fig. 4.4). No mode higher than mode five described more than 5% of the total variance observed in the shape of each type of bone.

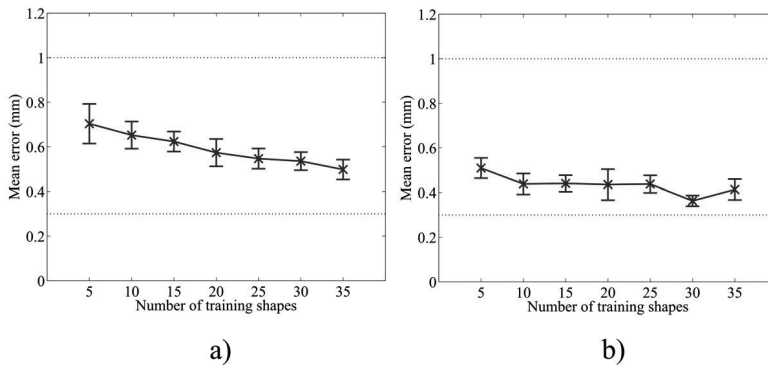


Figure 4.3: Mean distance between surfaces estimated by the SSMs of talus (a) and distal tibia (b) and those of test bones as a function of number of shapes in the models. The error bars describe 95% confidence intervals of 10 times repeated experiments. The dotted lines at 1 mm and at 0.3 mm represent maximum and minimum voxel size of the CT scans, respectively.

Distribution of shape parameters for the first five modes of variation, i.e. the first five components of vector $\boldsymbol{b}$ (mean $\pm$ 95% CI, in Appendix C), did not significantly differ from a Gaussian distribution (p values > 0.05).
The average surface distances (i.e. the segmentation reproducibility) were 0.13 mm, 0.06 mm, and 0.08 mm for the talus and 0.13 mm, 0.04 mm, and 0.07 mm for the tibia. As for the reproducibility of the method, the mean surface distances were less than 0.5 mm for all three samples in each first five modes of shape variation.
Neither age nor gender predicted variations in bone shape parameters (see Appendix C for the details of the analysis). No adjustment with respect to age or gender was therefore deemed necessary.
Mode 5 of the talus (Fig. 4.4) ($p = 0.004$) and Mode 1 of the tibia (Fig. 4.4) ($p < 0.0001$) showed statistically significant difference between control and OCD groups (Table 4.1).
Mode 5 of the talus primarily describes a change in the vertical neck angle (VNA) [167] (Fig. 4.4). VNA is the angle (i.e. in the sagittal plane of the talus) between

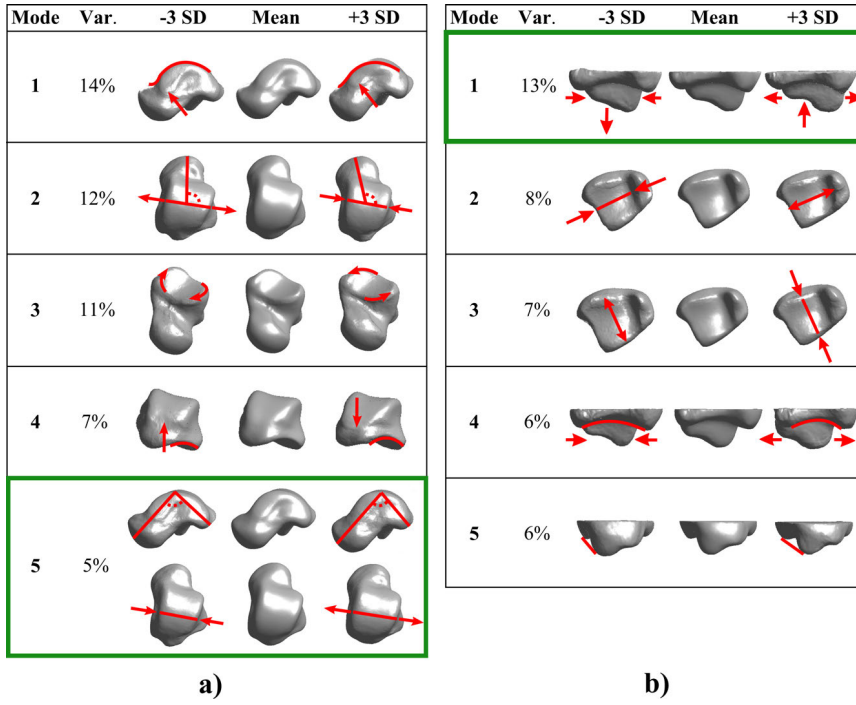


Figure 4.4: First five modes of talus (a) and distal tibia (b). These five modes describe 49% and 40% of the total variance of talus and tibia, respectively. Red lines and arrows point to the main changes in bone shapes for each mode of variation. Modes of variation within the green boxes are the ones showing significant differences between control and OCD groups.

two lines extending from the summit of trochlear surface to the mid-point of the convexity of the articular surface of the talar head and to the posterior tubercle of the talus. Besides a change in the VNA, alteration in talar trochlea was observed for this mode. Most of the OCD group talus shapes showed an increase in the VNA and a narrowing of the talar trochlea (Fig. 4.5). The shapes in the control group mostly lied within a positive standard deviation (Fig. 4.5) away from the mean shape, which describes a decrease in the VNA and an enlargement of the talar trochlea. As for the tibia, the means of the OCD and control groups shapes parameters for the first mode of variation were respectively -0.6 and 0.9 (Fig. 4.5). Deviations from the mean shape in the negative mode direction mainly represent enlargement and narrowing of the medial malleolus respectively in the superior–inferior and in the anterior–posterior directions.

Table 4.1: p values resulting from ANOVA test. Bonferroni p-value corresponds to $p = 0.005$, values under this threshold are marked with *.

Bone	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Talus	0.039	0.062	0.535	0.066	0.004*
Tibia	0.000*	0.115	0.768	0.060	0.347

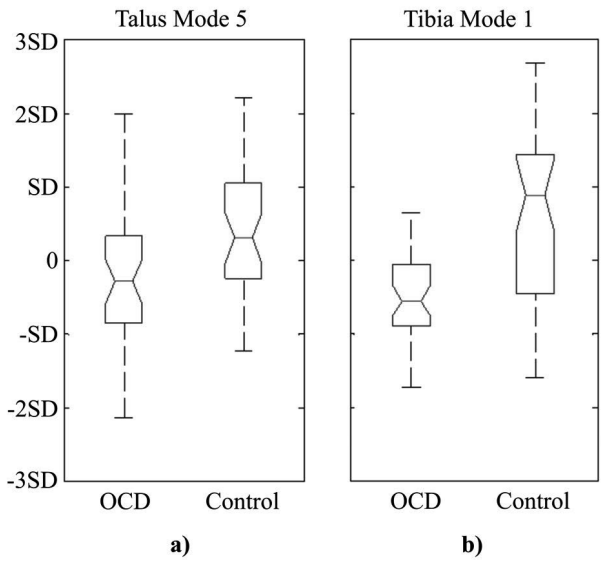


Figure 4.5: Box plots of shape parameter distributions for OCD and control groups. These distributions are presented for the Mode 5 of talus (a) and Mode 1 of tibia (b).

4.4 Discussion

The goal of this study was to investigate whether the morphology of the ankle (talocrural joint) bones of an OCD group is different from those of a control group. To this end, 3D SSMs of distal tibia and talus were built based on a population comprising subjects with and without an OCD. The 3D SSMs made it possible to quantitatively describe the complex 3D geometry of the bones and to identify specific shape modes that may be related to an increased risk of sustaining an OCD.

Mode 5 of the talus and Mode 1 of the tibia were significantly different between control and OCD groups. The shape of the talus in the OCD group primarily shows an increase in the VNA (Fig. 4.5), which implies flattening of the talus and in turn possible increase in the surface area on the talar trochlea for the tibia. However, change in VNA alone does not explain all the variations seen in Mode

5. In addition to an increase in the VNA, narrowing of the superior articular surface of the talus was observed. Before drawing any conclusions as to whether these observed variations in talus shape may have something to do with the development of the lesions, it is important to take the widely accepted injury mechanisms into account. Based on the earlier described mechanisms, a lateral OCD can be reproduced by strongly inverting a dorsiflexed ankle, while a medial OCD can be caused by exposing a plantarflexed ankle to a combination of inversion and rotation of the talus on the tibia [33, 168]. These injury mechanisms basically imply that the relative position and orientation of the bones as well as the biomechanical loads play specific roles in development of the lesions. It is also known that contact areas and their distributions at the tibiotalar articulating surfaces vary with respect to joint position and joint loading [169, 170, 171]. In case a decrease in the contact area at a specific joint position (e.g. dorsiflexion combined with inversion) occurs, contact stress may exceed the limits that can be tolerated by the joint structures and a lesion may result. As for the variations observed in the shape of the talus in the OCD group, it is reasonable to expect that the narrowing of the talar dome decreases the contact area, which might lead to an increased risk of sustaining an OCD during high impact events such as ankle sprains.

Regarding Mode 1 of the tibia shape, alterations seen at the medial side may especially contribute to development of the medial lesions, which are reported to be quite frequent ones [38, 172, 173]. This seems plausible given the fact that high prevalence of medial lesions was observed in the training dataset used for creating the SSM of the tibia. The shape of bones sustaining OCD usually deviated from the mean shape in the negative direction (Fig. 4.5). When comparing the changes seen at the medial malleolus at two extremes (i.e. $\pm 3SD$) qualitatively, it becomes clear that the tibiotalar contact area is relatively smaller at -3SD compared to those at the +3SD. Considering the facts that 10% to 20% of the joint loads (equal to 3 to 4 times the body weight) are transferred through the medial and lateral talar facets of the talocrural joint [174], and the fact that the medial malleolus is an important region in maintaining normal tibiotalar contact area and pressure, relatively small changes in the contact area could lead to increased peak loading in the contact area. This in turn can be associated with an increased risk of sustaining an OCD.

One of the limitations of the study is that information on malalignment in the lower limbs of subjects was not available. Due to lack of information, it was not possible to evaluate the contributions of leg malalignment to development of OCDs.

An important issue related to the dataset used in this study is that the subjects without an OCD (i.e. control group) are not guaranteed to never sustain a lesion. There is always a chance that the control group subject(s) may develop an OCD in the future. This basically means that there can be an overlap of bone shapes of subjects with and without an OCD.

Another issue associated with the dataset is that the subjects with and without an OCD were CT-scanned with slightly different techniques. As position and orientation differences among all bones of the same type were minimized dur-

ing generation of SSMs, the observed shape variations are not expected to be contaminated with shape variations caused by differently oriented coordinate systems of the CT scans. As for the differences in the voxel sizes of the CT scans, some relatively small details in shape of the bones may have been lost in CT scans with higher voxel sizes (i.e. 1.0 mm $\times$ 1.0 mm $\times$ 0.75 mm). However, these missing details are not expected to substantially influence the shape variations described by the first five modes of shape variation.

It is also worthy indicating that the shape differences observed in this study are consistent with the most commonly known trauma mechanism (i.e. inversion trauma), which causes intra-articular pressure impact [158]. There might be other shape differences, which could not be captured in this study due to relatively small sample size and can be explained with other mechanisms. The assessment of injury mechanisms due to changes in the bone morphology was beyond the scope of this study. To evaluate the effects of the talus and tibia morphologies on the mechanical characteristics of the talocrural joint and on formation of OCDs, computational models of the ankle [175] may be helpful.

The current study has revealed shape differences in the ankle bones between an OCD group and a control group. To the best of the authors' knowledge, no study has previously proposed the morphology of ankle bones as a possible risk factor for OCD. Therefore, it is difficult to directly compare the findings of the present study with those available in the literature. However, it may be feasible to compare the observed shape variations, which showed significant difference between the control and OCD groups, with the results of studies on the morphology of the ankle bones and their biomechanical significance.

In conclusion, the current study used 3D SSMs to assess the plausibility of our hypothesis that the shape of bones in the ankle joint is different between the OCD group and a control group. Two shape modes of the talus and tibia were found to be significantly different between the OCD group and a control group. The specific shapes of the tibia and talus associated with the OCD group seem to be factors increasing contact stresses by decreasing contact area. Therefore, the observed shape differences could be explained by biomechanical factors and indicate that specific distal tibia and talus shapes may contribute to increased risk of OCD.

5

Typical Shape Differences in Talus and Calcaneus Between Subjects with Chronic Ankle Instability and Controls

Submitted as:
N. Tümer, G. Vuurberg, L. Blankevoort, G.M.M.J. Kerkhoffs, G.J. Tuijthof, A.A. Zadpoor. Typical Shape Differences in Talus and Calcaneus Between Subjects with Chronic Ankle Instability and Controls.

5.1 Introduction

The injury surveillance data collected by National Collegiate Athletic Association (NCAA) in the USA, from 15 sport activities over a 16-year time period show that ankle sprains are the most common type of injury in all sports (e.g. basketball, soccer and volleyball with high injury rates) and account for 15% of all reported injuries [176]. An inversion injury or lateral ankle sprain (LAS), which comprises approximately 80% of all ankle sprains [177, 178], is not only a problem of young and physically active population, but is also seen in the elderly [177, 179]. An average of 2 million acute ankle sprains in the United States [180] and an estimated 5% of emergency room visits in the United Kingdom [181] were reported. The true incidence of LAS is even higher as it is estimated that approximately half of people sustaining a LAS do not seek professional help for evaluation and treatment [181, 182, 183].

Most patients experiencing LAS can be successfully treated and regain functional ankle stability by means of conservative treatment [184, 185]. However, despite adequate conservative treatment, up to 40% of patients still suffer from residual complaints [186, 187, 188, 179, 189] and may progress to chronic ankle instability (CAI) [190]. CAI is defined as instability of the ankle with the feeling of giving-way, episodes of recurrent ankle sprains, with or without the presence of joint laxity [177, 191].

Limited evidence is available indicating that surgical treatment is relatively better in decreasing recurrent ankle sprains [192]. Despite its potential success, surgical treatment is generally preserved for athletes to enable quick return to play, and it is not considered to be the preferred option to treat LAS due to the increased costs and risk of complications [193] without knowing whether the LAS will progress to CAI. Considering the clinical evidence that CAI can lead to an early onset of post-traumatic osteoarthritis [179, 194], it is important to identify patients at risk of developing CAI who might benefit more from surgical treatment than from conservative treatment.

A great deal of effort has been put into identifying the factors playing a role in the etiology of CAI. Ligament laxity, proprioceptive deficits, muscle weakness, postural control deficits [177, 195, 196], joint congruency [185], fibular position [183, 197, 198], cavus foot deformity [184, 197, 199] and varus ankle or hindfoot [200] are some of the factors associated with CAI. However, potential factors with bone origin and their correlations with CAI have been mostly overlooked. A few studies that have considered factors with bone origin [185] are limited to simple measurements on two-dimensional (2D) images (e.g. the radius of the talar surface measured on lateral X-rays) that cannot fully reflect bone's three-dimensional (3D) nature.

As the morphology of the articulating bones contributes to the stability of the joints and determines their kinematics [201, 175, 202], morphological variations are expected to change the mechanical environment of the joints and modify the risk of CAI. Therefore, we aimed to determine whether statistically significant shape differences exist between tali and calcanei of patients with CAI and their contralateral controls as well as healthy controls.

One of the reasons for focusing on variations in the bones forming the subtalar joint (STJ) is that instability does not solely appear in the talocrural joint, but in up to 58% of the cases is also present in the STJ of patients with CAI [177]. Moreover, despite the important role of the STJ in the hindfoot complex and its possible contributions to CAI [184], not much attention has been paid to the factors that may alter the mechanical environment of the STJ.
In this study, 3D shape variations of the calcaneus and talus were systematically described and quantitatively compared in threefold (CAI vs. CAI contralateral controls, CAI vs. healthy controls, and CAI contralateral controls vs. healthy controls) using the statistical shape modelling (SSM) technique [162, 203].

5.2 Materials and Methods

Two 3D SSMs, one for the talus and one for the calcaneus, were generated in a way similar to what described previously [203]. In the first step, computed tomography (CT) scans of subjects with unilateral CAI and without CAI were collected. In the second step, calcanei and tali were segmented from all CT scans. In the third step, all calcanei (or all tali) surfaces obtained as a result of segmentation process were aligned together (i.e. registration process), so that the differences between bones due to position, orientation and scaling were minimized. In the final step (Fig. 5.1), shape variations were extracted by performing multivariate analysis (i.e. principal component analysis) on the aligned bone data. The SSM of the talus and that of the calcaneus enabled us to quantitatively compare the shape of tali and calcanei between the three groups, respectively. Detailed descriptions of each step are provided in the following sub-sections.

5.2.1 Data Collection

CT scans of patients with a confirmed CAI and healthy controls without reported ankle joint pathology were collected from Academic Medical Center (AMC, Amsterdam). CT scans were discarded from the study if, patients were 16 years of age or younger at the time of scanning, other joint pathology or concomitant ankle injury existed or the entire volume of the calcanei were not visible. The final dataset included 52 CT scans from 52 subjects (26 patients with unilateral CAI and 26 healthy controls) (Fig. 5.1). Six CT scans were unilateral (i.e. 6 healthy controls with left ankles), the rest were bilateral (i.e. 26 patients with unilateral CAI and 20 healthy controls with both left and right ankles). As all scans were retrospectively acquired, a waiver was provided by the medical ethical committee.

5.2.2 Segmentation of Bones from CT Scans

98 calcanei and 98 tali were semi-automatically segmented using the scripting module of Mimics (Materialise, Leuven, Belgium). All segmentations were visually checked to ensure the proper definition of the contour of the bones and

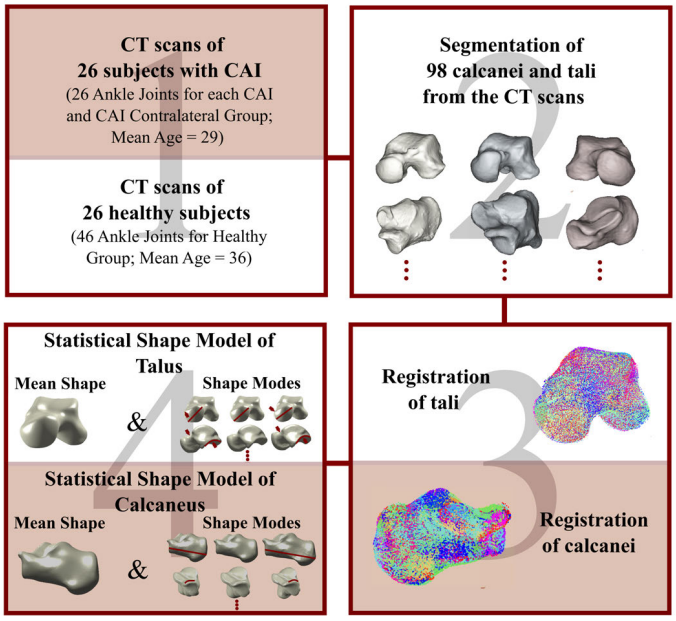


Figure 5.1: A scheme representing the main steps in generation of the SSM of talus (or calcaneus).

manual corrections were made if necessary. Triangulated bone surfaces were extracted from the segmentation results and all right side tali (or calcanei) were mirrored in the sagittal plane.

5.2.3 Registration of Bones

Position, orientation, and scaling of a triangulated bone (talus or calcaneus) surface could be different from the others. As the main goal is to describe variations in bone shape only, differences due to position, orientation, and scaling among bone surfaces should be minimized. Therefore, all tali (or calcanei) were aligned with each other using an unbiased registration algorithm [162, 203].

5.2.4 Extraction of Bone Shape Variations

Following the registration process, corresponding points on all aligned tali (or calcanei) were automatically determined [203, 166]. The covariance matrix of the data vectors [159] containing the ordered coordinates (x, y, z) of corresponding points across all tali (or calcanei) was computed. Then, principal component analysis (PCA) was performed on the covariance matrix. As a result of the PCA [159], modes of shape variation (i.e. eigenvectors) with a descending order of variance (i.e. eigenvalues) were obtained. Modes of shape variation mainly describe the directions of shape changes, while the variance describes how much

variation is present in the corresponding direction. A new or existing talus (or calcaneus) shape ($\boldsymbol{x}$) can be represented, using the mean shape of the talus (or the calcaneus) ($\bar{\mathbf{x}}$) and a weighted summation of the modes of shape variation for the talus (or for the calcaneus) ($\boldsymbol{\phi}$), as [159]:

$$\mathbf{x} = \bar{\mathbf{x}} + \sum_{s=1}^{c} b_s \boldsymbol{\phi}_s \tag{5.1}$$

in which, b_s values (i.e. weights or shape parameters) describe the contributions of the first c modes of shape variation to the mean bone shape.

5.2.5 Comparison of Bone Shapes Between the Groups

The SSM of the talus and that of the calcaneus were built on 98 tali and 98 calcanei, respectively, divided in three groups: CAI (26 tali and 26 calcanei from ankles that have a confirmed CAI), CAI contralateral controls (26 tali and 26 calcanei from the contralateral healthy ankles of patients with a confirmed CAI) and healthy controls (46 tali and 46 calcanei of healthy subjects without reported ankle joint pathology) (Fig. 5.1).
As the SSM of each bone was developed based on a mixture of all three groups, the mean talus (or calcaneus) shape ($\bar{\mathbf{x}}$) and the modes of shape variation for the talus (or for the calcaneus) ($\boldsymbol{\phi}$) were the same for each of 98 tali (or 98 calcanei) regardless of their group [203]. Only the shape parameters (i.e. b_s values in Eq. 5.1) were different for each of 98 tali (or 98 calcanei). Therefore, the shape parameters could be used to quantitatively compare bone shape variations within the subjects that had sustained a CAI (i.e. CAI vs. CAI contralateral controls) and between the patients with CAI and healthy controls (i.e. CAI vs. healthy controls and CAI contralateral controls vs. healthy controls).
The distributions of shape parameters across all tali and calcanei were evaluated by using histograms. To assess whether the shape parameters conform to a normal distribution, a Kolmogorov-Smirnoff test was performed.
For each bone type, the shape parameters for the first c modes of shape variation were compared between three groups with an ANOVA test. For pairwise comparisons of the groups (i.e. CAI vs. CAI contralateral controls and CAI vs. healthy controls and CAI contralateral controls vs. healthy controls), Bonferroni was used as a post-hoc test. The number of modes of shape variation c was determined such that all modes of shape variation who represented more than 5% of the total shape variation were included [203].
To evaluate whether observed bone shape variations were affected by age and gender of subjects, an ANCOVA test was performed by considering age and gender as covariate factors. All the statistical comparisons were performed using SPSS (version 22, Chicago, IL).

5.3 Results

5.3.1 Cohort Characteristics

Patients with CAI were almost equally distributed in terms of gender (14 males and 12 females). The mean age of these patients was 29 years (SD = 11). There were 12 males and 14 females in the healthy control group, with the mean age of 35 years (SD = 11). In 62% of the patients, the right ankle was affected by CAI ($n = 16$).

5.3.2 General Bone Dominant Shape Variations

Shape modes higher than six for the talus (Fig. 5.2a) and five for the calcaneus (Fig. 5.2b) described less than 5% of the total variance observed. The first six modes of variation for the talus and the first five modes of variation for the calcaneus described 49% (Fig. 5.2a) and 45% (Fig. 5.2b) of the total variance in the shapes of the talus and calcaneus, respectively. Higher modes of variations were not considered in the subsequent analyses. The distributions of the shape parameters for the first six modes of the talus and the first five modes of the calcaneus did not significantly differ from a Gaussian distribution (p values > 0.05).

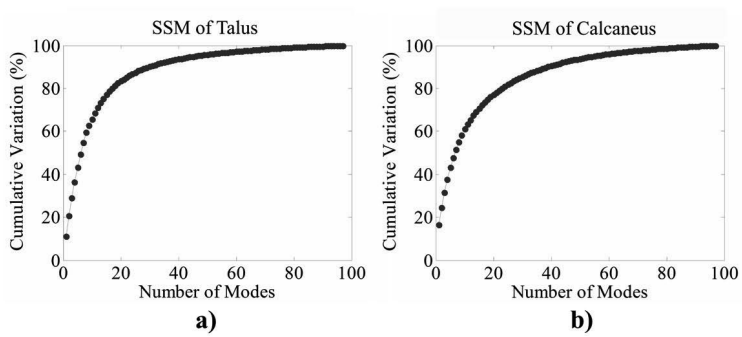


Figure 5.2: Contribution (%) of each shape mode to the total shape variation in (a) talus and (b) calcaneus.

5.3.3 Description of Bone Shape Variations

Descriptions of bone shape variations are only provided for the first three shape modes of the talus (Fig. 5.3) and calcaneus (Fig. 5.4). The reasoning behind this presentation is that the modes of shape variation higher than the mode three of the talus and calcaneus explain relatively small shape variations distributed over the bone surfaces and none of them were found to be significantly different between the three groups (see the next subsection for the details of statistical analyses).

The first three modes of shape variation of the talus (Fig. 5.3) and calcaneus (Fig. 5.4) described the following changes in:

- Talus Mode 1: the lateral rotation and the size of the talar head,
- Talus Mode 2: the curvature of the talar lateral process and the inclination angle of the talar neck relative to the body [204],
- Talus Mode 3: the lateral projection of the talar process and the length of the talus,
- Calcaneus Mode 1: the length of the calcaneus,
- Calcaneus Mode 2: the inclination of the sustentaculum tali,
- Calcaneus Mode 3: the medial and lateral tuberosity and the contour of the anterior articular surface.

5.3.4 Comparison of Bone Shape Variations Between Three Groups

The shapes of the talus and calcaneus did not significantly ($p > 0.05$) vary within both sides of the subjects that had had an unilateral CAI (i.e. CAI vs. CAI contralateral controls).
The shape variations of the talus described by Mode 2 (Fig. 5.3) were significantly different between the groups CAI vs. healthy controls ($p = 0.015$) as well as CAI contralateral controls vs. healthy controls ($p = 0.035$).
The shapes of the tali of the CAI and CAI contralateral groups deviated from the mean talus shape in the positive direction of Mode 2 of the talus (i.e. median = 0.03 SD, 25th percentile = -0.40 SD, 75th percentile = 1.06 SD for the CAI group; median = 0.44 SD, 25th percentile = -0.65 SD, 75th percentile = 1.08 SD for the CAI contralateral group), whereas those of the healthy controls did in the negative direction (i.e. median = -0.24 SD, 25th percentile = -0.94 SD, 75th percentile = 0.20 SD) (Fig. 5.5a). Positive deviations from the mean talus shape describe a decrease in the inclination angle of the talar neck relative to the talar body and a change in the curvature of the talar lateral process adversely affecting the congruency of the articular facets at the posterior side (Fig. 5.3). Observed changes are vice versa in the tali with negative deviations away from the mean talus shape (Fig. 5.3). Fig. 5.6a shows the tali and their calcanei counterparts extracted from three subjects. The talus of first subject is close to the mean talus shape (Fig. 5.6a, middle column). The tali of second (i.e. patient with CAI) and third (i.e. healthy individual) subjects deviate from the mean talus shape in the positive (i.e. +2.3 SD away from the mean talus shape) and negative (i.e. -2.5 SD away from the mean talus shape) direction of Mode 2 of the talus, respectively.
Shape variations represented by Mode 3 of the calcaneus (Fig. 5.4) were significantly different between the following groups: CAI vs. healthy controls ($p = 0.003$), and CAI contralateral controls vs. healthy controls ($p = 0.001$). The calcanei of the patients with CAI and healthy controls deviated from the mean calcaneus shape in the negative and positive direction of Mode 3 of the calcaneus

(i.e. median = -0.46 SD, 25*th* percentile = -1.06 SD, 75*th* percentile = 0.32 SD for the CAI group; median = -0.49 SD, 25*th* percentile = -1.08 SD, 75*th* percentile = 0.30 SD for the CAI contralateral group; median = 0.55 SD, 25*th* percentile = -0.42 SD, 75*th* percentile = 1.12 SD for the healthy controls) (Fig. 5.5b), respectively. In the calcanei with positive deviation away from the mean calcaneus shape, the medial tuberosity extended more distally as compared to the lateral tuberosity (Fig. 5.4).
Contrarily, the lateral tuberosity stood at the same horizontal line with the medial tuberosity or relatively more distally in the calcanei with negative deviations away from the mean calcaneus shape (Fig. 5.4). Additionally, the contour of the anterior articular surface was relatively flat in the calcanei with negative deviations as compared to those with positive deviations (Fig. 5.4). Fig. 5.6b shows the calcanei and their tali counterparts derived from three subjects whose calcaneus shape is close to the mean calcaneus shape (Subject 4), is at the positive side of the standard deviation (i.e. healthy subject, Mode 3 of calcaneus, +2.3 SD away from the mean, Subject 5) and is at the negative side of the standard deviation (i.e. patient with CAI, Mode 3 of calcaneus, -2.1 SD away from the mean, Subject 6).
Consideration of the age and gender of individuals in the statistical analysis as covariates (i.e. ANCOVA) caused no changes in the statistical significance of the results reported above.

5.4 Discussion

In this study, a 3D SSM technique was used to quantitatively compare the talus and the calcaneus shape differences within the two sides of individuals with a unilateral CAI as well as between the CAI patients and healthy controls. The results of the statistical analyses indicated that there are some specific shape differences in the talus and calcaneus, which make individuals with CAI distinguishable from healthy individuals with no known history of ankle joint pathology.
Considering the combination of statistically different shape variations in the talus and the calcaneus overall, CAI patients seem to be showing less subtalar joint congruency and flattened calcaneal ground-contact surface (Fig. 5.6b). Variations in the medial and lateral calcaneal tuberosity (Figs. 5.4 and 5.6b) can alter the loading moment, which is formed by a pair of ground reaction force (GRF) and joint reaction force (JRF), and acts on the STJ (Figs. 5.7a-b). In healthy subjects, the two axes along which GRF and JRF act do not coincide with each other (Fig. 5.7a) and a pronation exorotation moment occurs [205]. If the GRF slides to the lateral side (Fig. 5.7b), which is highly possible if both medial and lateral tuberosities extend to the ground, the distance (D) between the GRF and JRF axes will increase. In turn, the pronatory external moment will be higher and cause extra strain on the medial muscles and ligaments. A patient may minimize the effects of the GRF axis shift to the lateral side and avoid losing balance by means of an inversion movement (Fig. 5.7c) which may lead to a recurrent LAS if not countered on time.

Additionally, shape variations in the talus (i.e. changes in the inclination angle of the talar neck relative to the body, Mode 2 of talus) (Figs. 5.3 and 5.6a) and in the calcaneus (i.e. variations in the medial and lateral tuberosity, Mode 3 of calcaneus) (Figs. 5.4 and 5.6b) may affect the orientation of the STJ axis (Fig. 5.7d). A decrease in the inclination angle of the talar neck relative to the body (Fig. 5.7e) or/and a more distally extension of the lateral tuberosity (Fig. 5.7f), which is more likely to be seen in individuals with CAI compared to healthy controls, can cause the STJ to be more vertical (Fig. 5.7d, $\beta > \alpha$). In the case of vertically aligned STJ axis (Fig. 5.7d, $\beta > \alpha$), less rotation around the horizontal axis of the talus in relation with the calcaneus is observed for a given rotation around the vertical axis (an analogy can be made to a gear mechanism, Fig. 5.7g). Limitation in rotation around the horizontal axis may increase the risk of losing balance quicker in an inversion motion, as the control of the ankle joint muscles (i.e. mainly the peroneal muscles) ends as soon as the maximal range of motion is exceeded. The tendency of STJ axis to be more vertical due to the talus and calcaneus shape variations in individuals with CAI may explain why cavus foot, related to a smaller ground contact surface and a more vertical STJ axis [205], is associated with CAI [184, 197].

Whatever the mechanism behind the development of CAI may be, the most important finding of this study is that some specific talus and calcaneus shape variations appear to be significantly different between individuals with CAI and healthy controls. Whether these findings represent post-traumatic changes or whether recurrent LAS originate from these bone shape variations would require further research in a prospective setting. However, the fact that the talus and calcaneus shapes did not vary within individuals with one unstable ankle suggests these shape variations play a role on the development of CAI after a first-time ankle inversion injury and are not the consequence of CAI as only one ankle is affected.

The main limitation of this study is the retrospective nature of the data collection. Although there is certainty on the presence of instability, more concomitant problems (e.g. osteoarthritis, a fracture in history) may be the cause of the observed shape differences in the talus and calcaneus. Additionally, information on the activity level of the patients was not available, which may also be a contributing factor. Contrarily, the age and gender of subjects included in this study were known. Therefore, effects of age and gender could be controlled in the statistical analyses.

The bone shapes of subjects with and without CAI overlap. Some of the healthy subjects have the same geometric features as those identified as being typical for subjects in the CAI group. These individuals may not have developed CAI yet or sustained an injury that leads to the onset of CAI, but may nevertheless be at risk of developing CAI. Contrarily, some of the subjects with CAI do not answer to the features defined in this study as specific for CAI, but did sustain an event that lead to the genesis of CAI, potentially due to other risk factors. In order to evaluate the prognostic value of these geometric variations, and to define which patients require follow-up to monitor development of CAI, future studies may provide insight. Despite the complexity of the findings, it is highly important to

recognize bone shape differences found in CAI patients, as these may contribute to the increased risk at sustaining new sprains. Although CAI is a multifactorial condition, this study is a next step in identifying risk factors with bone shape origin.
Regeneration of the SSMs of the talus and calcaneus based on a prospective dataset could help to assess the prognostic values of the bone shape variations on the development of CAI. Additionally, significantly different shape variations in the talus and calcaneus could be translated to measurements for conventional radiology. If these measurements can reliably be assessed and have prognostic value, they can be included in prediction models. With an implementation of well-designed prediction models in clinical settings, patients at higher risk of developing CAI can be identified at an earlier stage. More accurate predictions of the risk of developing CAI can assist clinicians by providing more insight into which treatment options may work best with a patient who has sustained a LAS.

5.5 Conclusion

The 3D statistical shape models of the talus and calcaneus built based on the mixed data of patients with unilateral chronic ankle instability (CAI) and healthy controls enabled us to quantitatively compare shape variations between the affected side and contralateral control in CAI patients, between the affected side of CAI patients and healthy controls, and between contralateral controls of CAI patients and healthy controls. We found two specific statistically significant shape differences between CAI patients and healthy controls. In the talus, it was the curvature of talar lateral process combined with the inclination angle of talar neck relative to the body. In the calcaneus, it was the medial and lateral tuberosity combined with contour of anterior articular surface.

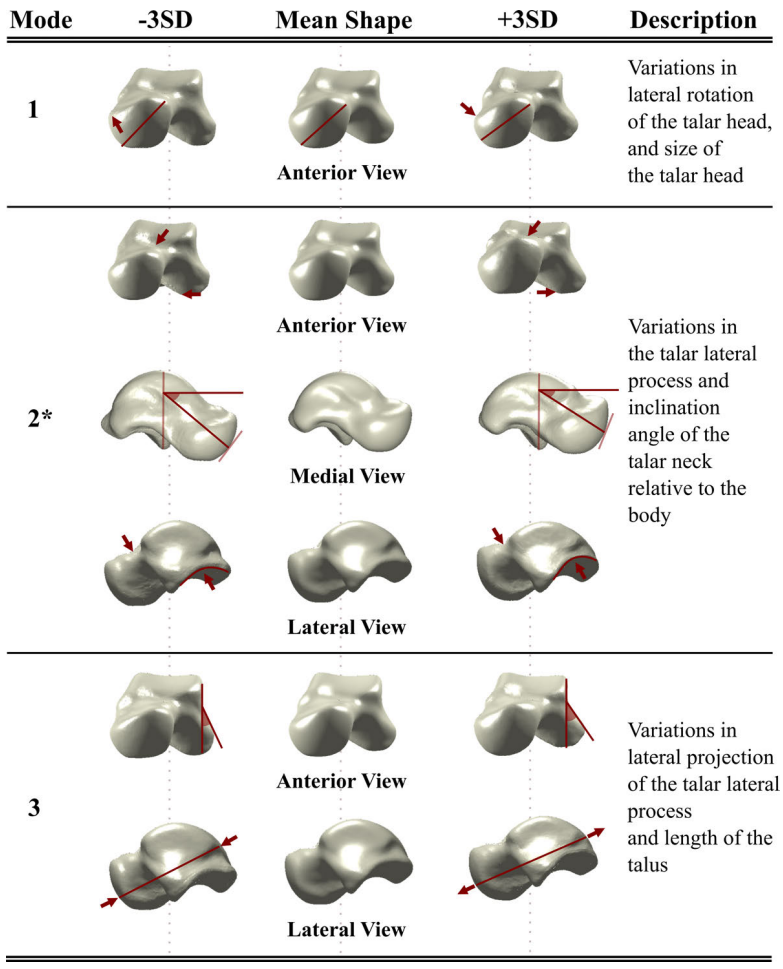


Figure 5.3: Talus shape variations described by the first three shape modes of talus. Red lines and arrows point to the shape changes. The shape mode, which represented significantly different talus shape variations between patients with CAI and healthy controls (i.e. CAI contralateral controls vs. healthy controls and CAI vs. healthy controls), is marked with *.

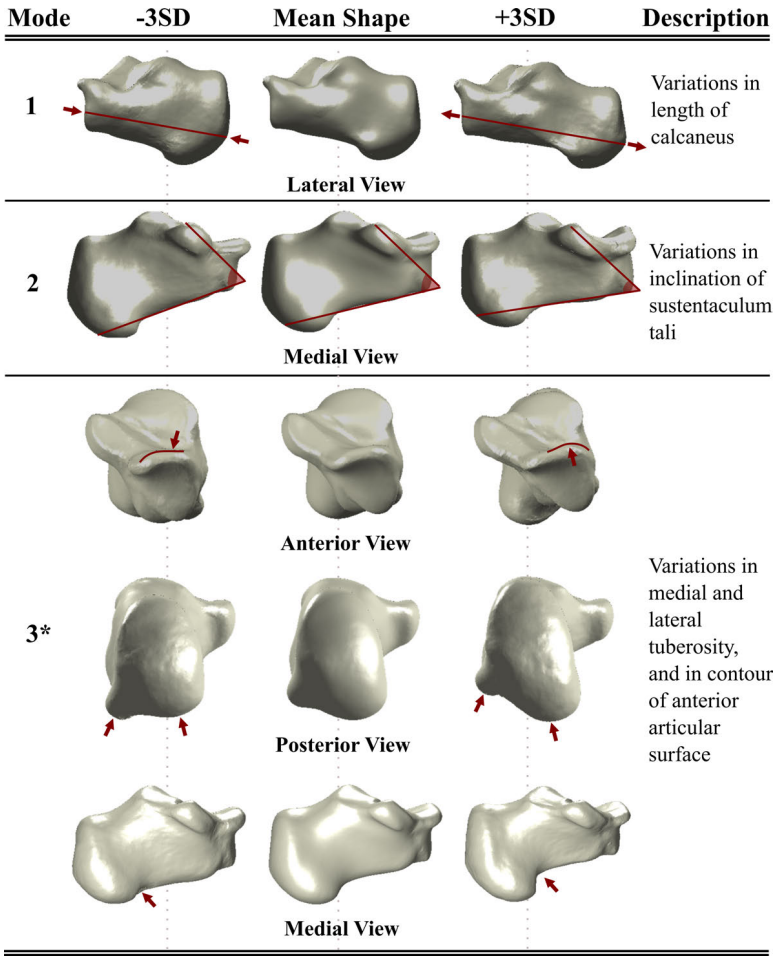


Figure 5.4: Calcaneus shape variations described by the first three shape modes of calcaneus. Red lines and arrows point to the shape changes. The shape mode, which represented significantly different calcaneus shape variations between patients with CAI and healthy controls (i.e. CAI contralateral controls vs. healthy controls and CAI vs. healthy controls), is marked with *.

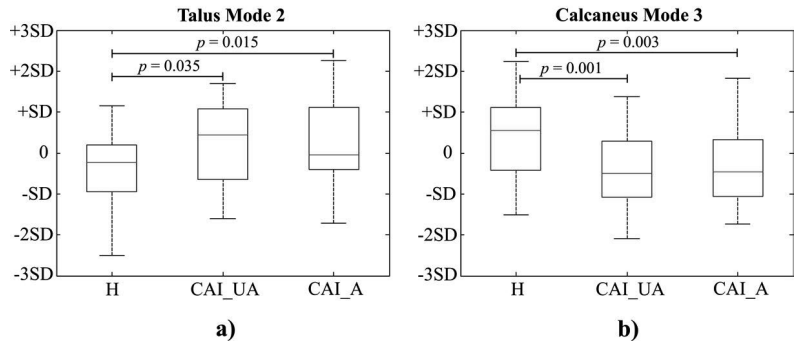


Figure 5.5: Box plots of shape parameters distributions for the three groups: CAI (CAI_A), CAI contralateral controls (CAI_UA) and healthy controls (H). These distributions are presented for the Mode 2 of talus (a) and Mode 3 of calcaneus (b). *p*-values resulting from the analysis of variance (ANOVA) test are indicated for the given shape mode.

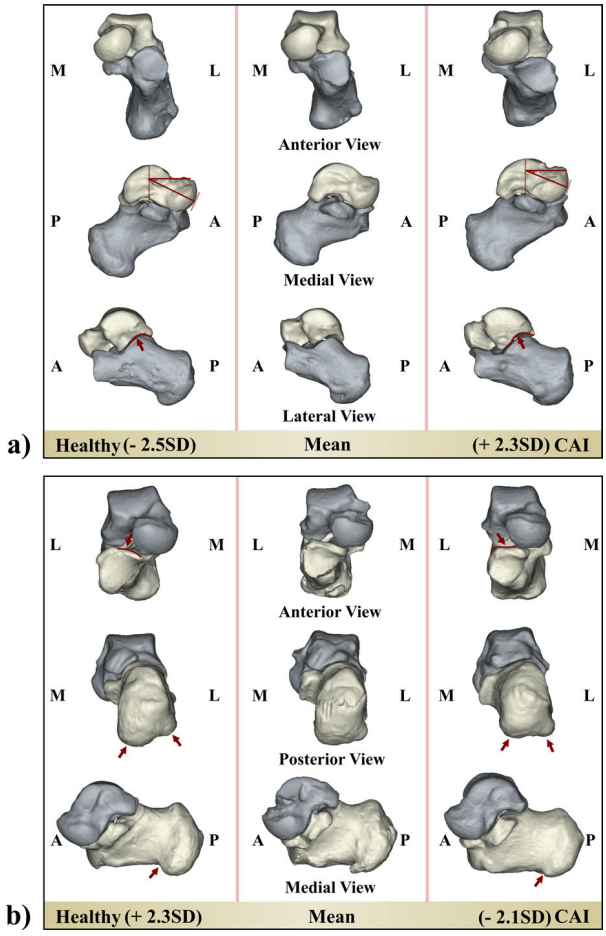


Figure 5.6: (a) The tali and their calcanei counterparts of the subjects whose talus shape is close to the mean talus shape (middle column; Subject 1), at the positive (right column; patient with CAI, Mode 2 of talus, +2.3 SD away from the mean talus shape, Subject 2) and negative extremities (left column; healthy subject, Mode 2 of talus, -2.5 SD away from the mean talus shape, Subject 3) (b) The calcanei and their tali counterparts of the subjects whose calcaneus shape is close to the mean calcaneus shape (middle column; Subject 4), at the positive (left column; healthy subject, Mode 3 of calcaneus, +2.3 SD away from the mean calcaneus shape, Subject 5) and negative bounds (right column; patient with CAI, Mode 3 of calcaneus, -2.1 SD away from the mean calcaneus shape, Subject 6).

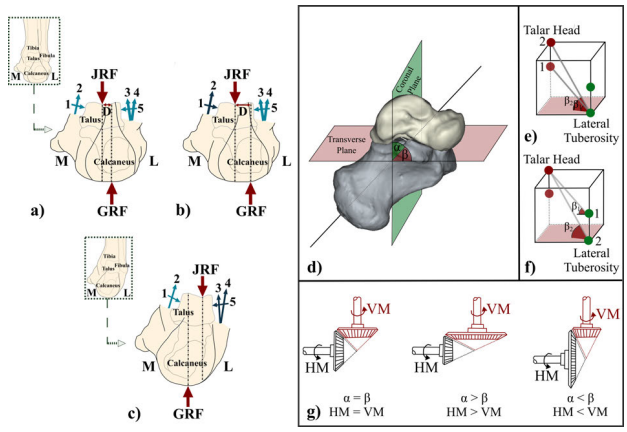


Figure 5.7: Mechanisms that can be affected by the observed shape variations in talus and calcaneus and may partly explain why some individuals are more prone to develop CAI. (a) JRF and GRF represent joint reaction force and ground reaction force, respectively. The arrows in light and dark blue display the forces acting on the subtalar joint due to: (1) Medial malleolar, (2) Deltoid ligament, (3) Anterotalofibular ligament, (4) Peroneal muscles, (5) Fibula. Dark blue represents relatively higher forces as compared to those indicated in light blue. In normal foot, there is a distance (D) between the axes along which GRF and JRF act. A pair of GRF and JRF mainly generates moment that have an effect on subtalar joint (STJ). (b) A slide of the GRF axis to the lateral side, which can be seen in case lateral tuberosity of calcaneus extend more distally, results in increase of distance (D) between a pair of GRF and JRF. In turn, effect of moment acting on STJ is higher. (c) With an inversion movement, the effects of the GRF axis shift to the lateral side may be minimized. (d) Positioning of the STJ axis. α and β are the angles from the coronal and transverse plane to the STJ axis, respectively. (e) Assuming that there is no change in the lateral tuberosity of calcaneus, alterations in the talar head/neck (e.g. Points 1 and 2) can lead different positioning of the STJ axis. $\beta 1$ and $\beta 2$ are the angles from transverse plane to the STJ axis passing through point 1 and point 2, respectively. (f) Similar configuration. This time, it is assumed that there is no change in the talar head/neck; but, the lateral tuberosity of calcaneus varies. $\beta 1$ and $\beta 2$ are the angles from transverse plane to the STJ axis passing through point 1 and point 2, respectively. (g) An analogy can be made to a gear mechanism. HM and VM stand for horizontal and vertical motion. If the STJ axis passes at 45° ($\alpha = \beta$), rotation around the vertical axis and the horizontal axis is equal to each other. If the STJ axis is vertically aligned ($\beta > \alpha$), rotation around the horizontal axis is less than the given one around the vertical axis. In case $\beta < \alpha$ (horizontal alignment of the STJ axis), it is vice versa.

6

Three-Dimensional Analysis of Shape Variations and Symmetry of the Fibula, Tibia, Calcaneus, and Talus

Accepted as:
N. Tümer, V. Arbabi, W.P. Gielis, P.A. de Jong, H. Weinans, G.J. Tuijthof, A.A. Zadpoor. Three-Dimensional Analysis of Shape Variations and Symmetry of the Fibula, Tibia, Calcaneus, and Talus. Journal of Anatomy. (2018).

6.1 Introduction

There has been long-standing interest in the geometric [206, 207, 208, 160, 209] and non-geometric (e.g. bone mineral density, structural stiffness, moment of areas) [210, 211] bilateral symmetry of the lower extremities. This interest is partly due to the symmetry assumption that is frequently made in clinical assessments and research studies. Some examples are summed. First, a common clinical practice is to use the contralateral unaffected side as a template for planning corrective osteotomy [212, 213]. Second, the contralateral unaffected side is often used as a reference in arthroplasty surgeries to determine the size of an implant and its position, when the limb of interest is deformed by a fracture or a degenerative disease (e.g. osteoarthritis) [210, 208, 214]. Third, the unaffected contralateral side usually serves as an intra-subject control or as a shape template in research studies that assess whether a bone shape can be a risk factor for the onset of an injury (e.g. acute knee injury) [215, 216] or a lesion (e.g. osteochondral defect) caused by an injury (e.g. lateral ankle sprain) [203].

For the symmetry assumption to be valid, it is necessary to establish that the differences in the geometric and non-geometric features of the left and right extremities are sufficiently small. Nevertheless, limited information is available regarding the (a)symmetry of the lower extremities within and across populations. In particular, little is known about the shape variations and (a)symmetry of the bones forming the talocrural (TC) and subtalar (ST) joints (i.e. the fibula, tibia, calcaneus and talus). The few studies that are available [206, 217, 160, 209] are limited due to small sample sizes, two-dimensional (2D) data, or limited number of points representing the three-dimensional (3D) shape. Therefore, the (in)validity of the symmetry assumption for the fibula, tibia, calcaneus, and talus is not yet established.

To gain insight in shape variations and (a)symmetry of the TCJ and STJ bones, we first analyzed whether similar shape patterns exist in left and right bones of the same type, and whether side-bias (i.e. directional asymmetry) [218, 219] appears. We then evaluated intra- and inter-subject bone shape variations in principal directions, which expressed the most variance in the shapes of bone samples. Moreover, the effects of gender on age-adjusted bone shape variations were assessed, as gender can lead to anatomical differences in lower extremities [206, 220, 221, 217, 80, 208, 222]. Unlike previous studies that have reduced the bone shape to a few anatomical landmarks [206, 217], we used 3D spatially dense descriptions of the TCJ and STJ bones together with advanced statistical techniques: a previously applied 3D statistical shape modeling method [162, 203] and distance based permutational statistics [219, 223].

6.2 Materials and Methods

The flow-chart presented in Fig. 6.1a outlines the main steps followed in this study. Briefly, in the first step, bilateral computed tomography (CT) images of individuals were collected. In the second step, both left and right TCJ and STJ

bones were segmented from all CT images, and triangulated bones surfaces were extracted from the segmentation results. In the third step, all bone surfaces of the same type were aligned into a common coordinate frame in such a way that the differences due to position, orientation, and scaling among bone instances were minimized. In the final step, the shape variations and (a)symmetry of the TCJ and STJ bones were analyzed. For each bone type, left and right bone samples were first considered as two separate groups and compared with each other on multiple aspects (i.e. group location, variance-covariance scale and orientation) using permutational statistics. The group location test (Fig. 6.1b) was performed to assess side-difference in the mean fibula (or tibia or calcaneus or talus) shape, in other words, the difference in central tendency. The variance-covariance scale (Fig. 6.1c) and orientation (Fig. 6.1d) tests were employed to analyze side-differences in the dispersion (i.e. the magnitude of shape variance only) and the shape variance directions around the mean fibula (or tibia or calcaneus or talus) shape, respectively. Following these analyses, all left and right bones of the same type were pooled into one group. Shape variations were extracted and statistically (i.e. an analysis of covariance, ANCOVA) compared between females and males after adjusting them for the effects of age. Moreover, intra- and inter-subject shape variations were assessed using the intraclass correlation coefficient (ICC). Each step is explained in detail in the following sub-sections.

6.2.1 Data Collection

Bilateral CT images of patients who had undergone CT scanning due to unrelated medical reasons (i.e. vascular indications) were collected from the Utrecht Medical Center (UMC, Utrecht, The Netherlands). Each CT image was evaluated by a trained medical doctor for signs of radiological osteoarthritis (OA) in the TCJs. Any CT scan exhibiting signs of moderate or severe ankle OA (unilateral or bilateral) [224] was excluded from the study. The final dataset consisted of 66 CT scans collected from 66 out of 99 individuals. The scans were anonymized. Only the gender and age (55 male, mean age: 61 and SD = 10 years; 11 female, mean age: 53 and SD = 15 years) of the patients were available to us. For the use of CT scans in this study, an approval from the Medical Ethical Committee of UMC was not necessary.
All CT scans were acquired either using Philips Brilliance 64 or Philips iCT scanner (Philips Medical Systems, Best, The Netherlands). The acquisition parameters were: tube voltage 120 kVp, effective dose 150 mAs, and slice thickness 1 mm. Tomographic reconstructions were made with a field of view covering both legs, a slice increment of 0.7 mm, and a matrix size of 512 $\times$ 512 pixels. The iDOSE4 reconstruction algorithm was used. Voxel sizes varied between 0.63 mm $\times$ 0.63 mm $\times$ 0.70 mm and 0.98 mm $\times$ 0.98 mm $\times$ 0.70 mm.

6.2.2 Segmentation of Bones from CT Scans

All left and right bones of the same type (i.e. 132 fibulae, 132 tibiae, 132 calcanei, and 132 tali) were segmented from the CT scans (Fig. 6.1a) using the freely

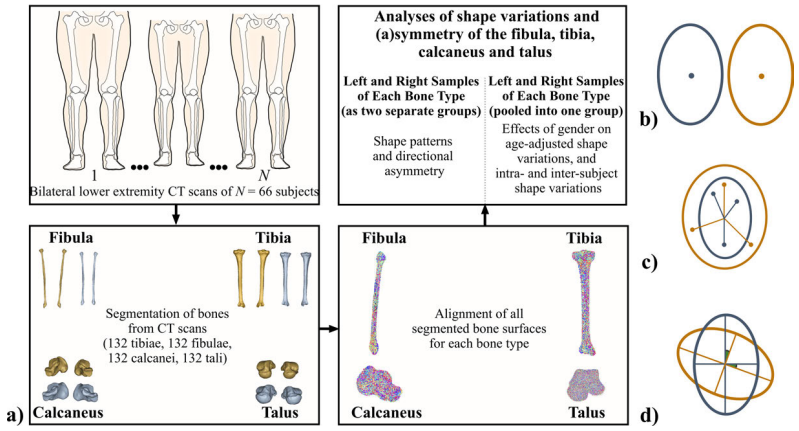


Figure 6.1: (a) A flow diagram of the study. Bilateral CT images were collected from 66 subjects. Left and right fibulae, tibiae, calcanei and tali were segmented from each CT scan. All bone samples of the same type were aligned into a common coordinate frame using an unbiased registration algorithm. For each bone type, shape variations and (a)symmetry were evaluated. (b-d) Multiple aspect analysis: group location, variance-covariance scale and orientation. Two groups that show difference in (b) their locations (c) variance-covariance scale and (d) variance-covariance orientation only. In the group location, variance-covariance scale and orientation tests, the features to concentrate on are (b) the sample mean, (c) the sample dispersion from the centroid and (d) the sample subspace described using eigenvectors and the principal angles between them, respectively.

available interactive graph-cut segmentation software MITK-GEM [225]. The triangulated bone surfaces were extracted from the segmentations. All right bones were mirrored in the sagittal plane.

6.2.3 Registration of Bones

All bone surfaces of the same type were brought into a common coordinate frame (Fig. 6.1a) using an unbiased registration algorithm [162, 203], which enabled us to minimize differences due to position, orientation, and scaling of bones. The registration parameters including the scale parameter for the mixture of Gaussians, σ , the number of points in the mean cloud, n_m, and the trade-off parameter, λ [162] were retrieved from [203] for all tali ($\sigma = 3mm$, $n_m = 2000$, $\lambda = 10^{-6}$). The registration parameters needed for alignment of the fibulae ($\sigma = 3mm$, $n_m = 2000$, $\lambda = 10^{-4}$), tibiae ($\sigma = 3mm$, $n_m = 2000$, $\lambda = 10^{-5}$) and calcanei ($\sigma = 3mm$, $n_m = 2000$, $\lambda = 5 \times 10^{-4}$) were defined based on the outcomes of numerical experiments performed in a way described previously [162, 203].

6.2.4 Statistical Analyses Ipsi- and Contralateral Sides as Separate Groups

Following the registration process, dense correspondence between aligned bone surfaces of the same type was automatically established [166, 203]. The number of corresponding points settled on all aligned fibulae ($n = 12465$), tibiae ($n = 31496$), calcanei ($n = 7717$), and tali ($n = 5541$) was approximately 0.7 times of the bone surface area averaged over all bone samples of the same type.
For the first set of comparisons, the left and right fibulae (or tibiae or calcanei or tali) were considered as two separate groups and compared with each other on multiple aspects including group location, variance-covariance scale, and orientation using distance-based permutational approaches in a similar manner as described in [219, 223]. The Euclidean distance between the means of two groups was employed as D(istance)-statistic in the group location test [219, 223] (see Appendix D). The left and right fibulae (or tibiae or calcanei or tali) were permutated 10000 times (N_{perm}) across groups and D-statistic was calculated at each permutation ($D\text{-}statistic_{perm}$). A p-value assessed under permutation (p_{perm}) was determined by:

$$p_{perm} = \frac{N_i}{N_{perm}} \tag{6.1}$$

where, N_i represents the number of cases in which permutated values are higher or equal to D-statistic (i.e. $D - statistic_{perm} \geq D - statistic$).
D-statistic used in the variance-covariance scale test was the absolute difference in the average residual of the two groups [219, 223] (see Appendix D). The permutations were realized and permutated values (i.e. $D\text{-}statistic_{perm}$) were obtained in exactly the same way as described for the group location test. A p_{perm}-value was calculated using Eq. 6.1.
In the variance-covariance orientation test, two shape subspaces represented with eigenvectors (i.e. principal components (PC) or modes of shape variation) and the principal angles between them were compared. To obtain the shape subspace of the left bones of the same type, a principal component analysis (PCA) was performed on the covariance matrix of the data vectors that consisted of the 3D coordinates of the corresponding points established on all left bones. The shape subspace of the right bones of the same type was obtained in the same manner. The D-statistic used in the orientation test was the projection metric [219, 223, 226]:

$$D_k = \sqrt{k - \sum_{i=1}^{k} cos^2\theta_i} \tag{6.2}$$

in which θ_i $i = 1, \ldots, k$ are the principal angles [227, 228] and k is the number of PCs to be kept. To determine the number of PCs (k) to retain, a parallel analysis (PA) [229] (see Appendix D) was carried out (Fig. 6.2). A set of k distances (D_k, D-statistic) was calculated based on Eq. 6.2 by incrementally increasing

the number of principal angles from 1 to k. Then, the left and right bones of the same type were permutated 10000 times across groups. At each permutation, the shape subspaces of the left and right bones of the same type were re-determined using the permutated data, and a set of k distances (D_k, $D\text{-}statistic_{perm}$) was computed. All p_{perm} values were determined based on Eq. 6.1. The group location, variance-covariance scale and orientation tests were carried out using the statistical routines developed [219] in Matlab (Matlab R20013b, The Mathworks, Inc., Natick, MA).

6.2.5 Statistical Analyses Ipsi- and Contralateral Sides Pooled into One Group

To describe the shape variations in the bones of the same type and analyze the side-shape differences in given shape variance directions, ipsi- and contralateral bones of the same type were pooled into one group and PCA was applied on the covariance matrix of the combined data vectors [203]. As a result, the shape parameters were obtained for each bone type, which expressed how much each bone sample deviated from the mean bone shape in given shape variance directions (i.e. PCs, eigenvectors, or modes of shape variation) [203, 159]. A Kolmogorov-Smirnoff test was carried out to evaluate whether the shape parameters given for each of the k PCs conformed to a normal distribution. The shape variations in each bone type adjusted for the effects of age were compared between males and females using an ANCOVA. Moreover, the intra-subject shape variation was compared to that of inter-subject for each of the k PCs using the ICC. A single measurement, absolute-agreement, and two-way random effects model [230] was employed for the latter analyses. All ICC estimates and their 95% confidence intervals (CIs) were reported. An ICC estimate of 1 indicated perfect symmetry within an individual. In other words, the total bone shape variation was described by the inter-subject shape difference only. In contrast, an ICC estimate of 0 implied that the intra- and inter-subject shape variations were equal to each other. When the 95% CI of the ICC for a PC included zero, we deemed the PC not significantly symmetrical. All the statistical analyses were conducted using SPSS (Version 22, Chicago, IL).

6.3 Results

6.3.1 Shape Patterns in Ipsi- and Contralateral Sides as Separate Groups

The p_{perm} values were higher than the statistical significance level of 0.05 for the group location, variance-covariance scale and orientation tests (Table 6.1). Therefore, the side-differences in the mean shape of the fibula, tibia, calcaneus and talus were not statistically significant (Table 6.1, group location test). Moreover, variations (Table 6.1, variance-covariance scale) and differences in shape variance directions (Table 6.1, variance-covariance orientation) around the mean

shape of each bone type were not significantly different between left and right.

Table 6.1: D_{stat} and p_{perm} values resulting from the group location (1st row), variance-covariance scale (2nd row) and orientation (3rd – 17th rows) tests performed for the fibula, tibia, calcaneus and talus. D_{stat} represents the Euclidean distance between the means of left and right groups (the group location test), the absolute difference in the average residual of the left and right groups (the variance-covariance scale test), and the projection metric (the variance-covariance orientation test). p_{perm} describes a p-value obtained under $N_{perm} = 10000$ permutations.

	Fibula		**Tibia**		**Calcaneus**		**Talus**	
	D_{stat}	p_{perm}	D_{stat}	p_{perm}	D_{stat}	p_{perm}	D_{stat}	p_{perm}
Location	0.53	0.84	0.68	0.74	0.62	1.00	0.87	0.98
Scale	0.08	0.60	0.00	0.99	0.04	0.77	0.16	0.42
Orientation	0.13	0.98	0.11	1.00	0.20	1.00	0.22	0.92
1:1:15	0.20	0.99	0.16	1.00	0.35	1.00	0.35	0.97
principal	0.28	0.99	0.24	1.00	0.47	1.00	0.44	1.00
angles	0.37	0.95	0.30	1.00	0.58	1.00	0.57	0.99
	0.46	0.98	0.37	1.00	0.68	1.00	0.69	0.98
	0.57	0.99	0.46	1.00	0.79	1.00	0.80	1.00
	0.77	0.99	0.56	1.00	0.91	1.00	0.92	1.00
	1.25	0.94	0.73	1.00	1.02	1.00	1.04	1.00
	-	-	-	-	1.15	1.00	1.17	1.00
	-	-	-	-	1.29	1.00	1.34	1.00
	-	-	-	-	1.44	1.00	1.51	1.00
	-	-	-	-	1.60	1.00	1.70	1.00
	-	-	-	-	1.77	1.00	1.95	1.00
	-	-	-	-	2.01	1.00	2.19	1.00
	-	-	-	-	2.23	1.00	-	-

6.3.2 Shape Patterns in Ipsi- and Contralateral Sides Pooled into One Group

The number of PCs kept for each bone type (Fig. 6.2) (i.e. $k = 8$ for the fibula and tibia, $k = 15$ for the calcaneus, and $k = 14$ for the talus) were used to describe shape variations. These PCs explained 79%, 84%, 67%, and 72% of the total shape variation in the fibula, tibia, calcaneus, and talus, respectively. None of the p-values resulting from a Kolmogorov-Smirnoff test were lower than the statistical significance level of 0.05, meaning that the shape parameters of all 132 fibulae (or tibiae or calcanei or tali) given for a specific PC came from a normal distribution. For each bone type, changes observed along the first three PCs were in:

- (PC 1 of the fibula, Fig. 6.3) the length of the fibula,
- (PC 2 of the fibula, Fig. 6.3) the curvature of the fibula shaft (bending around the anterior-posterior axis),
- (PC 3 of the fibula, Fig. 6.3) the curvature of the fibula shaft (bending around the medial-lateral axis),

- (PC 1 of the tibia, Fig. 6.3) the anterior border along the tibia shaft, and in the lateral and medial condyles,
- (PC 2 of the tibia, Fig. 6.3) the length of the tibia,
- (PC 3 of the tibia, Fig. 6.3) the distal and proximal epiphyses,
- (PC 1 of the talus,Fig. 6.4) the lateral rotation of the talar head,
- (PC 2 of the talus, Fig. 6.4) the medial tubercle and the body of the talus,
- (PC 3 of the talus, Fig. 6.4) the length of the talus, and in the lateral projection of the talar lateral process,
- (PC 1 of the calcaneus, Fig. 6.4) the length (i.e. anterior-posterior direction) and height (i.e. superior-inferior direction) of the calcaneus,
- (PC 2 of the calcaneus, Fig. 6.4) the inclination of the sustentaculum tali,
- (PC 3 of the calcaneus, Fig. 6.4) the medial and lateral tuberosity, and in the talar articulating surfaces.

Shape changes described by other remaining PCs are not presented here. The reasoning behind the choice made is that PCs higher than PC 3 explained relatively small shape variations distributed over the bone surfaces and most of them, except a few mentioned in the coming two subsections (see sections 6.3.3 and 6.3.4), did not express significantly different bone shape variations between males and females, nor did they describe higher intra-subject shape variations than inter-subject.

6.3.3 Effects of Gender on Age-Adjusted Shape Variations

After adjusting the statistical significance level of 0.05 with Bonferroni, one PC for the tibia (i.e. PC 1 of the tibia, $p = 0.003$), one PC for the talus (i.e. PC 8 of the talus, $p = 0.001$) and two PCs for the calcaneus (i.e. PC 1, $p < 0.001$ and PC 7 of the calcaneus, $p = 0.001$) expressed significant shape differences between the tibiae, tali and calcanei of male and female subjects, respectively (Table 6.2). Shape changes observed along PC 1 of the tibia (Fig. 6.3) and PC 1 of the calcaneus (Fig. 6.4) are presented in the previous section. Regarding PC 7 of the calcaneus (Fig. 6.4), shape changes were mainly in the posterior height, and in the posterior and anterior talar articular surfaces. PC 8 of the talus (Fig. 6.4) expressed changes in the posterior aspect of the talus. The shape variations in the fibula, tibia, calcaneus and talus described by the other remaining PCs did not significantly differ based on gender (Table 6.2).
The box plots showing the distribution of shape parameters observed within the studied population along PC 1 of the tibia, PC 1 and PC 7 of the calcaneus, and PC 8 of the talus are provided in Fig. 6.5a. The means of the shape parameters (i.e. along PC 1 of the tibia) for the tibiae of females and males were -0.568 and 0.114, respectively. Accordingly, females had relatively smaller lateral and medial condyles (deviation from the mean tibia shape in the negative direction of

the PC 1 of the tibia, Fig. 6.3) in average as compared to those of males (deviation from the mean tibia shape in the positive direction of the PC 1 of the tibia, Fig. 6.3). The means of the shape parameters for the calcanei of females along PC 1 and PC 7 of the calcaneus were 0.678 and -0.627, respectively (Fig. 6.5a). With reference to the calcanei of males, the means of the shape parameters were -0.1355 and 0.1253 along PC 1 and PC 7 of the calcaneus, respectively (Fig. 6.5a). Deviation from the mean calcaneus shape in the negative direction of the PC 1 of the calcaneus (Fig. 6.4) expressed the shortening and enlargement of the calcaneus in length (i.e. anterior-posterior direction) and height (i.e. superior-inferior direction), respectively. Observed changes in the calcanei having positive shape parameters along PC 1 of the calcaneus (Fig. 6.4) were vice-versa. Therefore, calcanei of females were on average longer in length (i.e. anterior-posterior direction) and shorter in height (i.e. superior-inferior direction) as compared to those of males. Deviation from the mean calcaneus shape in the negative direction of the PC 7 of the calcaneus, like observed in the calcanei of females (i.e. the mean of the shape parameters = -0.627), expressed the enlargement of an angle located between the posterior and anterior talar articular surfaces, and the lateral process of calcaneal tuberosity. The means of the shape parameters observed along PC 8 of the talus were 0.492 and -0.098 for females and males, respectively (Fig. 6.5a). Changes observed along PC 8 of the talus (Fig. 6.4) were in the posterior aspect of the talus. Deviation from the mean talus shape in the negative direction of this PC displayed enlargement of the posterior talar articular contour (Fig. 6.4) suggesting that tali of males had larger posterior aspect on average.

Table 6.2: *p*-values resulting from ANCOVA tests. The *p*-values under Bonferroni adjusted significance level of 0.006 (= 0.05/8) for the fibula and tibia, 0.003 (= 0.05/15) for the calcaneus and 0.004 (= 0.05/14) for the talus are marked with *.

Shape Mode	Fibula	Tibia	Calcaneus	Talus
1	0.341	0.003*	0.000*	0.219
2	0.406	0.777	0.589	0.426
3	0.735	0.971	0.159	0.476
4	0.032	0.045	0.013	0.521
5	0.056	0.278	0.005	0.043
6	0.901	0.844	0.022	0.248
7	0.047	0.583	0.001*	0.657
8	0.086	0.061	0.068	0.001*
9	-	-	0.236	0.703
10	-	-	0.273	0.036
11	-	-	0.661	0.922
12	-	-	0.336	0.135
13	-	-	0.911	0.024
14	-	-	0.143	0.957
15	-	-	0.334	-

Point-to-surface distance is presented in Fig. 6.5b, which was calculated between the female and male tibiae deviating from each other the most along PC 1 of the tibia. Similarly, three other point-to-surface distances calculated for the cases: (1) PC 1 and (2) PC 7 of the calcaneus, and (3) PC 8 of the talus, are given in Fig. 6.5b. Referring to Fig. 6.5b, point-to-surface distances exceeding 9 mm, 4 mm and 2 mm were observed between the tibiae (i.e. varied along PC 1 of the tibia), calcanei (i.e. varied along PC 1 or PC 7 of the calcaneus) and tali (i.e. varied along PC 8 of the talus) of females and males, respectively.

6.3.4 Intra- and Inter-Subject Shape Variations

For all of the forty-five PCs, except two for the fibula (i.e. PC 3 and PC 6 of the fibula) and one for the tibia (i.e. PC 8 of the tibia), intra-subject shape variations were smaller than inter-subject variations (Table 6.3). The ICC estimates and their 95% CI (Table 6.3) were 0.03 (-0.21-0.27), -0.04 (-0.28-0.20), and 0.06 (-0.19-0.29) for the PC 3 and PC 6 of the fibula, and PC 8 of the tibia, respectively. These PCs described changes in the curvature of the fibula shaft and the diameter of the fibula (Fig. 6.3, PC 3 and PC 6 of the fibula), and in the tibial tuberosity together with the diameter of the tibia (Fig. 6.3, PC 8 of the tibia).
The distributions of shape parameters observed along PC 8 of the tibia, and PC 3 and PC 6 of the fibula for both left and right sides are presented in Fig. 6.5c. The means of the shape parameters for left and right tibiae (i.e. varied along PC 8 of the tibia, Fig. 6.5c) were -0.028 and 0.028, respectively. Regarding to the left fibulae, the means of the shape parameters observed along PC 3 of the fibula and PC 6 of the fibula (Fig. 6.5c) were -0.082 and 0.098, respectively. Similarly, the means of the shape parameters were 0.082 (i.e. along PC 3 of the fibula, Fig. 6.5c) and -0.098 (i.e. along PC 6 of the fibula, Fig. 6.5c) for the right fibulae.
Point-to-surface distance calculated between a pair of tibia is presented in Fig. 6.5d, which deviated from each other the most along PC 8 of the tibia. In the same way, point-to surface distances are provided in Fig. 6.5d for two pairs of fibulae varying the most along PC 3 and PC 6 of the fibula. Referring to Fig. 6.5d, distances exceeding 2 mm were observed along the surfaces of each pair.

6.4 Discussion

The bilateral symmetry of the TCJ and STJ bones is often assumed in clinical practice and research studies. Nevertheless, the validity of the symmetry assumption is not yet established due to limited information on the shape variations and (a)symmetry of the fibula, tibia, calcaneus and talus. In this study, using detailed 3D bone shape data and advanced statistical techniques, we addressed whether: (1) both sides of each bone type exhibit similar shape patterns, and a side bias (i.e. directional asymmetry) [218, 219] exists, (2) gender has an influence on bone shape differences, and (3) intra-subject shape variations are smaller than those of inter-subject for given shape variance directions.
The ipsi- and contralateral sides of the TCJ and STJ bones (Table 6.1) had similar

Table 6.3: Intra-class correlation (ICC) estimates and their 95% confidence intervals (CIs) for the fibula, tibia, calcaneus and talus. Shape variance directions in which intra-subject shape variations were comparable to those of inter-subject are marked with *.

Shape Mode	Fibula		Tibia		Calcaneus		Talus	
	ICC	95% CI	ICC	95% CI	ICC	95% CI	ICC	95% CI
1	0.72	0.58-0.82	0.85	0.77-0.91	0.95	0.92-0.97	0.56	0.37-0.71
2	0.74	0.60-0.83	0.50	0.29-0.66	0.84	0.75-0.90	0.35	0.12-0.54
3	0.03*	-0.21-0.27	0.48	0.27-0.64	0.85	0.77-0.91	0.55	0.35-0.69
4	0.47	0.25-0.64	0.43	0.22-0.61	0.82	0.72-0.88	0.43	0.22-0.61
5	0.67	0.51-0.78	0.72	0.58-0.82	0.85	0.76-0.90	0.40	0.18-0.58
6	−0.04*	-0.28-0.20	0.87	0.80-0.92	0.72	0.58-0.82	0.27	0.03-0.48
7	0.66	0.49-0.77	0.59	0.41-0.73	0.90	0.85-0.94	0.67	0.51-0.78
8	0.40	0.18-0.58	0.06*	-0.19-0.29	0.74	0.61-0.83	0.67	0.51-0.78
9	-	-	-	-	0.80	0.70-0.88	0.45	0.23-0.62
10	-	-	-	-	0.82	0.72-0.88	0.59	0.41-0.73
11	-	-	-	-	0.83	0.74-0.90	0.49	0.28-0.65
12	-	-	-	-	0.82	0.73-0.89	0.29	0.06-0.50
13	-	-	-	-	0.81	0.70-0.88	0.37	0.14-0.56
14	-	-	-	-	0.80	0.69-0.87	0.46	0.25-0.63
15	-	-	-	-	0.67	0.51-0.79	-	-

shape patterns. There was no indication for left or right bias in any bone type (Table 6.1). Behavioral studies on the lower limb laterality [231, 232], have found right-footedness to be more prevalent. Nevertheless, it has been also stated that contralateral non-preferred foot supports the activities (e.g. kicking, stamping) of the dominant foot by contributing to postural stability [231, 232, 206]. Based on these studies, it seems plausible that contralateral non-preferred extremity is subjected to more or less the same mechanical loads as the dominant limb. Therefore, left-side or right-side bias may not be present for the fibula, tibia, calcaneus, and talus.

Gender led to tibial, calcaneal and talar shape differences in four shape variance directions (PC 1 of the tibia, Fig. 6.3; PC 1 and PC 7 of the calcaneus, Fig. 6.4; PC 8 of the talus, Fig. 6.4). Considering PC 1 of the tibia (Figs. 6.3 and 6.5a), females had relatively smaller lateral and medial condyles. Our findings are in agreement with the outcomes of previous studies [220, 222]. In [220], large variations in mediolateral dimensions were observed. Moreover, smallest tibiae were predominantly found in females [220]. In [222], smaller tibial head widths were reported for females as compared to those of males. Regarding PC 8 of the talus (Figs. 6.4 and 6.5a), tali of males had a relatively larger posterior aspect. This outcome is in line with previous studies [233, 234], which analysed the morphology of the talus and its sexual dimorphism. Furthermore, relatively larger values for the talar breadth and surface area have been reported for male tali [233, 234]. On average, female calcanei in our study seem to be longer in length (i.e. anterior-posterior direction) and shorter in height (i.e. superior-inferior direction). This observation is not in full agreement with other studies [235, 236, 233], since the average length and height of the calcaneus have been reported to be larger in males.

In this study, the sample size and the number of females involved in the dataset

is limited. The disproportion of males and females could affect the generalizability of all the analyses, except the one performed to assess the effects of gender on age adjusted shape variations. In the latter analysis, the small dataset could have an effect on the statistical power. To analyze the effects of gender on age adjusted shape variations, multiple comparisons were performed. Therefore, the statistical significance level of 0.05 was adjusted according to Bonferroni with the aim of reducing type I errors. The cost of this correction is an increased probability of type II errors (i.e. reduced power). Another limitation of this study is that bone shape differences between females and males, and bilateral (a)symmetry could not be studied based on different age groups. The inclusion of younger and older individuals might impact the generalizability of the findings reported in this study. Future studies should aim for a larger and more representable population sample (e.g. more females), to increase statistical power and generalizability of the results.

Intra-subject shape variations in the talus and calcaneus along each of k PCs (fourteen and fifteen shape variance directions for the talus and calcaneus, respectively) were smaller than those of inter-subject (Table 6.3). These results suggest that the shapes of the calcaneus and talus were more symmetric within an individual than between subjects. Due to a scarcity of information on bilateral shape (a)symmetry of the talus and calcaneus, it is not easy to compare the findings of this study with those of others. The study presented by Islam et al. [161] is one of the references that can be referred to. Although the methods followed in [161] are different than those presented here, their observations made using CT data of 11 intact tali (eight male and three female subjects) imply that the shape of talus is bilaterally symmetric. Regarding to the calcaneus, intra-subject variations within the anatomy were reported to be smaller than those of inter-subject in [237] for the area and 3D orientation of the joint surfaces of the calcaneus. The outcomes of this study on the bilateral (a)symmetry of the talus and calcaneus, and those of the studies [161, 237] imply that the shapes of the talus and calcaneus are bilaterally symmetric, and the shape of contralateral side can be used as a control during a surgery (e.g. anatomical reconstruction of the calcaneus of a patient with a calcaneus fracture) or as a shape template for implant design.

Intra-subject shape variations were in general smaller than those of inter-subject for the tibia and fibula (Table 6.3), 95% CIs did not include zero for thirteen out of sixteen shape variance directions). Nevertheless, the curvature of the fibula shaft (bending around the medial-lateral axis, PC 3 of the fibula, Fig. 6.3), the diameter of the fibula (PC 6 of the fibula, Fig. 6.3), and the tibial tuberosity together with the diameter of the tibia (PC 8 of the tibia, Fig. 6.3) varied within a subject as much as between individuals (Table 6.3). One of the explanations for observing different (a)symmetry level in the cross-sectional dimensions of the fibula and tibia as compared to those seen in other bone features, such as the length (i.e. PC 1 of the fibula and PC 2 of the tibia, Fig. 6.3) may be that different structural features within the same bone exhibit independent development. For example, the subperiostal growth of bone cortices could endure throughout life, although a long bone stops growing in length after the closure of the epiphyseal growth plate [206]. The cross-sectional dimensions of skeletally mature weight bear-

ing bones could be more sensitive to mechanical loadings. As bone cortices can grow, changes in the cross-sectional dimensions could be observed while bone adapts to its mechanical environment. Referring to [238], the threshold at which bone deposition/reposition is stimulated is not constant, but varies with respect to several intrinsic (e.g. genetic factors, and the age and hormonal status of individuals) and extrinsic (e.g. loading history, the frequency of loading) factors. Considering these aspects, it seems plausible that bilateral differences could exist in bone stimulation threshold, bone deposition/reposition within individuals depending on their foot preference and physical activities that can affect the contribution of none-preferred limb to their postural stability. The cross-sectional dimensions of the fibula and tibia may influence the determination of implant size and its placement in arthroplasty surgery, while the curvature of the fibula can be relevant for planning corrective osteotomy. Therefore, side-shape differences in the fibula (PC 3 and PC 6 of the fibula, Fig. 6.3) and tibia (PC 8 of the tibia, Fig. 6.3) may adversely affect the success of an arthroplasty surgery and corrective osteotomy performed under shape symmetry assumption.
A strength of this study is that all bone samples were spatially dense sampled in 3D. This enabled us to cover bony regions that cannot be described with a set of conventional 2D or 3D measurements. Using 3D spatially dense data, we analyzed for the first-time shape variations and (a)symmetry of all the bones forming the TCJ and STJ. Considering the nature of the PCA to describe shape variations, it is wise to mention that isolated locations of asymmetry may not be sufficiently captured [239]. An independent component analysis (ICA) could be used as an alternative to PCA to describe shape variations in a more localized way. We refer interested readers to [240, 241] for the details on ICA. Although ICA has potential extracting substantially distinct features, it has been less often used in the area of statistical shape analysis [242]. Therefore, PCA was preferred in this study, which is the most commonly used technique to describe shape variations [242, 243].

6.5 Conclusions

We observed that both sides of the bones forming the talocrural (TC) and subtalar (ST) joints exhibited similar shape patterns, and directional asymmetry did not exist in any bone type (i.e. fibula, tibia, calcaneus, and talus). Gender did not explain, in general, significantly different shape variations in all above-mentioned bones. Nevertheless, four shape variance directions described statistically significant shape differences between the tibiae (i.e. changes in the anterior border along the tibia shaft, and in the lateral and medial condyles), calcanei (i.e. changes in the length and height of the calcaneus, and in the posterior height, and in the posterior and anterior talar articular surfaces) and tali (i.e. changes in the posterior aspect of the talus) of females and males, after Bonferroni adjustment. The shape symmetry assumption was in general valid. However, intra-subject shape variations were as high as those of inter-subject in the shape variance directions describing changes in the curvature of the fibula shaft, the

diameter of the fibula, and the tibial tuberosity together with the diameter of the tibia. These observations indicate that the symmetry assumption may be violated. Deviation from symmetry in the fibula and tibia may adversely affect the outcomes of studies using the contralateral side as a shape template or intra-subject control, and the success of an arthroplasty surgery or corrective osteotomy performed with shape symmetry assumption.

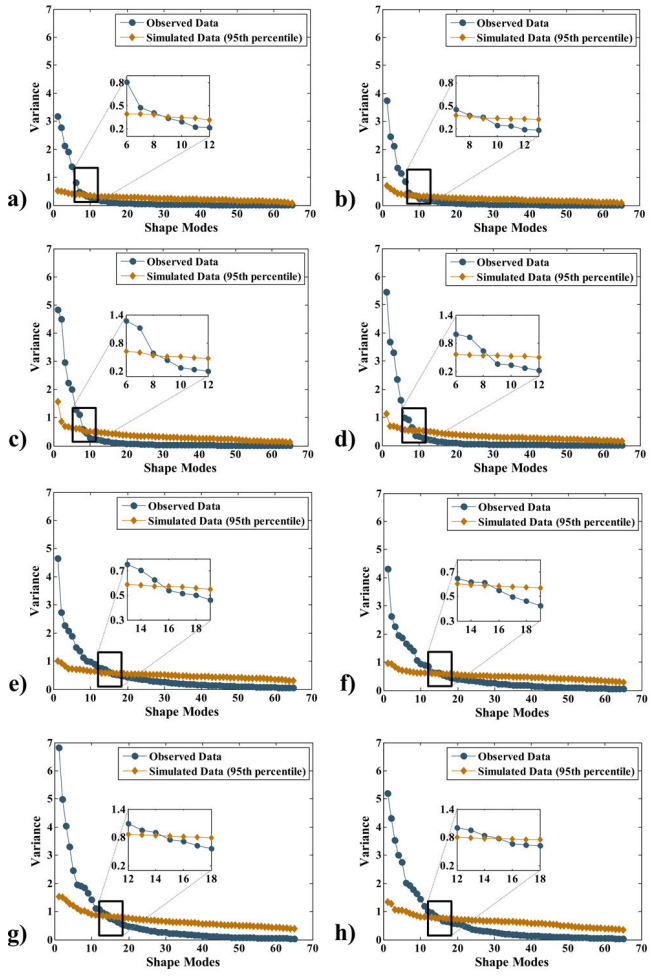


Figure 6.2: The scree plots with parallel analyses are given for (a) left fibulae, (b) right fibulae, (c) left tibiae, (d) right tibiae, (e) left calcanei, (f) right calcanei, (g) left tali and (h) right tali. Blue and dark gold markers stand for observed and simulated data, respectively. All the principal components up to the one found at the intersection of two lines (lines represented with blue and dark gold colors) were retained and used in the variance-covariance orientation test and description of bone shape variations within a studied population. Accordingly, the number of principal components kept is (a-b) 8 for left and right fibulae (c-d) 8 for left and right tibiae (e-f) 15 for left and right calcanei and (g-h) 14 for left and right tali.

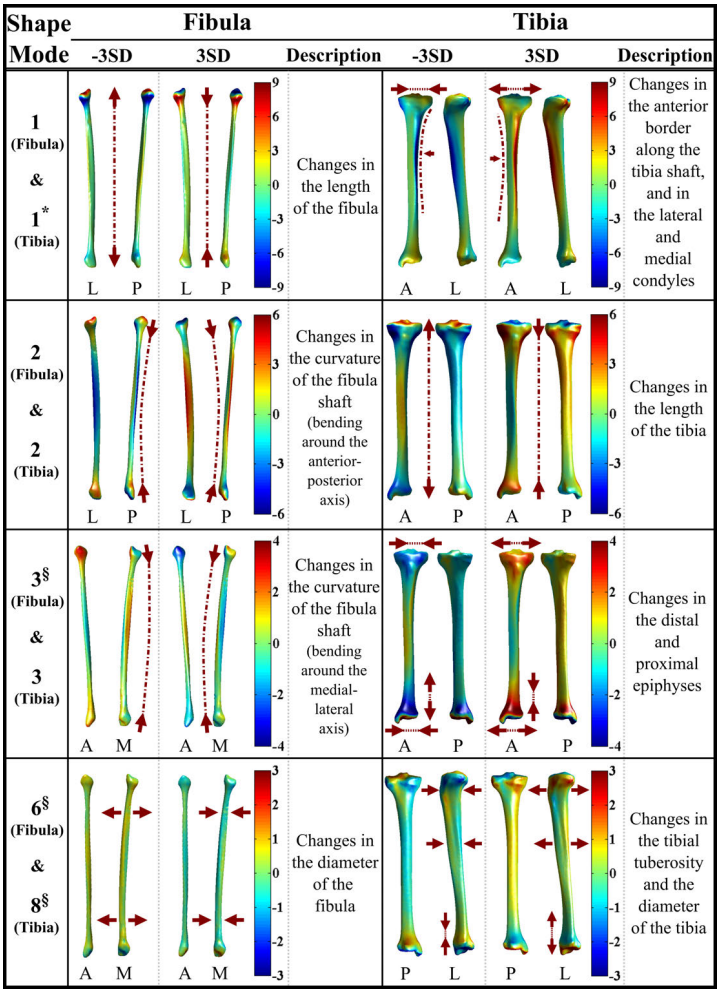


Figure 6.3: The first three rows display deviations (mm) of the fibula and tibia from the mean fibula shape (left column) and the mean tibia shape (right column) in the positive (+3SD) and negative (-3SD) directions of the first three PCs (i.e. principal components) of the fibula and tibia, respectively. Shape variations of the fibula and tibia explained by PC 6 of the fibula and PC 8 of the tibia, respectively, are shown in the fourth row. The shape variance directions that expressed significantly different shape variations between females and males are marked with '*'. The marker '§' is used to indicate shape variance directions for which intra- and inter-subject shape variations were comparable to each other.

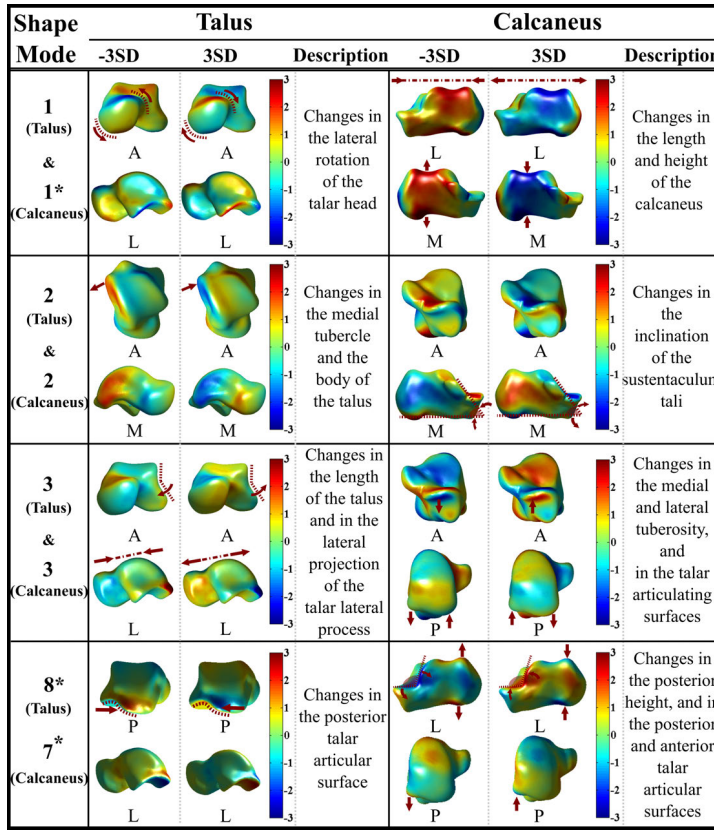


Figure 6.4: The first three rows display deviations (mm) of the talus and calcaneus from the mean talus shape (left column) and the mean calcaneus shape (right column) in the positive (+3SD) and negative (-3SD) directions of the first three PCs (i.e. principal components) of the talus and calcaneus, respectively. Shape variations in the talus and calcaneus explained by PC 8 of the talus and PC 7 of the calcaneus are shown in the fourth row. Shape variance directions that expressed significantly different shape variations between females and males are marked with '*'.

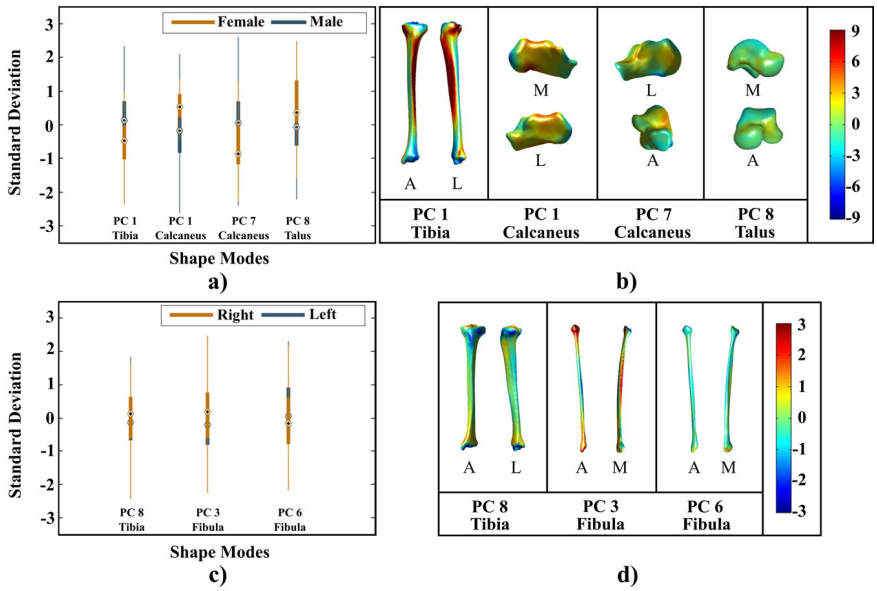


Figure 6.5: (a) Box plots showing the distributions of shape parameters observed along PC 1 of the tibia, PC 1 and PC 7 of the calcaneus, and PC 8 of the talus. These PCs described statistically significant shape variations between males (blue color) and females (dark gold color). (b) Point-to-surface distances (mm) calculated between female and male (1) tibiae deviating from each other the most along PC 1 of the tibia, (2) calcanei deviating from each other the most along PC 1 of the calcaneus and (3) PC 7 of the calcaneus, and (4) tali deviating from each other the most along PC 8 of the talus. (c) The distributions of shape parameters observed for the left (blue color) and right (dark gold color) sides along PC 8 of the tibia, PC 3 and PC 6 of the fibula. Intra-subject shape variations in the tibia and fibula described by these PCs were similar to those of inter-subject. (d) Point-to-surface distances (mm) calculated between left and right (1) tibiae deviating from each other the most along PC 8 of the tibia, (2) fibulae deviating from each other the most along PC 3 of the fibula and (3) PC 6 of the fibula.

7

Three-Dimensional Registration of Freehand-Tracked Ultrasound to CT Images of the Talocrural Joint

Published as:
N. Tümer, A.C. Kok, F.M. Vos, G.J. Streekstra, C. Askeland, G.J. Tuijthof, A.A. Zadpoor. Three-Dimensional Registration of Freehand-Tracked Ultrasound to CT Images of the Talocrural Joint. Sensors. 18 (2018) 2375.

7.1 Introduction

(Osteo)chondral defects (OCDs) of the talocrural joint predispose patients to premature osteoarthritis (OA), if left untreated [38, 244, 157]. Adequate treatment of OCDs is highly essential, since both OCD and OA adversely affect the patients' quality of life with joint pain and dysfunction [244].
There are various surgical treatment options for talar OCDs including "the excision of the lesion, excision and curettage, excision combined with curettage and drilling/microfracturing, placement of an autogenous bone graft, antegrade drilling, retrograde drilling, osteochondral transplantation and autologous chondrocyte implantation" [36]. The failure rates (i.e. number of unsuccessfully treated patients $\times$ 100 / number of treated patients) of these treatment strategies can go up to 70% (range 0-70%) [36, 173]. The exact reasons for the failure of the treatment are often not completely clear, preventing formulation of effective individual treatment strategies. This could, at least partially, be due to a lack of knowledge on cartilage tissue regeneration in-vivo [245]. To understand the nature of cartilage healing in-vivo and in turn to devise treatment scenarios, longitudinal monitoring of OCDs of patients in short intervals could be helpful [41, 246].
Ultrasound (US) imaging facilitates frequent evaluation of patients [247] in time, because it is non-invasive, cost-effective, and easily accessible [248]. Recent studies have shown that US could be used to image cartilage and bony abnormalities (e.g. cortex irregularities) [41, 249]. Reproducible evaluation of OCDs with US would be, however, challenging. This is because, US images cannot so easily be interpreted and compared over time without having access to the talocrural joint anatomy provided by other modalities, such as computed tomography (CT). To monitor the post-operative tissue regeneration and evaluate any changes in the OCDs as a result of the treatment, US images of the patient could therefore be registered to and used in conjunction with pre-operative CT scans of the same patient.
In the literature, numerous US to CT bone registration schemes have been previously presented, which include feature- and/or intensity-based techniques [250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262]. Referring to feature-based techniques, the registration of bone surfaces extracted from both CT and US volumes using a variant of the iterative closest point (ICP) [263] is a common approach. In the study carried out by Muratore et al. [256], ICP technique was used to register vertebral bone surfaces derived from both CT and US volumes. Barratt et al. [250] presented a surface-based registration algorithm that can simultaneously update the freehand tracked US system calibration parameters. Another point based US to CT registration scheme was described by Moghari et al. [254]. In [254], Unscented Kalman Filter was used to estimate rigid transformation parameters needed for registering two point sets extracted from CT and US images. In the study presented by Brounstein et al. [251], point clouds derived from CT and US volumes were represented as Gaussian Mixture Models (GMM). The registration of two point clouds was realized by minimizing the dissimilarity measure (i.e. L2 distance metric) [251]. In the study [252], the previ-

ously proposed GMM surface registration algorithm was improved. Regarding to intensity-based registration techniques, Penney et al. generated probability images from both the CT and US volumes. A normalized cross correlation metric was used as a similarity measure to register the probability images [255]. Winter et al. [257, 264] described a surface-volume registration scheme to register US volume with enhanced bone contours to bone surface points culled from CT volume. Nagpal et al. [260] presented a multi-body registration technique, which maximizes the similarity between CT and US data using features derived from the two modalities and voxel intensity information.

All these previously presented registration schemes have been applied to long bones [255, 259], vertebrae [257, 260, 261], and pelvis [254, 255, 262]. Although, there is no limitation in theory for applying previously described registration techniques to the talocrural joint, none of the studies have focused on this anatomical location. It is therefore not clear how these registration techniques perform individually and relative to each other when dealing with US images of the talocrural joint. Accordingly, a rigid surface-volume registration scheme is presented in this study to match CT and freehand-tracked US images of the talocrural joint. The robustness of the registration algorithm to different initialization positions and to the sampling of US data was evaluated on a dataset collected from 12 cadaveric ankles. The presented scheme was also used in clinical practice to evaluate the potential of US imaging as an alternative to a CT scan in pre-operative planning of arthroscopic access to anterior talar OCDs [265]. The contributions of this study are twofold: (1) the talocrural joint is the main focus as an anatomical location, and (2) the question about the potential use of US imaging in clinics has been addressed for the first time, while assessing the performance of the registration scheme.

7.2 Materials and Methods

An overview of the registration scheme is presented in Fig. 7.1. Initially, freehand-tracked 2D US images and a CT scan of the talocrural joint of the same patient are acquired. In the pre-processing step of the CT scan, a point set is extracted from the distal tibial and talar bone surfaces. Simultaneously, bone contours in 2D US data are enhanced based on the monogenic signal representation of 2D US images [266]. Subsequently, 3D US image with enhanced bone contours is reconstructed from the preprocessed 2D US data and using the position of the conventional US probe that was recorded with an optical tracking system. During the registration of the surface extracted from the CT scan to the monogenic signal feature volume, six transformation parameters (three for rotation and three for translation) are estimated so as to optimize the sum of monogenic signal features over the transformed surface extracted from the CT data. The CT and US data are considered as "fixed" and "moving" images, respectively, during the registration.

In the following paragraphs, the methods used for acquisition of freehand-tracked 2D US images and CT scan of the talocrural joint are provided first. Subse-

quently, the steps in the proposed registration scheme are extensively described. Then, the details of the tests performed on the acquired data to evaluate the precision of the registrations are provided. Finally, the methodology applied in a clinical study is presented.

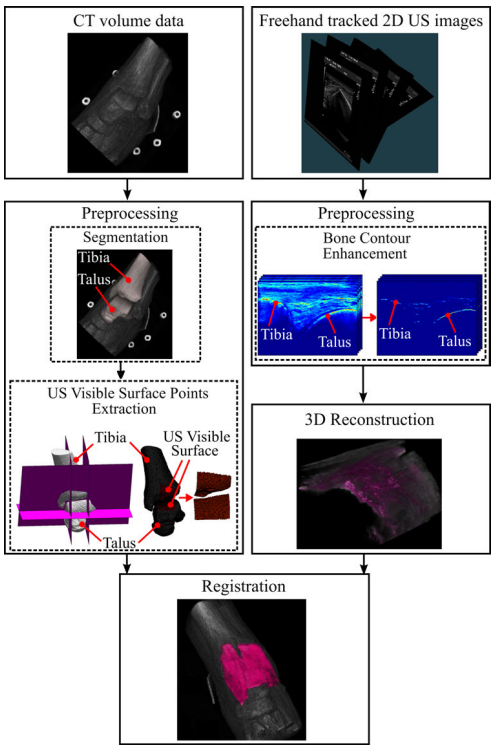
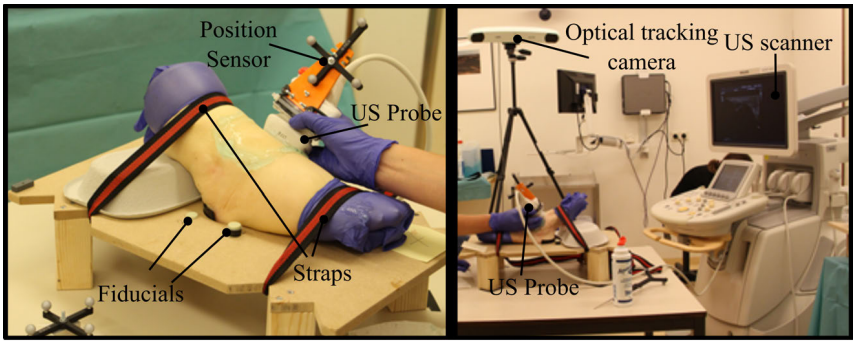


Figure 7.1: Overview of the rigid surface-volume registration scheme to match CT and US images of the talocrural joints. CT and US images are pre-processed before the registration. During the pre-processing step, surfaces of bones (i.e. tibia and talus) that can be visualized with US in maximal-plantar flexed ankle are extracted from the CT image and bone contours in freehand tracked 2D US images are enhanced using intensity invariant local-phase based approach and bone shadow information. 3D bone response data is reconstructed from 2D enhanced US images using the position of the US probe that had been recorded with an optical tracking system. Registration is initialized at a location defined using the position of the six fiducials in the US and CT spaces.

7.2.1 Test Data Acquisition

A test dataset consisting of CT scans and freehand-tracked 2D US images of 12 cadaveric ankles was collected at Academic Medical Center (AMC, Amsterdam,

The Netherlands). For the use of the cadaveric ankles, permission from the Medical Ethical Committee of AMC was not needed.
Each cadaveric ankle was tightly fixed on a foot-plate in maximal plantar flexion with straps to avoid any movement between acquisition of CT scan and US sweeps (Fig. 7.2a). Six fiducial markers were attached to the foot-plate (i.e. four markers on top of the foot-plate and two markers underneath), which are visible in the CT scan. These markers were pointed out with an external pointer and recorded during the US data acquisition (Fig. 7.2a).



a) b)

Figure 7.2: Experimental settings in the echo room. (a) Prior to CT imaging and US sweeping, each cadaveric ankle was placed on a foot plate in maximal plantar flexion and was tightened using straps. The US probe on which the position sensor mounted was slowly swept over the cadaveric ankles and the position of the probe was recorded using (b) the optical tracking camera.

7.2.1.1 CT Scans

CT scans were acquired using a Philips Brilliance 64 CT scanner (Philips Healthcare, Best, The Netherlands). The acquisition parameters were: effective dose 150: mAs/slice, rotation time: 0.75 s per 360°, pitch: 0.875, slice thickness: 0.6 mm, and ultra-high resolution mode. Tomographic reconstructions were made with a field of view of 154 mm, a slice increment of 0.3 mm, and a matrix of 512 $\times$ 512 pixels. The voxel sizes were 0.3 mm $\times$ 0.3 mm $\times$ 0.3 mm.

7.2.1.2 Freehand Tracked 2D US Images

For the US data acquisition, a iU22 xMatrix scanner (Philips Healthcare, Best, The Netherlands) (Fig. 7.2b) was used together with a 17- to 5-MHz broadband linear array probe (Philips Healthcare, Best, The Netherlands) (Fig. 7.2a). The position of the US probe was recorded by the means of a Polaris optical tracking system (Northern Digital Inc., Waterloo, Ontario, Canada) (Fig. 7.2b), by mounting a passive reflective marker to the US probe that was followed by the tracking

camera (Fig. 7.2a) (3D root mean square volumetric accuracy of the position sensor $\leq$ 0.25 mm).
The navigation system CustusX (SINTEF, Trondheim, Norway) [267], which allows for connections between the US scanner and the optical tracking system, was employed for data acquisition and reconstruction of the 3D US data from 2D US images. Prior to US data acquisition, the US probe was calibrated (calibration error = 0.23 mm) using a point target phantom [268].
The US imaging parameters (e.g. depth, gain) were set by an experienced radiologist specialized in musculoskeletal ultrasound imaging. For each cadaveric ankle, two different sweep types, namely "Sweep Type I" and "Sweep Type II", were performed. In the "Sweep Type I", the US acquisition started at the medial side of the talocrural joint with the probe placed perpendicular to the foot, on the deltoid ligament. The US probe was moved from the medial side to the lateral rim of the talus in such a way that the tibial rim and the dorsal surface of the talus were constantly visualized. The "Sweep Type II" involved positioning the probe slightly more proximal, supplemented with a return sweep back to the starting position on the medial side with the probe positioned more distally at the lateral rim of the talus. During each acquisition, the US probe was moved slowly (i.e. minimum of twenty seconds spent) over a cadaveric ankle while aiming to keep complete contact between the US probe and the cadaveric skin. After both sweeps, the position of the six fiducials was recorded using a calibrated pointer and the CustusX software.

7.2.2 Data Preprocessing

7.2.2.1 Surface Point Extraction from CT Data

Both talus and distal tibia were segmented from each CT scan using Mimics (version 14.01, Materialise, Leuven, Belgium) to yield triangulated surfaces. During the segmentation of the bones, a similar procedure as described by Tümer [203] was followed.
Regions on the triangulated bone surfaces that can be visualized with US (Fig. 7.1) were defined and isolated from the rest using custom-made code developed in Matlab (Matlab 2013b, The Mathworks Inc., Natick, MA) (see Appendix E). The vertices of the triangles located on the isolated surfaces were extracted.

7.2.2.2 Bone Surface Enhancement in Ultrasound

Bone contours in 2D US images were enhanced using an intensity-invariant local-phase technique [269, 270, 271, 272]. Following [271], a local phase-based feature detector "Phase Symmetry ($PS(x, y)$)", which is sensitive to the ridgelike edges observable at the bone boundaries in the US images was used. The PS was defined based on the monogenic signal representation [273] of 2D US images (Fig. 7.3a) (see Appendix E). As PS has high response at symmetry locations [274], other features such as soft tissue-muscle interface were also enhanced together with bone contours (Fig. 7.3b). To highlight the bony contours further, bone shadow information [274] was taken into account. Briefly, the weighted sum of

intensity values of pixels extending from a pixel of interest was used to quantify the shadow:

$$SH(x,y) = \frac{\sum_{j=x}^{H} G(j,y)U(j,y)}{\sum_{j=x}^{H} G(j,y)} \tag{7.1}$$

where $SH(x,y)$ is the shadow value for a pixel at row x and column y of the US image. H and $G(.)$ represent the total number of rows of the image $U(x,y)$ and the Gaussian weighting function, respectively. Combining PS and shadow information, the bone responses $BR(x,y)$ [274] (Fig. 7.3c) that represent 2D US image with enhanced bone contours were obtained:

$$BR(x,y) = PS(x,y) \cdot SH(x,y) \tag{7.2}$$

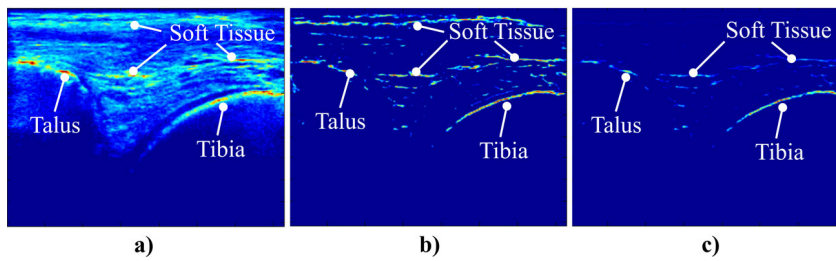


Figure 7.3: Bone contour enhancement in a 2D US image. (a) An original 2D US image (b) Phase Symmetry PS map calculated based on monogenic signal representation of the 2D US image. (c) Bone response map obtained based on the product of PS and the shadow values (SH).

7.2.3 3D Bone Response Data and US to CT Registration

3D bone response data was reconstructed from 2D enhanced US images (i.e. BR) using the position information of the US probe.
For each of 24 US sweeps (i.e. 12 cadaveric ankles $\times$ 2 sweep types), the registration was initialized at a location defined using the position of the six fiducials in the US and CT spaces.
The objective function, which was optimized during the registration, was the sum of the bone response values over the transformed surface extracted from the CT data:

$$f(\alpha,\beta,\gamma,\Delta_x,\Delta_y,\Delta_z) = f(\mathbf{x}) = \sum_{i=1}^{N} BR(\boldsymbol{R}_{\alpha\beta\gamma} \cdot \boldsymbol{p_i} + (\Delta_x,\Delta_y,\Delta_z)^T) \tag{7.3}$$

where N is the total number of sampling points $\boldsymbol{p}$ lying on the surface extracted from the CT data, $\boldsymbol{R}_{\alpha\beta\gamma}$ is the rotation matrix that depends on the angles α, β, and

γ (i.e. rotation parameters expressed in rad), and Δ_x, Δ_y and Δ_z are the translational parameters in mm.
Six transformation parameters (i.e. three for rotation and three for translation) were estimated using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [275], as the technique can be used for medical image registration tasks [264] and evolutionary algorithms have been reported to be robust [264, 276]. During the optimization, a box constraint was applied in a way that the transformation parameters stayed in the range of ± 0.2 rad ($\pm 11.5°$) and ± 5 mm surrounding the initialization location (i.e. α, β, γ, Δ_x, Δ_y and Δ_z were all equal to zero at the start of the registration).
A similar stopping criterion was applied as described by Winter et al. [264]. Registrations were stopped, when the relative progress of the best individuals over the last 10 generations dropped below the threshold, $f_{thr} = 10^{-6}$. With the aim of decreasing the possibility of registration algorithm to get stuck in local maxima, multistart optimization was performed. The optimization algorithm was restarted once the stopping criterion was met from the same initial position together with all the strategy parameters of the CMA-ES reset to their initial values [264].

7.2.4 Evaluation of the Registrations

A reference registration for each cadaveric ankle was identified in a similar way as described by Winter et al. [264]. For each of the 24 US sweeps, the registration algorithm was run 20 times, following an initialization at a location defined using the position of the six fiducials in the US and CT spaces. Among the 20 registrations per sweep, the one having the highest fitness value was picked as the reference registration. The validity of the reference registrations was visually checked.
The registration algorithm was run 100 times for each US sweep with initial misalignments (i.e. Test I). Taking each of the 24 US sweeps into account, a set of 100 initialization positions was determined by transforming the reference registration to a position away from its assumed optimum. The transformations were randomly created using a uniform distribution with rotation and translation parameters ranging between ± 0.2 rad ($\pm 11.5°$) and ± 5 mm, respectively. The initial position and the position of a volume after registration were compared to the position in the defined optimum (i.e. di_{RMS}, df_{RMS}) by calculating the root-mean-square (RMS) target registration error (TRE) [255]:

$$df_{RMS} = \sqrt{\frac{1}{N} \sum_{x} \| \boldsymbol{T_{REG}} \boldsymbol{x} - \boldsymbol{T_{REF}} \boldsymbol{x} \|^2} \tag{7.4}$$

where $\boldsymbol{x}$ is a point set extracted from 3D US data and N is the number of points in the set. $\boldsymbol{T_{REG}}$ and $\boldsymbol{T_{REF}}$ represent transformation matrices obtain as a result of a registration and as a result of the reference registration.
To assess the effects of the 3D US volume "density" on the performance of the registration algorithm (i.e. Test II), the amount of collected 2D US slices in each

US sweep was decreased 25% by randomly leaving out slices. The registration algorithm was run 100 times for each US sweep starting from the same initial position using 75% of the full data. The position of a volume after registration were compared to the position in the defined optimum (df_{RMS}) according to Eq. 7.4.
During evaluation of the Test I and II results, registrations having df_{RMS} values higher than 2 mm were considered as failures.

7.2.5 Clinical Study

To understand whether US imaging could be an alternative to a CT scan in pre-operative planning of arthroscopic access to anterior talar OCDs, the percentage of the talar cartilage surface that could be visualized with US imaging was determined following a similar approach as described by van Bergen [277]. Two observers analyzed the US data of 12 cadaveric ankles that were registered to the CT scans of the same cadaveric ankles using the proposed registration scheme. Each observer, using a custom-made code [277] developed in Matlab (Matlab 2016b, The Mathworks Inc., Natick, MA), defined the percentage of the visible talar dome contour (i.e. the arc length of the anterior talar dome, α divided by the arc length of the complete talar dome, $\alpha + \beta$) (Fig. 7.4) on 2D slices extracted from the lateral, central, and medial aspects of the talus. To evaluate the intraobserver reliability, the data was assessed one more time by one of the observers. Average percentages of the visible talar dome contour were calculated for each location and for each observer.

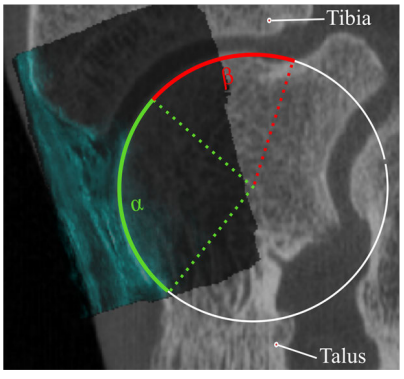


Figure 7.4: Sagittal view of one of the US volumes registered to the CT scan. The visible cartilage and cartilage covered by the tibia are represented, respectively, with α and β. The percentage of the visible cartilage was defined as the ratio of α to $\alpha + \beta$ (i.e. the total cartilage surface).

7.3 Results

An example from Test I, which depicts the initial misalignment and the final registration is given in Fig. 7.5. The visualization shows that it is possible to register US to CT images of the talocrural joint quite well using the proposed registration scheme.

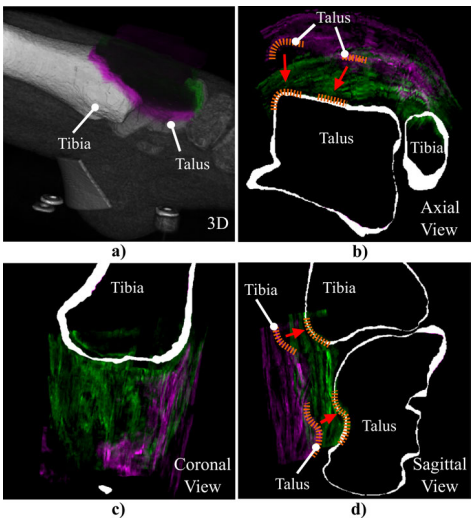


Figure 7.5: A visualization of one of the successful registrations achieved in Test I. (a) 3D view (b) Axial view (c) Coronal view and (d) Sagittal view. US volume in purple color represents the data at initialization position prior to registration. US volume in green depicts the result of the registration. Dashed lines in yellow highlight the contours of the tibia and talus observed in both axial and sagittal views. Arrows in red shows that US volume goes from its initial position (i.e. US volume in purple) to its final position (i.e. US volume in green).

Table 7.1 shows the number of successful registrations achieved in Test I and the mean of the di_{RMS} values together with df_{RMS} averaged over all successful registrations. The success rate across all the US sweeps in "Sweep Type I" and "Sweep Type II" were, respectively, 46% and 59%. The mean of the df_{RMS} values averaged over all successful registrations in "Sweep Type I" and "Sweep Type II" were, respectively, 0.9 mm $\pm$ 0.5 mm and 0.8 mm $\pm$ 0.3 mm.

Results of the Test I are displayed as plots of df_{RMS} vs. di_{RMS} for the US sweeps showing the best (i.e. 99% for both the "Sweep Type I" and "Sweep Type II") and worst success rate (i.e. 1% for the "Sweep Type I" and 0% for the "Sweep Type II") at a total of 100 registrations each (Figs. 7.6a-b). No clear relationship could be detected between the registration failure and the distance of the initial point

Table 7.1: Number of successful registrations and the mean RMS TRE (i.e. di_{RMS} and df_{RMS} values for each US sweep averaged over all successful registrations in Test I. A measure di_{RMS} describes how far a volume is away initially from its assumed optimum. Besides, df_{RMS} is based on the comparison of the position of a volume after registration to that in the defined optimum.

US Sweep Type	Number of Successful Registrations	Mean(and STD)(mm)	
		di_{RMS}	df_{RMS}
I	99	6.96 (3.27)	0.46 (0.17)
	80	7.50 (2.87)	1.24 (0.62)
	71	6.99 (2.49)	1.80 (0.18)
	65	8.72 (3.78)	0.62 (0.28)
	46	9.66 (4.66)	0.51 (0.17)
	39	7.43 (2.46)	0.57 (0.35)
	37	10.60 (3.79)	0.70 (0.54)
	35	6.97 (2.12)	0.76 (0.68)
	29	6.95 (2.73)	0.57 (0.36)
	24	10.04 (3.64)	0.65 (0.67)
	20	8.09 (2.66)	0.81 (0.62)
	1	7.04 (0.00)	1.69 (0.00)
II	99	7.28 (3.06)	0.66 (0.32)
	97	8.04 (3.45)	0.93 (0.30)
	91	8.55 (3.30)	0.26 (0.41)
	89	8.33 (3.42)	1.21 (0.57)
	78	8.55 (3.88)	0.95 (0.36)
	77	8.67 (3.66)	0.43 (0.44)
	76	7.35 (3.01)	1.33 (0.14)
	35	9.50 (2.95)	0.82 (0.43)
	28	7.97 (2.84)	0.62 (0.45)
	24	9.73 (3.96)	0.64 (0.43)
	16	9.41 (4.03)	0.68 (0.35)
	0	-	-

from the defined reference. Registered bone surfaces and 3D bone response volumes given for the US sweeps with the best and worst success rate in Test I (Figs. 7.6c-f) show that similar regions in each sweep type were scanned during the US data acquisition. Regarding the raw 2D US slices extracted from the "Sweep Type I" and "Sweep Type II" sweeps with the worst success rate (Figs. 7.6h and j), soft tissues seem to stand relatively closer to the bony contours as compared to those seen in the ones obtained from the US sweeps with the best success rates (Figs. 7.6g and i). The scanline profiles presented in Figures 7.6h and j show the intensity values of the pixels crossed by the two vertical lines (red and blue). Based on these profiles, it could be noted that soft tissues surrounding the bone may have higher intensity values as compared to those of the bony contours.
Table 7.2 shows the number of successful registrations achieved in Test II and the mean of the df_{RMS} values averaged over all successful registrations for each of 24 US sweeps. The average success rate was 47% and 58%, respectively, in "Sweep Type I" and "Sweep Type II". The mean of the df_{RMS} values averaged over all successful registrations in "Sweep Type I" and "Sweep Type II" were, respectively, 0.7 mm $\pm$ 0.4 mm and 1.0 mm $\pm$ 0.6 mm. For 1 out 12 US sweeps with

reduced "density" in "Sweep Type II" registrations consistently failed, while it was the case for 5 US sweeps in "Sweep Type I" (Table 7.2).

Table 7.2: Number of successful registrations and the mean RMS TRE (i.e. df_{RMS}) values for each US sweep averaged over all successful registrations in Test II. df_{RMS} is based on the comparison of the position of a volume after registration to that in the defined optimum.

US Sweep Type	Number of Successful Registrations	Mean(and STD)(mm) df_{RMS}
I	100	0.31 (0.19)
	100	1.38 (0.28)
	99	0.62 (0.33)
	98	0.63 (0.26)
	90	0.94 (0.47)
	77	0.75 (0.29)
	2	0.40 (0.33)
	0	-
	0	-
	0	-
	0	-
	0	-
II	100	0.35 (0.27)
	100	0.47 (0.29)
	100	0.56 (0.20)
	100	0.60 (0.25)
	99	0.61 (0.32)
	62	0.47 (0.36)
	52	1.63 (0.18)
	31	1.42 (0.41)
	26	1.05 (0.32)
	22	1.91 (0.08)
	3	1.88 (0.06)
	0	-

The results of the clinical case in terms of the mean percentage of visible talar surface at the lateral, central, and medial part of the talus, are presented in Table 7.3. On average, 51% $\pm$ 3% of the talar surface of each maximum plantar flexed cadaveric ankle could be imaged using US (Table 7.3).

7.4 Discussion

The goal of this study was to register freehand-tracked 2D US images of the talocrural joint to CT images. To this end, a rigid-surface volume registration scheme was presented and the performance of the registration algorithm was tested on a dataset collected from 12 cadaveric ankles. The registration algorithm was also used for a practical problem to evaluate whether US imaging can be a substitute of CT in assessment of accessibility to an OCD by anterior arthroscopy.

Visual inspection of the reference registrations and the example use of the registration algorithm confirm that multi-modality image registration is possible for

Table 7.3: The average percentage of the talar surface that could be visualized with US from two observers for the lateral, central and medial part of the talus. Observer 1 served as an internal control by performing the analysis twice: Observer 1 (1) and Observer 1 (2).

	Visible Talar Surface (% of the Total Talar Surface)		
	Lateral	**Central**	**Medial**
Observer 1(1)	55.3 (47.2-69.1)	47.9 (42.3-58.2)	51.7 (42.0-62.7)
Observer 1(2)	55.4 (46.3-68.1)	47.3 (38.2-59.0)	52.4 (42.2-61.6)
Observer 2	53.6 (42.8-66.1)	47.3 (36.8-60.0)	51.4 (42.5-59.0)

a talocrural joint and satisfactory registration results (Fig. 7.5) could be achieved with the proposed scheme.
Concerning the Test I, the registration algorithm had varying levels of success (Table 7.1). The large variance observed in the registration rate cannot directly be attributed to various initialization positions with a certain distance away from the defined reference. Differences in the performance of the registration algorithm to compensate similar mean di_{RMS} values (e.g. the success rate of 99% and 1% achieved in "Sweep Type I" of two different cadaveric ankles with the mean di_{RMS} values of 7 mm) (Table 7.1) suggest that the success rate of the algorithm is dependent on the acquired US data. Comparing raw 2D US images of the US sweeps showing the best (i.e. 99% for both the "Sweep Type I" and "Sweep Type II") and worst success rate (i.e. 1% for the "Sweep Type I" and 0% for the "Sweep Type II") in the Test I (Figs. 7.6g-j), it was observed that soft tissues are positioned relatively closer to the bone contours and show more bony-like features in the most failed US sweeps (Figs. 7.6h and j). With a close look at the scanline profiles of both vertical lines (red and blue) shown in Figures 7.6h and j, it becomes clearer that dominant ridge edge responses do not always correspond to the expected bony location, but may represent the high intensity soft tissue interfaces. In the current study, commonly-used local phase based descriptor optimized for ridge detection (i.e. *PS*) [251, 271] has been used to extract bony contours present in US images. As the *PS* detector does not provide or correct the responses at locations where the assumed feature model, i.e. ridgelike edge for bone, is violated and the registration algorithm is not capable of distinguishing between soft and bony tissues, it is possible that the registrations could not always end in correct positions due to the existence of bony-like information in some of the 3D volumes (Figs. 7.6d and f). In the future, more sophisticated descriptors, such as the local phase tensor [266], are suggested to be used to determine whether an improvement in bone detection and in the performance of the registration algorithm could be achieved in the US sweeps consisting of different feature types.

The results of the Test II indicate that the quality of the acquired US data plays an important role in the success of the registration process. Decreasing the number of slices in US sweeps seems to adversely affect the performance of the registration algorithm (Table 7.2). Reduction in the amount of information on bones in 3D bone response volumes could well explain the relatively lower performance

of the registration algorithm. Moreover, in both tests, the mean success rate was higher in "Sweep Type II" as compared to those in "Sweep Type I". This finding can be explained by relatively larger scanned area of the talar bone in "Sweep Type II". Increase in the amount of data on bony contours could help the registration algorithm to more easily find the optimum position.

Both "Sweep Type I" and "Sweep Type II" seem to be adequate to image approximately 50% of the anterior talar dome (Table 7.3). This is comparable to the area defined previously as the part of the talar dome accessible with anterior arthroscopy [277]. The results of the clinical test therefore suggest that US could be used to assess whether an OCD can be accessed by anterior arthroscopy.

Although, the freehand-tracked 2D US images of 12 cadaveric ankles were acquired by the experienced operator and using a protocol that had been included in previous studies [41, 246], the results of the tests and the observations made on US sweeps suggest that there is a room for improvements in the US data acquisition phase. One of the changes that could be made is the use of an alternative good quality US probe specially designed/fine-tuned for musculoskeletal applications. A US probe with a relatively smaller head size (e.g. L15-7io, Philips Healthcare, Best, The Netherlands) may allow the operator to more easily orient the probe to maintain it at a perpendicular position with respect to the bones. With an enhancement in the positioning of the US probe on the ankle joint, the chance of getting sharp bone boundary interfaces and decreasing the amount of unwanted information on soft tissues around the bones may be increased. A closer study of the 3D bone response volumes of the US sweeps that displayed the best and worst success rate in the Test I (Figs. 7.6c-f) showed that the variances observed in the success rates of the registration algorithm do not originate from any specific region on talus or distal tibia. Moreover, the success rate is expected to increase, if the US sweeps contain more information regarding the edges of the target bones. The scanned area is recommended to be increased in the future studies and includes the regions beyond the visible cartilage surfaces, e.g. by starting and finishing the data acquisition until the malleoli are visualized.

In addition to enhancing the US data acquisition protocol, the presented registration scheme could be further improved by providing real-time feedback to the operator, thereby guaranteeing successful registration in the clinical settings. In the current study, the computation time for the pre-processing of a single 2D US slice and the registration of a 3D US bone response volume to a surface extracted from a CT scan is $\approx$4 seconds and $\approx$3 minutes, respectively. To provide real-time guidance, the computational time has to be decreased, for example, through optimizing the code and/or implementing the methods on a multi-processor graphic processing unit (GPU).

To determine "gold standard" registrations and quantitatively evaluate the accuracy of the registrations, data could be collected from bones into which fiducials have been implemented. Fiducials were not attached to the bones to keep the conditions as close to the clinical reality as possible. To delineate a "gold standard", another solution could be the use of anatomical landmarks. Although we made an attempt to define points on the talus that could be consistently found

in all individuals and be used as anatomical landmarks, no such points could be ultimately found. Neither could we find any studies in the literature that could be referred to for the use of anatomical landmarks on the talus. It is therefore not clear to what extent the idea of using anatomical landmarks is practically feasible.

7.5 Conclusions

To the best of our knowledge, this is the first study proposing a scheme to register freehand-tracked 2D US images of a talocrural joint to the CT scan of the same joint. Results of the study showed that multi-modality image registration is possible for the talocrural joint and satisfactory registration results could be achieved with the proposed method. This could enable anatomical correlation of US images based on CT scan, thereby improving the possibilities for retrospective analysis and prospective follow-up of patients with diseases at the talocrural joint. With the performance of retrospective and prospective clinical studies, it is highly probable to make a progress in understanding the nature of cartilage healing in-vivo and in turn in the formulation of effective patient-specific treatment strategies. Further research, however, needs to be performed before clinical implementation of the proposed method is possible. The presented registration scheme could be enhanced in particular by using a more sophisticated approach to the pre-processing of US images, modifying the protocol established for the US data acquisition and implementing the methods in such a way that real-time feedback could be provided to guide the operator.

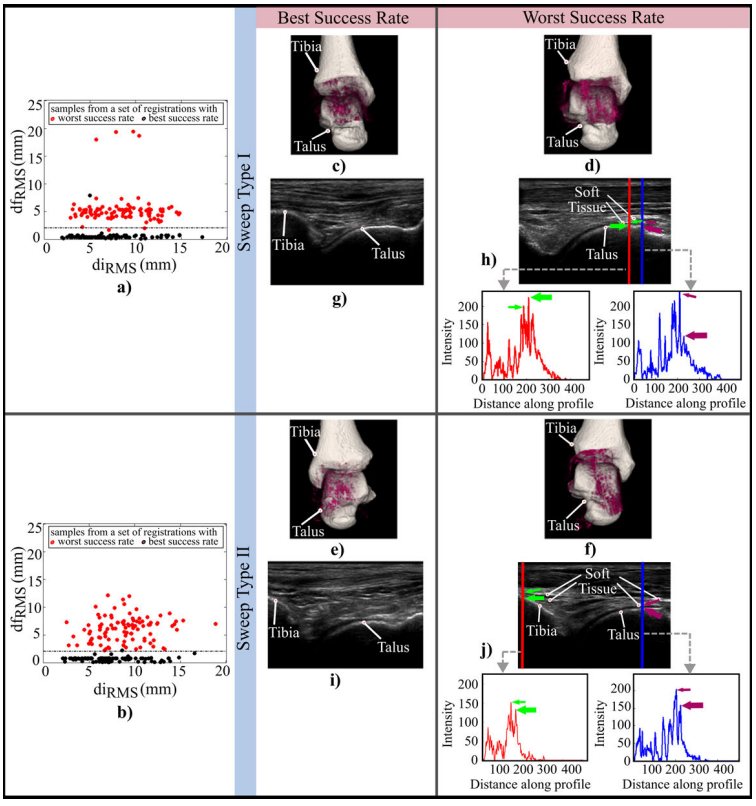


Figure 7.6: Plots of the df_{RMS} values vs. the di_{RMS} values calculated over 100 registrations are given for (a) the 'Sweep Type I' US sweeps (b) the 'Sweep Type II' US sweeps showing the best and worst success rate in Test I. Black ($n = 100$) and red ($n = 100$) dots in (a) are samples retrieved from a set of registrations performed for the 'Sweep Type I' US sweeps with the best and worst success rate, respectively. In a similar manner, those (n_{black} = 100 and n_{red} = 100) in (b) are related to a set of registrations run for the 'Sweep Type II' US sweeps with the best and worst success rate.Visualizations of 3D bone response volumes of the 'Sweep Type I' US sweeps (c) with the best and (d) the worst success rate, and of the 'Sweep Type II' US sweeps (e) with the best and (f) the worst success rate in the Test I are shown. Original 2D slices taken from the US sweeps (c-f) are presented in a corresponding manner (g-j). The intensity profiles of the two vertical lines (i.e. intensity vs. distance along profile graphs shown in red and blue correspond to the lines drawn in red and blue, respectively) displayed in (h) and (j) indicate that the ridge-like edge feature corresponding to the expected bone boundary can be weaker than those related to the soft tissue interface. Thin and thick arrows shown in green and purple point to the ridge edge features linked, respectively, to soft tissue interface and bone boundary.

8

A Novel Ultrasound Technique for Detection of Osteochondral Defects in the Ankle Joint: A Parametric and Feasibility Study

Published as:
N. Sarkalkan, A.J. Loeve, K.W. van Dongen, G.J. Tuijthof, A.A. Zadpoor. A novel ultrasound technique for detection of osteochondral defects in the ankle joint: A parametric and feasibility study. Sensors. 15 (2014) 148-165.

8.1 Introduction

The ankle joint is one of the common locations for (osteo)chondral defects (OCDs) representing disruption of articular cartilage together with or without disintegration of the subchondral bone [278, 37, 33]. These defects can occur due to acute trauma, repetitive micro-trauma, and torsional joint loading during sport [33, 279, 38] and domestic activities (e.g. a fall from stairs) [180]. If OCDs are left untreated, there is a potential risk of developing early osteoarthritis (OA) [278, 34]. As OA is a disease decreasing individuals' quality of life and imposes a huge socioeconomic burden on society [280], early detection and treatment of OCD is highly important.

The techniques currently used for detection of OCDs are computed tomography (CT), magnetic resonance imaging (MRI), and arthroscopy [33, 39, 40, 41]. Although these techniques are sensitive in detection of defects [41], they are not favorable to be used in longitudinal follow-up of individuals due to ionizing radiation (e.g. in CT), long acquisition time (e.g. in MRI), and invasiveness (e.g. arthroscopy) [40] as well as the involved costs. Ultrasound (US) imaging, which is known to be non-invasive, fast, and cost-effective [281], is not extensively used in diagnosis of OCDs. That is primarily because US imaging is not capable of successfully imaging all cartilage defects [282] because of the limited ability of US waves to penetrate through bone [247].

Keeping cost-effectiveness and non-invasiveness of US in mind, an acoustic wave method which does not require penetration of sound waves through bone might be a good alternative to the currently-used diagnostic techniques. We propose a new diagnostic technique, which is based on the propagation of US waves through the entire joint space of the ankle and on definition of changes in acoustic wave response properties ensued as a result of defect. Specification of acoustic parameters that can describe the entire joint space morphology and be robust to individual variations will be necessary for successful identification of OCDs. However, determination of optimal parameters is a major challenge, as little is currently known about the US propagation in the joint space of the ankle and effects of many parameters on the acoustic wave characteristics.

The aim of this work was to study US propagation in the joint space of the ankle and to determine the feasibility of the new concept for detection of defects. To reduce the complexity of the problem, simplified 2D finite-difference time-domain (FDTD) models of ankle joint were developed and effects of dominant variables relevant to the ankle joint (i.e. joint space width), US transducer (i.e. translation and rotation of US transducer acting as transmitter) and defect (i.e. width, depth and location) on acoustic wave response were evaluated. An experiment was performed to validate a FDTD model, which represents the healthy state of the joint. To the authors' knowledge, this is the first study to evaluate propagation of US in the joint space of the ankle and to determine the feasibility of such a concept for detection of OCDs.

8.2 Materials and Methods

Throughout this study, two series of simulations and one series of experiments were performed. The first set of simulations incorporates a reference model (Fig. 8.1a), which mimics the healthy state of the joint and was validated with experimental measurements. The second series includes models generated to evaluate the effects of main variables (Fig. 8.2) related to the ankle joint (i.e. joint space width), US transducers (i.e. translation and rotation of US transducer acting as transmitter), and defect (i.e. width, depth and location) on the output signal of the reference model.

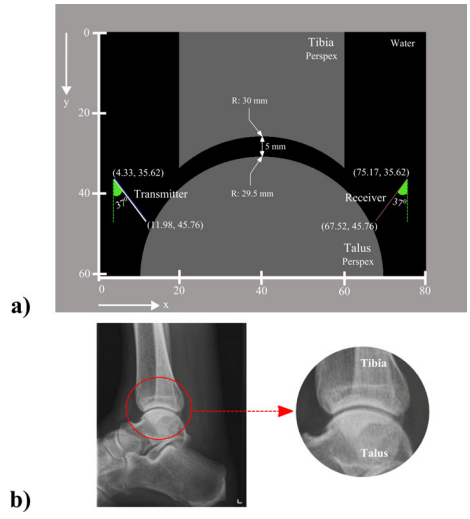


Figure 8.1: (a) Representation of the reference simulation model. Joint space width of the ankle is 5 mm. The radii of the tibial and talar arcs respectively are 30 mm and 29.5 mm. US probes are positioned in a nearly identical manner to the experiment. (b) Radiograph of a healthy ankle joint - Lateral view [283].

8.2.1 Simulations

All the simulations were performed using the US simulation software Wave2000 (Cyberlogic Inc., New York, NY, USA). To find an approximate solution to the 2D acoustic wave equation, the software uses the algorithm presented by Schechter et al. [284]. In brief, this algorithm assumes that a heterogeneous medium is composed of homogenous linear isotropic regions and imposes continuity of stresses and displacements across boundaries of four homogenous regions. Within each homogeneous grid element, an acoustic differential equation is solved and the displacement vector is computed at the intersection of four grid elements at each

time step of the simulation:

$$p\frac{\partial^2}{\partial t^2} = |\mu + \eta\frac{\partial}{\partial t}|\nabla^2 \boldsymbol{w} + |\lambda + \mu + \phi\frac{\partial}{\partial t} + \frac{\eta}{3}\frac{\partial}{\partial t}|\nabla(\nabla \cdot \boldsymbol{w}) \tag{8.1}$$

where $\boldsymbol{w}$ is a 2D vector whose components are the x and y components of displacement of the medium at location (x, y), ρ = material density $[\frac{kg}{m^3}]$, λ = first Lamé constant $[\frac{N}{m^2}]$, μ = second Lamé constant $[\frac{N}{m^2}]$, η = shear viscosity $[\frac{(N \cdot s)}{m^2}]$, ϕ = bulk viscosity $[\frac{(N \cdot s)}{m^2}]$, t = time [s], ∂ = the partial difference operator, ∇^2 = the Laplace operator, and $\nabla\cdot$ = the divergence operator.

8.2.1.1 Reference Model

The first step in generation of the model was definition of the geometry. As the reference model would be used in assessment of US propagation in the joint space of a healthy ankle joint, a simplified geometry (Fig. 8.1a) which is similar in size and shape (Fig. 8.1b) to the ankle joint was considered. The study of Stagni et al. was consulted for determination of a realistic size for the tibia and the talus [285]. Additionally, a joint space width of 5 mm (Fig. 8.1a) was used, considering the fact that the ankle joint space can be increased up to 5 mm upon distraction loading [286, 287]. A geometry file representing the simplified ankle joint was prepared using Matlab R2013b (The MathWorks, Inc.) and imported into Wave2000.
The second main step in preparation of the model was the definition of the material properties. Following White et al. [288], Perspex was chosen to prepare the experimental samples (i.e. tibia and talus) due to its ease of milling. Plexiglas material properties (Table 8.1), which were derived from the material library embedded in the Wave2000, were assigned to the parts representing the tibia and the talus (Fig. 8.1a). In addition, the simplified ankle joint was considered to be submerged in water, representing the synovial fluid. Material properties of water (Table 8.1) were defined using the material library of the Wave2000 and incorporated in the simulation model (Fig. 8.1a). All four sides of the simulation model boundaries were taken as expanding to infinity.

Table 8.1: Material properties used in simulations.

	Water	**Plexiglas/Lucite**
$\rho[\frac{kg}{m^3}]$	1000	1150
$\lambda[MPa]$	2241	5601
$\mu[MPa]$	0	1392
$\phi[Pa \cdot s]$	$9.998 \cdot 10^{-8}$	0.01
$\eta[Pa \cdot s]$	0.001	0.5
$C_L[\frac{m}{s}]$	1497	2700
$C_S[\frac{m}{s}]$	3.54491	1100.2
Longitudinal attenuation coefficient [dB/cm]	$6.81479 \cdot 10^{-4}$	$5.12405 \cdot 10^{-2}$
Shear attenuation coefficient [dB/cm]	153953	0.559761

Two non-focused 1 MHz US transducers (diameter: 12.7 mm) were positioned (Fig. 8.1a) and oriented similarly to the transducers in the experimental setup

that was used to validate the simulation result. A uniform apodization factor was used for both the transmitter and the receiver. Default settings of the Wave2000 for gain (i.e. 0 dB), blanking (i.e. 0 μs), and duration (i.e. 0 μs) were used for the receiver. As the simulations were intended to represent the experimental setup as closely as possible, the source signal (i.e. time function of US transducer acting as transmitter) was defined using the procedures presented in the section describing the experimental setup and was used for all numerical simulations.
To determine the grid size, the prepared model was run a total of 11 times while gradually decreasing the grid size of the model from 100 μm to 16.7 μm. Subsequently, simulations results were compared to each other and the similarity between two subsequent simulations was computed using normalized root mean square errors (NRMSE):

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_{obs,i}-x_{ref,i})^2}{n}}}{x_{obs,max} - x_{obs,min}} \times 100\% \tag{8.2}$$

where $x_{obs,i}$ is the observed value at time i, $x_{ref,i}$ = reference value at time i, $x_{obs,max}$ = maximum of the observed data, $x_{obs,min}$ = minimum of the observed data, and n = number of data points.
Simulation results were assumed to have converged if the NRMSE was less than 1%. Based on the simulation results, a grid size of 20 μm was found to result in convergence and was used throughout the study. The time step of 6.5 ns was automatically calculated by the software. The total simulation time was 100 μs.

8.2.1.2 Parametric Models

Before going through the preparation of all the models that would be necessary to evaluate the effects of the parameters (e.g. variables related to the ankle joint) on the acoustic wave response, experiments were performed. The experimental results were used to investigate whether the reference simulation can represent reality sufficiently accurate. Reassured by a good agreement between the reference simulation results and experimental findings, all the parametric models were generated using the same procedures. Parameters expected to be most dominantly influencing the wave response were studied. These parameters were classified into three main categories: factors related to (i) the ankle joint geometry, (ii) US transducers, and (iii) defect (Fig. 8.2). These categories consisted of one (i.e. joint space width), two (i.e. translation and rotation of US transducer acting as transmitter) and two (i.e. size and location of the defect) categories, respectively. In addition, categories relevant to size and location of the defect and to translation of US transducer were all split into two subcategories (Fig. 8.2).
Only one parameter of the model was varied at each simulation to gain insight into the influence of each parameter on the output signals. With respect to the factors related to the ankle joint (Fig. 8.2), the joint space width values were varied from 2 mm to 5 mm (i.e. reference model) [286, 289] in intervals of 1 mm. Moreover, OCDs (Fig. 8.3a) were represented with a rectangular shape

(Fig. 8.3b) in the parametric models simulating parameters related to defect. The values used to describe depth and width of defects were respectively varied from 2 mm to 4 mm and 2 mm to 6 mm, thereby remaining within the range of values reported for OCDs [290]. In addition to the defects with negative depth values, positive defects (Fig. 8.2), which may occur in the stage IV lesions (i.e. detached and displaced osteochondral fragment) [291] were considered.
Throughout the simulations, the size of temporal and spatial domain, the grid size, the boundary conditions, and the material properties were kept identical to the reference model.

8.2.2 Experiments

The experimental setup used is depicted in Fig. 8.4. It consists of two unfocused 1 MHz US transducers with one being used as transmitter and one as receiver (Olympus V303, transducer diameter: 12.7 mm, Panametrics Inc.), an arbitrary waveform generator (Agilent Technologies, 33250A), an oscilloscope (Agilent Technologies, DSO7054A) and the simplified ankle model made of Perspex (Fig. 8.4b). The ankle model was immersed in a water tank at room temperature. To hold the model tibia and model talus in their proper positions, as well as the US transducers, a custom-made Perspex frame was used (Fig. 8.4b).
The transmitter was excited by a short normally distributed pulse with a peak-to-peak voltage of 9 V, a center frequency of 1 MHz, and a width of 4 μs. Before going through the measurements to validate the reference simulation results, the transmitter and the receiver were positioned facing each other in a straight line at a known distance. Having only water (i.e. no joint model placed) in between the transmitter and the receiver, the signal was recorded and used as a time function for the source (Fig. 8.5a) in the simulation models. Measurements performed for the validation were repeated 20 times, each time repositioning both US transducers.

8.2.3 Validation of the Reference Model

To evaluate the similarity between the results of the reference simulation and those of the experiment, the mean of the 20 measured output signals was calculated in the time domain and compared to the output signal of the simulation. The normalized maximum cross-correlation (NMCC) was used as a similarity measure:

$$NMCC = \frac{max|(f * g)[i]|}{\sqrt{\sum_{i=1}^{n} f[i]^2} \cdot \sqrt{\sum_{i=1}^{n} g[i]^2}} \tag{8.3}$$

where f and g are functions with a length of n, and $f * g$ is the cross correlation of functions f and g. The reference simulation was assumed to represent reality sufficiently, if the calculated NMCC was within a margin of 5% (i.e. NMCC $\geq$ 0.95).

8.2.4 Assessment of Parameters Effects

To assess effects of individual parameters on the output signals, results of each parametric simulation were compared to those of the reference simulation. For each case, NRMSE (Eq. 8.2) values were calculated and used for quantitative evaluation of changes in the reference output signal ensued as a result of variations in parameters. To discover whether variations in parameters mainly alter the shape of the reference output signal, the NMCC was determined as well, as it is invariant to time shift.

8.3 Results

The output signal of the reference simulation was compared to the mean of 20 measured output signals in the time (Fig. 8.5a) and in the frequency domain (Fig. 8.5b). The shapes of both output signals were highly similar, apart from a comparatively high flattening of the output signal of the reference simulation after 60 μs (Fig. 8.5a). The similarity measure (i.e. NMCC) between the output signals was found to be 0.97.
The amplitude of the output signal progressively increased when increasing the joint space (Fig. 8.6a). NRMSE and NMCC ranged from 6.29% (for joint space width: 4 mm) to 65.25% (for joint space width: 2 mm), and from 0.976 (for joint space width: 4 mm) to 0.391 (for joint space width: 2 mm), respectively (Table 8.2).

Table 8.2: Normalized root mean square error (NRMSE) and normalized maximum cross correlation (NMCC) for parametric models simulating the parameters related to the ankle joint (i.e. joint space width) and to the US transducers (i.e. translation and rotation of US transducer acting as transmitter). In each parametric simulation, only one parameter of the model was changed, and the other parameters were kept identical to those of the reference model.

Variable	NRMSE (%)	NMCC
Joint space width (2 mm)	65.25	0.391
Joint space width (3 mm)	29.75	0.643
Joint space width (4 mm)	6.29	0.976
Transducer x-translation (2 mm)	9.65	0.998
Transducer x-translation (4 mm)	9.57	0.996
Transducer y-translation (2 mm)	3.96	0.992
Transducer y-translation (4 mm)	11.72	0.929
Transducer rotation(37°)	8064.2	0.629
Transducer rotation(25°)	78.02	0.632
Transducer rotation(12°)	19.59	0.924

No noticeable change in the amplitude of the output signal was present when translating the transducer in x-direction (Fig. 8.7a), but, there was a clear phase

shift.
The amplitude of the output signal does considerably decrease when translating the transducer in y-direction (Fig. 8.7b). The NRMSE measures support the finding that translation of the transducer in x-direction does not generate noticeable changes on the output signal in terms of amplitude, whereas translation in y-direction does (Table 8.2). The amplitude of the output signal is almost zero when the transducer is rotated 37° clockwise, which implies perpendicular alignment to the talar surface (Fig. 8.7c). Analogous to the joint space width, variations in the transducer's rotational position cause high NRMSE (e.g. 78.02% for the transducer rotated 25° clockwise) and low NMCC (e.g. 0.632 for the transducer rotated 25° clockwise), supporting that apparent changes in the amplitude of the output signal occur.
The defects having depths of -4 mm and -2 mm (i.e. representing an actual hole in the bone) cause only small changes in the output signal (Figs. 8.8 and 8.9) for any location on the tibia or the talus. The so-called positive defects with depths of 2 mm and 4 mm cause large differences in the output signal compared to the reference output as is clear from their higher NRMSE values (Fig.8.8). The defect in the tibia with a depth of 2 mm and width of 6 mm shows the maximum deviation (i.e. 36.16% NRMSE) from the reference signal (Fig. 8.8). The maximum NRMSE for a talus defect (i.e. 25.71%) is found for a depth of 4 mm and width of 6 mm (Fig. 8.8).

8.4 Discussion

FDTD models of a joint space of the ankle were generated to determine the effects of variations of the joint space width, US transducer orientation and defect size and location on the propagation of a 1 MHz US pulse. The reference simulation results were validated against experimental measurements (Fig. 8.5). From 60 μs, shape of output signal of the reference simulation differed slightly from that of the mean of 20 measured output signals (Fig. 8.5a). Such a variation in shape may be results of reflections due to objects present in the actual setup, which were not included in simulations. However, considering the NMCC value of 0.97, which remains within a margin of 5%, it is admitted that the reference simulation agrees well with the experimental data.
A decline in the amplitude of the output signal was seen as a result of decreasing the joint space width (Fig. 8.6). These findings indicate that to support the propagation of 1MHz US pulse within the joint space of the ankle, it is recommended to distract the ankle joint. This is feasible for the majority of the population for up to 5 mm of joint space, as distraction is a standard procedure when performing ankle arthroscopy for treatment of OCD [286, 287]. However, an increase in the joint space width does not mean that defects would relatively more easily be detected. Having a narrower joint space width might be more beneficial to detect defects, because the part of the signal ending up in any defect would be relatively larger in comparison to the part of the signal traveling through the joint space.

The NRMSE (i.e. 3.96% - 8064.2%) and NMCC (i.e. 0.998 - 0.629) values of the models simulating the translation and rotation of the transducer showed no noticeable effect in x-direction, but a strong decrease of the output signal with translation in y-direction and rotation (Fig. 8.7). Taking the high NMCC (i.e. 0.998 and 0.996) and NRMSE (i.e. 9.65% and 9.57%) for the models with transducer' translations in x-direction into account, one could conclude that the changes in the output signal mainly increased due to the shift of US pulse in time. By increasing the y-translation or rotation, a major part of US pulse will be sent through the bone into which the US beam has a limited ability to penetrate. This is not desired in the application of the proposed concept. From these simulations, it is concluded that the positioning of the transducers should be done with great care, and a device should be developed that can position the transducers with high reliability around the ankle joint of patients at multiple occasions, giving an optimal field of view for the transducers into the joint space.
Analysing the results of variations in the defect size, it seems that positive defects, which may occur in the stage IV lesions (i.e. detached and displaced osteochondral fragment) [291], made noticeable alterations to the output signal, whereas, negative defects (Fig. 8.2) created relatively small changes (Figs. 8.8 and 8.9).
Increasing the width for negative talus defects did not seem to vary the output signal. However, the NRMSE values for increasing widths of the negative tibial defects did show what seems to be a linear trend (Fig. 8.8) that might be helpful to determine tibial defect sizes with the proposed concept of wave propagation. No changes in the output signal due to variations in width of talus defects can be explained from Fig. 8.10. The US pulse is guided along the tibia rim whereas it is less or not guided along the talus rim. Therefore, changes in the talus rim will not alter the wave propagation, but variations in the width of the tibia defects can be seen as a change in the amplitude of the signal (i.e. the NRMSE values). This observation indicates that using the current configuration (e.g. frequency of the US transducers) detection of tibial defects would be easier than those of talus defects.
No distinguishable effect of the defect location (i.e. defects located at 60°, 90° and 120°, Fig. 8.2) was found on the output signal. Considering the fact that an accurate description of size and location of defect could help orthopedic surgeons in operative planning and in monitoring response to therapy [38], it would be more valuable if the location of defect besides its size could be determined using the proposed concept. In this respect, there is a need for further research on parameters (e.g. other frequencies), which might allow one to estimate the location of the defect.

One of the limitations of this study is the considerable simplification of the ankle joint. In the present model, cartilage and other soft tissues were not considered. As water may not be able to sufficiently replicate the absorption of US by soft tissues, the current results may somewhat differ from simulations including soft tissues in terms of US pulse amplitude. However, the general trends due to variations of the joint space width, and the translation and rotation of the US transducer are not expected to change. Furthermore, the alterations in the reference

output signal due to defects are expected to be less noisy if the cartilage is also incorporated into the model, because the losses due to reflection and refraction at the water-cartilage interface would be smaller compared to the current water-Perspex interface, as water and cartilage have very similar densities.
Another limitation of this study is the omission of several parameters (e.g. the wave frequency, radius of the tibia and talus). To have a complete picture of US propagation within the joint space of the ankle, it is necessary to take also the omitted parameters into account and to evaluate their effects.
In this study, numerical modelling was limited to 2D to avoid relatively excessive computational cost of 3D models. Considering the other computational studies conducted to assess US interactions with cartilage [292], bone [293, 294, 295] and to evaluate US propagation within the joint space of a human knee [296], it is believed that 2D models are good enough for obtaining preliminary insight in the new concept.
In the present models, bony parts were represented by Perspex shapes so that simulations could be easily validated with experiments. A concern might be raised about use of a homogeneous joint mock-up to describe highly heterogeneous bony tissues. In the proposed concept, the focus is on the US wave sent through the ankle joint space and not through the bone itself, that is why both cortical bone and cancellous bone can be assumed homogeneous as suggested by White et al. [296]. Taking studies of White et al. [296, 297, 288] into account, the use of Perspex models therefore seems to be a reasonable starting point prior to the development of more complex models. In the future, simulations considering more realistic composition and structure of bony tissues will be developed and used to further investigate US wave travelling through the ankle joint and its use in detection of defects.
The current study has shed some lights on effects of variables (e.g. joint space width) on acoustic wave response and gives an idea on the feasibility of the new concept. Further research is required to determine acoustic parameters that can describe the entire joint space morphology and be robust enough to cope with individual variations. A next step will be the evaluation of the proposed concept using human ankle cadaver models as suggested by Tuijthof et al. [246]. For each human ankle cadaver model, size and location of defect will be estimated using US simulations and results will be compared with the real case (i.e. the human ankle cadaver model).
The performance of the proposed diagnostic technique in identification of defects is expected to be better than that of the traditional US imaging modality. That is because the proposed technique has the potential to increase the percentage of the scanned articular surface to 100%, allowing US wave to propagate through the ankle joint. In comparison, only an estimated 50% of the talar surface could be visualized using traditional US imaging when the foot is in maximum plantar flexion [246]. Since only 80% of defects [38] are located in the area visualized by traditional US imaging, it cannot be used for diagnostic assessment. Moreover, the proposed diagnostic technique is more suitable for long-term follow of individuals having OCDs as compared to CT and MRI due to its non-invasiveness, compact set up and cost-effectiveness.

In summary, the current study provided a good start to investigate the feasibility of the concept to use acoustic waves traveling through the entire joint space for characterisation of changes in the joint shape due to the presence of OCDs. To the authors' best knowledge, this is an entirely new concept and the results indicate that it could be feasible provided that the effects of transducer orientation and joint space width are reduced. This requirement obliges development of a device that could fixate transducer orientation and joint space width in a reliable and repeatable manner.

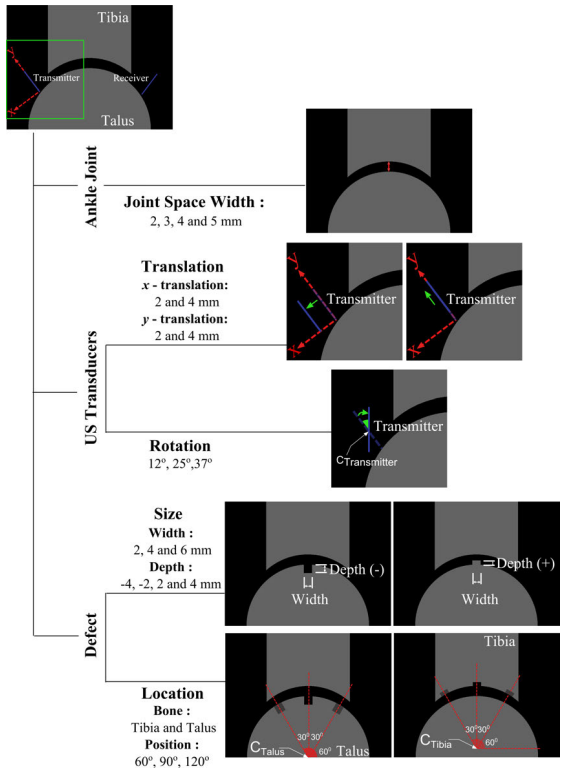


Figure 8.2: Representation of parameters related to the ankle joint (i.e. joint space width), US transducers (i.e. translation and rotation of US transducer acting as transmitter) and defect (i.e. width, depth and location). Joint space width values are varied from 2 mm to 5 mm (i.e. reference simulation model) in intervals of 1 mm. Translations of US transducer were ranged from 2 mm to 4 mm. The US transducer rotated 37° around its center implies that it is perpendicularly aligned to the talar surface, whereas the other two angles were arbitrarily chosen: 12°, and 25°.

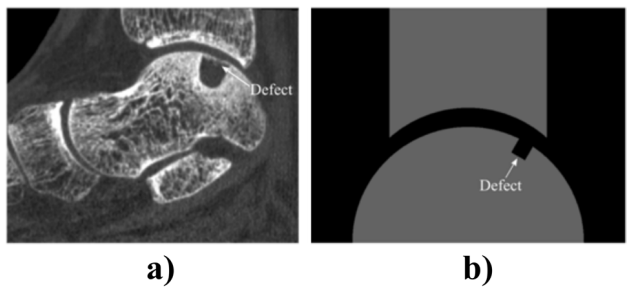


Figure 8.3: (a) Sagittal reconstructed image of CT scans of an ankle with OCD (b) Representation of a defect in simulations.

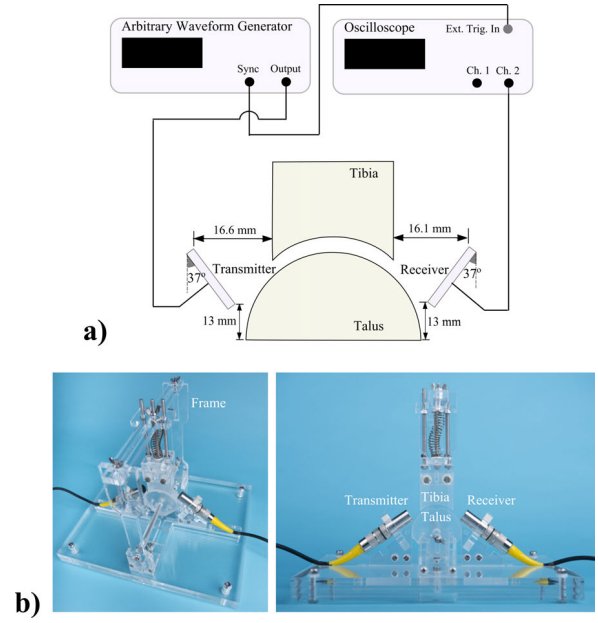


Figure 8.4: (a) Schematic presentation of the experiment setup; (b) Perspex frame designed to hold the tibia and talus parts. Holders for transducers are included to fixate them at the desired positions.

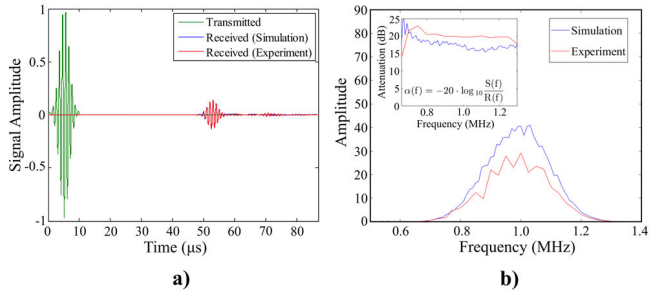


Figure 8.5: (a) Time function of the US transducer acting as transmitter, mean of 20 measurements and output of the reference simulation. (b) Amplitude spectra of the simulated signal and the mean signal recorded in experiment. In small window, attenuations [dB] in the simulated model and in the experiment are shown. Attenuations were calculated based on the provided formula, where $\alpha(f)$ is the attenuation [dB], log10 is the logarithm to the base 10, $R(f)$ and $S(f)$ respectively are the amplitude spectrum of the reference waveform (i.e. input signal) and the amplitude spectrum of a receiver waveform (i.e received signal in experiment or in the reference simulation).

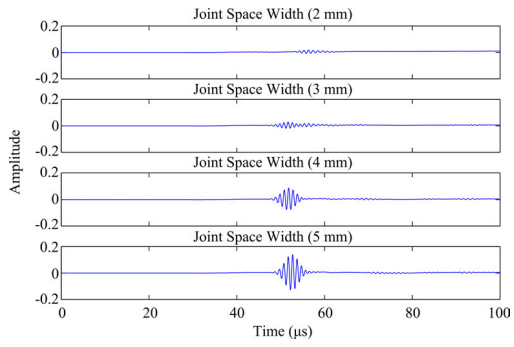


Figure 8.6: Outputs of the parametric models simulating the joint space width.

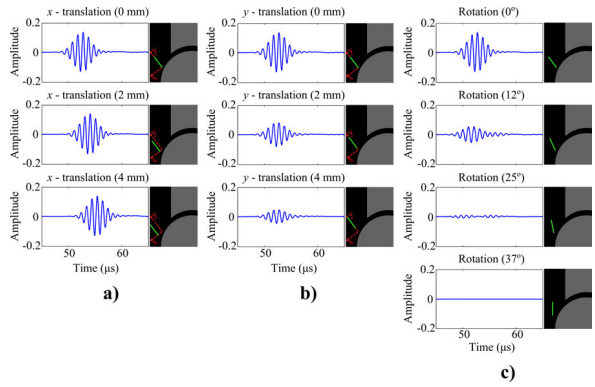


Figure 8.7: (a) Outputs of the parametric models simulating translation of US transducer (i.e. transmitter) in x-direction. Orientations of the transmitter and the receiver were kept identical to those in the reference model (i.e. 37° from vertical) (b) Outputs of the parametric models simulating translation of US transducer (i.e. transmitter) in y-direction. Orientations of the transmitter and the receiver were kept identical to those in the reference model (i.e. 37° from vertical) (c) Outputs of the parametric models simulating rotation of US transducer (i.e. transmitter). Position and orientation of the receiver were not changed and kept identical to those in the reference model.

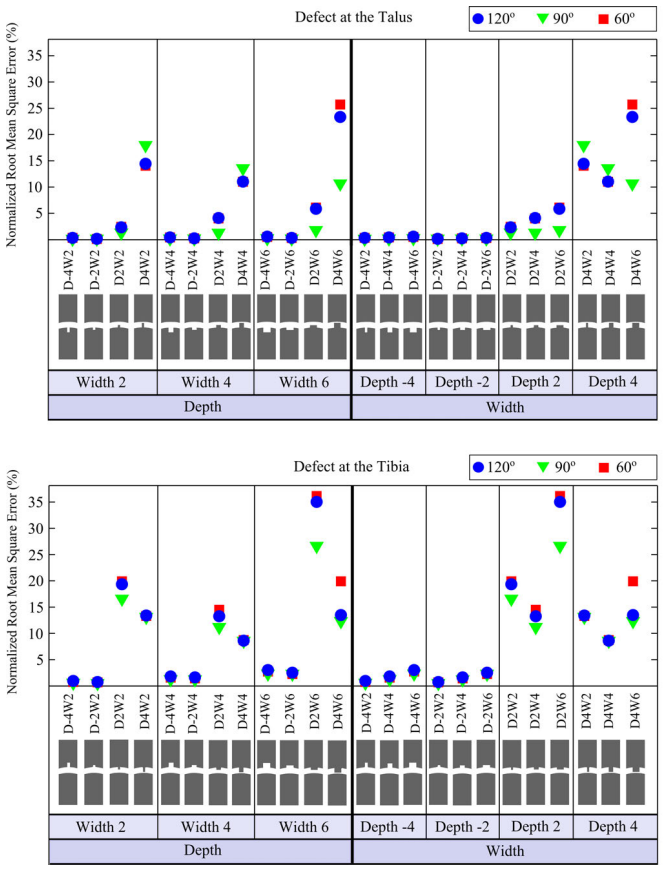


Figure 8.8: Normalized root mean square errors (NRMSE) for models simulating defects located on the talus (i.e. first graph) and on the tibia (i.e. second graph). The first three columns of the first and second graphs are provided to clearly represent the effects of varying defects depths when the defect width was fixed at 2 mm, 4 mm or 6 mm. The last four columns are presented to easily capture the changes related to the varying defects widths when the defect depth was fixed at -4 mm, -2 mm, 2 mm or 4 mm. Negative values for defect depth are used to represent actual holes in the bone, while positive values are used to describe positive defects, which may occur in the stage IV lesions (i.e. detached and displaced osteochondral fragment). In the first and second graphs, depth and width of defect are represented by a 'D' and a 'W', respectively. The minus sign '-' is used to describe negative defect, an actual hole in the bone (e.g. D-2W6 represents negative defect, i.e. a hole, with a depth of 2 mm and a width of 6 mm).

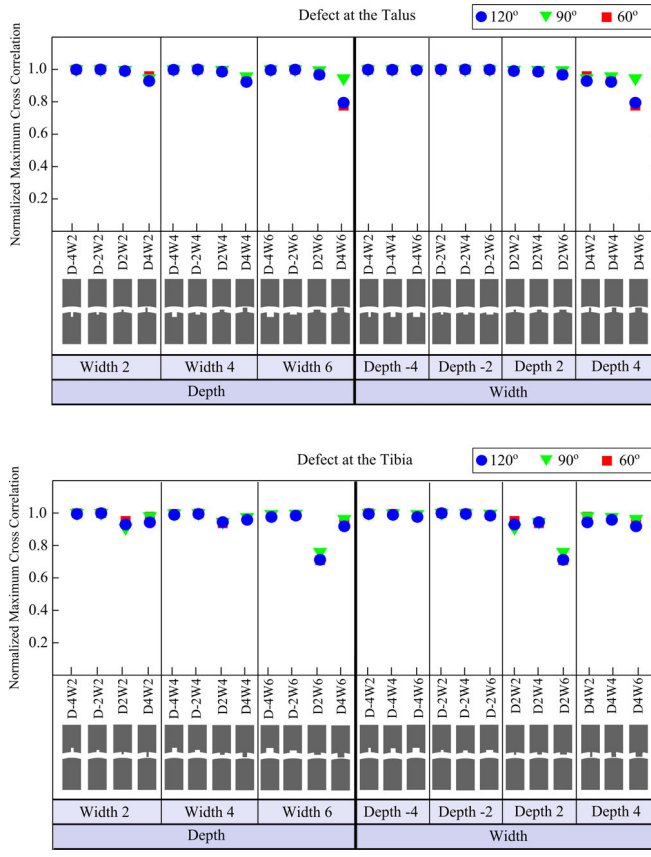


Figure 8.9: Normalized maximum cross correlation (NMCC) for models simulating defects located on the talus (i.e. first graph) and on the tibia (i.e. second graph). The first three columns of the first and second graphs are provided to clearly represent the effects of varying defects depths when the defect width was fixed at 2 mm, 4 mm or 6 mm. The last four columns are presented to easily capture the changes related to the varying defects widths when the defect depth was fixed at -4 mm, -2 mm, 2 mm or 4 mm. Negative values for defect depth are used to represent actual holes in the bone, while positive values are used to describe positive defects, which may occur in the stage IV lesions (i.e. detached and displaced osteochondral fragment). In the first and second graphs, depth and width of defect are represented by a 'D' and a 'W', respectively. The minus sign '-' is used to describe negative defect, an actual hole in the bone (e.g. D-2W6 represents negative defect, i.e. a hole, with a depth of 2 mm and a width of 6 mm).

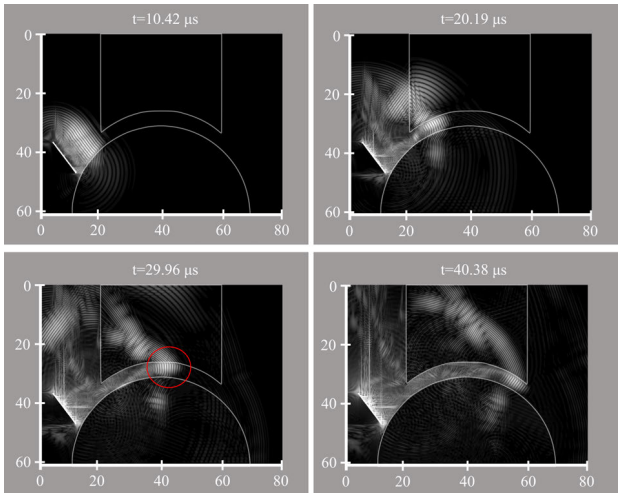


Figure 8.10: Simulated ultrasound pulse at t=10.42 μs, 20.19 μs, 29.96 μs and 40.38 μs. The white bar is the transducer acting as transmitter. Red circle is used to emphasize US pulse guided along the tibia rim.

9

General Discussion and Summary

9.1 General Discussion

The main focus of this thesis was bone shape and its relationship with skeletal diseases. A wide range of research questions were addressed within this context, which are either explicitly (Chapters 3-6) or implicitly (Chapters 7-8) related to bone shape. The applications are in detection of the risk factors of skeletal diseases (Chapters 3-5), the longitudinal monitoring of patients (Chapter 7), and early diagnosis (Chapter 8). In general, the results of the thesis show that bone shape has a role to play in various aspects of skeletal diseases including determining the risk factors, diagnosis of the diseases, and monitoring of the onset and progression of the diseases. However, advanced approaches such as SSMs, SAMs, and FEM models are needed to unravel such relationships.
The review on SSMs and SAMs in Chapter 2 demonstrates that these models can be used to describe variations in bone shapes (i.e. SSM) or densities (i.e. SAM) or in a combination of both (i.e. SSAM) within a studied population. Numerous successful applications [23, 13, 1, 3] have established the role of statistical models in investigating bones and their relations with skeletal diseases. As statistical models are well-suited for the use in conjunction with engineering tools such as FEM technique, it is expected that their use in bone research will continue to grow in the coming decades.
One example of a combined remedy to predict osteoporosis-related hip fracture is presented in Chapter 3, which has been developed based on SSAM and FEM techniques. The generation of patient-specific 2D FE model based on DEXA images [298, 299, 300, 128] for estimation of fracture risk is not a new idea. Nevertheless, this study contributes to the field by presenting an automated platform that requires user-interaction only at the beginning (i.e. extraction of bone geometry and density from a DEXA image) with a mouse click to generate such a model. The results of the study suggest that the models generated using the presented scheme could be a good alternative to those built manually, as strong correlation was found between the fracture loads estimated using manually and automatically generated models, $R^2 = 0.82$. Although the proposed scheme is a viable option to be used in clinics due to its associated time and costs that are comparable to those of clinical methods employed at present, its ability to predict fracture risk is currently unknown. The specificity and sensitivity of the presented scheme in fracture risk estimation could not be determined because of the nature of dataset available to us at the time. To estimate the utility of the proposed technique in clinical settings, it would be necessary to perform another study with a large dataset including a mixture of DEXA scans of healthy individuals and osteoporotic patients for whom long-term follow-up data is available. Another issue worth mentioning is that the 2D FE models are limited in describing the distribution of material properties and force components in the third dimension. To overcome these limitations, the presented scheme could be modified in a way that 3D FE models are developed. As 3D imaging modalities such as CT and QCT are rarely used at present in clinics [10] and are not expected to be a common option either in the coming decade for osteoporosis management, it is believed that generation of 3D FE models based on these imaging modalities

would not be the best approach. Alternatively, 3D-2D registration algorithms [301, 302, 64] could be employed to model 3D bony structures from 2D DEXA image(s). Nevertheless, due to the high computational time (i.e. up to 40 hours [302]), the description of 3D-DEXA bony structures was not clinically feasible until recently. Lately, a method has been proposed for estimating 3D bony structures from an anteroposterior DEXA projection in approximately 2 minutes [10].
In Chapters 4 and 5, SSMs were used to study 3D shape variations in the bones forming the talocrural (i.e. talus and distal tibia) and subtalar (i.e. talus and calcaneus) joints and their interactions with the OCD and CAI. These studies contribute to the clinical field mainly with their findings suggesting that shape differences exist between patients with an injury and healthy controls. Based on the results of Chapter 4, an increase in the vertical neck angle (VNA) of the talus and the sharpness of the medial malleolus may make individuals more prone to sustain an OCD. Moreover, changes in the curvature of the talar lateral process combined with the inclination angle of the talar neck relative to the body (Chapter 5), and the medial and lateral tuberosity of the calcaneus combined with the contour of the anterior articular surface may contribute to the development of CAI in individuals who have sustained a lateral ankle sprain. It should be noted that the dataset used in both studies were retrospectively acquired. Therefore, it is unclear whether the findings represent post-traumatic changes or whether these indicated shape variations have contributed to the onset of the disease. To clarify this issue, the SSMs should be rebuilt based on a prospective dataset, and bone shape variations and their prognostic values for these specific diseases should be analyzed. Based on the experience gained in two studies, it is noteworthy to indicate that the utility of 3D statistical models in clinical settings is open to discussion. SSMs enable one to identify shape variations that cannot be captured with a simple set of 3D measurements. However, clinicians are more used to interpreting radiographs, as they are widely available and analyzed to gain insight in patient's conditions. It may therefore be worth to translate the shape variations explained by 3D statistical model to (digitally constructed) 2D radiographic measurements to strengthen the adoptability of these models. We are currently working on the translation of 3D statistical-model based shape variations in the talus and distal tibia to 2D parameters that can potentially contribute to the development of CAI and be measured using standard plain radiographs. If the translated parameters can be reliably measured and have prognostic value, they could be used as a part of a risk prediction model for CAI.
When building 3D SSMs of the talus and distal tibia in Chapter 4, the contralateral unaffected talocrural joints of OCD group were included assuming that bilateral symmetry exists in the bones forming the ankle joint. This assumption was made referring to a few studies [160, 161]. During the literature search, the scarcity of documentation on (a)symmetry of lower extremity bones caught our attention, although ipsi- and contralateral bone shape of an individual is often assumed to be symmetric in clinical assessments and research studies. The added value of Chapter 6 to bone research field have become evident at this point. Contrarily to previous studies that have investigated symmetry by reducing bone shape to a few anatomical landmarks [206, 217], the bones forming the talocrural

and subtalar joints (i.e. the tibia, fibula, talus and calcaneus) were described in a spatially dense manner in our study. It was found that the bone shapes were in general bilaterally symmetric. Nevertheless, symmetry assumption may not be valid in cases for which, (1) the curvature of the fibula shaft; (2) the diameter of the fibula; or (3) the tibial tuberosity together with the diameter of the tibia are particularly important. For instance, the diameter of the fibula and tibia could be pertinent to identifying the correct implant size and proper placement in an arthroplasty surgery. At the same time, the curvature of the fibula shaft could be relevant for planning corrective osteotomy. Accordingly, the outcomes of the surgical procedures (i.e. arthroplasty, corrective osteotomy) performed with shape symmetry assumption may be influenced by side-shape variations in the fibula and tibia. In this study, the isolated locations of asymmetry may have been overlooked, as principal component analysis (PCA) was used to describe the shape variations. This issue needs to be clarified in future studies.

The problems handled in the following two chapters were more linked to the longitudinal monitoring of a patient with a disease (i.e. OCD, Chapter 7) and early diagnosis of a disease (i.e. OCD, Chapter 8). In Chapter 7, a rigid surface-volume registration scheme was presented to match CT and freehand-tracked US images of the talocrural joint. Although the registration of CT and US images of the same individual is not a new concept and there exist several registration algorithms that have been proposed, none of the previous studies have worked on the talocrural joint. Previously presented registration schemes have been applied to various anatomical sites including long bones [255, 259], vertebrae [257, 261, 260] and pelvis [255, 254, 262]. The main contribution of this chapter is that the registration scheme is presented for the talocrural joint and its competence was tested on the joint itself. Another added value of the study is that the potential of US imaging as an alternative to CT in pre-operative planning of arthroscopic access to anterior talar OCDs was shown based on a set of US and CT data merged using the presented scheme. The results of the study implied that freehand-tracked 2D US images of a talocrural joint could successfully be registered to the CT scans of the same joint with a mean registration error of 0.9 ± 0.5 mm. Yet, the robustness of the registration algorithm against different initialization positions and the sampling of US data needs to be improved before its use in clinical settings. Based on the visual inspection of the pre-processed US data with the worst registration success rate, it was found that soft tissues in the nearby region of bones could not be totally filtered out due to their bony-like appearances. It is possible that the presence of such unwanted features worsens the outcomes. Therefore, it is suggested that more-sophisticated descriptors such as local phase tensor [266] be tested in the future to assess their potential in discarding the responses of undesired soft tissues, while enhancing bone contours in US images. The results of the robustness tests and our observations made during the US data acquisition implied that there is also room for improvements in the US data acquisition protocol. It is advised that the effects of an alternative high-quality probe (i.e. US probe with a relatively smaller head size and specially designed for musculoskeletal applications) and different ankle sweeping techniques be studied in the future. Another important issue worth stating

is that the presented scheme requires the use of a combination of different software (e.g. Mimics) and custom-made code developed in Matlab. To increase the user-friendliness of the registration scheme and its utility in clinical setting, all the steps should run in a single platform and in a way that real-time feedback could be provided to guide the operator. We are currently working to improve the presented methods and to generate an embeddable code to be used in a single environment.

To diagnose an OCD, we presented a new technique based on US wave propagation through the ankle joint and studied its feasibility in Chapter 8. Experimentally-validated 2D finite-difference time-domain models of the ankle joint gave an idea of the feasibility of the new concept and helped us gain insights into the effects of some variables related to the ankle joint (i.e. joint space width), defect (i.e. width, depth, and location) and transducer (i.e. translation and rotation of a transducer acting as transmitter) on the acoustic wave response. The results suggested that the technique could be feasible provided that the effects of the width of the joint space and the transducer orientation on wave response could be reduced. Yet, the study has several limitations. Not all of the parameters that influence wave propagation within the joint space of the ankle were included in the study. One such parameter is wave frequency. In the study, one single wave frequency of 1 MHz was employed. Following the feasibility study, a set of computational simulations with frequencies varying from 40 kHz to 2 MHz were performed to assess the effects of frequency in identification of a defect. Preliminary results indicated that a frequency of around 100 kHz may make the technique more sensitive for identifying a defect. As the results are not concrete yet and our investigations on this issue will continue in the future, they were not covered within the scope of this work.

Like in the given example of homemade cookies in Chapter 1, bones of the same type are not absolutely identical to each other and their shapes vary within and between individuals. The studies presented in this thesis clearly show that the shape of bones interacts with skeletal diseases (i.e. osteoporosis, OCD and CAI) in multiple aspects. Although it cannot be stated whether bone shape is the most dominant factor in the determination of the relationship between bone and diseases, the outcomes of the studies and our experiences gained through the work suggest that it is an integral part of the process. It is believed that a large economic burden of diseases on society can be decreased by improving patient-specific treatment and prevention strategies. Accordingly, inclusion of risk factors with bone shape origin in prediction models and use of diagnostic and monitoring techniques, such as the ones presented in Chapters 3, 7 and 8, are expected to positively contribute to the development of patient-specific treatment and prevention strategies. Therefore, we highly encourage researchers to continue performing the studies on the shape of bones and its association with skeletal diseases.

9.2 Summary

Despite their high structural and functional likeness, bones of the same type are not exactly the same. This is mainly because of intra- and inter-individual variations in the shapes and constituents of bones. Deviations in the variables of a bone from their optima could adversely affect bone functions and may have severe implications for individuals. There is increasing evidence suggesting that the shape of bones could be both a cause and consequence of skeletal diseases. Considering the serious economic and social burden caused by skeletal diseases and the potential of bone shapes in interacting with them, the focus of this thesis is mainly on the analysis of relationship between bone shape and various skeletal diseases.

Within the scope of this thesis, clinically important problems that are directly (Chapters 3-6) or indirectly (Chapters 7-8) associated with the morphology of bones, and either have been overlooked (Chapters 4-6) or entail technical improvements (Chapters 3, 7 and 8) have been handled. The studies, except two of them (Chapters 2 and 6), can also be classified into one of three following categories: (1) risk factors (Chapters 3-5), (2) long-term monitoring of patients (Chapter 7) and (3) early diagnosis of a disease (Chapter 8).

While seeking solutions to problems stated in Chapters 3-6, either a combination of statistical shape models (SSM) and statistical appearance models (SAM) or SSM alone have been used. To gain insight into these statistical models prior to their use in the studies presented, an overview on them and their applications in bone research is presented in Chapter 2. Accordingly, the first part of Chapter 2 is dedicated to the main concepts and algorithms relevant to SSM and SAM. In the second part of Chapter 2, the applications of SSM and SAM in the study, diagnosis, and treatment of skeletal diseases have been discussed. An appendix of this chapter reviews the most important technical aspects of SSM and SAM.

A scheme based on a combination of SSAM and finite element modeling (FEM) techniques has been presented in Chapter 3 for fast and automated prediction of the risk of osteoporosis-related hip fractures. While SSAM has been used for extracting the geometry of a bone and its density distribution from an unseen DXA image of a patient in an accurate and automated manner, FEM enables considering the factors (i.e. the shape and density-based mechanical properties of a bone and loading conditions) that play important roles in causing the onset of a fracture. The fracture loads estimated using automatically generated models did not significantly differ from those obtained manually.

In Chapters 4 and 5, 3D SSMs have been built for all of the bones forming the talocrural (i.e. the talus and distal tibia, TCJ, Chapter 4) and subtalar (i.e. the talus and calcaneus, STJ, Chapter 5) joints using the CT scans of a mixed population consisting of individuals with no known ankle joint pathology and patients with a disease, namely an osteochondral defect (OCD) (Chapter 4) or chronic ankle instability (CAI) (Chapter 5). Using the SSMs and statistical tools, 3D shape variations in the above-mentioned bones have been analyzed and compared between the disease and control groups. The results of two studies showed that there are some specific shape variations that may help us distinguish the bones

of individuals who suffer from an OCD (Chapter 4) or CAI (Chapter 5) from those of healthy controls. Statistically significantly different shape modes were observed between the groups were mainly in: (1) the vertical neck angle of the talus and (2) the medial malleolus in the case of the OCD patients (Chapter 4). In Chapter 5, changes in: (1) the curvature of the talar lateral process and the inclination angle of the talar neck relative to the body and (2) medial and lateral tuberosity, and in the contour of anterior articular surface of the calcaneus were found to be statistically significantly different between the groups.

While performing the study described in Chapter 4, we assumed that the talus and distal tibia are bilaterally symmetric. This assumption was made based on a few available studies. In Chapter 6, we assessed 3D shape variations, and bilateral (a)symmetry in the bones forming the TCJ and STJ (i.e. the fibula, tibia, talus and calcaneus). The effects of gender on shape variations were also evaluated. The methodology followed in the study mainly consisted of 3D SSM technique that had been previously applied in Chapters 4 and 5, and distance-based permutational statistics. The results of the study suggested that the shapes of bones were in general symmetric. However, the symmetry assumption did not hold for changes in the tibial tuberosity together with the diameter of the tibia, and the curvature of the fibula shaft and the diameter of the fibula.

Ultrasound (US) imaging is a preferred alternative to computed tomography (CT), magnetic resonance imaging (MRI), and arthroscopy techniques in post-operative long-term follow-up of patients with a disease (e.g. OCD), as it is a cost-effective and non-invasive method and does not expose patients to ionizing radiation. Nevertheless, US images cannot not be easily interpreted and compared with each other over time without having access to anatomical data provided by other modalities such as CT. In Chapter 7, we presented a rigid surface-volume registration scheme based on evolutionary approaches to use freehand-tracked 2D US images of the talocrural joint in conjunction with CT scans of the same joint. The performance of the registration scheme and its robustness with respect to different initialization positions and to the sampling of US data was evaluated on a dataset collected from 12 cadaveric ankles. A practical use of the presented scheme was demonstrated in a case study, in which the potential of US imaging as an alternative to CT for pre-operative planning of arthroscopic access to the anterior talar OCDs was studied. 2D freehand-tracked US images and CT scans of the talocrural joint could be registered with a mean registration error of 0.9 ± 0.5 mm using the presented scheme. The results of the practical case study demonstrated that US imaging might be a promising candidate for evaluating the accessibility of OCD by anterior arthroscopy.

Due to its inherent properties (e.g. non-invasiveness), US imaging is the preferred method in diagnosing skeletal diseases such as OCD. Nevertheless, not all OCDs could be discriminated using US imaging. That is because the articulating surfaces of the talus and distal tibia cannot be completely swept due to the limited ability of US wave to penetrate through bones. In Chapter 8, we propose a new diagnostic technique based on US wave propagation through the talocrural joint. The feasibility of the proposed technique was studied using experimentally-validated 2D finite-difference time-domain models of the talocru-

ral joint. The effects of variables relevant to the talocrural joint (i.e. joint space width), US transducers (i.e. translation and rotation of US transducer acting as transmitter) and OCDs (i.e. width, depth and location of a defect) on wave propagation were studied. The results suggested that an OCD could be detected using the proposed technique, if the effects of some parameters including joint space width, and the orientation of US transducer particularly could be reduced.

9.3 Samenvatting

Ondanks de hoge structurele en functionele gelijkenis zijn botten van hetzelfde type nooit precies hetzelfde. Dit komt voornamelijk door de variatie van zowel de vorm als de bestanddelen van botten binnen en tussen personen. Deviaties van deze variabelen in een bot ten opzichte van hun optima kan de functionaliteit van botten nadelig beïnvloeden en kan ernstige gevolgen hebben voor individuen. Er is een toenemende mate van bewijs die impliceert dat de vorm van botten zowel de oorzaak als het gevolg kan zijn van skeletale afwijkingen. Gezien de hoge socio-economische belasting die gepaard gaat met skeletale afwijkingen en de mogelijke interactie met variatie in bot vorm, is deze thesis hoofdzakelijk gericht op de analyse van de relatie tussen variatie in botvorm en verschillende skeletale afwijkingen.
In deze thesis zullen klinisch belangrijke problemen die direct (Hoofdstukken 3-6) of indirect (Hoofdstukken 7-8) geassocieerd zijn met de morfologie van botten, en die in het verleden zijn overzien (Hoofdstukken 4-6) of betrekking hebben op technische verbeteringen (Hoofdstukken 3, 7 en 8) worden behandeld. De studies, afgezien van twee (Hoofdstukken 2 en 6), kunnen ook in de volgende drie categorieën worden geclassificeerd: (1) risico factoren (Hoofdstukken 3-5), (2) lange termijn monitoring van patiënten (Hoofdstuk 7) en (3) vroege diagnose van een ziekte (Hoofdstuk 8).
Tijdens het zoeken naar oplossingen voor problemen zoals beschreven in Hoofstukken 3-6, zijn 'statistical shape models (SSM)' of een combinatie van SSM en 'statistical appearance models (SAM)' gebruikt. Om inzicht te krijgen in deze statistische modellen, voordat de studies worden toegelicht, is er een overzicht van de modellen en de toepasbaarheid in bottenonderzoek gepresenteerd in Hoofdstuk 2. Het eerste deel van Hoofdstuk 2 is gewijd aan de hoofdconcepten en algoritmen die relevant zijn ten aanzien van SSM en SAM. In het tweede deel van Hoofdstuk 2 worden de applicaties van SSM en SAM in de studie, diagnose en behandeling van skeletale afwijkingen besproken. In de appendix van dit hoofdstuk worden de meest belangrijk technische aspecten van SSM en SAM besproken.
In het kader van snelle en geautomatiseerde voorspelling van risico van osteoporose gerelateerde heupfracturen is er een schema gepresenteerd in Hoofdstuk 3 gebaseerd op een combinatie van SSAM en 'Finite Element Modeling' (FEM) technieken. Hoewel SSAM is gebruikt voor het extraheren van de botgeometrie en densiteitsverdeling van een DEXA-beeld van een patiënt op een nauwkeurige en geautomatiseerde manier, maakt FEM het mogelijk om de factoren (d.w.z. de

vorm en op dichtheid gebaseerde mechanische eigenschappen van een bot en belastingcondities) die een belangrijke rol spelen bij ontstaan van een fractuur. Met behulp van de automatisch gegeneerde modellen werd er geen significant verschil gevonden ten aanzien van de geschatte fractuurbelasting in vergelijking met de handmatig verkregen metingen.
In Hoofdstuk 4 en 5 zijn 3D SSM's ontworpen voor alle botten betrokken bij het talocrurale (d.w.z. de talus en distale tibia, TCJ, Hoofdstuk 4) en subtalaire (d.w.z. de talus en calcaneus, STJ, Hoofdstuk 5) gewricht met behulp van CT-scans van een gemengde populatie bestaande uit individuen zonder bekende enkelgewrichtspathologie en patiënten met een skeletale afwijking, namelijk osteochondraal defecten (OCDs) (Hoofdstuk 4) of chronische enkelinstabiliteit (CEI) (Hoofdstuk 5). Met behulp van de SSM's en statistische hulpmiddelen zijn 3D-vormvariaties in de bovengenoemde botten geanalyseerd en vergeleken tussen de aangedane patiënten en controlegroepen. De resultaten van twee onderzoeken tonen aan dat er enkele specifieke vormvariaties zijn die kunnen helpen de botten van individuen die lijden aan een OCD (Hoofdstuk 4) of CEI (Hoofdstuk 5) te onderscheiden van die van gezonde enkels. Statistisch significante vorm verschillen werden waargenomen tussen de groepen in: (1) de hoek van de nek van de talus en (2) de mediale malleolus in het geval van de OCD-patiënten (hoofdstuk 4). In hoofdstuk 5 werden verschillen gezien in: (1) de kromming van processus lateralis tali en de inclinatiehoek van de talaire nek ten opzichte van corpus tali en (2) de mediale en laterale tuberositas, en in de contouren van het anterieur gewrichtsoppervlak van de calcaneus waren statistisch significant verschillend tussen beide groepen.
Tijdens het uitvoeren van het in Hoofdstuk 4 beschreven onderzoek hebben we aangenomen dat de talus en distale tibia bilateraal symmetrisch zijn. Deze aanname is gedaan op basis van een aantal beschikbare onderzoeken. In Hoofdstuk 6 hebben we 3D vormvariaties en bilaterale (a) symmetrie van de botten die het talocrurale en subtalaire gewricht vormen (d.w.z. de fibula, tibia, talus en calcaneus) beoordeelt. De effecten van geslacht op vormvariaties zijn ook geëvalueerd. De methodologie bestond voornamelijk uit de 3D SSM-techniek die eerder in de Hoofdstukken 4 en 5 was toegepast, en op afstand gebaseerde permutatiestatistieken. De resultaten van de studie suggereerden dat de vormen van botten over het algemeen symmetrisch waren. De veronderstelling van symmetrie gold echter niet voor veranderingen in de tuberositas tibiae, voor de diameter van de tibia, de kromming van de fibula-schacht of voor de diameter van de fibula.
Echografie is een alternatief die de voorkeur heeft ten opzichte van computertomografie (CT), magnetische resonantiebeeldvorming (MRI) en artroscopietechnieken bij postoperatieve lange termijn follow-up van patiënten met een skeletale afwijking (zoals OCD), omdat het een kosteneffectieve en niet-invasieve methode is en patiënten niet blootstelt aan ioniserende straling. Desondanks kunnen echobeelden moeilijk over de tijd worden geïnterpreteerd en met elkaar worden vergeleken zonder anatomische gegevens die worden verstrekt door andere modaliteiten zoals CT. In Hoofdstuk 7 presenteerden we een rigide oppervlaktevolume-registratieschema gebaseerd op evolutionaire benaderingen

om handgemaakte 2D-echobeeldenvan het talocrurale gewricht in combinatie met CT-scans van hetzelfde gewricht. De prestaties van het registratieschema en de robuustheid ervan met betrekking tot verschillende initialisatieposities en de samenstelling van echo-gegevens werden geëvalueerd op basis van een dataset van 12 kadaver enkels. Het praktisch gebruik van het gepresenteerde schema werd aangetoond in een case study, waarin de potentie van echografische beeldvorming als een alternatief voor CT voor preoperatieve planning van arthroscopische benadering van anterieure talaire OCDs werd bestudeerd. De 2D met de hand verkregen echobeelden en CT-scans van het talocrurale gewricht kunnen worden geregistreerd met een gemiddelde registratiefout van 0,9 $\pm$ 0,5 mm met behulp van het gepresenteerde schema. De resultaten van de praktische case study toonden aan dat echografische beeldvorming een veelbelovend zou kunnen zijn om de benaderbaarheid van OCDs d.m.v. anterieure artroscopie te evalueren.
Vanwege zijn inherente eigenschappen (niet invasief), heeft echografische beeldvorming de voorkeur bij het diagnosticeren van skeletale afwijkingen zoals OCDs. Niettemin kunnen niet alle OCDs worden onderscheiden met behulp van echografische beeldvorming. Dat komt omdat de articulerende oppervlakken van de talus en distale tibia niet volledig kunnen worden afgebeeld vanwege het beperkte vermogen van ultrasonografie (US)-golf om door botten te dringen. In Hoofdstuk 8 stellen we een nieuwe diagnostische techniek voor die gebaseerd is op de golfexpansie van de US door het talocrurale gewricht. De haalbaarheid van de voorgestelde techniek werd bestudeerd met behulp van experimenteel gevalideerde '2D finite-difference time-domain' modellen van het talocrurale gewricht. De effecten van variabelen die relevant zijn voor het talocrurale gewricht (d.w.z. breedte van de gewrichtsspleet), echo-transducers (d.w.z. translatie en rotatie van de echo-transducer die als zender werken) en OCDs (d.w.z. breedte, diepte en locatie van een defect) werden op golfvoortplanting bestudeerd. De resultaten suggereerden dat een OCD kon worden gedetecteerd met behulp van de voorgestelde techniek, als de effecten van sommige parameters, waaronder de breedte van de gewrichtsspleet, en de oriëntatie van de echo-transducer in het bijzonder, konden worden gereduceerd.

Appendices

Appendix A

The material presented in this supplementary document is generally a summary of the methods and techniques presented in the excellent texts written by several prominent scholars in the area of statistical shape and appearance modeling particularly the report written by the co-inventor of statistical shape models Tim Cootes [42], and a few excellent review papers [43, 55, 56].

A.1 Statistical Models of Shape and Appearance

The main aim of statistical models is to capture possible variations within a class of objects [105]. For this purpose, statistical analysis of a training dataset that includes shape examples of an identical object is necessary. The procedure for generation of statistical shape models mainly includes shape representation and statistical analysis of the shape data. To build proper statistical models, correspondence should be established between all shapes of the training dataset. Moreover, all shapes need to be in a common coordinate system [303]. As for statistical appearance models, generation of models includes shape representation, texture representation around and/or within the part of image covered by the shape, and statistical analysis of the available data. In this section, we review the most important techniques that are used for generation of statistical models of shape and appearance with a focus on the techniques that are particularly useful for orthopedic applications. Other techniques can be found in more specialized review papers, e.g. [43].

A.1.1 Statistical Shape Models

The first step in generation of statistical shape model is shape representation. The available techniques for shape representation can be classified according to the shape descriptors that are used for representing the shape. They include landmarks, dense surface meshes such as techniques that use Fourier surfaces and spherical harmonics, skeleton-based representations (medial models), deformation fields, and distance maps [43, 304]. As indicated by Golland et al. [304], there are two essential characteristics of shape descriptors that can notably affect the quality of the statistical models. One factor is related to the model's sensitivity to the noise of the images and the other one pertains to its ability to align all the shapes in a common coordinate system.
The most generic method used in orthopedics is based on landmarks [49, 73, 74, 305, 53, 26]. In the literature, statistical shape models represented by landmarks are called Point Distribution Models (PDMs) [43, 62]. Landmarks are occasionally named nodes, vertices, or model points [138]. As it is indicated by Cootes et al., the best landmarks are the points that can be consistently located on each image of the training dataset [42]. The methods for defining anatomically consistent points and establishing the correspondence are discussed in the following subsection, but the simplest manual method is to define the landmarks based on certain anatomical bony landmarks that could be relatively easily identified by experienced anatomists on different images. Once a number of anatomically

well-defined landmarks are identified, the rest of landmarks could be defined with reference to those anatomically well-defined landmarks.
If the shape of a bone is defined with n points in d dimensions, a vector $\boldsymbol{x}$ with nd elements can be composed as:

$$\boldsymbol{x} = (x_1, \cdots, x_n, y_1, \cdots, y_n, z_1, \cdots, z_n)^T \tag{A.1}$$

where (x_n, y_n, z_n) represents the coordinates of the nth point. If the training dataset contains N images, there will be N number of vectors $\boldsymbol{x}$.

A.1.1.1 Correspondence

To build proper statistical models, landmarks should be defined as the points that can be consistently found on each image. In other words, all the points of a shape should correspond with those of the other remaining images of the training dataset. There are numerous methods to establish correspondence between all the shapes of a training dataset [43].
Manual definition of landmarks over the surface of the examined object is a general concept for 2D models [49, 73, 74, 27, 70]. However, this method is subjective and decreases the reproducibility of the results. Furthermore, manual annotation is time consuming and error-prone. Due to these drawbacks, manual landmarking is becoming impractical and unpopular for 3D and even for 2D models with large number of training images. To overcome the difficulties encountered in manual "landmarking", algorithms that automatically define correspondences are being developed.
The problem of establishing correspondence between the landmarks defined on different images effectively reduces to the problem of finding proper relationships between the elements of the training dataset. The input data may be the structures that are derived from the images including sets of points, contours, surfaces and volumes and may also be the intensity of the raw pixels as well as the intensity gradient [56]. Correspondence may directly be obtained by optimizing a similarity measure between the above-mentioned elements of images. Another alternative is to first align the shapes by using the geometric transformations and then establish correspondence between the aligned elements. Finally, correspondence can be also established by iterating between both previously mentioned approaches [56, 306].
The similarity measure that is also referred to as the objective function, cost function, or error measure is an assessment measure for the quality of the correspondence. It is formulated based on the type of the input data (e.g. surface) and the problem (e.g. alignment of the shapes with rigid transformation). Similarity measures that try to establish correspondence without applying rigid or non-rigid transformation generally include two terms. The first term tries to maximize the similarity between the corresponding elements, whereas the second one attempts to minimize the distortions that might occur in the shapes.
The cost functions of the correspondence methods that are based on rigid body alignments try to represents the alignment error. When the input data is composed of point sets, this error can be expressed in terms of the sum of squared dis-

tances between the corresponding point pairs. When non-rigid transformation is carried out for alignment of shapes, the cost function is generally composed of at least two terms. One of the terms normally represents the alignment error whereas the other term (also known as regularization or penalty term) constrains the deformation to satisfy the similarity of the transformations at the neighboring vertices [56, 306, 307].

Different optimization algorithms including the Downhill simplex method, the gradient ascent or descent, and the Levenberg-Marquart are used for optimizing the similarity measures [56]. Given all possible variations in the data type (e.g. points), problem type (e.g. rigid alignment of the shapes), similarity measures, and optimization algorithms, it is not surprising that there is a wide range of correspondence algorithms. These are classified into different groups by some authors [43, 115].

One of the most important methods is the Iterative Closest Point (ICP) algorithm. Josephson et al. use the ICP algorithm proposed by Besl et al. [263, 308]. The proposed method uses two surfaces as input data. While one triangulated surface (source) is kept as reference, a corresponding triangulation on the other surface (target) is found in an iterative way. The iterative process mainly designates the nearest points at the target surface for each vertex of the source triangles. It refines the rigid transformations (three translational and three rotational parameters) in such a way that the mean square error between both point sets is minimized. Afterwards, the transformation found in the previous step is applied. These processes are applied repeatedly until a convergence is reached [308].

The main drawback of the standard ICP algorithm is its limitations to rigid transformations. These transformations are not sufficient to perfectly align the objects that do not satisfy the rigidity criteria and the objects of the same class with different geometries. Non-rigid transformations can be used in order to overcome the above-mentioned limitations [43, 306]. Similarity (translation, rotation and uniform scaling), affine (translation, rotation, scaling and shear), or the curved transformations that are sub-classes of non-rigid transformations can be considered for the non-rigid case [56]. By using free-form deformation, i.e. a type of curved transformation, it is possible to allow for all types of deformation. Many free-form deformation models require a definition of a grid of points. Following the creation of the grid, control points are moved separately to optimize the similarity measure thereby allowing for description of local deformations. The found transformations are conveyed to the neighborhoods of the control points by interpolation [56].

Fleute et al. used a multi-resolution approach based on octree splines as proposed by Szeliski et al. [47, 309]. The method basically aims to find minimum non-rigid deformation between two 3D surfaces. For this purpose, the method enforces the least squares minimization of the distances between both sets of points. Moreover, transformations between 3D surfaces are defined based on the notions from free-form deformations, octree-splines, and hierarchical basis functions [47].

A different alternative to establish correspondence between the shapes of the training data is based on pattern recognition techniques. Baka et al. used a corre-

spondence algorithm that is originally proposed by Ferrarini and his colleagues in 2007 [44, 310]. In this method, a topology-learning unsupervised clustering is used to define the interesting parts of the surface and the correspondence problem among the shapes of the training data is considered as a classification issue. By using a classifier and a topology learning unsupervised clustering, corresponding parts between the shapes are detected [43, 310].
Another method to provide correspondence between the shapes has been suggested by Frangi et al. [311] and used by several authors [79, 52]. The introduced method uses a free-form registration technique based on maximization of a similarity measure, i.e. the normalized mutual information (NMI). The method requires construction of an atlas from the examined objects. Following the previously mentioned step, the landmarks are automatically extracted from the atlas. The extracted landmarks are matched to every shape of the training dataset by using a quasi-affine transformation and elastic transformation [311, 43].
A different approach that allows one to register an atlas, not to the other segmented features (e.g. surfaces) but to the original gray-value images is suggested by Rueckert et al. [312]. The proposed method and its variations are used by several authors [45]. Within the method, a deformation field is described instead of modeling the shape. Statistical analysis is directly performed on the deformation field, which is obtained based on a non-rigid registration. For the non-rigid transformation, free-form deformation model with B-splines interpolation is chosen. Moreover, the used similarity measure is NMI [46].
With the aim of establishing correspondence between two topologically equivalent 3D shapes, described as triangular meshes, Lamecker and his colleagues have used a method that allows mapping between the shapes and an appropriate base domain (a disc) [303]. This method was first proposed by Zöckler et al. [313]. In this method, surfaces are divided into patches by using a number of manually selected points on the surfaces. Each corresponding pair of patches that represent similar anatomical regions is mapped to a base domain, a unit disc. As there are many patches, a special care is taken to achieve continuity across the patch borders. The patches are then parameterized independently from each other, by using a complex parameterization method, shape preserving mapping. Following the parameterization, one to one correspondence between the input patches is described by overlaying the two previously found parameterizations [303, 313].
Many of the previously described approaches solve the correspondence problem by defining relation between each shape pairs (pair-wise approaches) of the training set. Moreover, a particular form of correspondence algorithms takes into account a group of shapes (group-wise approaches) at the same time. Unlike the pair-wise approaches, group-wise correspondences aim to optimize a cost function over the training set in a way that the resulting statistical model shows desirable properties (e. g. compact model) [43, 310].
One of the known group-wise methods, Minimum Description Length (MDL), is described by Davies et al. [314] and used by several authors [315, 316, 48, 114]. MDL principally aims to find the parameterization for each shape of the training dataset in such a way that the selected set of parameters lead to a compact

model, eligible to efficiently describe the training data. The main procedures of the method include the following: generation of parameterization for each shape in a recursive manner, generation of a model from the sampled shapes in agreement with the correspondence defined by the parameterization, evaluation of the model functioning through an objective function and generation of a new model using new parameters. The new parameters are found in a way that the defined objective function is optimized. The main steps are iteratively applied until a converged solution is obtained. In the original MDL approach, a piecewise linear, recursive parameterization scheme was used to locate landmarks. Moreover, due to high non-linearity of the cost function, a genetic algorithm was involved in the optimization of the objective-function [43, 310, 314].
Davies et al., later on, used the Nelder-Mead algorithm for the optimization and a re-parameterization with Cauchy Kernels [317]. Furthermore, an improved version of the MDL of the model is described by Heimann et al. [318]. The enhanced method uses a gradient descent optimization instead of the Nelder-Mead algorithm. Besides, for re-parameterization, locally constrained Gaussian Kernel is preferred rather than the Cauchy Kernels [318]. An application of the enhanced method that leads to a faster convergence compared to the method involving the Nelder-Mead algorithm can be found in the study of Bredbenner et al. [25].

A.1.1.2 Shape Model Alignment

Once the correspondence is established between the shapes of the training data, it is possible to find a mean shape and a number of modes of variation from the training set. Before the performance of statistical analysis, it is essential not to take the shape changes resulting from the transformations into consideration, since the shape does not change under similarity transformations. Keeping the statistical model as specific as possible, all the shapes of the training set need to be aligned in a common coordinate system. There are numerous methods for this purpose. The most popular method for the alignment process is probably Generalized Procrustes Analysis (GPA) [49, 53, 52, 319, 51]. The main idea behind the GPA is to align each shape of the training set in a way that the sum of distances of each shape to the mean, D, is minimized [42] where D is defined as:

$$D = \sum_{i=1}^{N} |\boldsymbol{x_i} - \bar{\boldsymbol{x}}|^2 \tag{A.2}$$

Since a group of shapes is aligned to their unknown mean, GPA is simply an iterative method. In the iterative approach, it is necessary to choose a shape and assume it as the mean shape. Subsequently, all the remaining shapes of the training set are aligned to the mean shape. The new mean shape is calculated from the aligned shapes. If the calculated mean is different from the previously estimated mean shape, the procedure is repeated again. The iterations continue until the mean shape does not significantly change [138].
Furthermore, the methods used during the alignment have an effect on the resultant model. The general approach used for alignment is to position each shape

such that their centroids are on the origin, and to modify each shape in order to have equal sizes and to define the rotation that minimizes D in Eq. A.2 [42, 138]. The limitation of this approach is that nonlinearities caused by large shape differences within the training dataset may exist [42]. The nonlinearities may lead to inaccurate results in the determination of linear modes of variation [43].
Two other approaches for the alignment are also described by Cootes et al. [42]. First, it is possible to vary both scaling and rotation parameters when minimizing D (Eq. A.2). This approach may also introduce nonlinearities. In the second approach, each shape is transformed into the tangent space of the mean with the aim of preserving the linearity between the shapes. The proposed approach aligns the shapes with the mean, allows the variations of both scaling and rotation parameters, and projects each shape into the tangent space of the mean by scaling $\boldsymbol{x}$ with $1/(\boldsymbol{x} \cdot \bar{\boldsymbol{x}})$ [42].

A.1.1.3 Shape Variation

After alignment of all the shapes of the training dataset in a common coordinate system, the intra-class shape variation is extracted. A common approach is to use a statistical method called Principal Component Analysis (PCA) [74, 138, 27, 70, 59]. In the reminder of this subsection, application of PCA to the training dataset is briefly explained based on the treatment presented in [42, 43].
If there are N shapes in the training set, there are N sets of vectors $\boldsymbol{x}$ describing the coordinates of landmarks that are placed on each of the aligned shapes.
The mean shape can be calculated as:

$$\bar{\boldsymbol{x}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x_i} \tag{A.3}$$

The covariance matrix is then given as:

$$\boldsymbol{S} = \frac{1}{N-1} \sum_{i}^{N} (\boldsymbol{x_i} - \bar{\boldsymbol{x}})(\boldsymbol{x_i} - \bar{\boldsymbol{x}})^T \tag{A.4}$$

Once the covariance matrix is obtained, it is possible to calculate eigenvectors (principal modes of variation) $\boldsymbol{\phi_s}$ and their corresponding eigenvalues (variances) λ_s by eigen-decomposition of the covariance matrix, $\boldsymbol{S}$. The resulting eigenvectors are sorted accoording to their corresponding eigenvalues, in a descending order $(\lambda_s > \lambda_{s+1})$ [42]. Subsequently, a shape instance can be described by adding the contributions of c most important modes of variation to the average shape:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \sum_{s=1}^{c} b_s \boldsymbol{\phi_s} \tag{A.5}$$

where b_s is the shape model parameter and describes the contribution of the sth mode of variation. There are numerous ways to determine the number of modes

of variation that should be retained. The decision as to how many modes are retained may affect the accuracy and the compactness of the resulting model [138]. The common approach in defining the number of modes c is to increase the number of modes until the ratio of the accumulated variance to the total variance, r , reaches a certain threshold value:

$$r = \frac{\sum_{s=1}^{c} \lambda_s}{\sum_{s=1}^{N-1} \lambda_s} \tag{A.6}$$

The generally accepted range for r is 0.9-0.98 [43]. An alternative approach, Horn's parallel analysis, can be used to determine the number of modes [65]. In this method, Monte-Carlo based simulations are used to compare the eigenvalues extracted from the correlation matrix with those retrieved from a random dataset. A component is generally accepted to be significant only if its corresponding eigenvalue is larger than the 95th of the eigenvalues that are extracted from the uncorrelated data [229].
Another important issue when generating a shape using Eq. A.5 is limiting the variation to obtain plausible shapes. Therefore, vector $\boldsymbol{b_s}$ that contains the shape model parameters has to be confined within a certain range. A common approach for generating shapes similar to those found in the original training dataset is to limit the model parameters b_i [42, 43, 49, 308, 47, 51, 116] to following limits:

$$-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i} \tag{A.7}$$

Another approach for defining valid regions for plausible shape generation is to restrict the Mahalanobis distance of $\boldsymbol{b_s}$:

$$D_m^2 = \sum_{s=1}^{c} (\frac{b_s^2}{\lambda_s}) \leq D_{max}^2 \tag{A.8}$$

Assuming that all modes represent one multivariate distribution, the limit D_{max} can be determined from the χ^2 distribution [73].

A.1.2 Statistical Models of Appearance

Statistical shape models can be used for image segmentation. However, only the shape of objects may not be sufficient to locate them on an image. While the sharpest image edges had been employed to adapt the first version of shape models (see section A.2.2), more specific models (i.e. statistical appearance models) were developed later on [43]. Similar to the case of statistical shape models, the statistical appearance models are built using a training set of images. The latter models capture the shape and the image texture variations across the training set. The image texture that might mean pixel intensities or other descriptors (e.g. image gradients) can be modeled around and/or within the part of the image covered by the shape of the object [72]. In the review written by Heimann et al., statistical appearance models are divided into two categories: the models based

on boundary features and those based on region-based features [43]. In order to make ASM (see section A.2.2) and AAM (see section A.2.3) easier to understand, here we only cover the statistical models of appearance that are often used in orthopedic applications and are usually based on region-based features with a focus on the statistical appearance model originally described by Cootes et al. [140]. The presented method and its variants were used by several authors in their studies [70, 50, 13, 320].

A.1.2.1 Statistical Models of Texture

With the aim of building the statistical models of texture as described by Cootes et al., there are three main steps to be followed. These steps include the capture of pixel information, the normalization of the data to discard the global illumination effects and the modeling of the normalized texture variation [138]. As Stegmann indicates it, the texture data acquisition is not a straightforward process. To capture the texture information, there is a need for a consistent method. Image warping that transforms one spatial configuration of an image into another can be used in the process of collecting the texture information. As the shape of the objects is generally represented in orthopedics with landmarks, image warping can simply be considered as the process of transforming one-point set's configuration $(\boldsymbol{x_i})$ into another $(\boldsymbol{x'_i})$ one [138]:

$$\boldsymbol{x'_i} = f(\boldsymbol{x_i}) \tag{A.9}$$

where f is a continuous mapping function. Two different forms of f, namely piece-wise affine and thin plate spline, are explained by Cootes [42].
To derive the texture information from the training set, each training example is warped to a standard shape that is generally the mean shape of the SSM [42, 43]. Following the warping process, the intensity information is sampled from the shape-normalized image and is concatenated into a texture vector, $\boldsymbol{g_{im}}$. Furthermore, the normalized version of intensities is necessary to reduce the effects of global lighting variations. The normalization of intensities can be done by scaling the texture vectors with α while applying an additional offset β:

$$\boldsymbol{g} = \frac{(\boldsymbol{g_{im}} - \boldsymbol{\beta 1})}{\alpha} \tag{A.10}$$

where $\mathbf{1}$ is a unit vector and the parameters α and β can be chosen in such a way that all grey values are transformed to zero-mean and unit-variance [42, 43]:

$$\alpha = \boldsymbol{g_{im}} \cdot \bar{\boldsymbol{g}} \tag{A.11}$$

$$\beta = \frac{(\boldsymbol{g_{im}} \cdot \mathbf{1})}{n} \tag{A.12}$$

where $\bar{\boldsymbol{g}}$ is the mean of the normalized texture and n is the number of elements in the vectors. As α is described with respect to the mean, the estimation of the mean of the normalized data is a recursive process [42, 138]. Following the

normalization of the texture data, as in the case of SSM, the texture variation can be modeled by applying the PCA to the normalized data. The covariance matrix of the textures can be written as:

$$S_g = \frac{1}{N-1}\sum_{i=1}^{N}(g_i - \bar{g})(g_i - \bar{g})^T \tag{A.13}$$

where $\bar{g}$ is the mean normalized texture vector:

$$\bar{g} = \frac{1}{N}\sum_{i}^{N} g_i \tag{A.14}$$

In a similar manner as in shape instance generation, a texture instance can be created by deforming the mean texture by a linear combination of eigenvectors [138]:

$$g = \bar{g} + \boldsymbol{\phi}_g \boldsymbol{b}_g \tag{A.15}$$

where vector $\boldsymbol{b}_g$ is a set of texture parameters and $bm\phi_g$ is a set of eigenvectors.

A common approach to describing the texture model is to use intensity values sampled from the image. However, texture models that are built based on the image intensities alone are sensitive to changes in imaging parameters. Therefore, these texture models may not perform adequately during the matching phase of the model with new unseen images [320]. In order to decrease the sensitivity of the model to imaging conditions, Scott et al. enhanced the texture model by using gradient orientation, corner, and edge strength [320]. In another attempt, Roberts et al. enhanced the texture models by utilizing the measures of corner and edge strength as described in Scott et al. [70, 320].

A.1.2.2 Combined Appearance Models

With the aim of obtaining one complete compact model, the shape and texture models that were briefly presented, can be unified. As it is indicated before, a shape instance can be generated by using a set of shape model parameters, $\boldsymbol{b}_s$, while texture instance can be constructed using a set of texture model parameters, $\boldsymbol{b}_g$. The set of shape and texture parameters, $\boldsymbol{b}_s$ and $\boldsymbol{b}_g$, can be concatenated into a vector $\boldsymbol{b}$:

$$\boldsymbol{b} = \begin{pmatrix} \boldsymbol{W}_s \boldsymbol{b}_s \\ \boldsymbol{b}_g \end{pmatrix} = \begin{pmatrix} \boldsymbol{W}_s \boldsymbol{\phi}_s^T (\boldsymbol{x} - \bar{\boldsymbol{x}}) \\ \boldsymbol{\phi}_g^T (\boldsymbol{g} - \bar{\boldsymbol{g}}) \end{pmatrix} \tag{A.16}$$

where $\boldsymbol{\phi}_s$ and $\boldsymbol{\phi}_g$ are the eigenvector matrices of shape and texture, respectively. Moreover, $\boldsymbol{W}_s$ is a diagonal matrix of weights for the shape parameters. As the shape parameters $\boldsymbol{b}_s$ and the texture parameters $\boldsymbol{b}_g$ have different units (distance and intensity, respectively), the use of $\boldsymbol{W}_s$ matrix is important to commensurate these parameters [42, 138]. A simple approach to calculate the $\boldsymbol{W}_s$ matrix is:

$$\boldsymbol{W}_s = r\boldsymbol{I} \tag{A.17}$$

where $\boldsymbol{I}$ is a unit matrix and r is the ratio of the total intensity variation to the total shape variation as seen in the training set [42, 138]:

$$r = \frac{\sum \lambda_{g_i}}{\sum \lambda_{s_i}} \tag{A.18}$$

where λ_{g_i} and λ_{s_i} are respectively the variances of the ith texture parameter b_{g_i} and that of the ith shape parameter b_{s_i}.
With the aim of obtaining a more compact model, a 3rd PCA can be applied on $\boldsymbol{b}$:

$$\boldsymbol{b} = \boldsymbol{\phi_c c} \tag{A.19}$$

where $\boldsymbol{c}$ is a set of combined model parameters and $\boldsymbol{\phi_c}$ is the eigenvector matrix of the combined shape and texture parameters. Furthermore, a complete model instance that covers the shape, $\boldsymbol{x}$, and the texture, $\boldsymbol{g}$, can be created by using the combined model parameters, $\boldsymbol{c}$:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{\phi_s W_s^{-1} \phi_{c,s}} \tag{A.20}$$

$$\boldsymbol{g} = \bar{\boldsymbol{g}} + \boldsymbol{\phi_g \phi_{c,g} c} \tag{A.21}$$

where

$$\boldsymbol{\phi_c} = \begin{pmatrix} \boldsymbol{\phi_{c,s}} \\ \boldsymbol{\phi_{c,g}} \end{pmatrix} \tag{A.22}$$

A.2 Search Algorithms

As previously mentioned, SSM and SAM can be used to search images in order to detect objects of interest. The process of applying a SSM or SAM to an image can simply be considered as an optimization problem. In order to match the model to the image, the set of parameters that refer to position, shape and conceivably the texture should be found in such a way that the cost function, representing the quality of the fit, is optimized [42]. There are several methods that can be used for solving the optimization problem for deformable models. The global search methods, such as the Genetic algorithms, Simulated Annealing and the Marquardt-Levenberg, are suitable for finding the parameters that optimize the fit, in case there is no information about the position of the object of interest in the image [42, 138]. On the other hand, local optimization techniques such as Powell's method and Nelder-Mead Simplex can be used in the presence of an initial guess for the correct solution. As Heimann et al. indicate it, methods that use local search algorithms are the preferred ones [43]. Since the use of local search methods is common, two popular local search algorithms, i.e. Active Shape Models (ASM) and Active Appearance Models (AAM) are presented here [43]. As indicated above, local search algorithms require initial estimation of the model parameters. Therefore, we first review several approaches that aim to acquire a rough estimation of the correct solution and initialize the search process before going through the ASM and AAM local search methods.

A.2.1 Initialization

The easiest way to initialize the search process is with user input. A common approach in user interaction is to roughly position the mean shape of the SSM close to the target object. The study of Lamecker et al. is one of the examples that involve user interaction to initialize the search process. As an alternative, some authors propose manual definition of a few specific points on the object of interest in the image [49, 73, 72, 75]. Furthermore, in order to (semi) automate the initialization, several image-processing techniques are deployed in the literature. Benjelloun et al. used Harris corner detector that is not sensitive to illumination variation and image noise for semi-automatic initialization. In the proposed method, a series of filters are used to reduce the high number of corners that result from Harris corner detector [49]. On the other hand, some authors had preferred to use the variants of Hough Transform, which detect parametric (e.g. circle) objects in an image. Pilgram et al. use Canny edge detection and Hough lines together with local constraints, related to anatomical object of interest [316]. As it is not always possible to find a parametric description for objects in an image, Generalized Hough Transform (GHT) is proposed by Ballard et al. in order to overcome that limitation [321]. The GHT technique has been used, for example, in [76, 322]. Furthermore, Fripp et al. automatically initialized the search procedure by using an affine registration to an atlas that was chosen from the training datasets [51].

A.2.2 Active Shape Models (ASM)

Active shape model is a local search algorithm based on a point distribution model. It was first proposed by Cootes et al. [323]. The introduced method and its variants are used by many authors [73, 305, 53, 308, 316, 13, 15, 134]. The main principle of ASM is to find the proper pose and shape parameters to best fit a model to an image. For this purpose, following the initialization of the ASM by using one of the initialization techniques, a recursive process starts. The main steps of the iterative process are as follows [42]:

- Generation of model instance and its expression in the image frame,
- Examination of the image around each model point to define a displacement that moves the model point to a better location,
- Updating the shape and pose parameters of the model with respect to the displacement found in the previous step.

These steps are performed iteratively, until a certain specified convergence criterion is satisfied [42]. In the following lines, the procedure will be briefly described.

A.2.2.1 Generation of Model Instance and Its Expression in the Image Frame

As previously described, a shape instance can be generated by deforming the mean shape by a linear combination of eigenvectors:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{\phi_s} \boldsymbol{b_s} \tag{A.23}$$

where $\boldsymbol{\phi_s}$ is the matrix of the eigenvectors and $\boldsymbol{b_s}$ is a vector of shape parameters. Furthermore, the generated shape instance can be expressed in the image frame by:

$$\boldsymbol{X}(\boldsymbol{p}, \boldsymbol{b_s}) = T(\boldsymbol{p})[\boldsymbol{x}] \tag{A.24}$$

where T is the similarity transformation matrix that imposes the rotation, scaling and translation. The transformation matrix is used for converting the shape space coordinates $\boldsymbol{x}$ into image coordinates $\boldsymbol{X}$ [116].

A.2.2.2 Calculation of a Proper Movement for Each Model Point

The examination of the image around each model point, in practice, is realized by looking along the profiles normal to the boundary. Several methods are used in literature, in order to find the proper movement for each model point. One approach is to locate the strongest edge along the profile, assuming that the model boundary corresponds to an edge. Using the position of the edge along the profile, the new suggested location of the landmark can be defined. Unfortunately, it is not always possible to locate model points on the strongest edge in the locality, since the strongest edge close to the model points might be related to another image structure or might represent a weaker secondary edge [42]. To lessen the possible wrong determination of landmark positions, an alternative approach is described by Cootes et al. [323]. The proposed method is mainly composed of two phases. In the first step (i.e. training phase), the statistical models of the image structure at the neighborhood of contour points are built by using the training set:
For a given landmark j in the ith training image, sampling is done along a profile normal to the boundary with a length of $2n_p + 1$ pixels centered at the landmark j. The $2n_p + 1$ samples are put in a vector $\boldsymbol{g_i}$. This procedure is repeated for each training image with the aim of obtaining a set of samples for the given landmark, j. In order to obtain the statistical model of the image structure for the given landmark j, the mean profile $\bar{\boldsymbol{g}}_{\boldsymbol{i}}$ and the covariance matrix $\boldsymbol{S_{g_i}}$ are calculated assuming that the data has a Gaussian distribution. To obtain one statistical model for each landmark, these procedures are repeated [42, 43, 116]. The extracted profiles might contain absolute grey-level (pixel intensity) values, the derivatives of the pixel intensities, or the normalized versions of those [43, 116]:

$$\boldsymbol{g_i'} = \frac{\boldsymbol{g_i}}{\sum_{k=1}^{2n_p+1} |g_{ik}|} \tag{A.25}$$

$$\boldsymbol{g_i''} = \frac{\boldsymbol{g_i} - \boldsymbol{min}(\boldsymbol{g_i})}{max(\boldsymbol{g_i}) - min(\boldsymbol{g_i})} \tag{A.26}$$

where g_{ik} is the kth element of the profile vector $\boldsymbol{g_i}$. Sampling the derivative along the profile, instead of sampling the absolute grey-level values, decreases the influence of global intensity changes [42]. On the other hand, in the second step (i.e. during search), the image structure around the contour point of the generated model instance is examined to find the new landmark position. For a contour point of the generated model instance, j, the image is investigated at distances $\Delta = -n_s \cdots + n_s$ pixels along the profile normal to the boundary. This can be considered as there are $2n_s + 1$ candidate target point positions (i.e. candidate new landmark position) [116]. For each candidate point, c, the profile that is centered at c and that is identical to $\boldsymbol{g_i}$, resulting in $\boldsymbol{g_{ci}}$ is obtained. A cost function that represents the quality of fit of the jth landmark of the model instance to its cth candidate point is constructed [116]. One of the cost functions that can be used for the assessment of the quality of the fit is the Mahalanobis distance of the sample $\boldsymbol{g_{ci}}$ from the model mean $\bar{\boldsymbol{g}}_{\boldsymbol{i}}$ that is separately estimated for each landmark, in the first step of the method:

$$f(\boldsymbol{g_{cj}}) = (\boldsymbol{g_{cj}} - \bar{\boldsymbol{g}}_l)^T \boldsymbol{S_{gi}^{-1}} (\boldsymbol{g_{cj}} - \bar{\boldsymbol{g}}_l) \tag{A.27}$$

The optimal target point (i.e. the best new location of the landmark of the generated model) is defined as the candidate point at which the Mahalanobis distance gets its lowest value [43]. To find the new locations of all of the model points, the explained procedure is repeated for each contour point.

A.2.2.3 Updating the Shape and the Pose Parameters of the Model

Following the determination of new optimal positions of all of the model points, the pose and shape parameters of the model are updated to best fit the model to the target points. In the original method by Cootes et al., the pose parameters of the model are first modified by keeping the shape parameters fixed. After the adjustment of the pose parameters, the shape parameters are updated [116].
As it is stated before, the generated instance of the model is expressed in the image frame by Eq. A.24, where $\boldsymbol{X}$ represents the image coordinates of the generated model instance with a set of pose parameters $\boldsymbol{p}$ and a set of shape parameters $\boldsymbol{b_s}$. On the other hand, the coordinates of the defined target points in the image frame can be given by:

$$\boldsymbol{X} + \boldsymbol{dX} \tag{A.28}$$

The pose parameters are updated in a way that the new shape minimizes the expression of:

$$\|(\boldsymbol{X} + \boldsymbol{dX}) - \boldsymbol{X'}\|^2 \tag{A.29}$$

where $\boldsymbol{X'}$ is the new shape expressed by:

$$\boldsymbol{X'}(\boldsymbol{p'}, \boldsymbol{b_s}) = T(\boldsymbol{p'})[\boldsymbol{x}] \tag{A.30}$$

Following the adjustment of pose parameters, there remain residuals:

$$[\boldsymbol{dx}] = (T(\boldsymbol{p'}))^{-1} \cdot [\boldsymbol{X} + \boldsymbol{dX} - \boldsymbol{X'}] \tag{A.31}$$

where $\boldsymbol{dx}$ is the residual expressed in the model coordinate frame. This residual can only be minimized by an update of shape parameters. With the aim of updating the shape parameters, $\boldsymbol{b_s}$, and the residuals are projected onto the model eigenmodes by:

$$\boldsymbol{db_s} = \boldsymbol{\phi_s^{-1}} \boldsymbol{dx} \tag{A.32}$$

Afterwards, the shape parameters are updated:

$$\boldsymbol{b'_s} = \boldsymbol{b_s} + \boldsymbol{db_s} \tag{A.33}$$

such that the new shape, $\boldsymbol{X''}(\boldsymbol{p'}, \boldsymbol{b'_s})$, minimizes the expression:

$$\|(\boldsymbol{X} + \boldsymbol{dX}) - \boldsymbol{X''}(\boldsymbol{p'}, \boldsymbol{b'_s})\|^2 \tag{A.34}$$

As the shape of the model instance should remain similar to those of the training set, limits can be enforced on shape parameters. Following the adjustments of parameters and application of limits on shape parameters, the updated positions of the model points are calculated [116].
Later on, to increase the robustness of the ASM algorithm, and to simultaneously reduce the required time for the search, a multi-resolution framework is presented [324]. Based on the proposed method, the image data (the training images and the image in which the search will be performed) is organized as a multi-resolution pyramid. For each different level of the pyramid, distinct statistical models of the image structure along normal profiles through each landmark are built in the way previously described. On the other hand, ASM search starts at the coarsest resolution level. If the model converges to a good solution, search at the next (finer) resolution starts. When the defined convergence criterion is satisfied at the finest resolution, the search is stopped [42]. As the multi-resolution framework enhances the speed and the quality of the model fitting to an image, several authors are reported to use the approach in their studies [308]. Furthermore, during the standard ASM search, the best new positions of the landmarks are estimated independently for each contour point of the model. It is, however, possible to encounter outliers that lead to erroneous positioning of the landmarks. Due to outliers such as multiple nearby edges in the image, the ASM contour may locally be detached from the correct boundary locations [316, 116]. To make ASM method more robust against such outliers, several modifications are suggested for the standard ASM. One of the modifications is proposed by Behiels et al. [116]. In this approach, Behiels et al. define a cost function that covers the whole set of target points which is minimized with a minimal cost path search. In the method, the outliers are penalized by imposing smoothness constraint on the displacement of adjoining model points towards their target points, in a way that the optimal solution is obtained regarding to the defined cost function [43, 116]. In a different approach, geometry-driven

scale space method that benefits from the geometry of a deformable model is used by Pilgram et al. [316]. Unlike the traditional linear scale space technique which blurs the boundary and accordingly makes the segmentation difficult, the geometric driven scale space blurs the image profiles along the shape boundary, but not across the object boundary. The scale space is constructed on the image profiles that are sampled from the image in a way that it is described in ASM. Afterwards, the statistical model of the image structure is built by using the features from the scale space. As Pilgram et al. indicate, features from the scale space are sampled in a coarse to fine fashion which allows the regularization of potential outliers [316, 325].

A.2.3 Active Appearance Models (AAM)

The previously described local search algorithm, ASM, locates the model points on a new image by using the constraints of the shape models together with additional information about the image structure around the model points. One limitation of this search algorithm is that it does not benefit from all the possible information that can be derived from the object of interest [42]. Therefore, an alternative local search algorithm, AAM, is described by Cootes et al. to also take advantage of more information [140]. AAM can be considered as an extension of ASM, since AAM uses texture information (i.e. the pixel intensities across the object) besides the shape information to search images for new instances of the object of interest [138].
An instance of an AAM can be generated by using the combined shape and texture parameters, $\boldsymbol{c}$. Furthermore, the generated model instance can be expressed in the image frame by the use of the transformation matrix. In order to assess the quality of fit, the image texture is warped to the mean shape of the model. Following the warping, the texture is normalized and can be written as a vector $\boldsymbol{g_{im}}$. The residuals between the normalized texture of image and the model can be expressed as:

$$\delta \boldsymbol{g} = \boldsymbol{g}_{image} - \boldsymbol{g}_{model} \tag{A.35}$$

Moreover, the error can be defined:

$$\boldsymbol{E} = \delta \boldsymbol{g}^2 \tag{A.36}$$

Like in the case of ASM, the search can be considered as an optimization problem. In order to match the model to the image in the best possible way, the error that is defined with $\boldsymbol{E}$ is aimed to be minimized by adjusting the pose parameters $\boldsymbol{p}$ and combined model parameters $\boldsymbol{c}$. As there are quite a number of parameters (model and pose parameters) that should be adjusted during fitting, the optimization problem is high-dimensional. It is a known fact that solving high-dimensional problems are computationally expensive. In order to lessen the complexity of the problem, Cootes et al. propose the use of a-priori knowledge on properly adjusting the parameters during search [140]. This "a-priori knowledge" can simply be obtained during the model building time. Thus, the

computational time during the search can be reduced. In order to get the "a-priori knowledge", the simplest linear model is considered by proposing that the spatial pattern in $\delta \boldsymbol{g}$ can foresee the adjustments of parameters in a way that the error, $\boldsymbol{E}$ is minimized. The simple linear model can be expressed as [138]:

$$\delta \boldsymbol{c} = \boldsymbol{R} \delta \boldsymbol{g} \tag{A.37}$$

where, $\delta \boldsymbol{g}$ is the texture residuals and $\delta \boldsymbol{c}$ is the parameter updates. As it is pointed out by Heimann et al., the success of the optimization scheme is mainly the subject of the derivative matrix $\boldsymbol{R}$ [43]. On the other hand, a suitable matrix $\boldsymbol{R}$ can be computed using a multivariate linear regression on a series of experiments. These experiments can be conducted by systematically displacing pose and model parameters with a known amount. Afterwards, the difference between the model and the part of the image that remain under the model can be measured [138]. Later on, numeric differentiation, a more reliable and faster method compared to regression, is proposed by Cootes et al. to compute the derivative matrix $\boldsymbol{R}$ [140].

Appendix B

B.1 Removing the Scaling and Rotational Effects

During the alignment of all shape examples of the training set in a common coordinate system, the scaling and rotational effects were removed by using the technique described by Du et al. and Arun et al. [135, 136], respectively. To find the rotational matrix, the described method basically covers the use of singular value decomposition (SVD) algorithm. The main steps involved in obtaining the rotational matrix can be given as the following [135]:

1. Calculate the matrix, $\boldsymbol{M}$, by using two set of points, $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$:

$$\boldsymbol{M} = \boldsymbol{x}_1 \boldsymbol{x}_2^T \tag{B.1}$$

   where, the superscript T represents the transpose of the vector.

2. Find the SVD of the matrix, $\boldsymbol{M}$:

$$\boldsymbol{M} = \boldsymbol{U} \Lambda \boldsymbol{U}^T \tag{B.2}$$

3. Calculate the rotation matrix that would be necessary for superimposing the point set $\boldsymbol{x_1}$ upon $\boldsymbol{x_2}$:

$$\boldsymbol{R} = \boldsymbol{V} \boldsymbol{U}^T \tag{B.3}$$

On the other hand, to filter out the scaling effects, the scaling can be calculated as [136]:

$$\boldsymbol{S} = \frac{trace(\boldsymbol{x_2} \boldsymbol{x_1^T} \boldsymbol{R^T})}{trace(\boldsymbol{x_1} \boldsymbol{x_1^T})} \tag{B.4}$$

In the alignment process, the point set, $\boldsymbol{x_2}$, can be considered as the vector representing the reference shape (i.e. mean shape).

B.2 Combined Appearance Model

With the aim of obtaining one compact model, the shape and texture models were unified as described by Cootes et al. [42]. Based on the described method, the set of shape and texture parameters, $\boldsymbol{b_s}$ and $\boldsymbol{b_g}$, can be concatenated into a vector $\boldsymbol{b}$:

$$\boldsymbol{b} = \begin{pmatrix} \boldsymbol{W_s} \boldsymbol{b_s} \\ \boldsymbol{b_g} \end{pmatrix} = \begin{pmatrix} \boldsymbol{W_s} \boldsymbol{\phi_s^T} (\boldsymbol{x} - \bar{\boldsymbol{x}}) \\ \boldsymbol{\phi_g^T} (\boldsymbol{g} - \bar{\boldsymbol{g}}) \end{pmatrix} \tag{B.5}$$

where $\boldsymbol{\phi_s}$ and $\boldsymbol{\phi_g}$ are the eigenvector matrices of shape and texture, respectively. Moreover, $\boldsymbol{W_s}$ is a diagonal matrix of weights for each shape parameter. The main steps to construct weight matrix $\boldsymbol{W_s}$ can be given as:
For each training example,

1. Sample the part of the image that remains under the untransformed shape and put the sampled values in a vector, $\boldsymbol{g}$.

2. Displace each shape model parameters, b_s from its optimum value and obtain displaced shape.

3. Sample the part of the image that remains under the displaced shape and put the sampled values in a vector, $\boldsymbol{g}'$.

4. Find the change in texture:

$$\boldsymbol{g_{ch}} = \boldsymbol{g}' - \boldsymbol{g} \tag{B.6}$$

5. Calculate the root mean square (RMS) change in texture. This gives the weight, w_s that will be applied to the related shape parameter.

With the intent of obtaining a more compact model and of discarding the correlation between the shape and texture parameters, a 3rd PCA can be applied on $\boldsymbol{b}$:

$$\boldsymbol{b} = \boldsymbol{\phi_c c} \tag{B.7}$$

where, $\boldsymbol{c}$ is a set of combined model parameters and $\boldsymbol{\phi_c}$ is the eigenvector matrix of the combined shape and texture parameters. Furthermore, a complete model instance that covers the shape, $\boldsymbol{x}$ and the texture, $\boldsymbol{g}$ can be created by using the combined model parameters, $\boldsymbol{c}$:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{\phi_s W_s^{-1} \phi c, sc} \tag{B.8}$$

$$\boldsymbol{g} = \bar{\boldsymbol{g}} + \boldsymbol{\phi_g \phi c, gc} \tag{B.9}$$

where,

$$\boldsymbol{\phi_c} = \begin{pmatrix} \boldsymbol{\phi_{c,s}} \\ \boldsymbol{\phi_{c,g}} \end{pmatrix} \tag{B.10}$$

B.3 Derivation of Derivative Matrix

Derivative matrix $\boldsymbol{R}$ is obtained based on the method described by Cootes et al. [140]. For one training image, I, the main steps of the method can be given as the following:

1. Warp the image I with the model parameters, $\boldsymbol{p}$ (combined model and pose parameters) to the mean shape.

2. Sample the intensities in the image and use these intensities to get combined model parameters.

3. Use the relation between the combined model parameters and the shape and texture parameters to obtain the model texture, $\boldsymbol{g_m}$.

4. Displace a parameter $\boldsymbol{p_i}$ with a small amount:

$$\boldsymbol{p_i'} = \boldsymbol{p_i} + \delta\sigma_i \tag{B.11}$$

   where, δ is a small value and σ_i is the standard deviation of the parameter.

5. Warp image I with the permuted parameters $\boldsymbol{p'}$ to the mean shape. Like in the step 1-3, get the model texture, $\boldsymbol{g_o}$.

6. Find the weighted sum of residual between the model texture, $\boldsymbol{g_m}$ and $\boldsymbol{g_o}$.

$$\frac{\delta \boldsymbol{r}}{\delta \boldsymbol{p_i}} = (\boldsymbol{g_m} - \boldsymbol{g_o})w_i \tag{B.12}$$

   where w_i is a weight,

$$w_i = \frac{exp(-\delta^2/2\sigma_i^2)}{\delta} \tag{B.13}$$

7. Repeat the above steps for a number of offset values δ (up to 0.5 standard deviations, σ_i for each parameter of $\boldsymbol{p_i}$) and find the mean of all $\frac{\delta \boldsymbol{r}}{\delta \boldsymbol{p_i}}$ approximations.

The described steps are repeated for each image of the training set. Afterwards, the mean of $\frac{\delta \boldsymbol{r}}{\delta \boldsymbol{p_i}}$ approximations can be found. Using this mean approximation, the derivative matrix $\boldsymbol{R}$ can be computed:

$$\boldsymbol{R} = (\frac{\delta \boldsymbol{r}^T}{\delta \boldsymbol{p}} \frac{\delta \boldsymbol{r}}{\delta \boldsymbol{p}})^{-1} \frac{\delta \boldsymbol{r}^T}{\delta \boldsymbol{p}} \tag{B.14}$$

B.4 Definition of Proximal Femur Thickness

Similar to the study of Naylor et al., the thickness of the proximal femur was defined by assuming that the proximal femur is a plate with a constant thickness and by considering that the cross section at femoral neck region is circular [128]. By using the equality of a plate's moment of inertia to those of the circle, the thickness can be defined:

$$\frac{\Pi D^4}{64} = \frac{tD^3}{12} \rightarrow t = \frac{3\Pi D}{16} \tag{B.15}$$

where, D is the width (cross section at the middle femoral neck).

B.5 Definition of Material Properties

Following the mesh generation on the segmented bone, material properties were defined. The main steps involved in material properties definition and assignment to one element can be given as:

1. Find the centroid of triangular element by using the element vertices, $\boldsymbol{x_1}$, $\boldsymbol{x_2}$ and $\boldsymbol{x_3}$.

2. Calculate the distances, d_1, d_2 and d_3 between the centroid of element and the vertices.

3. Find the maximum of the distances, d_{max} and create a grid by using it.

4. Search part of the original DXA image within the grid generated in the previous step and find the pixels that lie within the triangle.

5. Get the gray values of the pixels found in the previous step and find the average of all these values.

6. Calculate aBMD by using the average gray value and the calibration factor and the data packing factor given for the used DXA machine.

7. Divide the aBMD to the thickness of the proximal femur, t (see section, Definition of Proximal Femur Thickness) to get the volumetric BMD (vBMD):

$$vBMD = \frac{aBMD}{t} \tag{B.16}$$

8. Convert the vBMD to ash density (ρ_{ash}) and to apparent density (ρ_{app}) according to Naylor et al. and Schileo et al. [128, 131]:

$$\rho_{ash} = \frac{\rho_v}{1.14} \tag{B.17}$$

and

$$\rho_{app} = \frac{\rho_{ash}}{0.6} \tag{B.18}$$

9. Define the material properties by using the emprical equations presented by Morgan et al. [142]:

$$\begin{aligned} E &= 15010\rho_{app}^{2.18} \qquad \rho_{app} \leq 0.280\tfrac{g}{cm^3} \\ E &= 6850\rho_{app}^{1.49} \qquad \rho_{app} > 0.280\tfrac{g}{cm^3} \end{aligned} \tag{B.19}$$

B.6 Definition of Peak Impact Force

The peak impact force (N) was defined according to Naylor et al. [128]. To calculate the peak impact force, the weight, W (kg), and height, H (m), of the individuals are used:

$$F_{peak} = 9.806 m_e (n sin(w) - cos(w) + 1) \tag{B.20}$$

where, m_e is the effective mass, n is the impact number and v is the impact velocity:

$$m_e = 0.35W \tag{B.21}$$

$$n = \frac{v(\frac{71000}{m_e})^{0.5}}{9.806} \tag{B.22}$$

$$v = 2.72H^{0.5} \tag{B.23}$$

$$w = \Pi - tan^{-1}(n) \tag{B.24}$$

Appendix C

C.1 Determination of Subchondral Bone Surface

For each pair of adjacent bones (i.e. tibia and talus), subchondral bone surfaces were obtained as follows:

1. Triangulated bone surfaces (Fig. C.1) were reconstructed from segmentation results.

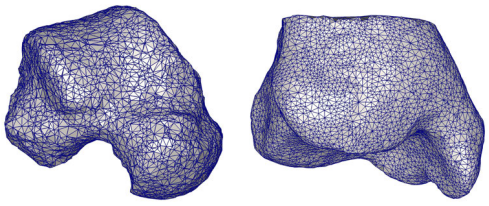


Figure C.1: Triangulated bone surfaces of talus (left) and tibia (right).

2. The normal vectors of the surface vertices (Fig. C.2) were calculated for both bones.

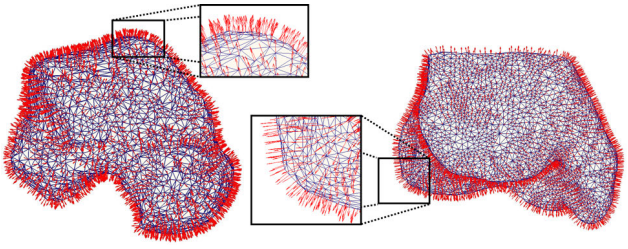


Figure C.2: Normal vectors of talus (left) and tibia (right) surface vertices.

3. The center of gravity of the triangles was calculated for both adjacent bone pairs.

4. A Point P located on the primary surface (e.g. tibia) was selected.

5. For Point P, the following steps were repeated for each triangle of the secondary surface (e.g. talus):

   5.1. A triangle T_n located on the secondary surface (e.g. talus) was selected.

   5.2. The distance between Point P and the center of gravity of the Triangle T_n was calculated.

5.3. The perpendicular distance between Point P and the secondary surface was found, if the calculated distance in Step 5.2 was less than the user-specified value of 4 mm. Otherwise, the Triangle T_n was skipped and Step 5.1 was restarted.

5.4. The distance value calculated in Step 5.3 was kept as a part of distance map, if the intersection of the perpendicular line was inside the triangle.

6. The Steps 4 - 5 were repeated for each point located on the primary surface.

C.2 Addition of Extra Points to Obtain Closed Form of Transected Tibia

Following the transection procedure of the tibiae, the upper parts of transected bones were open. Those open parts were closed by adding points to properly obtain triangulated surfaces of the transected bones in the next step (Fig. C.3).

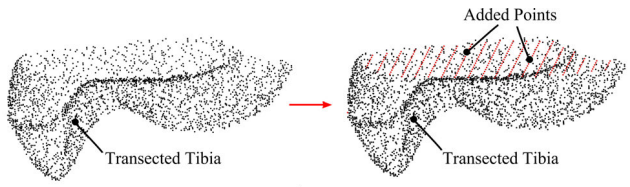


Figure C.3: An example of closing the transected bone's open parts with additional points.

The addition of extra points was performed as follows:

1. Edge points of the transected bone were determined that were located close to the cutting plane (Fig. C.4).

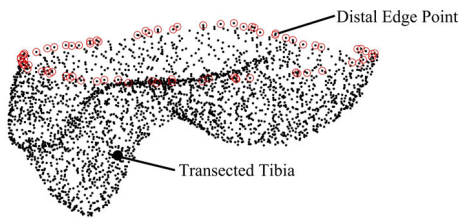


Figure C.4: Transected tibia. Distal edge points were represented with red circles.

2. Minimum and maximum coordinate values of the edge points were found.

3. Using the minimum and maximum coordinate values, a rectangle was generated and 19×19 equidistant points were created within the boundary of the rectangle (Fig. C.5).

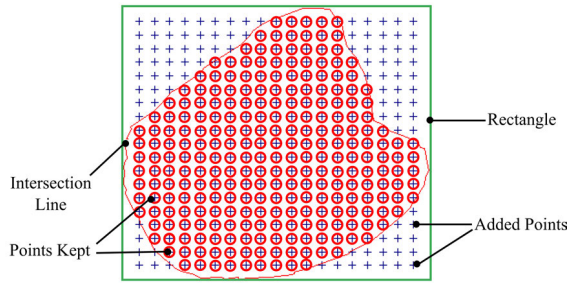


Figure C.5: Additional points were created within the boundary of a rectangle and points remaining inside the intersection line were kept.

4. Intersection of the rectangle and the edge points of the transected bone was found and the additional points were kept, if they were inside the intersection line (Fig. C.5).

C.3 Unbiased Point Registration Algorithm

The unbiased registration algorithm presented by van de Giessen [162] can be used to register a large number of shapes (e. g. talus bone, ulna bone) described by point clouds. The proposed method mainly fits an evolving mean shape to each of the point cloud. Simultaneously, each of the point clouds is rigidly aligned to the mean cloud, providing the stability of the method [162].
In this document, some of important points related to the registration algorithm will be provided. If more details are needed, readers are referred to the paper written by van de Giessen [162].
The registration algorithm establishes correspondence between N number of point clouds $C_i(i, 1, \cdots, N)$(Fig. C.6) and evolving mean cloud M with n_m points (Fig. C.6).
The algorithm basically consists of five steps:

1. Estimate an initial mean cloud M with n_m points using information of the surfaces from which the point clouds $C_i(i, 1, \cdots, N)$ were obtained (Fig. C.6).

2. Register the estimated mean cloud M to each cloud C_i by using the L_2 divergence as a similarity measure between C_i and deformed form of M, which results in M_i (Fig. C.6).

3. Update the current estimate of the mean shape M by computing the mean of all $M_i(i = 1, \cdots, N)$ (Fig. C.6).

4. Check the convergence. If the convergence is not reached, go to Step 2, else follow Step 5.

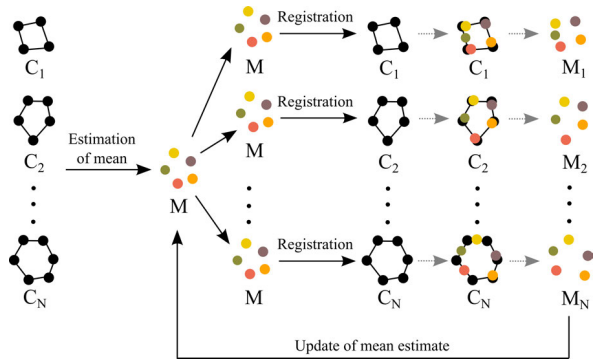


Figure C.6: Mean shape M is initially estimated using information of N number of point clouds $(C_1, C_2, \cdots, C_N)$. N number of copies of M are registered to each point cloud $(C_1, C_2, \cdots, C_N)$. $M_i (i = 1, \cdots, N)$ represents deformed forms of M and corresponding points between clouds M and $M_i (i = 1, \cdots, N)$ have the same colors. The new estimate of the mean shape M is computed based on the mean of M_i.

5. Use the correspondence relation between M_i and M to transform C_i to the mean cloud.

As it can be guessed from the main steps, an initial mean cloud estimation has an influence on the registration accuracy. Robustness of the technique to initial estimate of a mean cloud was evaluated by van de Giessen [162] and found that initialization of mean shape by drawing points from the 0-level set of the average signed distance transform gives the most accurate registration results. Therefore, the initial mean cloud was estimated as follows:

1. Each of the surfaces containing the points C_i was represented by a signed transform, with the surface as the 0-level.
2. The average of the signed transforms was obtained.
3. The surface of the mean shape was found by applying a marching cubes method on the 0-level of the average of the signed transforms.
4. The initial mean cloud was obtained by sampling the surface found in Step 3 with n_m points.

C.4 Optimization of Registration Parameters (Numerical Experiments)

An important issue to be mentioned is that a dense correspondence estimate between the point clouds C_i is realized using the registered surfaces [166]. Therefore, registration process and its accuracy play an important role in generation

of SSMs. The proposed registration algorithm requires a few parameters (i.e. the scale parameter σ for the mixture of Gaussians, the number of points in the mean cloud n_m, and the regularization term λ) to be set by the user. The parameter n_m together with the parameter λ defines the accuracy of the estimated correspondences during the registration process.
The optimal parameter settings (i.e. $n_m = 2000$, $\sigma = 1.5$ mm, $\lambda = 10^{-6}$) were determined by carrying out numerical experiments. The steps to perform the experiments are elaborately described by van de Giessen [326].
For optimization of parameters σ and n_m, the basic steps in this study were:

1. Take a triangulated surface of a single 3D shape.
2. Randomly sample the triangulated surface with n_m points.
3. Repeat the Step 2 two more times to obtain three different point clouds C_1, C_2 and C_3.
4. Do the registration of the clouds using the unbiased registration algorithm.
5. Transform the point clouds C_i (i=1, 2 and 3) to the mean cloud using the correspondence relation between M_i and M.
6. Compute the registration accuracy (E_{acc}) (i.e. the mean signed point-to-plane distance between all pairs of clouds V_i and V_j) using:

$$E_{acc} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1\setminus i}^{N} |\frac{1}{n_i} \sum_{k=1}^{n_i} \boldsymbol{u}_k \cdot (\boldsymbol{x}_k - \boldsymbol{x}_k^c)| \tag{C.1}$$

   where n_i is the number of points in the cloud V_i, $\boldsymbol{u_k}$ is the normal vector at point $\boldsymbol{x}_k$ in point cloud V_i and $\boldsymbol{x}_k^c$ is the point in cloud V_j closest to $\boldsymbol{x}_k$.

7. Compute the registration precision (i.e. the average point-to-plane distance between all pairs of clouds V_i and V_j) using:

$$E_{prec} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1\setminus i}^{N} \frac{1}{n_i} \sum_{k=1}^{n_i} |\boldsymbol{u}_k \cdot (\boldsymbol{x}_k - \boldsymbol{x}_k^c)| \tag{C.2}$$

This experiment was repeated for all combinations of $\sigma \in \{0.3, 0.6, 1.5, 3, 6, 15\}$, $n_m \in \{100, 200, 500, 1000, 2000\}$ and for 15 bone (i.e. talus and tibia) surfaces. Throughout the experiment, the regularization term λ, which is used to balance the data misfit and the deformation was kept as 10^{-9}.
For the determination of a suitable range of the regularization term λ, similar steps were followed:

1. Randomly select 3 different bone surfaces (i.e. bones from three different individuals).

2. Sample the surfaces with a fixed number of points ($n_m = 2000$) to obtain three point clouds C_1, C_2 and C_3.

3. Do the registration of the clouds using the unbiased registration algorithm.

4. Transform the point clouds C_i (i = 1, 2 and 3) to the mean cloud using the correspondence relation between M_i and M.

5. Compute the registration accuracy and the registration precision using the Eq. C.1 and Eq. C.2, respectively.

This experiment was repeated for all combinations of $\sigma \in \{0.3, 0.6, 1.5, 3, 6, 15\}$, $\lambda \in \{10^{-8}, 10^{-7}, \cdots, 10^{-3}\}$ and for 15 random selections of bone (i.e. talus and tibia) surfaces.

C.5 Robustness of Statistical Model(s)

To evaluate the robustness of the SSMs (i.e. SSM of tibia and SSM of talus), a bootstrap analysis was performed in a similar way described by Pedoia et al. [1]. In this study, five bootstrap replications were considered. At each replication, the SSM of talus and SSM of tibia were rebuilt, being independent from each other. These regenerated models were based on resampled versions of the surfaces (i.e. a resampling with replacement for 70% of the original surfaces) included in the training set of original SSMs. Using the rebuilt SSMs at each bootstrap, surfaces (i.e. mean $\pm$ 3SD) were modeled in each of first five modes of variation. These surfaces were compared to those obtained from the original SSMs by calculating the mean distance between the surface points. The mean distances were less than 1.4 mm for each case (Tables C.1 and C.2).

Table C.1: The average distances between the talus surfaces obtained from each bootstrap replication in each mode and the talus surfaces modelled from the original SSM of talus. BS stands for Boot Strap.

	Mean distance (mm)				
Mode	**1st BS**	**2nd BS**	**3rd BS**	**4th BS**	**5th BS**
Mode 1 (+3 SD)	1.36	1.31	0.45	0.49	0.24
Mode 1 (- 3 SD)	1.38	1.23	0.46	0.53	0.21
Mode 2 (+3 SD)	0.55	1.29	1.22	1.26	1.15
Mode 2 (- 3 SD)	0.53	1.21	1.12	1.14	1.24
Mode 3 (+3 SD)	0.49	0.95	0.92	0.22	1.00
Mode 3 (- 3 SD)	0.46	0.78	0.82	0.20	0.85
Mode 4 (+3 SD)	0.74	0.37	0.72	0.71	0.23
Mode 4 (- 3 SD)	0.87	0.35	0.81	0.69	0.24
Mode 5 (+3 SD)	0.35	0.92	0.37	0.86	0.38
Mode 5 (- 3 SD)	0.24	0.83	0.36	0.83	0.27

Table C.2: The average distances between the tibia surfaces obtained from each bootstrap replication in each mode and the tibia surfaces modelled from the original SSM of tibia. BS stands for Boot Strap.

	Mean distance (mm)				
Mode	**1st BS**	**2nd BS**	**3rd BS**	**4th BS**	**5th BS**
Mode 1 (+3 SD)	1.03	1.03	1.04	0.19	1.06
Mode 1 (- 3 SD)	0.96	0.96	1.05	0.20	1.01
Mode 2 (+3 SD)	0.23	0.24	0.77	0.96	0.20
Mode 2 (- 3 SD)	0.21	0.21	0.75	0.90	0.16
Mode 3 (+3 SD)	0.75	0.27	0.61	0.72	0.26
Mode 3 (- 3 SD)	0.77	0.25	0.66	0.74	0.21
Mode 4 (+3 SD)	0.22	0.65	0.56	0.24	0.61
Mode 4 (- 3 SD)	0.23	0.65	0.58	0.24	0.68
Mode 5 (+3 SD)	0.46	0.47	0.24	0.47	0.17
Mode 5 (- 3 SD)	0.50	0.51	0.20	0.48	0.16

C.6 Shape Parameters (Mean and 95% Confidence Interval)

The mean values of shape parameters together with 95% confidence interval for both control and OCD groups are presented in Table C.3.

Table C.3: The average and 95% confidence interval for shape parameters of both groups: OCD and control.

Bone Type	Mode	OCD	Control
Talus	**Mode 1**	0.24 (-0.08,0.55)	-0.25 (-0.60,0.10)
	Mode 2	-0.21 (-0.55,0.12)	0.23 (-0.10,0.56)
	Mode 3	-0.07 (-0.40,0.26)	0.08 (-0.28,0.43)
	Mode 4	0.21 (-0.17,0.60)	-0.22 (-0.49,0.04)
	Mode 5	-0.33 (-0.67,0.01)	0.35 (0.04,0.66)
Tibia	**Mode 1**	-0.47 (-0.67,-0.26)	0.49 (0.11,0.88)
	Mode 2	0.18 (-0.18,0.55)	-0.19 (-0.50,0.11)
	Mode 3	-0.03 (-0.30,0.24)	0.03 (-0.38,0.44)
	Mode 4	0.27 (-0.07,0.61)	-0.28 (-0.60,0.03)
	Mode 5	-0.01 (-0.42,0.40)	0.01(-0.24,0.26)

C.7 Effects of Age and Gender on Bone Shape Parameters

To test whether age and gender were predictors of shape parameters, a linear regression analysis was performed (Table C.4). The significance level ($p = 0.05$)

was adjusted for multiple comparisons using the Bonferroni correction, resulting in a significance threshold of $p = 0.005$ (i.e. $\frac{0.05}{2bones \times 5modes}$). Neither age nor gender predicted variations in bone shape parameters (p values > 0.005).

Table C.4: Linear regression analysis to test whether age and gender were predictors of shape parameters. β stands for the regression coefficient, while 95% CI represent 95% confidence interval for β.

		β	t	p -value	95% CI	
					Lower	**Upper**
Talus	**Mode 1**					
	Age	0.005	0.437	0.664	-0.019	0.029
	Gender	-0.222	-0.886	0.379	-0.724	0.279
	Mode 2					
	Age	-0.014	-1.257	0.213	-0.036	0.008
	Gender	0.451	1.929	0.058	-0.016	0.918
	Mode 3					
	Age	0.002	0.213	0.832	-0.020	0.025
	Gender	-0.419	-1.778	0.080	-0.890	0.052
	Mode 4					
	Age	0.032	2.675	0.010	0.008	0.056
	Gender	-0.470	-2.016	0.048	-0.936	-0.004
	Mode 5					
	Age	-0.019	-1.644	0.105	-0.043	0.004
	Gender	-0.222	-0.894	0.375	-0.716	0.273
Tibia	**Mode 1**					
	Age	-0.028	-2.455	0.017	-0.051	-0.005
	Gender	-0.069	-0.288	0.774	-0.550	0.411
	Mode 2					
	Age	0.017	1.432	0.157	-0.007	0.040
	Gender	-0.355	-1.448	0.153	-0.844	0.135
	Mode 3					
	Age	-0.013	-1.101	0.275	-0.036	0.010
	Gender	-0.088	-0.359	0.720	-0.579	0.402
	Mode 4					
	Age	0.018	1.619	0.110	-0.004	0.040
	Gender	-0.354	-1.519	0.134	-0.820	0.111
	Mode 5					
	Age	-0.001	-0.114	0.910	-0.025	0.023
	Gender	0.141	0.555	0.581	-0.366	0.648

Appendix D

D.1 Statistical Analyses of Ipsi- and Contralateral Sides as Separate Groups

This section covers some important details about the methods mentioned in the section "Statistical analyses ipsi- and contralateral sides as separate groups" of Chapter 6.

D.1.1 *D*-Statistic Used in Group Location Test

D(istance)-statistic used in the group location test was the Euclidean distance between the means of two groups [219, 223]. The main steps followed to calculate the *D*-statistic were:

1. Two $N \times K$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ consisting of the coordinates of m (i.e. $K = m \times 3$) landmarks established on all left (i.e. $N = 66$) and right (i.e. $N = 66$) side bones of the same type, respectively, were formed:

$$\boldsymbol{A} = \begin{pmatrix} Ax_{11} & Ay_{11} & Az_{11} & \cdots & Ax_{1m} & Ay_{1m} & Az_{1m} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ Ax_{N1} & Ay_{N1} & Az_{N1} & \cdots & Ax_{Nm} & Ay_{Nm} & Az_{Nm} \end{pmatrix} \tag{D.1}$$

$$\boldsymbol{B} = \begin{pmatrix} Bx_{11} & By_{11} & Bz_{11} & \cdots & Bx_{1m} & By_{1m} & Bz_{1m} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ Bx_{N1} & By_{N1} & Bz_{N1} & \cdots & Bx_{Nm} & By_{Nm} & Bz_{Nm} \end{pmatrix} \tag{D.2}$$

2. The mean of each column in the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, resulting in $\bar{\boldsymbol{A}}$ and $\bar{\boldsymbol{B}}$ is calculated:

$$\bar{\boldsymbol{A}} = \frac{1}{N} \times \sum_{i=1}^{N} \begin{pmatrix} Ax_{i1} & Ay_{i1} & Az_{i1} & \cdots & Ax_{im} & Ay_{im} & Az_{im} \end{pmatrix} \tag{D.3}$$

$$\bar{\boldsymbol{B}} = \frac{1}{N} \times \sum_{i=1}^{N} \begin{pmatrix} Bx_{i1} & By_{i1} & Bz_{i1} & \cdots & Bx_{im} & By_{im} & Bz_{im} \end{pmatrix} \tag{D.4}$$

3. The matrix $\bar{\boldsymbol{B}}$ is subtracted from the matrix $\bar{\boldsymbol{A}}$ in element-wise manner and the square of each difference is calculated.

$$\begin{aligned} &\bar{\boldsymbol{A}} - \bar{\boldsymbol{B}} \\ &= \begin{pmatrix} (\bar{A}_{x1} - \bar{B}_{x1})^2 & (\bar{A}_{y1} - \bar{B}_{y1})^2 & \cdots & (\bar{A}_{ym} - \bar{B}_{ym})^2 & (\bar{A}_{zm} - \bar{B}_{zm})^2 \end{pmatrix} \end{aligned} \tag{D.5}$$

4. All elements of the matrix obtained in Step 3 was summed and the square root of it was calculated.

D.1.2 *D*-Statistic Used in Variance-Covariance Scale Test

D-statistic used in the variance-covariance scale test was the absolute difference in the average residual of the two groups [219, 223]. The main steps followed to calculate the *D*-statistic were:

1. Steps 1-2 described in the previous section were followed.
2. The matrices $\bar{\boldsymbol{A}}$ and $\bar{\boldsymbol{B}}$ were subtracted from the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. The square of each difference is calculated, summed and the square root of the sum was calculated.

$$\begin{aligned}\boldsymbol{A} - \bar{\boldsymbol{A}} &= \boldsymbol{C} \\ &= \begin{pmatrix} \sqrt{(Ax_{11} - \bar{Ax}_1)^2 + (Ay_{11} - \bar{Ay}_1)^2 + \cdots + (Ay_{1m} - \bar{Ay}_m)^2 + (Az_{1m} - \bar{Ax}_m)^2} \\ \vdots \\ \sqrt{(Ax_{N1} - \bar{Ax}_1)^2 + (Ay_{N1} - \bar{Ay}_1)^2 + \cdots + (Ay_{Nm} - \bar{Ay}_m)^2 + (Az_{Nm} - \bar{Az}_m)^2} \end{pmatrix}\end{aligned} \tag{D.6}$$

$$\begin{aligned}\boldsymbol{B} - \bar{\boldsymbol{B}} &= \boldsymbol{D} \\ &= \begin{pmatrix} \sqrt{(Bx_{11} - \bar{Bx}_1)^2 + (By_{11} - \bar{By}_1)^2 + \cdots + (By_{1m} - \bar{By}_m)^2 + (Bz_{1m} - \bar{Bx}_m)^2} \\ \vdots \\ \sqrt{(Bx_{N1} - \bar{Bx}_1)^2 + (By_{N1} - \bar{By}_1)^2 + \cdots + (By_{Nm} - \bar{By}_m)^2 + (Bz_{Nm} - \bar{Bz}_m)^2} \end{pmatrix}\end{aligned} \tag{D.7}$$

3. The mean of the matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ was calculated, $\bar{\boldsymbol{C}}$ and $\bar{\boldsymbol{D}}$.
4. The square root of the squared difference between $\bar{\boldsymbol{C}}$ and $\bar{\boldsymbol{D}}$ was calculated.

D.2 Parallel Analysis

Parallel analysis (PA) [327, 229] was performed by generating a random dataset that has an equal dimension with that of the original data (i.e. same sample size and number of variables). Eigenvalues were derived from the random dataset. Generation of the random dataset and extraction of eigenvalues were repeated 50 times. The 95th of the distribution of the replicated eigenvalues (i.e. 95th percentile) for each component was determined. A principal component was kept if the associated eigenvalue (i.e. the one derived from the original data) was bigger than the calculated 95th percentile.

Appendix E

E.1 Monogenic Signal, Local Phase and Phase Symmetry Measure

To perform local analysis, a complex analytic signal $f_A(x)$ can be used and be formed by taken the original signal $f(x)$ (i.e. the real part) and its Hilbert transform $H[f(x)]$ (i.e. the imaginary part) into consideration:

$$f_A(x) = f(z) + iH[f(x)] \tag{E.1}$$

The local phase $\varphi(x)$ can be computed from the analytic signal:

$$\varphi(x) = tan^{-1}(\frac{H[f(x)]}{f(x)}) \tag{E.2}$$

Hilbert transform is mathematically restricted to a 1-D function. Felsberg and Sommer presented the monogenic signal [273], which is an extension of the analytic signal from 1-D to N-D by means of the Riesz transform.
The monogenic signal $f_M(x,y)$ in 2-D can be written in the same format as the 1-D analytic signal:

$$f_M(x,y) = f_e(x,y) + if_R(x,y) \tag{E.3}$$

where, f_e is the even component of the original 2-D image calculated by convolving the original image with an isotropic band-pass filter (e.g. log-Gabor filter), f_R represents the odd components of the band-pass filtered image:

$$f_e(x,y) = f(x,y) * g(x,y) \tag{E.4}$$

$$\begin{aligned} & f_R(x,y) \\ & = \sqrt{(f(x,y) * g(x,y) * h_1(x,y))^2 + (f(x,y) * g(x,y) * h_2(x,y))^2} \end{aligned} \tag{E.5}$$

where, $g(x,y)$ is the spatial domain representation of an isotropic band-pass filter $G(u,v)$; $h_1(x,y)$ and $h_2(x,y)$ are the spatial domain representation of the Riesz filter' fourier domain representation $H_1(u,v)$ and $H_2(u,v)$:

$$(H_1(u,v), H_2(u,v)) = (i\frac{u}{\sqrt{u^2+v^2}}, i\frac{v}{\sqrt{u^2+v^2}}) \tag{E.6}$$

Referring to the study of Hacihaliloglu et al. [271], a ridge-like edge appears at bone boundaries in US images. Therefore, in this study, a monogenic signal based multiscale ridge detector $PS(x,y)$ (i.e. Phase Symmetry) that is sensitive to bone surface localization in US images is calculated in a similar way described by Hacihaliloglu et al. [271, 266]:

$$PS(x,y) = \sum_{sc} \frac{\lfloor [|even_{sc}(x,y)| - |odd_{sc}(x,y)|] - T_{sc} \rfloor}{\sqrt{even_{sc}^2(x,y) + odd_{sc}^2(x,y)} + \epsilon} \tag{E.7}$$

where, sc stands for the scale variable and ϵ is a small constant to avoid division by zero. The $\lfloor \cdot \rfloor$ operator indicates that any negative values are replaced by zero. T_{sc} is a scale-specific noise threshold value [328].

E.2 Surface Point Extraction From CT Data

A point set from each pair of adjacent bones (i.e. distal tibia and talus) surfaces, which can partly be visualized using US, were obtained as follows:

1. Dicom images of a cadaveric ankle were imported in Matlab (Matlab 2013b, The Mathworks Inc., Natick, MA) and users were let scrolling down through slices (i.e. slices in sagittal view) to pick a slice in which the talar head – neck transition region (Fig. E.1) can clearly be seen.

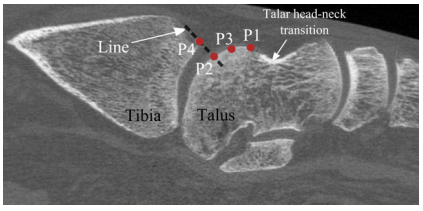


Figure E.1: Sagittal view of the ankle joint. Three points 'P1-P3' on talus and one point 'P4' on distal tibia are determined by users. Black dash line represents a fictive line passing through the utmost subchondral part of the tibia and it is used in determination of points 'P2' and 'P4'.

2. Users were let to pick four different points (i.e. three points on talus and one point on distal tibia) (Fig. E.1). The first point on talus is located close to the talar head-neck transition region. The second point on talus is positioned at the intersection point of a fictive line passing through the utmost subchondral part of the tibia with the talar dome (Fig. E.1). The third point on talar dome is defined in a way that it lies in between two previously created points (Fig. E.1). The fourth point is marked on the utmost subchondral part of the distal tibia (Fig. E.1).

3. Users were second time let to pick a slice by scrolling down through slices (i.e. slices in axial view). A slice is determined in a way that the talar dome width approximately gets its maximum (Fig. E.2).

4. Users were enabled to pick two different points at the outermost points of the talar dome (Fig. E.2).

5. Normals of the planes that pass through the points 'P1', 'P3', 'P5' and 'P6' are determined as follows:

   5.1. **Plane 1** (i.e. it passes through the point 'P1') (Fig. E.3): A circle is created using the points 'P1', 'P2' and 'P3'. Generated circle is discretized with 1000 points. One of the points, which is the nearest by the point 'P1' is chosen and named 'P7'. The normal of the 'Plane 1' points from the point 'P1' to the point 'P7'.

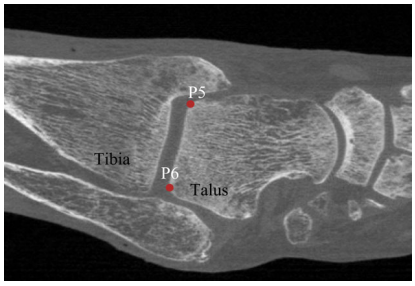


Figure E.2: Axial view of the ankle joint. Two points 'P5' and 'P6' are determined by at the outermost locations of the talar dome in the slice, in which the talar dome width roughly gets its maximum.

5.2. **Plane 2** (i.e. it passes through the point 'P3') (Fig. E.3): A line is drawn using the points 'P3' and 'P4' and it is rotated by 90°.

5.3. **Plane 3** (i.e. it passes through the point 'P5') (Fig. E.3) and **Plane 4** (i.e. it passes through the point 'P6') (Fig. E.3): Normals are defined in a way that they point outward from the respective face of talus.

6. Planes are drawn using the points 'P1', 'P3', 'P5' and 'P6' and the normals defined in the previous step (Fig. E.3). Triangulated bone (i.e. distal tibia and talus) surfaces, which have been extracted from the segmentation results, are cut using the four planes 'Plane 1 - Plane 4' and surfaces remaining between four planes are isolated (Fig. E.3).

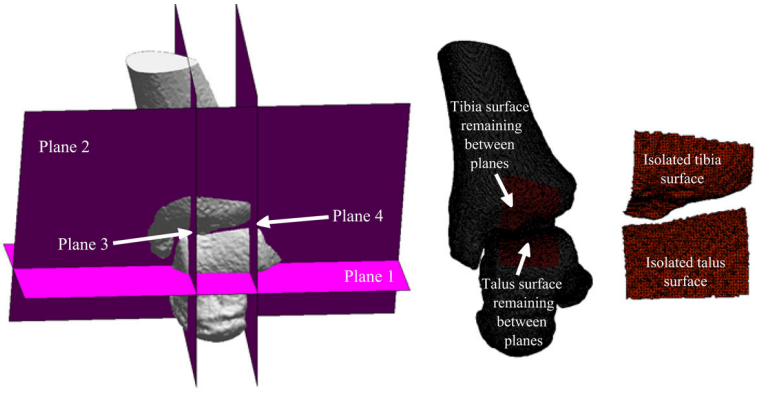


Figure E.3: (a) Four planes passing through the points 'P1', 'P3', 'P5' and 'P6' are generated. (b) Bone surfaces remaining between four planes are determined. (c) Surfaces remaining between four planes are isolated and vertices of the triangles located on isolated surfaces are used as sampling points in registration.

7. Vertices of triangles located on the isolated surfaces (Fig. E.3) are extracted

and used as sampling points during registration.

References

[1] V. Pedoia, D. A. Lansdown, M. Zaid, C. E. McCulloch, R. Souza, C. B. Ma, and X. Li, "Three-dimensional MRI-based statistical shape model and application to a cohort of knees with acute ACL injury," *Osteoarthritis and Cartilage*, vol. 23, no. 10, pp. 1695–1703, 2015.

[2] J. C. Baker-LePain and N. E. Lane, "Role of bone architecture and anatomy in osteoarthritis," *Bone*, vol. 51, no. 2, pp. 197–203, 2012.

[3] M. Reijman, S. Bierma-Zeinstra, J. Verhaar, H. Weinans, J. Waarsing, *et al.*, "Total hip replacement but not clinical osteoarthritis can be predicted by the shape of the hip: a prospective cohort study (CHECK)," *Osteoarthritis and Cartilage*, vol. 21, no. 4, pp. 559–564, 2013.

[4] T. Willson, S. D. Nelson, J. Newbold, R. E. Nelson, and J. LaFleur, "The clinical epidemiology of male osteoporosis: a review of the recent literature," *Clinical Epidemiology*, vol. 7, p. 65, 2015.

[5] M. Lorentzon and S. R. Cummings, "Osteoporosis: the evolution of a diagnosis," *Journal of Internal Medicine*, vol. 277, no. 6, pp. 650–661, 2015.

[6] A. Qaseem, M. A. Forciea, R. M. McLean, and T. D. Denberg, "Treatment of low bone density or osteoporosis to prevent fractures in men and women: a clinical practice guideline update from the American College of Physicians," *Annals of Internal Medicine*, vol. 166, no. 11, pp. 818–839, 2017.

[7] I. Castro-Mateos, J. M. Pozo, T. F. Cootes, J. M. Wilkinson, R. Eastell, and A. F. Frangi, "Statistical shape and appearance models in osteoporosis," *Current Osteoporosis Reports*, vol. 12, no. 2, pp. 163–173, 2014.

[8] J. S. Gregory and R. M. Aspden, "Femoral geometry as a risk factor for osteoporotic hip fracture in men and women," *Medical Engineering & Physics*, vol. 30, no. 10, pp. 1275–1286, 2008.

[9] J. Carballido-Gamio and D. P. Nicolella, "Computational anatomy in the study of bone structure," *Current Osteoporosis Reports*, vol. 11, no. 3, pp. 237–245, 2013.

[10] L. Humbert, Y. Martelli, R. Fonolla, M. Steghofer, S. di Gregorio, J. M. Sierra, J. Romera, and L. M. del Río Barquero, "3D-DXA: Assessing the femoral shape, the trabecular macrostructure and the cortex in 3D from DXA images.," *IEEE Trans. Med. Imaging*, vol. 36, no. 1, pp. 27–39, 2017.

[11] T. L. Bredbenner, R. L. Mason, L. M. Havill, E. S. Orwoll, D. P. Nicolella, and O. F. in Men (MrOS) Study, "Fracture risk predictions based on statistical shape and density modeling of the proximal femur," *Journal of Bone and Mineral Research*, vol. 29, no. 9, pp. 2090–2100, 2014.

[12] M. Taylor, E. Perilli, and S. Martelli, "Development of a surrogate model based on patient weight, bone mass and geometry to predict femoral neck strains and fracture loads," *Journal of Biomechanics*, vol. 55, pp. 121–127, 2017.

[13] S. Goodyear, R. Barr, E. McCloskey, S. Alesci, R. Aspden, D. Reid, and J. Gregory, "Can we improve the prediction of hip fracture by assessing bone structure using shape and appearance modelling?," *Bone*, vol. 53, no. 1, pp. 188–193, 2013.

[14] E. Dall'Ara, B. Luisier, R. Schmidt, M. Pretterklieber, F. Kainberger, P. Zysset, and D. Pahr, "DXA predictions of human femoral mechanical properties depend on the load configuration," *Medical Engineering & Physics*, vol. 35, no. 11, pp. 1564–1572, 2013.

[15] J. S. Gregory, D. Testi, A. Stewart, P. E. Undrill, D. M. Reid, and R. M. Aspden, "A method for assessment of the shape of the proximal femur and its relationship to osteoporotic hip fracture," *Osteoporosis International*, vol. 15, no. 1, pp. 5–11, 2004.

[16] J. C. Baker-LePain, K. R. Luker, J. A. Lynch, N. Parimi, M. C. Nevitt, and N. E. Lane, "Active shape modeling of the hip in the prediction of incident hip fracture," *Journal of Bone and Mineral Research*, vol. 26, no. 3, pp. 468–474, 2011.

[17] S. Glyn-Jones, A. Palmer, A. Price, T. Vincent, H. Weinans, A. J. Carr, *et al.*, "Osteoarthritis," *The Lancet*, vol. 386, no. 9991, pp. 376–387, 2015.

[18] V. Silverwood, M. Blagojevic-Bucknall, C. Jinks, J. Jordan, J. Protheroe, and K. Jordan, "Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis," *Osteoarthritis and Cartilage*, vol. 23, no. 4, pp. 507–515, 2015.

[19] V. Pedoia, M. A. Samaan, G. Inamdar, M. C. Gallo, R. B. Souza, and S. Majumdar, "Study of the interactions between proximal femur 3D bone shape, cartilage health, and biomechanics in patients with hip osteoarthritis," *Journal of Orthopaedic Research*, vol. 36, no. 1, pp. 330–341, 2018.

[20] M. Conconi, E. Halilaj, V. P. Castelli, and J. J. Crisco, "Is early osteoarthritis associated with differences in joint congruence?," *Journal of Biomechanics*, vol. 47, no. 16, pp. 3787–3793, 2014.

[21] H. G. Ahedi, R. M. Aspden, L. C. Blizzard, F. R. Saunders, F. M. Cicuttini, D. A. Aitken, G. Jones, and J. S. Gregory, "Hip shape as a predictor of osteoarthritis progression in a prospective population cohort," *Arthritis Care & Research*, vol. 69, no. 10, pp. 1566–1573, 2017.

[22] F. J. Blanco, N. Arden, F. Rannou, C. Cooper, F. W. Roemer, M. K. Javaid, A. Guermazi, J.-Y. Reginster, D. Hunter, and D. Hayashi, "Atlas of osteoarthritis," 2014.

[23] R. J. Barr, J. S. Gregory, D. M. Reid, R. M. Aspden, K. Yoshida, G. Hosie, A. J. Silman, S. Alesci, and G. J. Macfarlane, "Predicting OA progression to total hip replacement: can we do better than risk factors alone using active

shape modelling as an imaging biomarker?," *Rheumatology*, vol. 51, no. 3, pp. 562–570, 2011.

[24] M. Heijboer, S. Bierma-Zeinstra, J. Verhaar, H. Weinans, J. Waarsing, *et al.*, "Pincer deformities and mild acetabular dysplasia: the relationship between acetabular coverage and development of hip OA in the nationwide prospective CHECK cohort," *Osteoarthritis and Cartilage*, vol. 21, pp. S52–S53, 2013.

[25] T. L. Bredbenner, T. D. Eliason, R. S. Potter, R. L. Mason, L. M. Havill, and D. P. Nicolella, "Statistical shape modeling describes variation in tibia and femur surface geometry between control and incidence groups from the osteoarthritis initiative database," *Journal of Biomechanics*, vol. 43, no. 9, pp. 1780–1786, 2010.

[26] J. Waarsing, R. Rozendaal, J. Verhaar, S. Bierma-Zeinstra, and H. Weinans, "A statistical model of shape and density of the proximal femur in relation to radiological and clinical OA of the hip," *Osteoarthritis and Cartilage*, vol. 18, no. 6, pp. 787–794, 2010.

[27] J. S. Gregory, J. H. Waarsing, J. Day, H. A. Pols, M. Reijman, H. Weinans, and R. M. Aspden, "Early identification of radiographic osteoarthritis of the hip using an active shape model to quantify changes in bone morphometric features: can hip shape tell us anything about the progression of osteoarthritis?," *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, vol. 56, no. 11, pp. 3634–3643, 2007.

[28] R. Agricola, K. M. Leyland, S. M. Bierma-Zeinstra, G. E. Thomas, P. J. Emans, T. D. Spector, H. Weinans, J. H. Waarsing, and N. K. Arden, "Validation of statistical shape modelling to predict hip osteoarthritis in females: data from two prospective cohort studies (Cohort Hip and Cohort Knee and Chingford)," *Rheumatology*, vol. 54, no. 11, pp. 2033–2041, 2015.

[29] J. Lynch, N. Parimi, R. Chaganti, M. Nevitt, N. Lane, S. of Osteoporotic Fractures (SOF) Research Group, *et al.*, "The association of proximal femoral shape and incident radiographic hip OA in elderly women," *Osteoarthritis and Cartilage*, vol. 17, no. 10, pp. 1313–1318, 2009.

[30] T. Neogi, M. A. Bowes, J. Niu, K. M. de Souza, G. R. Vincent, J. Goggins, Y. Zhang, and D. T. Felson, "Magnetic resonance imaging–based three-dimensional bone shape of the knee predicts onset of knee osteoarthritis: data from the Osteoarthritis Initiative," *Arthritis & Rheumatism*, vol. 65, no. 8, pp. 2048–2058, 2013.

[31] D. J. Haverkamp, D. Schiphof, S. M. Bierma-Zeinstra, H. Weinans, and J. H. Waarsing, "Variation in joint shape of osteoarthritic knees," *Arthritis & Rheumatism*, vol. 63, no. 11, pp. 3401–3407, 2011.

[32] A. V. Pavlova, F. R. Saunders, S. G. Muthuri, J. S. Gregory, R. J. Barr, K. R. Martin, R. J. Hardy, R. Cooper, J. E. Adams, D. Kuh, *et al.*, "Statistical shape modelling of hip and lumbar spine morphology and their relationship in the MRC National Survey of Health and Development," *Journal of Anatomy*, vol. 231, no. 2, pp. 248–259, 2017.

[33] C. N. van Dijk, M. L. Reilingh, M. Zengerink, and C. J. van Bergen, "Osteochondral defects in the ankle: why painful?," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 18, no. 5, pp. 570–580, 2010.

[34] A. M. Bhosale and J. B. Richardson, "Articular cartilage: structure, injuries and review of management," *British Medical Bulletin*, vol. 87, no. 1, pp. 77–95, 2008.

[35] D. Blalock, A. Miller, M. Tilley, and J. Wang, "Joint instability and osteoarthritis," *Clinical Medicine Insights: Arthritis and Musculoskeletal Disorders*, vol. 8, pp. CMAMD–S22147, 2015.

[36] M. Zengerink *et al.*, "Osteochondral talar lesions and ankle biomechanics," 2017.

[37] P. F. O'loughlin, B. E. Heyworth, and J. G. Kennedy, "Current concepts in the diagnosis and treatment of osteochondral lesions of the ankle," *The American Journal of Sports Medicine*, vol. 38, no. 2, pp. 392–404, 2010.

[38] S. M. Raikin, I. Elias, A. C. Zoga, W. B. Morrison, M. P. Besser, and M. E. Schweitzer, "Osteochondral lesions of the talus: localization and morphologic data from 424 patients using a novel anatomical grid scheme," *Foot & Ankle International*, vol. 28, no. 2, pp. 154–161, 2007.

[39] T. Badekas, M. Takvorian, and N. Souras, "Treatment principles for osteochondral lesions in foot and ankle," *International Orthopaedics*, vol. 37, no. 9, pp. 1697–1706, 2013.

[40] A. Delle Sedie, L. Riente, and S. Bombardieri, "Limits and perspectives of ultrasound in the diagnosis and management of rheumatic diseases," *Modern Rheumatology*, vol. 18, no. 2, pp. 125–131, 2008.

[41] A. Kok, M. Terra, S. Muller, C. Askeland, C. van Dijk, G. Kerkhoffs, and G. Tuijthof, "Feasibility of ultrasound imaging of osteochondral defects in the ankle: a clinical pilot study," *Ultrasound in Medicine & Biology*, vol. 40, no. 10, pp. 2530–2536, 2014.

[42] T. F. Cootes, C. J. Taylor, *et al.*, "Statistical models of appearance for computer vision," 2004.

[43] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: a review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543–563, 2009.

[44] N. Baka, B. L. Kaptein, M. de Bruijne, T. van Walsum, J. Giphart, W. J. Niessen, and B. P. Lelieveldt, "2D–3D shape reconstruction of the distal femur from stereo x-ray imaging using statistical shape models," *Medical Image Analysis*, vol. 15, no. 6, pp. 840–850, 2011.

[45] D. C. Barratt, C. S. Chan, P. J. Edwards, G. P. Penney, M. Slomczykowski, T. J. Carter, and D. J. Hawkes, "Instantiation and registration of statistical shape models of the femur and pelvis using 3D ultrasound imaging," *Medical Image Analysis*, vol. 12, no. 3, pp. 358–374, 2008.

[46] C. S. Chan, P. J. Edwards, and D. J. Hawkes, "Integration of ultrasound-based registration with statistical shape models for computer-assisted or-

thopaedic surgery," in *Medical Imaging 2003: Image Processing*, vol. 5032, pp. 414–425, International Society for Optics and Photonics, 2003.

[47] M. Fleute, S. Lavallée, and R. Julliard, "Incorporating a statistically based shape model into a system for computer-assisted anterior cruciate ligament surgery," *Medical Image Analysis*, vol. 3, no. 3, pp. 209–222, 1999.

[48] K. T. Rajamani, M. A. Styner, H. Talib, G. Zheng, L. P. Nolte, and M. A. G. Ballester, "Statistical deformable bone models for robust 3d surface extrapolation from sparse data," *Medical Image Analysis*, vol. 11, no. 2, pp. 99–109, 2007.

[49] M. Benjelloun, S. Mahmoudi, and F. Lecron, "A framework of vertebra segmentation using the active shape model-based approach," *Journal of Biomedical Imaging*, vol. 2011, p. 9, 2011.

[50] N. Boukala, E. Favier, and B. Laget, "Active appearance model-based segmentation of hip radiographs," in *Medical Imaging 2005: Image Processing*, vol. 5747, pp. 443–453, International Society for Optics and Photonics, 2005.

[51] J. Fripp, S. K. Warfield, S. Crozier, and S. Ourselin, "Automatic segmentation of the knee bones using 3D active shape models," in *null*, pp. 171–174, IEEE, 2006.

[52] L. Humbert, T. Whitmarsh, M. de Craene, L. M. del Río Barquero, K. Fritscher, R. Schubert, F. Eckstein, T. Link, and A. F. Frangi, "3D reconstruction of both shape and bone mineral density distribution of the femur from DXA images," in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pp. 456–459, IEEE, 2010.

[53] H. H. Thodberg and A. Rosholm, "Application of the active shape model in a commercial medical device for bone densitometry," *Image and Vision Computing*, vol. 21, no. 13-14, pp. 1155–1161, 2003.

[54] G. Zheng, S. Gollmer, S. Schumann, X. Dong, T. Feilkas, and M. A. G. Ballester, "A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated x-ray images," *Medical Image Analysis*, vol. 13, no. 6, pp. 883–899, 2009.

[55] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, "A review of 3D/2D registration methods for image-guided interventions," *Medical Image Analysis*, vol. 16, no. 3, pp. 642–661, 2012.

[56] F. P. Oliveira and J. M. R. Tavares, "Medical image registration: a review," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 17, no. 2, pp. 73–93, 2014.

[57] M. Fleute and S. Lavallée, "Nonrigid 3D/2D registration of images using statistical models," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 138–147, Springer, 1999.

[58] G. Zheng, M. Á. Ballester, M. Styner, and L.-P. Nolte, "Reconstruction of patient-specific 3D bone surface from 2D calibrated fluoroscopic images

and point distribution model," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 25–32, Springer, 2006.

[59] R. Kurazume, K. Nakamura, T. Okada, Y. Sato, N. Sugano, T. Koyama, Y. Iwashita, and T. Hasegawa, "3D reconstruction of a femoral shape using a parametric model and two 2D fluoroscopic images," *Computer Vision and Image Understanding*, vol. 113, no. 2, pp. 202–211, 2009.

[60] A. Hurvitz and L. Joskowicz, "Registration of a CT-like atlas to fluoroscopic x-ray images using intensity correspondences," *International Journal of Computer Assisted Radiology and Surgery*, vol. 3, no. 6, p. 493, 2008.

[61] T. S. Tang and R. E. Ellis, "2D/3D deformable registration using a hybrid atlas," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 223–230, Springer, 2005.

[62] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Training models of shape from sets of examples," in *BMVC92*, pp. 9–18, Springer, 1992.

[63] T. Whitmarsh, K. D. Fritscher, L. Humbert, L. M. del Río Barquero, T. Roth, C. Kammerlander, M. Blauth, R. Schubert, and A. F. Frangi, "Hip fracture discrimination from dual-energy x-ray absorptiometry by statistical model registration," *Bone*, vol. 51, no. 5, pp. 896–901, 2012.

[64] T. Whitmarsh, L. Humbert, M. de Craene, L. M. del Rio Barquero, and A. F. Frangi, "Reconstructing the 3D shape and bone mineral density distribution of the proximal femur from dual-energy x-ray absorptiometry," *IEEE Transactions on Medical Imaging*, vol. 30, no. 12, pp. 2101–2114, 2011.

[65] T. Whitmarsh, L. Humbert, M. de Craene, M. Luis, K. Fritscher, R. Schubert, F. Eckstein, T. Link, and A. F. Frangi, "3D bone mineral density distribution and shape reconstruction of the proximal femur from a single simulated DXA image: an in vitro study," in *Medical Imaging 2010: Image Processing*, vol. 7623, p. 76234U, International Society for Optics and Photonics, 2010.

[66] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *Proc. European Conference on Computer Vision*, 1998.

[67] O. Ström, F. Borgström, J. A. Kanis, J. Compston, C. Cooper, E. V. McCloskey, and B. Jönsson, "Osteoporosis: burden, health care provision and opportunities in the EU," *Archives of Osteoporosis*, vol. 6, no. 1-2, pp. 59–155, 2011.

[68] S. P. Väänänen, J. S. Jurvelin, and H. Isaksson, "Estimation of 3D shape, internal density and mechanics of proximal femur by combining bone mineral density images with shape and density templates," *Biomechanics and Modeling in Mechanobiology*, vol. 11, no. 6, pp. 791–800, 2012.

[69] L. Ferrar, G. Jiang, J. Adams, and R. Eastell, "Identification of vertebral fractures: An update," 2005.

[70] M. G. Roberts, T. Oh, E. M. B. Pacheco, R. Mohankumar, T. F. Cootes, and J. E. Adams, "Semi-automatic determination of detailed vertebral shape from lumbar radiographs using active appearance models," *Osteoporosis International*, vol. 23, no. 2, pp. 655–664, 2012.

[71] M. Roberts, T. Cootes, E. Pacheco, and J. Adams, "Quantitative vertebral fracture detection on DXA images using shape and appearance models," *Academic Radiology*, 2007.

[72] M. G. Roberts, E. M. B. Pacheco, R. Mohankumar, T. F. Cootes, and J. E. Adams, "Detection of vertebral fractures in DXA VFA images using statistical models of appearance and a semi-automatic segmentation," *Osteoporosis International*, 2010.

[73] J. E. Iglesias and M. de Bruijne, "Semiautomatic segmentation of vertebrae in lateral x-rays using a conditional shape model," *Academic Radiology*, 2007.

[74] P. Mysling, K. Petersen, M. Nielsen, and M. Lillholm, "A unifying framework for automatic and semi-automatic segmentation of vertebrae from radiographs using sample-driven active shape models," in *Machine Vision and Applications*, 2013.

[75] P. P. Smyth, C. J. Taylor, and J. E. Adams, "Vertebral shape: automatic measurement with active shape models.," *Radiology*, 1999.

[76] B. Howe, A. Gururajan, H. Sari-Sarraf, and L. R. Long, "Hierarchical segmentation of cervical and lumbar vertebrae using a customized generalized hough transform and extensions to active appearance models," in *Image Analysis and Interpretation, 2004. 6th IEEE Southwest Symposium on*, pp. 182–186, IEEE, 2004.

[77] D. T. Felson and T. Neogi, "Osteoarthritis: Is it a disease of cartilage or of bone?," *Arthritis and Rheumatism*, 2004.

[78] J. H. Waarsing, M. Kloppenburg, P. E. Slagboom, H. M. Kroon, J. J. Houwing-Duistermaat, H. Weinans, and I. Meulenbelt, "Osteoarthritis susceptibility genes influence the association between hip morphology and osteoarthritis," *Arthritis and Rheumatism*, 2011.

[79] E. F. Chan, C. L. Farnsworth, J. A. Koziol, H. S. Hosalkar, and R. L. Sah, "Statistical shape modeling of proximal femoral shape deformities in legg-calvé-perthes disease and slipped capital femoral epiphysis," *Osteoarthritis and Cartilage*, vol. 21, no. 3, pp. 443–449, 2013.

[80] C. Lindner, S. Thiagarajah, J. M. Wilkinson, G. A. Wallis, and T. F. Cootes, "Development of a fully automatic shape model matching (FASMM) system to derive statistical shape models from radiographs : application to the accurate capture and global representation of proximal femur shape," vol. 21, pp. 1537–1544, 2013.

[81] S. P. Väänänen, H. Isaksson, J. H. Waarsing, A. A. Zadpoor, J. S. Jurvelin, and H. Weinans, "Estimation of 3D rotation of femur in 2D hip radiographs," *Journal of Biomechanics*, 2012.

[82] T. Alonso-Rasgado, D. Jimenez-Cruz, C. G. Bailey, P. Mandal, and T. Board, "Changes in the stress in the femoral head neck junction after osteochondroplasty for hip impingement: A finite element study," *Journal of Orthopaedic Research*, 2012.

[83] A. E. Anderson, B. J. Ellis, S. A. Maas, and J. A. Weiss, "Effects of idealized joint geometry on finite element predictions of cartilage contact stresses in the hip," *Journal of Biomechanics*, vol. 43, no. 7, pp. 1351–1357, 2010.

[84] A. A. Zadpoor, "Open forward and inverse problems in theoretical modeling of bone tissue adaptation," *Journal of the Mechanical Behavior of Biomedical Materials*, 2013.

[85] D. Kluess, R. Souffrant, W. Mittelmeier, A. Wree, K. P. Schmitz, and R. Bader, "A convenient approach for finite-element-analyses of orthopaedic implants in bone contact: Modeling and experimental validation," *Computer Methods and Programs in Biomedicine*, 2009.

[86] G. Campoli, M. Borleffs, S. Amin Yavari, R. Wauthle, H. Weinans, and A. Zadpoor, "Mechanical properties of open-cell metallic biomaterials manufactured using additive manufacturing," *Materials & Design*, 2013.

[87] C. Dopico-González, A. M. New, and M. Browne, "Probabilistic finite element analysis of the uncemented hip replacement-effect of femur characteristics and implant design geometry," *Journal of Biomechanics*, 2010.

[88] P. J. Prendergast, "Finite element models in tissue mechanics and orthopaedic implant design," 1997.

[89] R. B. Ruben, P. R. Fernandes, and J. Folgado, "On the optimal shape of hip implants," *Journal of Biomechanics*, 2012.

[90] D. Buchanan and A. Ural, "Finite element modeling of the influence of hand position and bone properties on the colles' fracture load during a fall," *Journal of Biomechanical Engineering*, 2010.

[91] D. Christen, D. J. Webster, and R. Müller, "Multiscale modelling and nonlinear finite element analysis as clinical tools for the assessment of fracture risk.," *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 2010.

[92] A. Tsouknidas, K. Anagnostidis, G. Maliaris, and N. Michailidis, "Fracture risk in the femoral hip region: A finite element analysis supported experimental approach," *Journal of Biomechanics*, 2012.

[93] G. Campoli, H. Weinans, F. van der Helm, and A. A. Zadpoor, "Subject-specific modeling of the scapula bone tissue adaptation," *Journal of Biomechanics*, 2013.

[94] P. Christen, B. van Rietbergen, F. M. Lambers, R. Müller, and K. Ito, "Bone morphology allows estimation of loading history in a murine model of bone adaptation," *Biomechanics and Modeling in Mechanobiology*, 2012.

[95] P. G. Coelho, P. R. Fernandes, H. C. Rodrigues, J. B. Cardoso, and J. M. Guedes, "Numerical modeling of bone tissue adaptation-a hierarchical approach for bone apparent density and trabecular structure," *Journal of Biomechanics*, 2009.

[96] J. Böhme, V. Shim, A. Höch, M. Mütze, C. Müller, and C. Josten, "Clinical implementation of finite element models in pelvic ring surgery for prediction of implant behavior: A case report," *Clinical Biomechanics*, 2012.

[97] G. Campoli, H. Weinans, and A. A. Zadpoor, "Computational load estimation of the femur," *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 10, pp. 108–119, 2012.

[98] A. A. Zadpoor, G. Campoli, and H. Weinans, "Neural network prediction of load from the morphology of trabecular bone," *Applied Mathematical Modelling*, vol. 37, no. 7, pp. 5260–5276, 2013.

[99] S. Poelert, E. Valstar, H. Weinans, and A. A. Zadpoor, "Patient-specific finite element modeling of bones," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 227, no. 4, pp. 464–478, 2013.

[100] A. A. Nikooyan, H. E. J. Veeger, P. Westerhoff, B. Bolsterlee, F. Graichen, G. Bergmann, and F. C. T. van der Helm, "An EMG-driven musculoskeletal model of the shoulder," *Human Movement Science*, vol. 31, no. 2, pp. 429–447, 2012.

[101] P. Saraswat, M. S. Andersen, and B. A. MacWilliams, "A musculoskeletal foot model for clinical gait analysis," *Journal of Biomechanics*, 2010.

[102] A. A. Nikooyan and A. A. Zadpoor, "An improved cost function for modeling of muscle activity during running," *Journal of Biomechanics*, vol. 44, no. 5, pp. 984–987, 2011.

[103] A. A. Nikooyan and A. A. Zadpoor, "Effects of muscle fatigue on the ground reaction force and soft-tissue vibrations during running: A model study," *IEEE Transactions on Biomedical Engineering*, 2012.

[104] A. A. Zadpoor and A. A. Nikooyan, "A mechanical model to determine the influence of masses and mass distribution on the impact force during running–a discussion.," *Journal of Biomechanics*, 2006.

[105] R. Bryan, P. B. Nair, and M. Taylor, "Use of a statistical model of the whole femur in a large scale , multi-model study of femoral neck fracture risk," *Journal of Biomechanics*, vol. 42, no. 13, pp. 2171–2176, 2009.

[106] D. P. Nicolella and T. L. Bredbenner, "Development of a parametric finite element model of the proximal femur using statistical shape and density modelling," *Computer Methods in Biomechanics and Biomedical Engineering*, 2012.

[107] C. Merle, W. Waldstein, J. S. Gregory, S. R. Goodyear, R. M. Aspden, P. R. Aldinger, D. W. Murray, and H. S. Gill, "How many different types of femora are there in primary hip osteoarthritis? an active shape modeling study.," *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*, vol. 32, no. 3, pp. 413–22, 2014.

[108] R. Bryan, P. Surya Mohan, A. Hopkins, F. Galloway, M. Taylor, and P. B. Nair, "Statistical modelling of the whole human femur incorporating geometric and material properties," *Medical Engineering & Physics*, 2010.

[109] L. Querol, P. Büchler, D. Rueckert, L. P. Nolte, and M. Á. G. Ballester, "Statistical finite element model for bone shape and biomechanical properties," in *Medical Image Computing and Computer-Assisted Interventation*, 2006.

[110] L. M. Specht and K. J. Koval, "Robotics and computer-assisted orthopaedic surgery.," *Bulletin (Hospital for Joint Diseases (New York, N.Y.))*, 2002.

[111] S. Benameur, M. Mignotte, S. Parent, H. Labelle, W. Skalli, and J. de Guise, "3D/2D registration and segmentation of scoliotic vertebrae using statistical models," *Computerized Medical Imaging and Graphics*, vol. 27, no. 5, pp. 321–337, 2003.

[112] Z. Zhu and G. Li, "Construction of 3D human distal femoral surface models using a 3D statistical deformable model," *Journal of Biomechanics*, 2011.

[113] G. Zheng, X. Dong, K. T. Rajamani, X. Zhang, M. Styner, R. U. Thoranaghatte, L. P. Nolte, and M. A. Gonzalez Ballester, "Accurate and robust reconstruction of a surface model of the proximal femur from sparse-point data and a dense-point distribution model for surgical navigation," *IEEE Transactions on Biomedical Engineering*, 2007.

[114] S. Schumann, M. Tannast, L. P. Nolte, and G. Zheng, "Validation of statistical shape model based reconstruction of the proximal femur-a morphology study," *Medical Engineering and Physics*, 2010.

[115] R. Davies, C. Twining, and C. Taylor, *Statistical models of shape: optimisation and evaluation*. 2008.

[116] G. Behiels, F. Maes, D. Vandermeulen, and P. Suetens, "Evaluation of image features and search strategies for segmentation of bone structures in radiographs using active shape models," *Medical Image Analysis*, 2002.

[117] Jianhua Yao and Taylor, "Assessing accuracy factors in deformable 2D/3D medical image registration using a statistical pelvis model," in *Proceedings Ninth IEEE International Conference on Computer Vision*, 2003.

[118] M. G. Roberts, T. F. Cootes, E. Pacheco, T. Oh, and J. E. Adams, "Segmentation of lumbar vertebrae using part-based graphs and active appearance models," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2009.

[119] M. a. Styner, K. T. Rajamani, L.-P. Nolte, G. Zsemlye, G. Székely, C. J. Taylor, and R. H. Davies, "Evaluation of 3D correspondence methods for model building.," *Information Processing in Medical Imaging*, 2003.

[120] M. B. Stegmann and D. D. Gomez, "A brief introduction to statistical shape analysis," *Informatics and Mathematical Modelling*, 2002.

[121] A. Gooßen, E. Hermann, G. M. Weber, T. Gernoth, T. Pralow, and R. R. Grigat, "Model-based segmentation of pediatric and adult joints for orthopedic measurements in digital radiographs of the lower limbs," in *Computer Science - Research and Development*, 2011.

[122] G. Guglielmi, F. Palmieri, M. G. Placentino, F. D'Errico, and L. P. Stoppino, "Assessment of osteoporotic vertebral fractures using specialized workflow software for 6-point morphometry," *European Journal of Radiology*, 2009.

[123] M. Roberts, T. F. Cootes, and J. E. Adams, "Vertebral morphometry: Semi-automatic determination of detailed shape from dual-energy x-ray absorp-

tiometry images using active appearance models," *Investigative Radiology*, 2006.

[124] M. Mahfouz, E. E. H. A. Fatah, L. S. Bowers, and G. Scuderi, "Three-dimensional morphology of the knee reveals ethnic differences," *Clinical Orthopaedics and Related Research*, vol. 470, no. 1, pp. 172–185, 2012.

[125] S. Amin Yavari, R. Wauthle, J. van der Stok, A. C. Riemslag, M. Janssen, M. Mulier, J. P. Kruth, J. Schrooten, H. Weinans, and A. A. Zadpoor, "Fatigue behavior of porous biomaterials manufactured using selective laser melting," *Materials Science and Engineering C*, 2013.

[126] J. van der Stok, O. P. van der Jagt, S. Amin Yavari, M. F. de Haas, J. H. Waarsing, H. Jahr, E. M. van Lieshout, P. Patka, J. A. Verhaar, A. A. Zadpoor, and H. Weinans, "Selective laser melting-produced porous titanium scaffolds regenerate bone in critical size cortical bone defects," *Journal of Orthopaedic Research*, 2013.

[127] J. van der Stok, H. Wang, S. Amin Yavari, M. Siebelt, M. Sandker, J. H. Waarsing, J. A. Verhaar, H. Jahr, A. A. Zadpoor, S. C. Leeuwenburgh, and H. Weinans, "Enhanced bone regeneration of cortical segmental bone defects using porous titanium scaffolds incorporated with colloidal gelatin gels for time- and dose-controlled delivery of dual growth factors," *Tissue Engineering Part A*, 2013.

[128] K. E. Naylor, E. V. McCloskey, R. Eastell, and L. Yang, "Use of DXA-based finite element analysis of the proximal femur in a longitudinal study of hip fracture," *Journal of Bone and Mineral Research*, 2013.

[129] D. Testi, M. Viceconti, A. Cappello, and S. Gnudi, "Prediction of hip fracture can be significantly improved by a single biomedical indicator," *Annals of Biomedical Engineering*, 2002.

[130] J. H. Keyak, S. A. Rossi, K. A. Jones, and H. B. Skinner, "Prediction of femoral fracture load using automated finite element modeling," *Journal of Biomechanics*, 1997.

[131] E. Schileo, F. Taddei, L. Cristofolini, and M. Viceconti, "Subject-specific finite element models implementing a maximum principal strain criterion are able to estimate failure risk and fracture location on human femurs tested in vitro," *Journal of Biomechanics*, 2008.

[132] G. J. Edwards, T. F. Cootes, and C. J. Taylor, "Advances in active appearance models," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 1, pp. 137–142, IEEE, 1999.

[133] J. Westerweel, "Digital particle image velocimetry," 1997.

[134] E. C. Pegg, S. J. Mellon, G. Salmon, A. Alvand, H. Pandit, D. W. Murray, and H. S. Gill, "Improved radiograph measurement inter-observer reliability by use of statistical shape models," *European Journal of Radiology*, 2012.

[135] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 5, pp. 698–700, 1987.

[136] D. Shaoyi, Z. Nanning, Y. Shihui, Y. Qubo, and W. Yang, "An extension of the ICP algorithm considering scale factor," in *Proceedings - International Conference on Image Processing, ICIP*, 2007.

[137] I. T. Jolliffe, *Principal Component Analysis, Second Edition*. 2002.

[138] M. B. Stegmann, "Active appearance models: Theory, extensions and cases," *Month*, 2000.

[139] T. F. Cootes and C. J. Taylor, "Anatomical statistical models and their role in feature extraction," *British Journal of Radiology*, 2004.

[140] T. Cootes, "Active appearance models," *Pattern Analysis and . . .* , 2001.

[141] D. J. Kroon, *Segmentation of the Mandibular Canal in Cone-Beam CT Data*. PhD thesis, 2011.

[142] E. F. Morgan, H. H. Bayraktar, and T. M. Keaveny, "Trabecular bone modulus-density relationships depend on anatomic site," *Journal of Biomechanics*, 2003.

[143] R. Scholz, F. Hoffmann, S. von Sachsen, W. G. Drossel, C. Klöhn, and C. Voigt, "Validation of density-elasticity relationships for finite element modeling of human pelvic bone by modal analysis," *Journal of Biomechanics*, 2013.

[144] M. L. Bouxsein and E. Seeman, "Quantifying the material and structural determinants of bone strength," 2009.

[145] R. Donner, M. Reiter, G. Langs, P. Peloschek, and H. Bischof, "Fast active appearance model search using canonical correlation analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006.

[146] S. Amin, D. L. Kopperdhal, L. J. M. 3rd, S. J. Achenbach, T. M. Therneau, B. L. Riggs, T. M. Keaveny, and S. Khosla, "Association of hip strength estimates by finite element analysis with fractures in women and men," *Journal of Bone and Mineral Research*, 2011.

[147] E. M. Lochmüller, J. B. Zeller, D. Kaiser, F. Eckstein, J. Landgraf, R. Putz, and R. Steldinger, "Correlation of femoral and lumbar DXA and calcaneal ultrasound, measured in situ with intact soft tissues, with the in vitro failure loads of the proximal femur," *Osteoporosis International*, 1998.

[148] S. P. Väänänen, H. Isaksson, P. Julkunen, J. Sirola, H. Kröger, and J. S. Jurvelin, "Assessment of the 3D shape and mechanics of the proximal femur using a shape template and a bone mineral density image," *Biomechanics and Modeling in Mechanobiology*, 2011.

[149] A. A. Zadpoor, J. Sinke, and R. Benedictus, "Elastoplastic deformation of dissimilar-alloy adhesively-bonded tailor-made blanks," *Materials and Design*, 2010.

[150] A. A. Zadpoor, J. Sinke, and R. Benedictus, "Global and local mechanical properties and microstructure of friction stir welds with dissimilar materials and/or thicknesses," *Metallurgical and Materials Transactions A: Physical Metallurgy and Materials Science*, 2010.

[151] S. Amin Yavari, J. van der Stok, H. Weinans, and A. A. Zadpoor, "Full-field strain measurement and fracture analysis of rat femora in compression test," *Journal of Biomechanics*, 2013.

[152] E. Verhulp, B. V. Rietbergen, and R. Huiskes, "A three-dimensional digital image correlation technique for strain measurements in microstructures," *Journal of Biomechanics*, 2004.

[153] P. Sztefek, M. Vanleene, R. Olsson, R. Collinson, A. A. Pitsillides, and S. Shefelbine, "Using digital image correlation to determine bone surface strains during loading and after adaptation of the mouse tibia," *Journal of Biomechanics*, 2010.

[154] F. M. Colacino, R. Emiliano, and B. R. Mace, "Subject-specific musculoskeletal parameters of wrist flexors and extensors estimated by an EMG-driven musculoskeletal model," *Medical Engineering and Physics*, 2012.

[155] D. G. Lloyd and T. F. Besier, "An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo," *Journal of Biomechanics*, 2003.

[156] P. Pulkkinen, C. C. Glüer, and T. Jämsä, "Investigation of differences between hip fracture types: A worthy strategy for improved risk assessment and fracture prevention," 2011.

[157] J. D. Orr, L. K. Dawson, E. J. Garcia, and K. L. Kirk, "Incidence of osteochondral lesions of the talus in the United States Military," *Foot & Ankle International*, 2011.

[158] M. Reilingh, C. van Bergen, and C. van Dijk, "Diagnosis and treatment of osteochondral defects of the ankle," *South African Orthopaedic Journal*, 2009.

[159] N. Sarkalkan, H. Weinans, and A. A. Zadpoor, "Statistical shape and appearance models of bones," 2014.

[160] S. Radzi, M. Uesugi, A. Baird, S. Mishra, M. Schuetz, and B. Schmutz, "Assessing the bilateral geometrical differences of the tibia - are they the same?," *Medical Engineering and Physics*, 2014.

[161] K. Islam, A. Dobbe, A. Komeili, K. Duke, M. El-Rich, S. Dhillon, S. Adeeb, and N. M. Jomha, "Symmetry analysis of talus bone: A geometric morphometric approach.," *Bone & Joint Research*, 2014.

[162] M. van de Giessen, F. M. Vos, C. A. Grimbergen, L. J. van Vliet, and G. J. Streekstra, "An efficient and robust algorithm for parallel groupwise registration of bone surfaces.," *Medical Image Computing and Computer-Assisted Intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2012.

[163] N. Amenta, M. Bern, and M. Kamvysselis, "A new Voronoi-based surface reconstruction algorithm," in *Proceedings of the 25th annual conference on Computer graphics and interactive techniques - SIGGRAPH '98*, 1998.

[164] M. van de Giessen, M. Foumani, G. J. Streekstra, S. D. Strackee, M. Maas, L. J. van Vliet, K. A. Grimbergen, and F. M. Vos, "Statistical descriptions of scaphoid and lunate bone shapes," *Journal of Biomechanics*, 2010.

[165] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, 1995.

[166] M. van de Giessen, N. Smitsman, S. D. Strackee, L. J. van Vliet, K. A. Grimbergen, G. J. Streekstra, and F. M. Vos, "A statistical description of the articulating ulna surface for prosthesis design," in *Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, 2009.

[167] N. K. Mahato, "Morphology of sustentaculum tali: Biomechanical importance and correlation with angular dimensions of the talus," *Foot*, 2011.

[168] J. W. Stone, "Osteochondral lesions of the talar dome," *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*, vol. 4, no. 2, pp. 63–73, 1996.

[169] J. Bruns and B. Rosenbach, "Pressure distribution at the ankle joint.," *Clinical Biomechanics*, 1990.

[170] L. Wan, R. J. de Asla, H. E. Rubash, and G. Li, "Determination of in-vivo articular cartilage contact areas of human talocrural joint under weight-bearing conditions," *Osteoarthritis and Cartilage*, 2006.

[171] F. Corazza, R. Stagni, V. P. Castelli, and A. Leardini, "Articular contact at the tibiotalar joint in passive flexion," *Journal of Biomechanics*, 2005.

[172] R. D. Ferkel, R. M. Zanotti, G. A. Komenda, N. A. Sgaglione, M. S. Cheng, G. R. Applegate, and R. M. Dopirak, "Arthroscopic treatment of chronic osteochondral lesions of the talus: long-term results.," *The American Journal of Sports Medicine*, 2008.

[173] R. A. W. Verhagen, P. A. A. Struijs, P. M. M. Bossuyt, and C. N. van Dijk, "Systematic review of treatment strategies for osteochondral defects of the talar dome," 2003.

[174] C. R. Lareau, J. T. Bariteau, D. J. Paller, S. C. Koruprolu, and C. W. DiGiovanni, "Contribution of the medial malleolus to tibiotalar joint contact characteristics," *Foot and Ankle Specialist*, 2015.

[175] C. W. Imhauser, S. Siegler, J. K. Udupa, and J. R. Toy, "Subject-specific models of the hindfoot reveal a relationship between morphology and passive mechanical properties," *Journal of Biomechanics*, 2008.

[176] J. M. Hootman, R. Dick, and J. Agel, "Epidemiology of collegiate injuries for 15 sports: Prevention initiatives," *Journal of Athletic Training*, 2007.

[177] T. Kobayashi and K. Gamada, "Lateral ankle sprain and chronic ankle instability - a critical review," *Foot & Ankle Specialist*, 2014.

[178] M. N. Houston, B. L. Van Lunen, and M. C. Hoch, "Health-related quality of life in individuals with chronic ankle instability," *Journal of Athletic Training*, 2014.

[179] P. A. Gribble, C. M. Bleakley, B. M. Caulfield, C. L. Docherty, F. Fourchet, D. T. P. Fong, J. Hertel, C. E. Hiller, T. W. Kaminski, P. O. McKeon, K. M. Refshauge, E. A. Verhagen, B. T. Vicenzino, E. A. Wikstrom, and E. Delahunt, "2016 consensus statement of the International Ankle Consortium: Prevalence, impact and long-term consequences of lateral ankle sprains," *British Journal of Sports Medicine*, 2016.

[180] B. R. Waterman, P. J. Belmont, K. L. Cameron, T. M. DeBerardino, and B. D. Owens, "Epidemiology of ankle sprain at the United States Military Academy," *The American Journal of Sports Medicine*, 2010.

[181] P. Gribble, E. Delahunt, C. Bleakley, B. Caulfield, C. Docherty, D.-P. Fong, F. Fourchet, J. Hertel, C. Hiller, T. Kaminski, P. McKeon, K. Refshauge, P. van der Wees, W. Vicenzino, and E. Wikstrom, "Selection criteria for patients with chronic ankle instability in controlled research: A position statement of the international ankle consortium," *Journal of Athletic Training*, 2014.

[182] E. A. Wikstrom, T. Hubbard-Turner, and P. O. McKeon, "Understanding and treating lateral ankle sprains and their consequences: A constraints-based approach," *Sports Medicine*, 2013.

[183] T. J. Hubbard, J. Hertel, and P. Sherbondy, "Fibular position in individuals with self-reported chronic ankle instability," *Journal of Orthopaedic & Sports Physical Therapy*, 2006.

[184] K. E. Morrison and T. W. Kaminski, "Foot characteristics in association with inversion ankle injury," 2007.

[185] A. Frigg, O. Magerkurth, V. Valderrabano, H. P. Ledermann, and B. Hintermann, "The effect of osseous ankle configuration on chronic ankle instability," *British Journal of Sports Medicine*, 2007.

[186] O. Hershkovich, S. Tenenbaum, B. Gordon, N. Bruck, R. Thein, E. Derazne, D. Tzur, A. Shamiss, and A. Afek, "A large-scale study on epidemiology and risk factors for chronic ankle instability in young adults," *The Journal of Foot and Ankle Surgery*, vol. 54, no. 2, pp. 183–187, 2015.

[187] A. Anandacoomarasamy and L. Barnsley, "Long term outcomes of inversion ankle injuries," *British journal of sports medicine*, vol. 39, no. 3, pp. e14–e14, 2005.

[188] E. Delahunt, G. F. Coughlan, B. Caulfield, E. J. Nightingale, C.-W. C. Lin, and C. E. Hiller, "Inclusion criteria when investigating insufficiencies in chronic ankle instability," *Medicine & Science in Sports & Exercise*, vol. 42, no. 11, pp. 2106–2121, 2010.

[189] J. W. Peters, S. G. Trevino, and P. A. Renstrom, "Chronic lateral ankle instability," *Foot & ankle*, vol. 12, no. 3, pp. 182–191, 1991.

[190] Y. Tourné, J. L. Besse, and C. Mabit, "Chronic ankle instability. which tests to assess the lesions? which therapeutic options?," 2010.

[191] J. Hertel, "Functional anatomy, pathomechanics, and pathophysiology of lateral ankle instability," 2002.

[192] H. Pihlajamäki, K. Hietaniemi, M. Paavola, T. Visuri, and V. M. Mattila, "Surgical versus functional treatment for acute ruptures of the lateral ligament complex of the ankle in young men: a randomized controlled trial," *JBJS*, vol. 92, no. 14, pp. 2367–2374, 2010.

[193] A. C. Pijnenburg, C. N. Van Dijk, P. M. Bossuyt, and R. K. Marti, "Treatment of ruptures of the lateral ankle ligaments: A meta-analysis," *Journal of Bone and Joint Surgery - Series A*, 2000.

[194] M. Lee, J. W. Kwon, W. J. Choi, and J. W. Lee, "Comparison of outcomes for osteochondral lesions of the talus with and without chronic lateral ankle instability," *Foot & Ankle International*, 2015.

[195] T. J. Hubbard, L. C. Kramer, C. R. Denegar, and J. Hertel, "Correlations among multiple measures of functional and mechanical instability in subjects with chronic ankle instability," *Journal of Athletic Training*, 2007.

[196] B. Hintermann, A. Boss, and D. Schafer, "Arthroscopic findings in patients with chronic ankle instability," *American Journal of Sports Medicine*, 2002.

[197] M. J. Berkowitz and D. H. Kim, "Fibular position in relation to lateral ankle instability," 2004.

[198] C. T. LeBrun and J. O. Krause, "Variations in mortise anatomy," 2005.

[199] E. Larsen and P. Angermann, "Association of ankle instability and foot deformity," *Acta Orthopaedica*, 1990.

[200] A. B. van Bergeyk, A. van Younger, and B. van Carson, "CT analysis of hindfoot alignment in chronic lateral ankle instability," *Foot & Ankle International*, vol. 23, no. 1, pp. 37–42, 2002.

[201] R. P. Kleipool and L. Blankevoort, "The relation between geometry and function of the ankle joint complex: A biomechanical review," 2010.

[202] E. Barbaix, P. van Roy, and J. P. Clarys, "Variations of anatomical elements contributing to subtalar joint stability: Intrinsic risk factors for posttraumatic lateral instability of the ankle?," *Ergonomics*, 2000.

[203] N. Tümer, L. Blankevoort, M. van de Giessen, M. P. Terra, P. A. de Jong, H. Weinans, G. J. Tuijthof, and A. A. Zadpoor, "Bone shape difference between control and osteochondral defect groups of the ankle joint," *Osteoarthritis and Cartilage*, 2016.

[204] S. Sarrafian and A. Kelikian, "Functional anatomy of the foot and ankle.," in *Sarrafian's anatomy of the foot and ankle: descriptive, topographical, functional*, 2011.

[205] E. Maceira and M. Monteagudo, "Subtalar anatomy and mechanics," 2015.

[206] B. M. Auerbach and C. B. Ruff, "Limb bone bilateral asymmetry: Variability and commonality among modern humans," *Journal of Human Evolution*, 2006.

[207] J. Dargel, J. Feiser, M. Gotter, D. Pennig, and J. Koebke, "Side differences in the anatomy of human knee joints.," *Knee Surgery, Sports Traumatology, Arthroscopy*, 2009.

[208] E. Y. Young, J. Gebhart, D. Cooperman, and N. U. Ahn, "Are the left and right proximal femurs symmetric? basic research," *Clinical Orthopaedics and Related Research*, 2013.

[209] D. G. Eckhoff, D. J. Jacofsky, B. D. Springer, M. Dunbar, J. J. Cherian, R. K. Elmallah, M. A. Mont, and K. A. Greene, "Bilateral symmetrical comparison of femoral and tibial anatomic features," *Journal of Arthroplasty*, 2016.

[210] M. A. Pierre, D. Zurakowski, A. Nazarian, D. A. Hauser-Kara, and B. D. Snyder, "Assessment of the bilateral asymmetry of human femurs based

on physical, densitometric, and structural rigidity characteristics," *Journal of Biomechanics*, 2010.

[211] L. Cristofolini, M. Baleani, E. Schileo, S. van Sint Jan, M. M. Juszczyk, C. Öhman, I. Zwierzak, P. Lefèvre, J. M. Juszczyk, and M. Viceconti, "Differences between contralateral bones of the human lower limbs: a multiscale investigation," *Journal of Mechanics in Medicine and Biology*, 2014.

[212] J. G. G. Dobbe, K. J. Du Pré, P. Kloen, L. Blankevoort, and G. J. Streekstra, "Computer-assisted and patient-specific 3D planning and evaluation of a single-cut rotational osteotomy for complex long-bone deformities," *Medical and Biological Engineering and Computing*, 2011.

[213] D. Santoro, S. Tantavisut, D. Aloj, and M. D. Karam, "Diaphyseal osteotomy after post-traumatic malalignment," 2014.

[214] P. W. Ten Berg, J. G. Dobbe, G. van Wolfswinkel, S. D. Strackee, and G. J. Streekstra, "Validation of the contralateral side as reference for selecting radial head implant sizes," *Surgical and Radiologic Anatomy*, 2016.

[215] S. J. Shultz and A. D. Nguyen, "Bilateral asymmetries in clinical measures of lower-extremity anatomic characteristics," *Clinical Journal of Sport Medicine*, 2007.

[216] H. C. Smith, P. Vacek, R. J. Johnson, J. R. Slauterbeck, J. Hashemi, S. Shultz, and B. D. Beynnon, "Risk factors for anterior cruciate ligament injury: A review of the literature - part 1: Neuromuscular and anatomic risk," *Sports Health*, 2012.

[217] R. Daud, M. R. Abdul Kadir, S. Izman, A. P. Md Saad, M. H. Lee, and A. Che Ahmad, "Three-dimensional morphometric study of the trapezium shape of the trochlea tali," *Journal of Foot and Ankle Surgery*, 2013.

[218] A. R. Palmer, "Fluctuating asymmetry analyses: a primer," in *Developmental instability: its origins and evolutionary implications*, pp. 335–364, Springer, 1994.

[219] P. Claes, M. Walters, M. D. Shriver, D. Puts, G. Gibson, J. Clement, G. Baynam, G. Verbeke, D. Vandermeulen, and P. Suetens, "Sexual dimorphism in multiple aspects of 3D facial symmetry and asymmetry defined by spatially dense geometric morphometrics," *Journal of Anatomy*, 2012.

[220] J. Bellemans, K. Carpentier, H. Vandenneucker, J. Vanlauwe, and J. Victor, "The John Insall award: Both morphotype and gender influence the shape of the knee in patients undergoing TKA," in *Clinical Orthopaedics and Related Research*, 2010.

[221] A. Unnanuntana, P. Toogood, D. Hart, D. Cooperman, and R. E. Grant, "Evaluation of proximal femoral geometry using digital photographs," *Journal of Orthopaedic Research*, vol. 28, no. 11, pp. 1399–1404, 2010.

[222] B. L. Wise, F. Liu, L. Kritikos, J. A. Lynch, N. Parimi, Y. Zhang, and N. E. Lane, "The association of distal femur and proximal tibia shape with sex: The osteoarthritis initiative," *Seminars in Arthritis and Rheumatism*, 2016.

[223] P. Claes, J. Reijniers, M. D. Shriver, J. Snyders, P. Suetens, J. Nielandt, G. De Tré, and D. Vandermeulen, "An investigation of matching symmetry in

the human pinnae with possible implications for 3D ear recognition and sound localization," *Journal of Anatomy*, 2015.

[224] M. M. Cohen, N. D. Vela, J. E. Levine, and E. A. Barnoy, "Validating a new computed tomography atlas for grading ankle osteoarthritis," *Journal of Foot and Ankle Surgery*, 2015.

[225] Y. Pauchard, T. Fitze, D. Browarnik, A. Eskandari, I. Pauchard, W. Enns-Bray, H. Pálsson, S. Sigurdsson, S. J. Ferguson, T. B. Harris, V. Gudnason, and B. Helgason, "Interactive graph-cut segmentation for fast creation of finite element models from clinical CT data for hip fracture prediction," *Computer Methods in Biomechanics and Biomedical Engineering*, 2016.

[226] J. Hamm and D. D. Lee, "Grassmann Discriminant Analysis : a unifying view on subspace-based learning," *Proceedings of the 25th international conference on Machine learning*, 2008.

[227] A. V. Knyazev and M. E. Argentati, *Principal Angles Between Subspaces in An A-Based Scalar Product: Algorithms and Perturbation Estimates*. 2002.

[228] W. Krzanowski, "Between-groups comparison of principal components," *Journal of the American Statistical Association*, vol. 74, no. 367, pp. 703–707, 1979.

[229] R. D. Ledesma, C. Universidad, N. D. Mar, P. Valero-mora, and U. D. Valencia, "Determining the number of factors to retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis," *Practical Assessment, Research & Evaluation*, 2007.

[230] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *Journal of Chiropractic Medicine*, 2016.

[231] V. Gentry and C. Gabbard, "Foot-preference behavior: A developmental perspective," *Journal of General Psychology*, 1995.

[232] J. Bell and C. Gabbard, "Foot preference changes through adulthood," *Laterality*, 2000.

[233] S. M. Harris and D. T. Case, "Sexual dimorphism in the tarsal bones: Implications for sex determination," *Journal of Forensic Sciences*, 2012.

[234] U.-Y. Lee, I.-B. Kim, and D.-S. Kwak, "Sex determination using discriminant analysis of upper and lower extremity bones: New approach using the volume and surface area of digital model," *Forensic Science International*, 2015.

[235] T. Riepert, T. Drechsler, H. Schild, B. Nafe, and R. Mattern, "Estimation of sex on the basis of radiographs of the calcaneus," *Forensic Science International*, 1996.

[236] E. Gualdi-Russo, "Sex determination from the talus and calcaneus measurements," *Forensic Science International*, 2007.

[237] D. Stephan, S. Panzer, M. Göttlinger, and P. Augat, "Analysis of the intra-individual differences of the joint surfaces of the calcaneus," *Computer Methods in Biomechanics and Biomedical Engineering*, 2014.

[238] C. Ruff, B. Holt, and E. Trinkaus, "Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation," 2006.

[239] M. van De Giessen, S. de Raedt, M. Stilling, T. B. Hansen, M. Maas, G. J. Streekstra, L. J. van Vliet, and F. M. Vos, "Localized component analysis for arthritis detection in the trapeziometacarpal joint," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011.

[240] A. Hyvärinen, "Independent component analysis: recent advances.," *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 2013.

[241] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," 2000.

[242] J. Wu, K. G. Brigham, M. A. Simon, and J. C. Brigham, "An implementation of independent component analysis for 3D statistical shape analysis," *Biomedical Signal Processing and Control*, vol. 13, pp. 345–356, 2014.

[243] Q. Zhao, K. Okada, K. Rosenbaum, M. Summar, and M. Linguraru, "Constrained local model with independent component analysis and kernel density estimation: Application to down syndrome detection," in *Proceedings - International Symposium on Biomedical Imaging*, 2015.

[244] A. Heijink, M. Vanhees, K. van den Ende, M. P. van den Bekerom, R. P. van Riet, C. N. van Dijk, and D. Eygendaal, "Biomechanical considerations in the pathogenesis of osteoarthritis of the elbow," 2016.

[245] J. Farr, B. Cole, A. Dhawan, J. Kercher, and S. Sherman, "Clinical cartilage restoration: Evolution and overview," in *Clinical Orthopaedics and Related Research*, 2011.

[246] G. J. M. Tuijthof, A. C. Kok, M. P. Terra, J. F. A. Aaftink, G. J. Streekstra, C. N. van Dijk, and G. M. M. J. Kerkhoffs, "Sensitivity and specificity of ultrasound in detecting (osteo)chondral defects: A cadaveric study," *Ultrasound in Medicine and Biology*, 2013.

[247] I. Moller, D. Bong, E. Naredo, E. Filippucci, I. Carrasco, C. Moragues, and A. Iagnocco, "Ultrasound in the study and monitoring of osteoarthritis," *Osteoarthritis and Cartilage*, 2008.

[248] A. H. Spannow, M. Pfeiffer-Jensen, N. T. Andersen, E. Stenbøg, and T. Herlin, "Inter -and intraobserver variation of ultrasonographic cartilage thickness assessments in small and large joints in healthy children," *Pediatric Rheumatology*, 2009.

[249] C. J. van Bergen, "Diagnosing, planning and evaluating osteochondral ankle defects with imaging modalities," *World Journal of Orthopedics*, 2015.

[250] D. C. Barratt, G. P. Penney, C. S. Chan, M. Slomczykowski, T. J. Carter, P. J. Edwards, and D. J. Hawkes, "Self-calibrating 3D-ultrasound-based bone registration for minimally invasive orthopedic surgery," *IEEE Transactions on Medical Imaging*, 2006.

[251] A. Brounstein, I. Hacihaliloglu, P. Guy, A. Hodgson, and R. Abugharbieh, "Towards real-time 3D US to CT bone image registration using phase and

curvature feature based GMM matching," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011.

[252] I. Hacihaliloglu, A. Brounstein, P. Guy, A. Hodgson, and R. Abugharbieh, "3D ultrasound-CT registration in orthopaedic trauma using GMM registration with optimized particle simulation-based data reduction.," *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2012.

[253] A. Lang, P. Mousavi, S. Gill, G. Fichtinger, and P. Abolmaesumi, "Multi-modal registration of speckle-tracked freehand 3D ultrasound to CT in the lumbar spine," *Medical Image Analysis*, 2012.

[254] M. Moghari and P. Abolmaesumi, "Point-based rigid-body registration using an Unscented Kalman filter," *IEEE Transactions on Medical Imaging*, 2007.

[255] G. P. Penney, D. C. Barratt, C. S. Chan, M. Slomczykowski, T. J. Carter, P. J. Edwards, and D. J. Hawkes, "Cadaver validation of intensity-based ultrasound to CT registration," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2005.

[256] D. M. Muratore, J. H. Russ, B. M. Dawant, and R. L. Galloway, "Three-dimensional image registration of phantom vertebrae for image-guided surgery: a preliminary study.," *Computer Aided Surgery*, 2002.

[257] S. Winter, I. Pechlivanis, C. Dekomien, C. Igel, and K. Schmieder, "Toward registration of 3D ultrasound and CT images of the spine in clinical praxis: design and evaluation of a data acquisition protocol.," *Ultrasound in Medicine & Biology*, 2009.

[258] C. X. Yan, B. Goulet, J. Pelletier, S. J. S. Chen, D. Tampieri, and D. L. Collins, "Towards accurate, robust and practical ultrasound-CT registration of vertebrae for image-guided spine surgery," *International Journal of Computer Assisted Radiology and Surgery*, 2011.

[259] B. Brendel, S. Winter, A. Rick, M. Stockheim, and H. Ermert, "Bone registration with 3D CT and ultrasound data sets," *International Congress Series*, 2003.

[260] S. Nagpal, P. Abolmaesumi, A. Rasoulian, I. Hacihaliloglu, T. Ungi, J. Osborn, V. A. Lessoway, J. Rudan, M. Jaeger, R. N. Rohling, *et al.*, "A multi-vertebrae CT to US registration of the lumbar spine in clinical data," *International Journal of Computer Assisted Radiology and Surgery*, vol. 10, no. 9, pp. 1371–1381, 2015.

[261] A. Rasoulian, P. Mousavi, M. H. Moghari, P. Foroughi, and P. Abolmaesumi, "Group-wise feature-based registration of CT and ultrasound images of spine," in *Medical Imaging 2010: Visualization, Image-Guided Procedures, and Modeling*, vol. 7625, p. 76250R, International Society for Optics and Photonics, 2010.

[262] D. V. Amin, T. Kanade, A. M. Digioia, and B. Jaramaz, "Ultrasound registration of the bone surface for surgical navigation," *Computer Aided Surgery*, 2003.

[263] P. Besl and N. McKay, "A method for registration of 3D shapes," 1992.

[264] S. Winter, B. Brendel, I. Pechlivanis, K. Schmieder, and C. Igel, "Registration of CT and intraoperative 3D ultrasound images of the spine using evolutionary and gradient-based methods," *IEEE Transactions on Evolutionary Computation*, 2008.

[265] C. J. A. van Bergen, G. J. M. Tuijthof, L. Blankevoort, M. Maas, G. M. M. J. Kerkhoffs, and C. N. van Dijk, "Computed tomography of the ankle in full plantar flexion: A reliable method for preoperative planning of arthroscopic access to osteochondral defects of the talus," *Arthroscopy - Journal of Arthroscopic and Related Surgery*, 2012.

[266] I. Hacihaliloglu, A. Rasoulian, R. N. Rohling, and P. Abolmaesumi, "Local phase tensor features for 3D ultrasound to statistical shape+pose spine model registration," *IEEE Transactions on Medical Imaging*, 2014.

[267] C. Askeland, O. V. Solberg, J. B. L. Bakeng, I. Reinertsen, G. A. Tangen, E. F. Hofstad, D. H. Iversen, C. Våpenstad, T. Selbekk, T. Langø, T. A. Hernes, H. Olav Leira, G. Unsgård, and F. Lindseth, "CustusX: an open-source research platform for image-guided therapy," *International Journal of Computer Assisted Radiology and Surgery*, 2016.

[268] L. E. Bø, E. F. Hofstad, F. Lindseth, and T. A. Hernes, "Versatile robotic probe calibration for position tracking in ultrasound imaging," *Physics in Medicine and Biology*, 2015.

[269] E. M. A. Anas, A. Seitel, A. Rasoulian, P. S. John, T. Ungi, A. Lasso, K. Darras, D. Wilson, V. A. Lessoway, G. Fichtinger, M. Zec, D. Pichora, P. Mousavi, R. Rohling, and P. Abolmaesumi, "Registration of a statistical model to intraoperative ultrasound for scaphoid screw fixation," *International Journal of Computer Assisted Radiology and Surgery*, 2016.

[270] A. Belaid, D. Boukerroui, Y. Maingourd, and J.-F. Lerallut, "Phase-based level set segmentation of ultrasound images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 1, pp. 138–147, 2011.

[271] I. Hacihaliloglu, R. Abugharbieh, A. J. Hodgson, and R. N. Rohling, "Bone surface localization in ultrasound using image phase-based features," *Ultrasound in Medicine and Biology*, 2009.

[272] S. Shojaeilangari, W.-Y. Yau, J. Li, and E.-K. Teoh, "Multiscale analysis of local phase and local orientation for dynamic facial expression recognition," *Journal ISSN*, vol. 1, no. 1, 2014.

[273] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transactions on Signal Processing*, 2001.

[274] E. M. A. Anas, A. Seitel, A. Rasoulian, P. S. John, D. Pichora, K. Darras, D. Wilson, V. A. Lessoway, I. Hacihaliloglu, P. Mousavi, R. Rohling, and P. Abolmaesumi, "Bone enhancement in ultrasound using local spectrum variations for guiding percutaneous scaphoid fracture fixation pro-

cedures," *International Journal of Computer Assisted Radiology and Surgery*, 2015.

[275] N. Hansen and A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies.," 2001.

[276] G. C. Kagadis, K. K. Delibasis, G. K. Matsopoulos, N. A. Mouravliansky, P. A. Asvestas, and G. C. Nikiforidis, "A comparative study of surface- and volume-based techniques for the automatic registration between CT and SPECT brain images," *Medical Physics*, 2002.

[277] C. J. van Bergen, G. J. Tuijthof, M. Maas, I. N. Sierevelt, and C. N. van Dijk, "Arthroscopic accessibility of the talus quantified by computed tomography simulation," *American Journal of Sports Medicine*, 2012.

[278] I. M. Khan, S. J. Gilbert, S. K. Singhrao, V. C. Duance, and C. W. Archer, "Cartilage integration: Evaluation of the reasons for failure of integration during cartilage repair. a review," 2008.

[279] J. A. Buckwalter, "Articular cartilage injuries," in *Clinical Orthopaedics and Related Research*, 2002.

[280] R. Bitton, "The economic burden of osteoarthritis.," *The American Journal of Managed Care*, 2009.

[281] H. J. Nieminen, Y. Zheng, S. Saarakkala, Q. Wang, J. Toyras, Y. Huang, and J. Jurvelin, "Quantitative assessment of articular cartilage using high-frequency ultrasound: Research findings and diagnostic prospects," *Critical Reviews in Biomedical Engineering*, vol. 37, no. 6, 2009.

[282] H. I. Keen and P. G. Conaghan, "Ultrasonography in osteoarthritis," *Radiologic Clinics*, vol. 47, no. 4, pp. 581–594, 2009.

[283] V. Gouttebarge and M. H. Frings-Dresen, "Ankle osteoarthritis in former elite football players: What do we know?," in *The Ankle in Football*, pp. 311–322, Springer, 2014.

[284] R. S. Schechter, H. H. Chaskelis, R. B. Mignogna, and P. P. Delsanto, "Real-time parallel computation and visualization of ultrasonic pulses in solids," *Science*, 1994.

[285] R. Stagni, A. Leardini, A. Ensini, and A. Cappello, "Ankle morphometry evaluated using a new semi-automated technique based on x-ray pictures," *Clinical Biomechanics*, 2005.

[286] P. A. Dowdy, B. V. Watson, A. Amendola, and J. D. Brown, "Noninvasive ankle distraction: Relationship between force, magnitude of distraction, and nerve conduction abnormalities," *Arthroscopy - Journal of Arthroscopic and Related Surgery*, 1996.

[287] C. Niek Van Dijk, R. A. W. Verhagen, and H. J. L. Tol, "Resterilizable non-invasive ankle distraction device," *Arthroscopy - Journal of Arthroscopic and Related Surgery*, 2001.

[288] D. White, J. Evans, J. Truscott, and R. Chivers, "Simulation of ultrasound in the knee," in *Journal of Physics: Conference Series*, vol. 1, p. 231, IOP Publishing, 2004.

[289] B. Goker, E. Gonen, M. D. Demirag, and J. A. Block, "Quantification of the radiographic joint space width of the ankle," *Clinical Orthopaedics and Related Research*, 2009.

[290] W. Hembree, J. Wittstein, E. Vinson, R. Queen, C. LaRose, K. Singh, and M. Easley, "Magnetic resonance imaging features of osteochondral lesions of the talus," *Foot & Ankle . . .*, 2012.

[291] J. D. Orr, J. R. Dutton, and J. T. Fowler, "Anatomic location and morphology of symptomatic, operatively treated osteochondral lesions of the talus," *Foot & Ankle International*, 2012.

[292] E. Kaleva, J. Liukkonen, J. Toyras, S. Saarakkala, P. Kiviranta, and J. Jurvelin, "2D finite difference time domain model of ultrasound reflection from normal and osteoarthritic human articular cartilage surface.," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2010.

[293] E. Bossy, M. Talmant, and P. Laugier, "Effect of bone cortical thickness on velocity measurements using ultrasonic axial transmission: A 2D simulation study," *The Journal of the Acoustical Society of America*, 2002.

[294] S. P. Dodd, A. W. Miles, S. Gheduzzi, V. F. Humphrey, and J. L. Cunningham, "Modelling the effects of different fracture geometries and healing stages on ultrasound signal loss across a long bone fracture," *Computer Methods in Biomechanics and Biomedical Engineering*, 2007.

[295] V. C. Protopappas, D. I. Fotiadis, and K. N. Malizos, "Guided ultrasound wave propagation in intact and healing long bones," *Ultrasound in Medicine and Biology*, 2006.

[296] D. White, J. A. Evans, J. G. Truscott, and R. A. Chivers, "Modelling the propagation of ultrasound in the joint space of a human knee," *Ultrasound in Medicine & Biology*, vol. 36, no. 10, pp. 1736–1745, 2010.

[297] D. White, J. a. Evans, J. G. Truscott, and R. a. Chivers, "Can ultrasound propagate in the joint space of a human knee?," *Ultrasound in Medicine & Biology*, 2007.

[298] D. Testi, M. Viceconti, F. Baruffaldi, and A. Cappello, "Risk of fracture in elderly patients: A new predictive index based on bone mineral density and finite element analysis," *Computer Methods and Programs in Biomedicine*, 1999.

[299] L. Yang, N. Peel, J. A. Clowes, E. V. McCloskey, and R. Eastell, "Use of DXA-based structural engineering models of the proximal femur to discriminate hip fracture," *Journal of Bone and Mineral Research*, 2009.

[300] Y. Luo, Z. Ferdous, and W. D. Leslie, "A preliminary dual-energy x-ray absorptiometry-based finite element model for assessing osteoporotic hip fracture risk," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 2011.

[301] C. M. Langton, S. Pisharody, and J. H. Keyak, "Generation of a 3D proximal femur shape from a single projection 2D radiographic image," *Osteoporosis International*, 2009.

[302] S. P. Väänänen, L. Grassi, G. Flivik, J. S. Jurvelin, and H. Isaksson, "Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image," *Medical Image Analysis*, 2015.

[303] H. Lamecker, M. Seebaß, H.-C. Hege, and P. Deuflhard, "A 3D statistical shape model of the pelvic bone for segmentation," in *Medical Imaging 2004: Image Processing*, vol. 5370, pp. 1341–1352, International Society for Optics and Photonics, 2004.

[304] P. Golland, W. E. L. Grimson, M. E. Shenton, and R. Kikinis, "Detection and analysis of statistical differences in anatomical shape," *Medical Image Analysis*, 2005.

[305] J. M. Sotoca, J. M. Iñesta, and M. A. Belmonte, "Hand bone segmentation in radioabsorptiometry images for computerised bone mass assessment," *Computerized Medical Imaging and Graphics*, 2003.

[306] O. van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or, "A survey on shape correspondence," *Eurographics Symposium on Geometry Processing*, 2011.

[307] Z. Yaniv, "Rigid registration," in *Image-Guided Interventions*, pp. 159–192, Springer, 2008.

[308] K. Josephson, A. Ericsson, and J. Karlsson, "Segmentation of medical images using three-dimensional active shape models," in *Scandinavian Conference on Image Analysis*, pp. 719–728, Springer, 2005.

[309] R. Szeliski and S. Lavallée, "Matching 3D anatomical surfaces with non-rigid deformations using octree-splines," *International Journal of Computer Vision*, 1996.

[310] L. Ferrarini, H. Olofsen, W. M. Palm, M. A. van Buchem, J. H. Reiber, and F. Admiraal-Behloul, "GAMEs: Growing and adaptive meshes for fully automatic shape modeling and analysis," *Medical Image Analysis*, 2007.

[311] A. F. Frangi, D. Ruecker, t, J. A. Schnabel, and W. J. Niessen, "Automatic 3D asm construction via atlas-based landmarking and volumetric elastic registration," in *Information Processing in Medical Imaging*, 2001.

[312] D. Rueckert, A. F. Frangi, and J. A. Schnabel, "Automatic construction of 3D statistical deformation models using non-rigid registration," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2001.

[313] M. Zöckler, D. Stalling, and H. C. Hege, "Fast and intuitive generation of geometric shape transitions," *Visual Computer*, 2000.

[314] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," *IEEE Transactions on Medical Imaging*, 2002.

[315] J. Fripp, S. Crozier, S. Ourselin, and S. Warfield, "Automatic initialisation of 3D deformable models for cartilage segmentation," in *Proceedings of the Digital Imaging Computing: Techniques and Applications, DICTA 2005*, 2005.

[316] R. Pilgram and C. Walch, "Knowledge-based femur detection in conventional radiographs of the pelvis," *Computers in Biology and Medicine*, 2008.

[317] R. H. Davies, C. J. Twining, P. D. Allen, T. F. Cootes, and C. J. Taylor, "Shape discrimination in the hippocampus using an MDL model.," *Information processing in medical imaging : proceedings of the ... conference*, 2003.

[318] T. Heimann, I. Oguz, and I. Wolf, "Implementing the automatic generation of 3D statistical shape models with ITK," *Open Science Workshop . . .* , 2006.

[319] N. Boukala, E. Favier, B. Laget, and P. Radeva, "Active shape model based segmentation of bone structures in hip radiographs," in *Industrial Technology, 2004. IEEE ICIT'04. 2004 IEEE International Conference on*, vol. 3, pp. 1682–1687, IEEE, 2004.

[320] I. M. Scott, T. F. Cootes, and C. J. Taylor, "Improving appearance model matching using local image structure.," *Information processing in medical imaging : proceedings of the ... conference*, 2003.

[321] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[322] H. Seim, D. Kainmüeller, and M. Heller, "Automatic segmentation of the pelvic bones from CT data based on a statistical shape model.," *Eurographics Workshop on Visual Computing for Biomedicine*, 2008.

[323] T. F. Cootes and C. J. Taylor, "Active shape model search using local grey-level models: A quantitative evaluation," *Procedings of the British Machine Vision Conference 1993*, 1993.

[324] T. F. Cootes and C. J. Taylor, "Using grey-level models to improve active shape model search," in *Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on*, vol. 1, pp. 63–67, IEEE, 1994.

[325] S. Ho and G. Gerig, "Scale-space on image profiles about an object boundary," in *International Conference on Scale-Space Theories in Computer Vision*, pp. 564–575, Springer, 2003.

[326] M. van de Giessen, *Statistical Modeling of Shape and Motion of the Wrist Bones*. PhD thesis, TU Delft, Delft University of Technology, 2011.

[327] S. B. Franklin, D. J. Gibson, P. A. Robertson, J. T. Pohlmann, and J. S. Fralish, "Parallel Analysis: a method for determining significant principal components," *Journal of Vegetation Science*, 1995.

[328] P. Kovesi, "Image features from phase congruency," *Videre*, 1999.

Acknowledgement

I still remember the first time I was in the TU Delft campus. It was 8 years ago and I had a chance to talk with Dick Plettenburg and Amir Zadpoor about both the Master's Programme in Biomedical Engineering and any possible PhD continuation. After my visit, I wholeheartedly decided to apply for the MSc programme at TU Delft. At the time, I did not know that I would, one day, be accepted as a master student and continue working as a PhD afterwards. After all these years, I cherish my experience at TU Delft and the possibility to meet many amazing lecturers, researchers, colleagues and friends. Herewith, I would like to express my gratitude to all of them.

Amir Zadpoor. I have been fortunate enough to complete my MSc and PhD work under your supervision. Your leadership style, way of motivating people and energy for research always amazed me, and assisted me in completing the whole PhD trajectory. Thank you for giving me freedom within my research and yet again ensuring that I stay on course. Besides the academic coaching, you also took time to be my life coach on a few occasions, for which I am thankful.

Harrie Weinans. I am grateful for your continuous support. Thank you for introducing me to clinicians and researchers at the Utrecht Medical Center (UMC), and for building a collaborative team environment. I am impressed by your drive to perform research in biomechanics, your knowledge in this field and, especially, your memory skills in keeping names of researchers and their studies in your mind.

Gabrielle Tuijthof. I would like to thank you for your supervision and for providing me the opportunity to familiarize myself with the clinical setting at the Academic Medical Center (AMC). I am grateful that you often encouraged me to attend conferences with full of orthopaedic surgeons. Besides all, I will always fondly remember the midnight that we had to push your car to get it working, after working long tiring hours together in the echo room.

Jenny Dankelman, Jaap Harlaar, Bill Taylor and Rintje Agricola. I would like to thank you all for being in my dissertation committee, for your time and effort to read my thesis.

Leendert Blankevoort. Your contribution to the studies (Chapters 4 and 5) and our discussions on many other different topics are very valuable to me. Thank you for all the time you spent with me and also for teaching me Dutch expressions.

Pim de Jong and Gino Kerkhoff. It was great pleasure to meet both of you. Your door was always open whenever I needed help or had questions about clinical issues. Without your input and critical feedback, it would be hard to complete Chapters 4-6 of this thesis. My special thanks to Gino for providing me the opportunity to be one day in an operating room and to watch surgeons perform surgeries.

Maaike Terra and Geert Streekstra. Thank you for all your help and support whenever we needed to collect US sweeps and CT scans of cadaveric ankles. It was quite enjoyable to work with you.

Koen van Dongen and Frans Vos. I am grateful to you for generously sharing with me your time and expertise. I cannot thank you enough for always being there to help me.

Martijn van de Giessen. Your previous work on statistical modelling of wrist bones enabled the studies presented in Chapters 4-6 of this thesis. Your collaboration was truly valuable to me. Thank you for your assistance and cooperation.

Koen Bos, Erwin Waarsing and Willem Paul Gielis. I would like to thank you all for your valuable contribution to Chapter 3 (Koen and Erwin) and Chapter 6 (Willem) of this thesis.

Aimee Kok and Gwen Vuurberg. You both are amazing collaborators and friends. I was so lucky to have a chance to meet and work with you. I would like to express here all my respect to your hard work. Aimee, I am so honoured to be your motherboard queen…and Gwen, thank you for all your support and checking my summary translated into Dutch.

Christian Askeland and Lars Eirik Bø. I don't know how many times I contacted you to calibrate the US probe and have the optical tracking system up and running at AMC. You were always helpful and supportive, for which I am grateful.

Thomas Langø, Sebastien Muller, Ole Vegard Solberg, Geir Arne Tangen, Erik Smistad, Merete Rørvik, Nina Vanvik Hansen, Javier Perez de Frutos, Jeanne and many others I met during my visit to SINTEF Digital, Health Research. You are all wonderful people! I was happy to be in Norway with you for those two months. Thank you for making my stay highly productive and enjoyable.

Arno Volker. I was most impressed by your in depth knowledge in the field of acoustics and your constructive criticism throughout. Your guidance has been invaluable to me.

Henry den Bok and Reinier van Antwerpen. It would be hard for me to successfully perform experiments (Henry, Chapter 8) and collect the dataset (Reinier, Chapter 9) without your support, for which I am grateful. My special thanks to Reinier for helping me with the transportation of equipment between TU Delft and AMC. I hope soon that I will get my Dutch driving license too…

Mirjam Birhuizen. I would like to thank you for dealing with all the administrative work.

Mascha Toppenberg. I truly appreciate your time and patience in assisting me and many other PhDs at 3mE faculty (TU Delft) to proceed all administrative work of the graduate school. Thank you for everything.

TU Delft: Anouk de Goede, Sabrina Rodrigues, Nancy Kouters, Hanneke Hustinx, Angelique Timmerman, Adinda Kohler, and AMC: Rosalie van de Sant. I would like to thank you all. You were always there to help me organize things that I could not be able to do alone.

Lidy Fratila Apachitei, Iulian Apachitei and Jie Zhou. Since the beginning of my master study at TU Delft, I've always felt your positive energy and trust in me. Thank you for all the motivation instilled in me and all nice talks we had.

Arjo Loeve. I've never seen another multitasking person like you in my life. Answering emails, helping me, checking students reports, talking with colleagues and preparing innovative sandwiches (strawberries placed between two slices of bread ??!) at the same time... All successfully accomplished! Thank you for everything.

Thank you **Tim Horeman**, I've always enjoyed our talks. **Marjon Stijntjes and Kerrie Lee Colman**, it was nice working with you all at different stages of the project I was involved in. I wish we could continue working together. **Elise Buiter**, I always remember you whenever I need to decide between a stairway (I know so well, you would be forcing me to use it!) and an elevator. I miss the time we spent together when you were here at TU Delft. **Kim Opdam and Jari Dahmen**, it was nice meeting you both at the AMC. Thank you for the positive atmosphere you created.

To my old roommates and colleagues: **Parisa Moshtagh, Gianni Campoli, Saber Yavari, Mohammad Ahmadi, Reza Hedayati, Budi Arifvianto, Vahid Arbabi and Helene Clogenson.** We had many memories together when I think back. I've enjoyed every moment. Thank you for everything.

Thanks to my long-time colleagues **Shahram Janbaz and Sander Leeflang** for the great time we had. Your personalities, humour and special requests (baklava or chocolate or both at the same time is better!) made my days so enjoyable. **Teunis van Manen**, you were quite often exposed to our talk for long hours and you've never complained. Thank you for your patience.

To my (relatively) new roommates: **Sebastien Callens and Eline Kolken.** I don't know how to keep my beach body (!) when you both are around me. Thank you for keeping the room vivid and my special thanks for your feedback on the cover page design of this thesis. **Ingmar van Hengel**, it was so nice sharing with you the same office, although it was for a short period of time. I've had fun talking with you.

Mahya Ganjian, Mahdiyeh Nouri, Sara Amiri, Mohammad Mirzaali and Khashayar Modaresifar. Thank you for always being so thoughtful, friendly and supportive. I am grateful to be surrounded by you. **Françoise Bobbert**, thank you for organizations that you carried enthusiastically. **Yageng Li, Marike Fokker, Gerwin Smit, Costanza Culmone, Juan Lopez, Marta Scali and Aimee Sakes**, thank you for all nice talks.

and my warm welcome to new researchers, PhDs: **Michelle Minneboo, Livia Angeloni, Helda Pahlavani, Jiahui Dong, Niko Putra, Mauricio Cruz.**

Last but not least, I would like to thank: my mother **Duygu Çalın**, my aunt and her family **Dilek, Melih, Mert and Aykut Aker**, and my parents in law **Tülay and Erdem Tümer**. Your support and belief in me were of great importance. My best friend and husband, **Can Tümer**, thank you for your patience, care and love. **Lokum and Enrico Tümer**, our lovely four legged boys, thank you for always being available as stress balls. My special thanks to our little brother, **Kaan Terün**, for translating my summary into Dutch.

About the Author

Nazlı Tümer was born in 1986 in Istanbul, Turkey. In 2010, she received her bachelor degree in Mechanical Engineering from Istanbul Technical University (ITU). In 2011, she completed a two-year master programme in Solid Mechanics at ITU. During her studies, she worked in the Strength of Materials Laboratory (Mechanical Engineering Faculty, ITU) as a teaching assistant for a three years period. At the end of her master studies in Turkey, she moved to Delft in 2011 and obtained her second master's degree in Biomechanical Engineering from Delft University of Technology (TU Delft) in 2013. Subsequently, she joined the Vibrant Vision Project as a PhD student, which aims to develop an ultrasound based technique that detects osteochondral defects at the earliest stage. Her research, from which this thesis was adapted, was conducted in close collaboration with the Academic Medical Center (AMC) and the Utrecht Medical Center (UMC). At the moment, she holds a post-doctoral position at the Department of Biomechanical Engineering (TU Delft), and continues research under the same project.

List of Publications

Peer Reviewed Journal Publications

- **N. Tümer**, V. Arbabi, W.P. Gielis, P.A. de Jong, H. Weinans, G.J.M. Tuijthof, A.A. Zadpoor, Three-dimensional analysis of shape variations and symmetry of the fibula, tibia, calcaneus and talus, J. Anat. (2018).
- **N. Tümer**, A. Kok, F. Vos, G. Streekstra, C. Askeland, G. Tuijthof, A. Zadpoor, Three-dimensional registration of freehand-tracked ultrasound to CT images of the talocrural joint, Sensors. 18 (2018) 2375.
- R. Hedayati, S.M. Ahmadi, K. Lietaert, **N. Tümer**, Y. Li, S. Amin Yavari, A.A. Zadpoor, Fatigue and quasi-static mechanical behavior of bio-degradable porous biomaterials based on magnesium alloys, J. Biomed. Mater. Res. Part A. 106 (2018) 1798–1811.
- S.M. Ahmadi, R. Hedayati, Y. Li, K. Lietaert, **N. Tümer**, A. Fatemi, C.D. Rans, B. Pouran, H. Weinans, A.A. Zadpoor, Fatigue performance of additively manufactured meta-biomaterials: The effects of topology and material type, Acta Biomater. 65 (2018) 292–304.
- S.M. Ahmadi, R. Kumar, E. V Borisov, R. Petrov, S. Leeflang, Y. Li, **N. Tümer**, R. Huizenga, C. Ayas, A.A. Zadpoor, From microstructural design to surface engineering: a tailored approach for improving fatigue life of additively manufactured meta-biomaterials, Acta Biomater. (2018).
- M.J. Mirzaali, I.C.P. Van Dongen, **N. Tümer**, H. Weinans, S.A. Yavari, A.A. Zadpoor, In-silico quest for bactericidal but non-cytotoxic nanopatterns, Nanotechnology. 29 (2018).
- Y. Li, J. Zhou, P. Pavanram, M.A. Leeflang, L.I. Fockaert, B. Pouran, **N. Tümer**, K.-U. Schröder, J.M.C. Mol, H. Weinans, Additively manufactured biodegradable porous magnesium, Acta Biomater. 67 (2018) 378–392.
- **N. Tümer**, L. Blankevoort, M. van de Giessen, M.P. Terra, P.A. de Jong, H. Weinans, G.J.M. Tuijthof, A.A. Zadpoor, Bone shape difference between control and osteochondral defect groups of the ankle joint, Osteoarthr. Cartil. 24 (2016) 2108–2115.
- **N. Sarkalkan**, A.J. Loeve, K.W.A. van Dongen, G.J.M. Tuijthof, A.A. Zadpoor, A novel ultrasound technique for detection of osteochondral defects in the ankle joint: A parametric and feasibility study, Sensors. 15 (2014) 148–165.
- **N. Sarkalkan**, J.H. Waarsing, P.K. Bos, H. Weinans, A.A. Zadpoor, Statistical shape and appearance models for fast and automated estimation of proximal femur fracture load using 2D finite element models, J. Biomech. 47 (2014) 3107–3114.
- **N. Sarkalkan**, H. Weinans, A.A. Zadpoor, Statistical shape and appearance models of bones, Bone. 60 (2014).

- H. Özben, L. Eralp, G. Baysal, A. Cort, **N. Sarkalkan**, T. Özben, Cisplatin loaded PMMA: mechanical properties, surface analysis and effects on Saos-2 cell culture., Acta Orthop. Traumatol. Turc. 47 (2013) 184–192.

- O. Hapa, H. Çakici, A. Kükner, H. Aygün, **N. Sarkalkan**, G. Baysal, Effect of platelet-rich plasma on tendon-to-bone healing after rotator cuff repair in rats: An in vivo experimental study, Acta Orthop. Traumatol. Turc. 46 (2012).

- O. Hapa, F.A. Barber, E. Sünbüloğlu, Y. Kocabey, **N. Sarkalkan**, G. Baysal, Tendon-grasping strength of various suture configurations for rotator cuff repair, Knee Surgery, Sport. Traumatol. Arthrosc. 19 (2011).

Submitted Journal Papers

- **N. Tümer**, G. Vuurberg, L. Blankevoort, G. M.M.J. Kerkhoffs, G. J. M. Tuijthof, A.A. Zadpoor, Typical shape differences in talus and calcaneus between subjects with chronic ankle instability and controls, 2018.

- S.J.P. Callens, **N. Tümer**, A.A. Zadpoor, Origami of minimal surfaces, 2018.

- D. Widyaratih, P. L. Hagedoorn, L. Otten, M. Ganjian, **N. Tümer**, I. Apachitei, C. W. Hagen, L. Apachitei, A.A. Zadpoor, Towards osteogenic and bactericidal nanopatterns, 2018.

Conference Papers, Posters and Presentations

- **N. Sarkalkan**, A.J. Loeve, K.W.A. van Dongen, G.J.M. Tuijthof, A.A. Zadpoor, A novel ultrasound technique for detection of osteochondral defects in the ankle joint, 5th Dutch Bio-Medical Engineering Conference 2015, Netherlands. (Poster)

- **N. Tümer**, L. Blankevoort, M. van de Giessen, M.P. Terra, P.A. de Jong, H. Weinans, G.J.M. Tuijthof, A.A. Zadpoor, Bone shape difference between control and osteochondral defect groups of the ankle joint, 12th IEEE EMBS International Summer School on Biomedical Imaging 2016, St Jacut de la Mer, France. (Poster)

- M. Stijntjes, **N. Tümer**, Motion for Life Pathology Lecture, TU Delft 2016, The Netherlands. (Presentation)

- **N. Tümer**, F. Vos, G. Streekstra, C. Askeland, G. J. M. Tuijthof, A. A. Zadpoor, Three-dimensional registration of freehand-tracked ultrasound to CT images of the talocrural joint, 23rd Congress of the European Society of Biomechanics 2017, Seville, Spain. (Presentation)

- **N. Tümer**, G. Vuurberg, L. Blankevoort, M. P. Terra, G. J. M. Tuijthof, A. A. Zadpoor, Talus bone shape difference between controls and chronic ankle instability patients, ESSKA-AFAS Open Meeting 2017, Bordeaux, France. (Presentation)

- G. Vuurberg, **N. Tumer**, L.M. Wink, I. Sierevelt, L. Blankevoort, M. Maas, G.M.M.J. Kerkhoffs, G.J.M. Tuijthof, Radiographic prognostic variables in chronic ankle instability: a reliability analysis, ESSKA-AFAS Open Meeting 2017, Bordeaux, France. (Presentation)
- **N. Tümer**, G. Vuurberg, L. Blankevoort, G. M.M.J. Kerkhoffs, G. J. M. Tuijthof, A.A. Zadpoor, Is there any bone shape difference between controls and patients with chronic ankle instability?, NOV Jaarcongres 2018, The Netherlands. (Presentation)