University of Birmingham

Ph.D thesis

Articulated Statistical Shape Models for the Analysis of Bone Destruction in Mouse Models of Rheumatoid Arthritis

Author:

James Brown

Supervisors:

Dr Amy Naylor

Dr Iain Styles

Dr Andrew Filer

Prof Ela Claridge

Articulated Statistical Shape Models for the Analysis of Bone Destruction in Mouse Models of Rheumatoid Arthritis

by

JAMES MARTIN BROWN

A thesis submitted to
The University of Birmingham
for the degree of
Doctorate of Philosophy

School of Chemistry
College of Engineering and Physical Sciences
University of Birmingham
July 2015

Abstract

Rheumatoid arthritis is an autoimmune disease that affects approximately 1% of the population, where chronic inflammation of the synovial joints can lead to active destruction of cartilage and bone. New therapeutic targets are discovered by investigating genes or processes that exacerbate or ameliorate disease progression. Mouse models of inflammatory arthritis are commonly employed for this purpose, in conjunction with biomedical imaging techniques and suitable measures of disease severity. This thesis investigated the hypothesis that a statistical model of non-pathological bone shape variation could be used to quantify bone destruction present in micro-CT images. A framework for constructing statistical shape models of the hind paw was developed, based on articulated registration of a manually segmented reference image. Successful registration of the reference towards ten healthy hind paw samples was followed by statistical shape analysis. Mouse models of inflammatory arthritis were then investigated and compared by identifying bone abnormalities as deviations from the model statistics. Validation of the model against digital phantoms and clinical scores indicates that the method is largely successful in this effort. Application of the method in a novel study of macrophage-mediated inflammation shows promising results that are supportive of previous findings.

Acknowledgements

I would first like to sincerely thank my lead supervisors Prof Ela Claridge and Dr Amy Naylor for their unwavering enthusiasm and guidance throughout this research. They have both given up many hours of their time to provide me with training, advice and encouragement without which I could not have hoped to succeed. In addition, I would like to extend my sincerest thanks to my co-supervisors Dr Andrew Filer and Dr Iain Styles, for their ongoing support and advice. This research could not have been undertaken without the funding provided by the Engineering and Physical Sciences Research Council (EPSRC), awarded through the Physical Sciences of Imaging in the Biomedical Sciences (PSIBS) Doctoral Training Centre, at the University of Birmingham.

My sincere thanks go to the members of the Medical Imaging and Image Interpretation Group, all of whom have been wonderful colleagues and friends. I also would like to thank the members of the Rheumatology Research Group (RRG) for accepting me as one of their own and embracing some unfamiliar topics. Special thanks go to Prof Andy Clark, Dr Ewan Ross and Dr Tim Smallie for providing the TTP transgenic mice and to Dr Guillaume Desanti for providing the CAIA mice. I cannot go without thanking all members of the PSIBS Doctoral Training Centre, both past and present, for their friendship and support over these past four years. Special thanks go to my house mates Alan Race and Alistair Bannerman for the many curries, beers and laughs we've shared.

To my parents Peter and Jennifer and my sister Louise, thank you from the bottom of my heart for supporting me in everything I've ever done in life. And to my girlfriend and fellow PSIBS student Elizabeth Randall, I thank you for your love and support over the past two and a half years.

Contents

List of Figures

Chapter 1

Introduction and aims

1.1 Introduction

The immune system is one of several defences that biological evolution has erected to ward off disease through a careful balance of resistance and tolerance. In some individuals, a combination of genetic and environmental factors can tip this balance unfavourably. The body becomes incapable of recognising itself, causing it to attack its own tissues and organs. This undesirable immune response is referred to as autoimmunity, and is responsible for a wide range of disorders.

Rheumatoid arthritis (RA) is an autoimmune disease that affects approximately 1% of the world's population [1]. Although it can affect multiple tissues and organs, it is generally regarded as a disease that affects the synovial joints. In severe cases, destruction of bone and cartilage can lead to clinical disability and increased mortality [1]. Pre-clinical studies attempt to uncover the underlying causes by emulating the disease in mice. In order to assess the degree of damage to bone, x-ray based techniques such as micro-CT are employed to reveal changes in bone structure. At present, researchers rely on subjective assessments that do not provide quantitative data about the nature of the destruction that has occurred. This thesis postulates that a statistical model of normal bone shape

can be used to quantify the bone destruction present in micro-CT images of mice.

The ultimate goal of this work is to provide the rheumatological research community with the necessary tools to characterise and quantify murine bone destruction in an entirely automated fashion. Using computer software, researchers would supply unprocessed micro-CT images and be provided with bone phenotype data that describes the type, severity and location of the bone destruction present. Such data would allow for reliable comparisons to be made between wild-type (normal) and diseased mice, or the effect of treatment on disease outcome. Furthermore, large volumes of image data could be assessed with minimal user interaction that that would otherwise be prohibitively time consuming, and subject to inter-operator error.

1.2 Rheumatoid arthritis

1.2.1 Inflammation and bone destruction

RA is one of a number of disorders that affects the musculoskeletal system. It is an erosive disease, characterised by localised loss of cartilage and bone. This process takes place in the synovial joints, which have a fluid-filled cavity to lubricate the bone surfaces and reduce friction as they articulate. The cavity is enclosed by a membranous lining called the synovium, which is further surrounded by a fibrous envelope called the joint capsule. Synovial joints are the most common type of joint found in the body, and can be found in the hand, wrist, elbow, shoulder, and knee. In RA, the autoimmune response mounted by the body gives rise to chronic inflammation of the synovial joints, characterised by a thickening of the synovial membrane, an increase in synovial size (hypertrophy) and increased synovial cell proliferation (hyperplasia). A diagram illustrating some of the differences between normal and diseased joints is shown in Figure 1.1.

Among those who suffer from RA, a proportion will develop persistent inflammation of the synovium leading to the destruction of both cartilage and bone. Bone resorption is

the breakdown of bone through the secretion of enzymes by osteoclasts that demineralise the bone matrix. The opposite process is bone formation, where osteoblasts deposit new bone mineral into the underlying bone matrix. In healthy individuals, the two mechanisms are carefully regulated to in order to maintain bone integrity. In patients with RA, bone erosion is thought to be due to increased osteoclast activity, in conjunction with normal or decreased osteoblast activity [2, 3]. Over time, destruction of the bone surface can impair normal joint function and lead to structural deformities. In approximately 90% of cases where inflammation is persistent, patients will be clinically disabled within 20 years [1].

1.2.2 Treatment: current and future

The exact cause of rheumatoid arthritis is unknown, which inevitably limits the number of available treatment options. Non-steroidal anti-inflammatory drugs (NSAIDs) are one class of drug employed for their analgesic and anti-inflammatory effects, which can alleviate both pain and inflammation when administered at high doses. These drugs may be used in conjunction with other treatments that regulate immune activity, such as glucocorticoids. In addition to non-specific immune regulation, there are several therapies available that target specific mechanisms associated with RA in order to slow disease progression. New therapeutic targets may be discovered by investigating genes or processes that exacerbate or ameliorate rheumatoid arthritis.

As scientific understanding of the disease improves, new drugs may be developed to specifically target the genes or proteins responsible for disease progression. Animal models of inflammatory arthritis are frequently employed for this purpose, in conjunction with suitable quantitative measures of disease severity. Recent developments in this respect come in the form of tumour necrosis factor (TNF) inhibitors such as infliximab, which

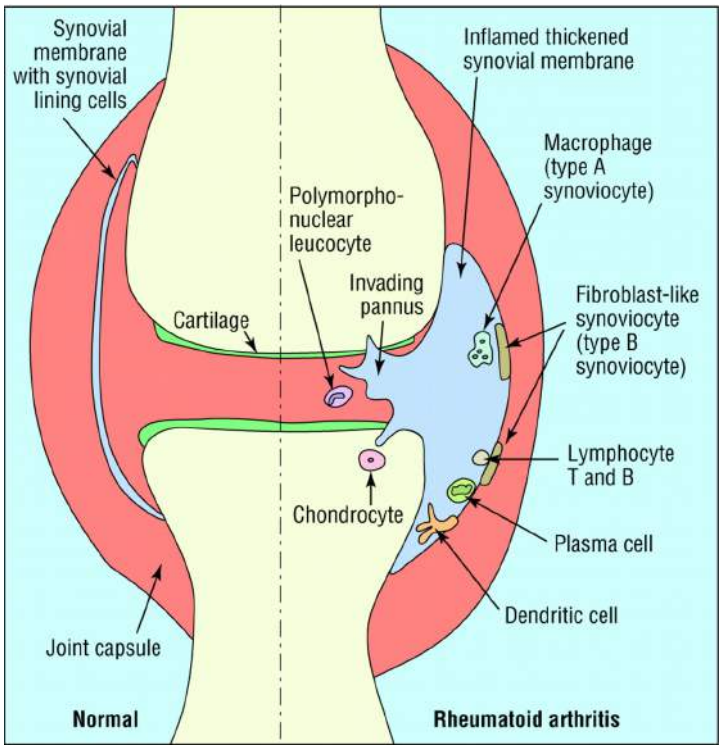
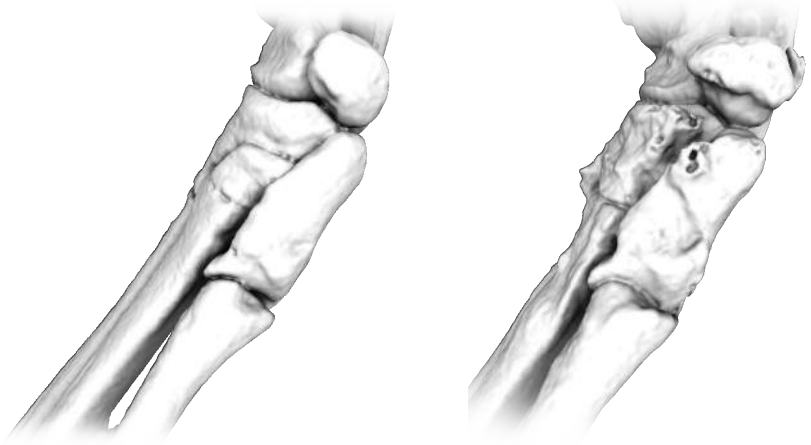


Figure 1.1: Diagram comparing typical healthy and RA-affected synovial joints. The left side of the diagram exemplifies a healthy joint, showing the joint capsule and synovial membrane are of normal size, with the joint surface showing no signs of damage. On the right, an example of a diseased joint shows an enlarged joint capsule and thickening of the synovial membrane. A portion of the synovial membrane referred to as a pannus is shown to be invading the joint surface. The pannus is composed of a localised accumulation of macrophage- and fibroblast-like cells that cause degradation of cartilage mediated by protease enzymes. The formation of osteoclasts later causes destruction of sub-chondral bone (beneath the cartilage). Figure reproduced from [1].

initiates apoptosis (programmed cell death) in TNF expressing T-cells [4]. This later led to the development of a mouse model of rheumatoid arthritis in which inflammation is caused by constitutive overexpression of TNF (TNF dARE) [5].

1.3 Mouse models of inflammatory arthritis

Animal models of disease allow biomedical researchers to investigate medical conditions through experimentation that would otherwise be infeasible or unethical to perform on human subjects. Development of an animal model often involves the alteration of a biological process that is implicated in disease in order to examine its effects. Mice are particularly well-suited for this task, due to their high genetic homology (number of genes in common) with humans, speed of breeding and inexpensive housing costs. A number of laboratory mouse strains have been developed and bred to be near genetically identical to one another, which allows for the roles of different genes to be investigated in a way that eliminates genetic variation as a factor. The most commonly used "background" strain is the C57BL/6 mouse, and is widely used in studies of RA. Disease induction may be achieved in a number of ways, including genetic modification, infection with a foreign antigen, or a surgical procedure. Although RA is predominantly an erosive disease, a number of mouse models show evidence of both erosion and formation [6, 7]. Figures 1.2 and 1.3 shows some examples of both kinds of destruction in micro-CT data. It is important to note that mouse models are, by definition, not truly representative of the disease and so their phenotypes may differ from that which is observed in human patients.



(a) Normal bone morphology (b) Abnormal bone loss

Figure 1.2: An example of bone erosion in the arthritic mouse hind paw. Bone erosions are small pits or depressions in cortical bone, usually situated around the synovial joints (periarticular bone). In severe cases, erosions can impair joint function and lead to structural deformities. Bone mineral can also be lost in a less localised fashion, becoming less dense and thinner in appearance in radiographic data.

(a) Normal bone morphology (b) Abnormal bone formation

Figure 1.3: An example of abnormal bone formation in the arthritic mouse hind paw. Although more prevalent in osteoarthritis, bumps or growths on the cortical surface can impair joint function and are variable in size and shape. Such growths are referred to as osteophytes if located on periarticular bone, or enthesophytes when found at entheses (attachment sites of ligaments and tendons).

1.3.1 Immune mediated models

Antibodies are proteins which are used by the immune system to recognise foreign bodies (antigens) and provoke an immune response. Immune mediated mouse models are those that give rise to inflammatory arthritis via an antibody-induced immune response, leading to the onset of inflammatory arthritis. One commonly used model is the K/BxN mouse model, which emerged following the development of a T-cell receptor transgenic mouse (KRN). KRN mice express a T-cell receptor that is specific to an epitope[1] of bovine pancreas ribonuclease, which is commonly used in studies of antigenic recognition. It was discovered that upon crossing these mice with non-obese diabetic (NOD) mice, they would develop spontaneous inflammatory arthritis [8]. This combined effect of having the KRN transgene and I-A(g7) MHC allele[2] expressed by NOD mice was unforeseen [9]. The resulting K/BxN mouse is known to produce antibodies to glucose-6-phosphate isomerase (G6PI), which brings about severe inflammatory arthritis. Serum extracted from the K/BxN mouse may be administered to naive mice to bring about short-term inflammation of the synovial joints and erosive inflammatory arthritis [10, 11].

In an alternative approach, antibodies to type II collagen (CII) may be used to induce the immune response. CII is a major component of articular cartilage, which is a primary site of destruction in patients with RA. This is due to breakdown of the collagen matrix by enzymes known as matrix metalloproteinases (MMPs) [12]. The model, known as collagen antibody induced arthritis (CAIA) [13], produces a transient form of inflammatory arthritis that may be exacerbated by lipopolysaccharide (LPS) stimulation [14].

Both the K/BxN serum-transfer and CAIA models are simple and easy to induce, and have up to 100% incidence rates. This reduces the number of mice required for a study,

[1] An epitope is the region on an antigen that is recognised by the immune system via an antibody.

[2] The major compatibility complex (MHC) is a class of molecules that have an important role in the immune system. Class II molecules are expressed by antigen-presenting cells (e.g. macrophages), which activate T-cells in order for them to regulate immune activity.

making it a cost-effective approach. However, the inflammation in both models is short-lived. This is due to clearing of the antibodies from the recipient mouse over time causing the immune response to resolve. In addition, the model may not be representative of the human disease as the mouse does not mount an autoimmune reaction (usually caused by a break of tolerance to the self-antigen), making the disease transient, rather than permanent as generally observed in humans. In studies where a longer lived inflammation is necessary, a supplementary dose of antibodies can be administered. Alternatively, short-term inflammation can be exploited to investigate the process by which rheumatoid arthritis can resolve in some patients.

In studies where self-sustaining, persistent inflammation is required, the collagen induced arthritis (CIA) model may be used. Type II collagen is administered in addition to Freund's adjuvant (an immunostimulator) causing immune cells to produce type II collagen antibodies [13]. This gives rise to autoimmunity in the mouse, resulting in interminable inflammatory arthritis that is more representative of the disease than the aforementioned models. The main drawback of this model is its significantly lower incidence rate. In susceptible mouse strains (e.g. DBA/1), the incidence rate is between 60% and 80%. In commonly used strains such as C57BL/6, the incidence rate is significantly lower (20% - 40%). Furthermore, inducing this model is significantly more time consuming than the K/BxN serum-transfer and CAIA models.

1.3.2 Cytokine/process mediated models

The roles of cytokines present in the synovium are of key interest in rheumatoid arthritis research [15], and knowledge about their roles can be used to model the disease. In cytokine or process mediated models, inflammatory arthritis is induced by altering the production or activity of cytokines that are known to be involved in the manifestation of rheumatoid arthritis. There are a number of cytokines that are known products of

inherently unstable mRNAs (i.e. quick to degrade) to ensure swift cessation of an immune response when necessary [16]. Regions of mRNA located in the 3' untranslated region (UTR) known as AU-rich elements (AREs) are partly responsible for this degradation, providing sites for RNA binding proteins such as tristetraprolin (TTP) to attach and accelerate the shortening of the poly(A) tail [17].

Deletion of the AU-rich elements on the tumour necrosis factor (TNF) gene leads to overexpression of TNF-α due to increased mRNA stability. Under this model known as TNF dARE, mice exhibit spontaneous inflammatory arthritis, as well as Crohn's-like inflammatory bowel disease [18]. It should be noted that this model does not cause autoimmunity, and is a gross simplification of the disease process. Despite this fact, the importance of TNF in the human form of the disease makes it a useful model in studies related to TNF therapy. Furthermore, the inflammatory arthritis is non-resolving, causing severe destruction of the joints.

Post-transcriptional regulation of cytokine mRNAs has gathered much interest as a potential therapeutic target. It is thought that by forcibly accelerating the degradation of pro-inflammatory cytokine mRNAs, the immune response may be suppressed in patients with inflammatory disorders such as RA. This has been demonstrated experimentally by modifying the gene for tristetraprolin, in such a way that the resultant protein is unable to be phosphorylated, and therefore permanently active [19]. This has been further shown to inhibit the expression of pro-inflammatory mediators. This model is discussed further in Chapter 7, Section 7.2.

1.4 Characterisation of disease in mouse models

1.4.1 Physiological disease index assessment

During the course of an experiment, mice may be examined for traits associated with inflammatory arthritis. Physiological disease index assessment is the process of assigning

scores to individual animals according to disease severity, based on observations and measurements made by one or more observers. For some traits, quantitative physical measurements can be taken in place of scores, such as paw and ankle thickness, body weight and grip strength. Other traits cannot be measured in this way, and are instead given qualitative scores. This is usually performed in a blinded fashion, in order to minimise the effects of confirmation bias. Mice are usually held supine by the scruff of the neck, exposing the fore and hind limbs as shown in Figure 1.4 [3]. This allows for closer examination of the paws for signs of inflammation such as swelling and redness, and identification of joint deformity.

Assessment of disease in this way is an effective method for determining the gross differences between individual animals, but only provides superficial data about the nature of the disease. Furthermore, assessment of traits such as limb deformity, colouration and gait are highly subjective. The frequency with which scoring is performed is critical to understanding the temporal nature of disease, making such assessments highly labour intensive.

1.4.2 Histopathology

Histological sections of the hind paw reveal differences in synovium and bone morphology at the cellular level. Upon termination of an experiment, hind limbs are dissected from sacrificed mice and chemically fixed to prevent decomposition. In order to make the cutting of sections easier, bones are usually decalcified using a chelating (metal-binding) agent such as EDTA, and embedded in a paraffin block. Sectioning is then performed using a microtome; an instrument capable of producing extremely thin sections of tissue. A number of different staining protocols can then be used to highlight certain cells and tissues

[3]This common method of handling is not harmful to the mouse. The scruff (or nape) of the neck is insensitive to pain, allowing it to be examined without experiencing discomfort.

(a) Wild-type mouse

(b) TNF dARE transgenic mouse

Figure 1.4: Observable differences between normal and arthritic mice. The wild-type mouse shown in (a) has paws that are a bright pink in colour, with the individual digits splayed out. In contrast, the TNF dARE mouse shown in (b) has distinct discolouration of the paws, with digits that appear bunched. Other indicators of disease include ruffled fur (piloerection) and loose bowel movements that are not observed in the wild-type.

under a bright-field microscope. Common stains include safranin O staining of cartilage, tartrate-resistant acid phosphatase (TRAP) staining of osteoclasts, and haemotoxylin and eosin (H & E) staining for visualisation of bone, cartilage and inflammation. Figure 1.5 gives an example of the latter, with localised bone erosions annotated.

Histopathology gives a deeper insight into the nature of disease than physiological scoring, which can improve understanding of disease mechanisms. However, analysis of histological sections is often qualitative, where scores are given based on the presence of different markers of inflammation and bone destruction. However, it only provides a snapshot of the destruction within the tissue; total destruction throughout the tissue cannot be easily measured. Furthermore, preparation of tissue sections is destructive, time-consuming and costly. In some studies, bone histomorphometry has been used to measure abnormal regions using semi-automated computer software [20]. Although such an approach provides comprehensive quantitative data about the nature of bone destruction, it is costly in terms of the hardware, software and training required.

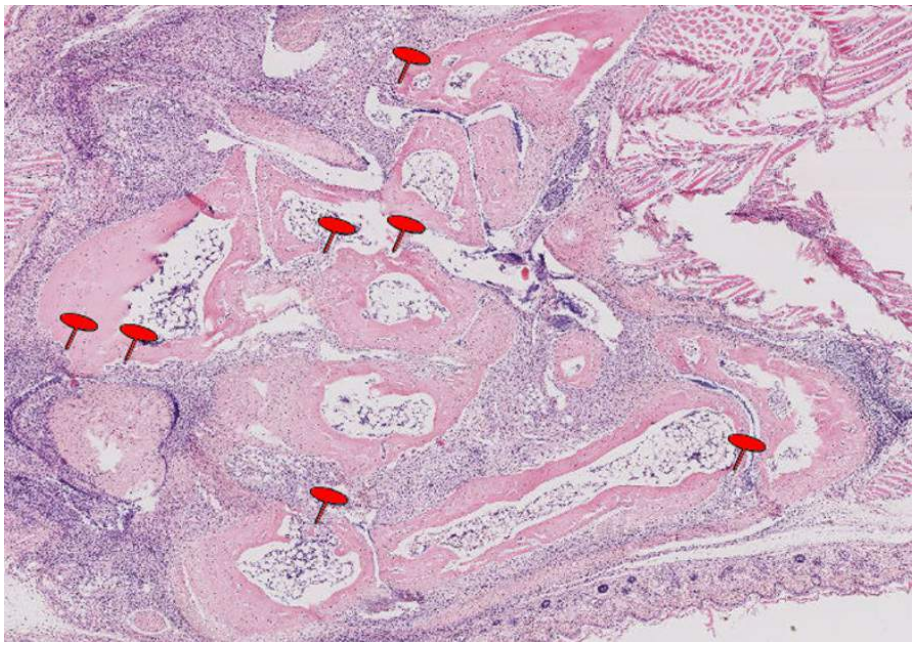


Figure 1.5: Histological section of an arthritic hind paw, stained with H & E. Bone and cartilage is shown in bright pink, with cell nuclei highlighted in purple. Inflammation can be found throughout the section, in some cases infiltrating the bone surface (pannus). Red flags are placed at locations where bone erosions have been identified, evidenced by pannus formation and discontinuities of the cortical surface.

1.4.3 X-ray microtomography (micro-CT)

Bone destruction is an important hallmark of pathology in mouse models of rheumatoid arthritis. The nature and severity of bone shape changes are useful in understanding disease mechanisms and the efficacy of new treatments. X-ray microtomography (micro-CT/μCT) is an imaging modality that can be used to image bone specimens in three dimensions at a spatial resolution of up to 1 μm. Although some instruments may be used to image whole small animals, others can only be used to image excised tissue [21]. The instrument available for this research fits the latter description, capable of imaging mouse extremities such as the fore and hind limbs (shown in Figure 1.6).

In micro-CT, a single 2D projection is obtained by directing a cone-shaped beam of x-ray radiation at the specimen, which is attenuated to varying degrees according to

Figure 1.6: SkyScan 1172 micro-CT scanner (Bruker micro-CT). This particular instrument is commonly used in the biomedical sciences to image small tissue samples (e.g. teeth, bone). Software is supplied with the instrument for image acquisition, post-processing, reconstruction and visualisation. (Photograph reproduced from www.microphotics.com)

tissue density and atomic number. X-rays that are transmitted through the specimen are collected by a discretised scintillation detector, where image contrast is determined by the distribution of hard and soft tissue in the specimen. A set of 2D projections is captured at discrete angles (between 0° and 180°) around the specimen. Individual slices can be reconstructed using the Feldkamp algorithm [22], and combined to produce a 3D volumetric image.

Micro-CT offers an affordable, non-destructive method for imaging murine bone architecture, that requires minimal sample preparation. The technique has been demonstrated in animal studies of RA as a suitable means for evaluating periarticular bone loss and osteophytosis, due to its high sensitivity to changes in bone structure. The role of the TRANCE/RANKL gene in bone resorption has previously been investigated using the K/BxN serum-transfer mouse model of rheumatoid arthritis [6]. Two-dimensional micro-CT slices were individually examined by experts to identify erosions, which were co-localised in histological sections to assess both hard and soft tissue damage. No quantification was performed on the micro-CT slices themselves, serving only as an aid to

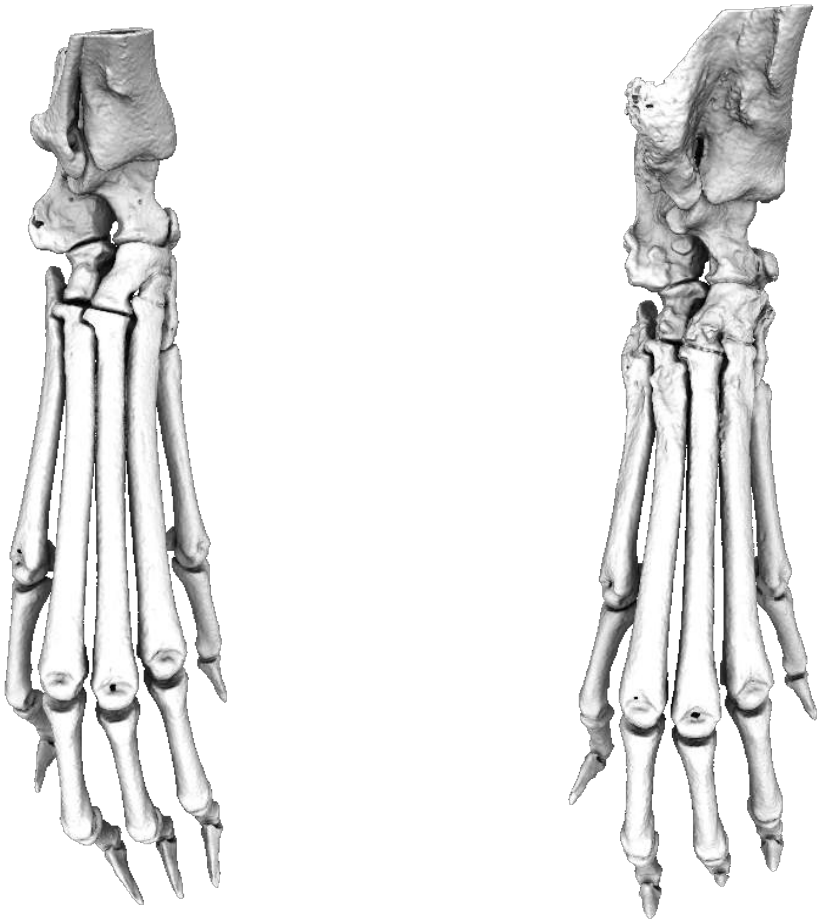
histological scoring. Micro-CT has since been used more resourcefully, showing that inhibition of the enzyme spleen tyrosine kinase (Syk) can suppress both inflammation and bone erosion in a collagen induced arthritis (CIA) mouse model [23]. Such scoring systems are subjective, and only provide a partial phenotypic assessment.

A comparison between normal and arthritic mice imaged by micro-CT is shown in Figure 1.7. The process of manually identifying bone abnormalities is aided by an observer's own experience of how such bones appear under normal (non-pathological) circumstances. Erosions are regions of localised cortical bone loss, typically observed as indentations or pits on the bone surface. Erosions may vary in size and shape depending on the degree and persistence of inflammation, but are most commonly found on periarticular bone (around the joint). In addition to erosive bone destruction, abnormal bone formation (ossification) may be observed as lumps or growths on the bone surface. Although not typically observed in RA, osteophytes (bone spurs) and other forms of abnormal bone formation have been observed in mouse models of arthritis [24, 25].

1.5 Computational analysis of tomographic data

Several methods have been proposed to automate the analysis of bone destruction in RA. Quantification of total bone volume is one possible approach, where gross differences are determined through segmentation of whole bone structures from image data. This approach has previously been used to assess mice with collagen induced arthritis (CIA), which were shown to have significantly lower bone volume and density than wild-type controls [26].

One of the main limitations of this approach is that it provides no data about bone morphology, or the nature of the bone destruction that has occurred. As such, nothing can be inferred with regards to the mechanisms involved. This problem has since been tackled using surface-based registration of bones [27]. In a longitudinal study of treatment efficacy,



(a) Normal mouse

(b) Arthritic mouse

Figure 1.7: Visual comparison of (a) normal and (b) arthritic hind paws. The mice are identical in terms of their age (sacrified at 12 weeks), gender (female) and genotype (C57BL/6). The normal hind paw has a smooth bone surface, with no irregularities other than those which occur naturally. The arthritic sample was excised from a mouse having been injected with K/BxN serum two weeks prior to being sacrificed. When compared to the normal sample, discernible bumps and indentations can be observed in the upper region of the paw.

three-dimensional CT images were acquired from patients with established RA. Individual bones were manually segmented, and surface representations generated as triangulated meshes. Surface registration of bones at different stages of treatment was followed by calculation of point-wise distances between them, and used to calculate local differences in bone volume. These differences were also visualised as colourised surface renderings, highlighting the presence of abnormalities. This approach is effective in assessing bone changes over time in individual patients, but is not well-suited to detecting abnormalities in subjects imaged at a single time point.

1.6 Aims

The primary aim of this research is to develop an automated method for acquiring quantitative measurements of bone destruction from micro-CT images of the mouse hind paw. The proposed method involves construction of a statistical model of normal bone morphology, based on a training set of healthy examples. The model is based on the highly popular active shape model (SSM) by Cootes *et al.* [28], for which a myriad of scientific literature exists detailing the various ways in which they can be constructed, adapted and applied to problems in medical image computing. In this research, a framework for articulated registration (based on Baiker *et al.* [29]) is proposed to establish point correspondence across the training set. The result is an articulated statistical shape model (ASSM), which may then be registered to arbitrary hind paw samples. Non-rigid deformation of the model bones towards the sample will allow for differences in bone shape to be identified as deviations from the model statistics. Abnormalities may then be characterised and quantified in terms of type (erosion/formation), severity (area/volume) and location, in a similar manner to Joshi *et al.* [27]. Ultimately, it is hoped that the developed method may be employed in pre-clinical studies, and help to answer questions about how different mechanisms contribute to RA pathogenesis.

1.7 Summary

The role of mouse models in RA research has been presented and discussed. Characterisation of bone destruction is key to understanding RA pathogenesis, and evaluating the efficacy of treatments in pre-clinical trials. Micro-CT has been demonstrated as being very well-suited to this task, but is hampered by a lack of automated methods for quantitative analysis of bone destruction. To solve this problem, a statistical model of non-pathological bone variation has been proposed, which can be both constructed and applied with negligible human interaction.

Chapter 2

Background and previous work

In this chapter, relevant theory and applications are presented and discussed. Specifically, concepts related to image registration, statistical shape modelling and abnormality detection are explored to identify the key considerations and caveats of building statistical shape models.

2.1 Image registration

Image registration (also known as spatial normalisation) is the process of aligning two or more images such that homologous structures overlap. When registering two images, transformations are applied to one image (the moving, or source image) whilst keeping the other image stationary (the fixed, or target image). The process of registering the moving image I_M onto the fixed image I_F involves finding a transformation $\mathcal{T}$ that minimises some cost function $\mathcal{C}$. The cost function measures the overall difference between the transformed moving image $I_M \circ \mathcal{T}$ and the fixed image I_F. The optimal transformation $\widehat{\mathcal{T}}$ is found by minimising this difference:

$$\widehat{\mathcal{T}} = \arg\min_{\mathcal{T}} \mathcal{C}(I_F, I_M \circ \mathcal{T}) \tag{1}$$

Using this general formulation, one can devise a registration algorithm that is suited

to the problem at hand. In biomedical imaging, this is largely dependent on the nature of the input images, the anatomy being studied, and the desired output of the analysis. Applications of registration within the field are numerous, and range from simple comparisons to computer assisted surgery. One of the key considerations when performing registration is how the image data should be represented. The two most common classes of registration algorithm are *intensity-based* and *point-based*, each having their benefits and drawbacks.

2.1.1 Intensity and point registration

An intensity image is a discrete grid of elements or pixels, each of which is assigned a numerical value within a specified range (e.g. 8-bit greyscale images can have $2^8 = 256$ possible values). In three dimensional images, individual grid elements are instead referred to as voxels which, depending on the imaging modality, may be uniformly (isotropic) or non-uniformly sized (anisotropic). The intensity values at pixel/voxel locations are the product of the subject being imaged and the imaging modality used. By contrast, point clouds describe only the surfaces of objects. The distance between individual points is not fixed, and so may take on non-integer coordinate values. In cases where the connectivity between points is known, they are together referred to as a mesh.

Intensity-based image registration is performed directly on greyscale or colour images, where values at pixel/voxel locations are compared and used to evaluate the dissimilarity between them. Minimisation of this dissimilarity (according to Equation (1)) is the objective of intensity-based registration, where the moving image is transformed to become aligned with the target image. If the objects in question are surface meshes, the objects are compared in terms of corresponding point coordinates (vertices). In either case, the objective is to minimise the dissimilarity between two images: a moving image I_M and a fixed image I_F. A dissimilarity metric $\mathcal{D}$ must satisfy the following requirements:

- Dissimilarity is commutative: $\mathcal{D}(I_F, I_M) = \mathcal{D}(I_M, I_F)$
- Dissimilarity between two images should be greater than zero: $\mathcal{D}(I_F, I_M) > 0$
- Dissimilarity may only equal zero if the images are identical: $I_F = I_M$

Minimisation of dissimilarity involves finding some optimal transformation that when applied to the moving image I_M aligns it with the fixed image I_F. There are numerous transformation models that exist for images and point clouds, each falling under one of three categories: rigid/inelastic (translation, rotation), affine (rigid + scaling and shearing) and non-rigid/elastic (image warping under some deformation model). Transformations applied to intensity images often require resampling via interpolation due to their discrete nature, and possible voxel anisotropy (e.g. MRI, which often has slices that are thicker than the in-plane pixel dimensions).

For point-based registration, transformations are applied to points as opposed to pixels. A number of point registration algorithms exist to solve for rigid, affine and non-rigid alignment of point clouds in two or three dimensions. A comprehensive review of 3D point cloud registration describes a number of different methods, as well as the practical considerations that should be taken into account [30]. The iterative closest point (ICP) algorithm is a registration algorithm that was originally conceived for the alignment of three-dimensional objects [31]. The algorithm minimises the sum of squared differences (SSD) between the two point sets, alternating between a *correspondence approximation* step and a *transformation* step. For each point in the moving set A, the closest point in the target set B is found. Here, the closest point is defined as that which has the smallest Euclidean distance E:

$$E(a, b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2} \tag{2}$$

For point sets with few points, a simple exhaustive search through all of the points

may suffice. For larger point sets, more efficient solutions may more appropriate, such as a kd-tree search [32]. Using this subset of corresponding points $B' \subseteq B$, a transformation is found which minimises their sum of squared differences, defined as:

$$SSD(A, B) = \sum_{j=1}^{n} \sum_{k=1}^{3} (A_{jk} - B'_{jk})^2 \tag{3}$$

This is formulated as an optimisation problem, which finds the transformation the best aligns A and B'. Depending on the implementation, this step may solve for both rigid (translation, rotation) and affine (scaling and shearing) transformations. The resulting transformation is applied to the moving point set B, and the process repeats until convergence. The algorithm is often defined to run until fixed number of iterations is reached, or when the difference in SSD between iterations falls beneath a specified threshold.

One of the main limitations of ICP is the necessity for a good initial alignment, as the algorithm will always converge on the local minimum. Furthermore, ICP is not particularly robust to noisy or missing data. A number of algorithms have been developed to overcome these limitations, such as robust point matching (RPM), which utilises "soft" correspondences between point sets rather than the binary assignment of correspondences by ICP [33]. An alternative method is coherent point drift (CPD), which utilises Gaussian Mixture Models to align two point sets [34].

2.1.2 Atlas-based registration

In medical image computing, the term atlas usually refers to single example image that has been segmented into a number of structures of interest. The result is a labelled image or volume, with each segment being assigned a unique label. Registration of the atlas towards an unlabelled sample image allows for the labels to be mapped (propagated) to provide a segmentation of the sample. The segmented image may then undergo some refinement to correct for small errors, using techniques such as mathematical morphology.

The first known application of atlas-based registration to the murine skeleton was by Baiker *et al.* [29]. The paper describes the construction of a whole mouse body atlas from micro-CT image data, and how it may be used as part of an articulated registration framework. The atlas was created by first segmenting the skeleton from the background via global thresholding. The skeleton was then manually segmented into its constituent structures (skull, spine, pelvis etc.) and the individual joints annotated. Following this, a surface representation was generated by the marching cubes algorithm [35]. Kinematic constraints were then established that defined how each component would move relative to its predecessor. Registration of the atlas involved an initial global alignment based on the curvature of the body, which accounted for any variations in overall orientation. Following this, an extension of the iterative closest point (ICP) algorithm was used to align the atlas components in a sequential, hierarchical fashion. The hierarchy itself defined the order in which components would be registered, where each successive registration initialised the next bone to undergo alignment. The concept of the whole mouse body atlas was later expanded upon, detailing how it may be applied to other anatomical structures, and to approximate organ locations in low contrast data [36, 37].

Registration using a single atlas is inherently biased towards the geometry of its structures. To counter this, multiple atlases may be used to account for the variations that exist among different subjects. Additionally, soft labelled atlases may be used where each voxel is given a probability of belonging to a particular structure. A hard consensus may then achieved using multiple registered atlases [38, 39]. The whole mouse body atlas discussed previously was extended to incorporate organ shape variations, by training a statistical shape model on 83 manually segmented micro-CT images. This enabled registration of the atlas with a planar x-ray and an optical photograph [40].

2.1.3 Identifying abnormalities

Imaging techniques are often used to identify structural (or indeed, functional) differences between different groups of subjects (e.g. wild-type and mutant, treated and untreated). Image registration yields spatial correspondences that allow for the direct comparison of homologous image regions. A simple subtraction (often referred to as a difference image) of two registered images highlights discrepancies that may be attributed to meaningful, physical differences between the two subjects being imaged. However, a certain proportion of these discrepancies is likely to be due to imperfections in the registration. Furthermore, a single image is unlikely to be representative of "normality". In order to reliably detect and quantify abnormalities, it is necessary to establish a representative normal reference onto which sample images may be registered. A reference could take the form of an average image, generated by co-registering a set of normal samples and then calculating the average image intensity at each corresponding pixel (or voxel) location.

Voxel-based and deformation-based morphometry are well-established methods used in neuroimaging for the assessment of differences in brain morphology [41]. The concepts, however, can theoretically be applied to other anatomical features. This approach has been used previously to detect differences between wild-type and mutant mouse embryos, imaged by micro-CT [42]. Images were deformably registered in a groupwise manner towards a labelled, wild-type average embryo [43]. Using the deformation fields produced by the B-spline registration, Jacobian determinants were computed to identify regions of local compression and expansion. Intensity values between the two groups of embryos were also analysed. Statistically significant differences between the two groups of embryos revealed morphological differences among mutants that recapitulated reported findings. The method relies on the abnormalities being consistently located at the same anatomical locations in multiple samples. Furthermore, variably penetrant abnormalities (i.e. those which do not affect all of the samples analysed) are unlikely to be detected. In mouse

models of RA, bone destruction can be highly variable in terms of location and penetrance and requires a different approach.

Novelty detection is a technique in computer vision that aims to identify anomalies through classification against a set of training examples. By constructing a representative model of "normality" from training data, irregularities may be identified as regions that are not recognised by the model. For biological structures, this approach may be used to identify instances of the model in medical image data.

2.2 Statistical shape models

2.2.1 Point distribution models

Point distribution models (PDMs) were first proposed as a method for analysing variations among a set of training shapes, and extracting the variation that exists among them [28]. The term *shape* refers to the geometric properties of an object in space that allows it to be distinguished from other objects. A more formal definition of shape is the "... geometrical information that remains when location, scale and rotational effects are filtered out from an object" [44]. In digital images, shapes can be delineated by boundaries represented by a finite number of points, or landmarks, described as "... a point of correspondence on each object that matches between and within populations." [44]. In building a PDM, each of the training shapes is represented by a set of n landmark points that capture the geometry of the objects being analysed. Corresponding landmarks must be located at the same anatomical location on each object, in order for their average positions to be determined. Additionally, the set of training shapes must be aligned such that they all have the same overall position (translation), orientation (rotation) and size (scale). This ensures that any non-shape variations are eliminated prior to shape modelling. One approach is Generalised Procrustes Analysis, defined as follows:

1. Arbitrarily select one shape as a reference.

2. Align each shape to the reference by translation, rotation and scaling. [1]

3. Calculate the average (mean) shape and set as the updated reference shape.

4. Go to 2, and repeat until the Procrustes distance falls beneath a specified threshold.

PDMs work by establishing linear trends in the data. By treating individual shapes as high-dimensional data points, the training data is modelled as a multivariate Gaussian function, which represents average shape and variance as a normal distribution. This is typically achieved through the application of principal component analysis (PCA) [45]. PCA is a technique that fits a collection of points in n-dimensional space to a set of orthogonal axes, called principal components. Each axis is effectively a line of best fit through the data, with each successive axis having lower variance than the next. PDMs are capable of generating feasible shapes (based on the data they are trained with) by examining how landmark points covary.

2.2.2 Constructing a point distribution model

We consider a set of aligned N three-dimensional training shapes, each described by a set of n corresponding landmarks. Individual shapes may be regarded as high-dimensional data points, each represented by a vector $\mathbf{x}$:

$$\mathbf{x} = (x_1, ..., x_n, y_1, ..., y_n, z_1, ..., z_n)^T \tag{4}$$

where (x_i, y_i, z_i) refers to a single landmark in three dimensions. The vector $\mathbf{x}$ is simply the concatenation of a shape's point coordinates into a single one-dimensional vector of length $3n$. The mean shape $\overline{\mathbf{x}}$ is calculated by averaging over the training set

[1] Other possible transformations include reflections and affine transformations such as shearing.

$$\overline{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{n}\mathbf{x}_i \tag{5}$$

Subtracting the mean shape from each training shape, a system matrix is formed

$$\mathbf{D} = (\mathbf{x}_1 - \overline{\mathbf{x}}, ..., \mathbf{x}_N - \overline{\mathbf{x}}) \tag{6}$$

The covariance matrix can then be computed as

$$\mathbf{S} = \frac{1}{N-1}\mathbf{D}\mathbf{D}^T \tag{7}$$

From this covariance matrix, its eigenvectors and eigenvalues are computed by PCA. The eigenvectors are lines of best fit through the data, with their corresponding eigenvalues describing the variance along them. The normalised eigenvectors $\mathbf{p_i}$ are sorted in order of descending variance $\lambda_\mathbf{i}$. In this context, the eigenvectors may also be referred to as *principal components*, *modes of variation* and *eigenmodes*, among others.

New shapes may be generated by the model as deviations from the mean, expressed as a linear combination of the components

$$\mathbf{x}' = \overline{\mathbf{x}} + Pb \tag{8}$$

where P is a matrix composed of the ordered set of principal components, and $\mathbf{b}$ is a vector of weights defining the contribution of each component. The term *shape instance* refers to any such deviation from the mean shape. In order to ensure that shape instances remain feasible, limits may be imposed such that $-3\sqrt{\lambda_i} < b_i < 3\sqrt{\lambda_i}$. In other words, new shape instances are constrained to lie within three standard deviations from the mean. This is commonly referred to as the "allowable shape domain". Additionally, it is common practice to only retain some proportion of the total variance explained by

the model, discarding some modes with low variance. Such modes may be attributed to factors such as noise or outliers caused by inaccuracies in landmarking.

2.2.3 Solving the correspondence problem

The correspondence problem remains one of the most challenging problems in computer vision and medical image computing. Unlike computers, humans have an innate ability to recognise objects within a scene, and automatically establish spatial correspondences between different instances of the same object. Landmark annotation is therefore a relatively trivial task for humans in principle. In practice however, this process is labour intensive and difficult to achieve without appropriate software.

A number of studies have attempted to overcome the need for manual annotation altogether, allowing for entirely automatic construction of statistical shape models. The correspondence problem can be tackled in several ways. A pairwise correspondence algorithm described by Brett *et al.* generates a dense polyhedral mesh of two input shapes A and B, composed of n_A and n_B vertices respectively [46]. The output of the algorithm is a pair of sparse subsets of A and B (A' and B'), each composed of n_ϕ corresponding vertices. Furthermore, the connectivity of the two polyhedra A' and B' is identical. A different approach involves manually landmarking a single reference, for example, and then propagating the landmarks onto the remaining shapes in a training set. For example, a study by Frangi *et al.* showed how statistical shape models of the heart can be constructed in an automated fashion using a pre-landmarked heart atlas, in which the segmented shapes to be landmarked are embedded [47]. The landmarks are then propagated onto the shape by iteratively deforming the mesh in a multi-resolution fashion, in order to achieve correspondence.

2.3 Computational anatomy

Computational anatomy is a well-established area of research which examines biological variability of structures in the human body. One purpose of computational anatomy is to assess differences between healthy and diseased populations based on statistics, which in itself presents a wide range of non-trivial technical challenges. Statistical shape modelling is particularly well-suited to examining anatomical variation within and between populations. This is evidenced by the widespread application of SSMs in medical image computing, and the continued efforts by the research community to refine and tailor them to different problems.

2.3.1 Statistical shape modelling of bone

Statistical shape modelling can be used to establish the relative contributions of each mode of variation to overall object shape. This is particularly useful when considering a set of pathological training shapes. Statistical shape models have been used to assess, for example, differences in wrist bone structure caused by ligament injury [48]. In the study, carpal bones from 50 healthy wrists were imaged by CT and segmented using level-set segmentation [49]. 3D surface meshes were constructed using the marching cubes algorithm [35], and surface landmarks were labelled semi-automatically to achieve correspondence. Prior to model construction, the set of wrist bone meshes were co-registered using an iterative closest point algorithm. A statistical shape model was then constructed, and the modes of variation were analysed to determine which features of the wrist bones they represented.

The deformable nature of SSMs allows them to be non-rigidly registered onto new shape examples, whilst regularising deformations to be biologically feasible. This has been exploited in the field of computer-assisted surgery, in order to locate the optimal drilling location during anterior cruciate ligament surgery [50]. Anatomical variation from

patient to patient makes this problem particularly difficult, where a restricted field of view of the knee can result in suboptimal placement of the ligament graft. This problem was addressed using a statistical shape model of a whole femur. A model was constructed from eleven healthy femurs, having been digitised by a 3D optical localiser. During surgery, points on the accessible surface of the knee were digitised in the same way to approximate the shape of the joint surface. The shape model was then non-rigidly registered onto the target shape using a simulated annealing based algorithm [51] that simultaneously finds the optimal rigid transformation and non-rigid shape deformation. The final fit of the model extrapolates from the accessible joint surface, estimating the shape of the obscured portion of the femur. This estimation allowed surgeons to make a more informed decision about where to place the ligament graft, despite the restricted field of view.

Although SSMs are particularly amenable to non-rigid registration problems, target shapes that exhibit large differences (due to pathology, for example) can cause common registration algorithms to fail. Statistical shape models have been used to segment and quantify osteophytes present in a rabbit femur model of osteoarthritis, having been imaged by micro-CT [52]. Initial pre-processing of the micro-CT data involved applying a simple threshold to segment the rabbit femur from the background, followed by despeckling and extraction of the largest component in the image. Any holes present in the binary images due to marrow space were then filled by morphological closing. In order to construct a model, a "fiducial landmarking system" was defined, whereby a set of 3D feature points were manually identified on each femur in a training dataset, using a purpose-built graphical user interface. A single femur was then taken from the training dataset, on which a more complete set of landmarks were identified. A mesh was generated from the surface points using the marching cubes algorithm, which was then refined to achieve uniformity, and to remove superfluous vertices and edges. This "prototype" mesh was then warped onto each femur from the training set, using the previously defined fiducial landmarks to

preserve correspondence.

A statistical shape model was then constructed from nine landmarked rabbit femurs. An algorithm to non-rigidly register the model to diseased examples was developed, based on Powell's conjugate gradient descent optimisation [53]. In order to quantify osteophytes in an input image, it was necessary to extrapolate the model over the diseased regions, rather than attempt to fit the model to them. This was achieved by computing a cost field from the input image, preventing the registration process from deforming the model to fit the osteophytes themselves. Upon convergence, quantification was achieved by subtracting the volume occupied by the model (corresponding to the healthy region) from the input image, leaving only the voxels belonging to osteophytes.

2.3.2 Dealing with non-linearity

In constructing a statistical model of shape, one must first consider the nature of the shape variation among the training examples. The PCA-based point distribution model as described in Section 2.2 assumes that a cluster of N landmarked training shapes can be reliably described by a $3N$-dimensional ellipsoid (or hyperellipse). Each axis of the hyperellipse is a mode of shape variation, where new shape instances can be generated as linear combinations of the modes. Put simply, any shape can be created by taking the average shape, and adding some contribution of each mode. In practice however, shape variation is to some degree non-linear, and falls under two categories; rotational non-linearity or high order non-linearity [54]. Rotational non-linearity is an inherent characteristic of objects that exhibit a large degree of bending, or have multiple parts that articulate with one another. Given a population of shapes with rotational variations, the cluster of training shapes in high-dimensional space will be curved in one or more dimensions, and so cannot be adequately represented by a linear system.

In the study of scoliosis (curvature of the spine), the orientations of vertebrae with

respect to one another are of interest, as well as their individual shapes. The articulated nature of the spine reveals the limitations of PCA-based shape models, where such attempts fail to represent variations in spinal anatomy satisfactorily. In one study, radiographic data of the spine was acquired by bi-planar radiography, with 16 surface landmarks annotated on each vertebra manually. Each vertebra was then registered onto its upper neighbour in turn, yielding a vector of inter-vertebral transformations. This shape representation does not belong in a vector space, and so cannot be analysed using conventional statistical methods. Instead, a Riemannian manifold geometry was adopted so that operations on the rigid transformations could be defined. A specialised multivariate technique named principal geodesic analysis (PGA) was developed to cope with the non-linear characteristics of spine shape [55].

An alternative approach to dealing with rotational non-linearity is to filter out differences in pose prior to shape modelling. This may be more appropriate in situations where rotational variation is less important than the shapes of individual structures. This approach was adopted in the construction a statistical shape model of the hip joint [56]. High order non-linearity is caused by inaccuracies in landmarking, usually in regions with highly variable curvature such as the sulcal folds of the brain. Such non-linearity poses a much greater problem in statistical shape modelling, and remains an active area of research.

Chapter 3

Materials and methods

This chapter details the animal models, protocols and procedures used in the mouse experiments undertaken as part of this research. As a number of different models were used in this research, only brief descriptions of the methods are given with citations to the original papers. Details of sample preparation, micro-CT imaging, reconstruction and processing are also provided.

3.1 Mouse models

3.1.1 Mouse acquisition and welfare

All murine experiments were performed in accordance with UK laws with the approval of the Home Office and local ethics committees. C57BL/6 wild-type were purchased from Harlan Laboratories, UK. Other strains of mice were bred and housed at the [redacted]. All mice were group-housed (3-6 animals per group) fed *ad libitum* standard chow diet and maintained on a 12 hour light/dark cycle. Mice were sacrificed either by cervical dislocation or exposure to carbon dioxide gas in a rising concentration.

3.1.2 K/BxN serum-transfer

Batches of serum from K/BxN mice were generous gifts from Dr Mohini Gray (University of Edinburgh, UK), Prof. Mauro Perretti (The William Harvey Research Institute, London, UK) and Prof. Harris Perlman (Northwestern University, Chicago, USA). 6-8 week old C57BL/6 mice were given a single intraperitoneal injection of 100 µl of K/BxN serum and developed arthritis 1-2 days later. Arthritis lasted for approximately 14 days. For micro-CT and histological analysis mice were culled at day 10 which represents the peak of inflammation. More details can be found in Monach et al. (2008) [9].

3.1.3 Collagen antibody induced arthritis (CAIA)

At day 0, eight to twelve week old C57BL/6 mice were intraperitoneally injected with 4 mg of mouse monoclonal anti-type II collagen 5-clone antibody cocktail (from Chondrex) [57]. At day 3, 25 µg of LPS (Chondrex) was injected intraperitoneally. The mice were monitored for paw inflammation, gait and weight from day 0. Inflammation reached peak severity at 7 - 9 days, and lasted for approximately 16 days.

3.1.4 TNF dARE

The TNFδARE/+ [B6.129-Tnftm2Gkl/Flmg] mouse (hereafter referred to as the TNF dARE mouse) was provided by [redacted] "Alexander Fleming", Greece [Kontoyiannis et al. 1999]. Mice overexpress TNF α due to the deletion of the regulatory AU-rich elements in the promotor of the gene. Mice develop a progressive erosive arthritis spontaneously, which is observable at approximately 6 weeks of age. Mice were scored for signs of arthritis from 4 weeks of age (after weaning) until they were culled for analysis at 12 weeks old.

3.1.5 CD248 knockout mice

C57BL/6 CD248 knockout (-/-) mice were maintained at the [redacted]. CD248 knockout mice were originally a gift from [redacted] and are described in Nanda et al. (2006) [58].

3.1.6 TTP transgenic mice

A transgenic mouse line with a mutant form of the tristetraprolin (TTP) gene was produced with the assistance of genOway. A targeting construct with serene to alanine substitutions at codons 52 and 78 was constructed and transfected into 129Sv embryonic stem (ES) cells. Recombinant ES cells were then injected into C57BL/6J blastocysts and re-implanted into pseudo-pregnant females. Two TTP $^{+/AA}$ males were then used as founders of a breeding colony, prior to back-crossing against a C57BL/6J background for ten generations. The mice were subsequently maintained at the [redacted]

3.1.7 Tissue dissection and preparation

Having removed the skin and fur from the hind limb (avoiding the paw), mouse hind limbs were dissected at the hip joint. Samples were fixed at room temperature in 4 % neutral buffered formalin over 24 hours, and subsequently stored at room temperature in 70 % ethanol.

3.2 Physiological disease index assessment

In this research, two types of scoring system were devised to characterise mouse phenotype: *in vivo* physiological scoring, and micro-CT bone destruction scoring. Using the *in vivo* system, scoring was performed over the course of several weeks to examine disease progression in the living mouse. Arthritis was scored daily per paw as follows: 0 = nor-

mal, 1 = 1 joint affected, 2 = 2 joints affected, 3 = multiple joints affected. Therefore a maximum score of 12 was possible per mouse. Gait was assessed by observing the mouse's response to an empty cage and scored as follows: 0 = normal, 1 = subtle abnormalities in gait, 2 = "paddling" (exaggerated movement of limb to the side when walking), 3 = reluctance to stand on hind legs/reluctance to move/reluctance to use a limb/limping. Degree of inflammation was assessed using callipers to measure swelling in the footpads and in the ankles of the hind-feet. TNF dARE mice do not show measurable swelling of the joints, therefore for this model the degree of inflammation was measured by calipers but not used. Blinded scoring was undertaken by two or three observers in order to minimise confirmation bias.

Upon termination of an experiment, the mice were sacrificed and hind limb samples prepared for imaging. The micro-CT scoring system was then used to assess bone destruction that could not be observed in live mice. Bone destruction was characterised according to type, location and severity by three independent observers in a double-blind fashion. The results acquired using the two scoring systems were used to validate the methodology proposed in this thesis. Details of the systems used and the results obtained can be found in Chapter 6.

3.3 Histology

Whole mouse hind limbs were fixed for 24 hours in 4 % paraformaldehyde and decalcified at room temperature using 10 % EDTA at pH 7.4, changed twice a week. Decalcification was confirmed by x-ray analysis, taking between 3 and 6 weeks to decalcify completely.[1] The decalcified samples were then paraffin embedded and 5 μm transverse paraffin sections were stained with haematoxylin and eosin.

[1]Some of the samples were imaged by micro-CT prior to decalcification for comparison with corresponding histological sections.

3.4 Micro-CT imaging

Hind limb samples were imaged using a Skyscan 1172 micro-CT scanner (Bruker). Formalin fixed samples were placed in an appropriately sized tube, and submerged in 70 % ethanol to minimise beam hardening artefacts. Polystyrene was used to pack the tube to prevent movement of the sample during the scan. The sample tube was mounted on the sample stage using dental wax, and aligned to be perpendicular to the base (Figure 3.1).

The x-ray beam was set to a source voltage of 60 kV and source current of 167 μA. A 0.5mm aluminium filter was used to minimise soft x-ray emission. Projections were taken every 0.45 degrees at 1000 ms exposure, with an image pixel size of 13.59 μm.

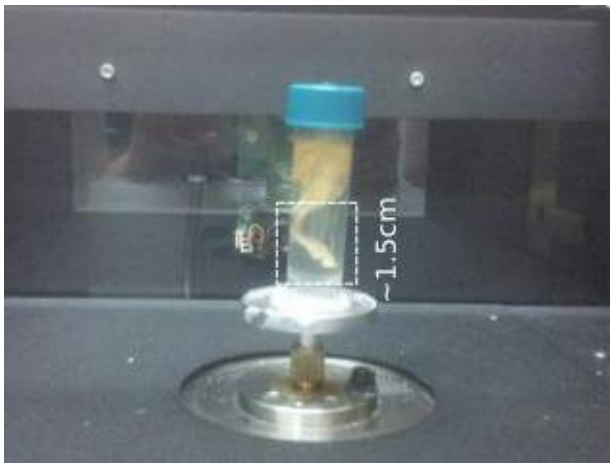


Figure 3.1: A hind paw sample, mounted in the micro-CT scanner. The dashed white box denotes the approximate field of view of the instrument.

3.5 Image and mesh pre-processing

3.5.1 Hardware and software specification

Image and mesh pre-processing was performed on a custom-built PC with an Intel Core i7 3.4 Ghz processor and 32 GB RAM, with Windows 7 Professional (64-bit) installed. In-house software was written in MATLAB 2010b (MathWorks), making use of the Image

Processing Toolbox, Optimization Toolbox and Statistics Toolbox. In addition, a number of freely available functions and scripts were obtained from the online File Exchange resource and used in development:

- `Kabsch`: calculation of rigid transformations
- `icp-finite`: iterative closest point (ICP) registration implementation
- `geom3d`: various functions for manipulating triangulated meshes
- `exportfig`: generation of figures from MATLAB

3.5.2 Image reconstruction

Flat field corrections were performed to remove any effects caused by varied pixel sensitivity or faulty detector elements. Image volumes (2000 x 2000 x 1187 isotropic voxels) were reconstructed using the Feldkamp algorithm [22] (NRecon 1.6.1.5, Bruker) having applied beam hardening correction to minimise cupping artefacts. Misalignment compensation was also performed as necessary to correct for "tail" artefacts (see Figure 3.2). A radiodensity range of -300.0 to 3000 Hounsfield units (HU) was chosen to isolate the bony structures from the imaging medium.

3.5.3 Segmentation of bony structures

CTAnalyser v1.12 (Bruker) image analysis software was used to extract an isosurface mesh representation of the mouse hind limb from the reconstructed micro-CT slices. Bony structures were first segmented from the background using a global threshold. The choice of threshold value was based on the requirement to delineate individual bones by maximising joint spacing, without introducing full thickness erosions where in reality, the cortical bone is intact but less dense. The threshold value was kept consistent for all experiments, ensuring that the same isosurfaces were being compared during registration

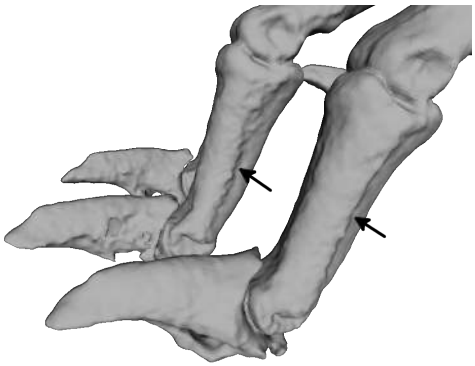


Figure 3.2: An example of a "tails" artefact, which occurs when the specimen is mounted off-centre in the micro-CT instrument. Although misalignment compensation can correct such artefacts to some extent, they are not always avoidable.

and statistical analysis. The marching cubes algorithm [35] was then used to extract a polygonal isosurface mesh from the binary image volume, and stored in Polygon File Format (PLY); a format suited to storing three-dimensional vertex and face data, as well as surface normals, colour, and transparency information for visualisation purposes [59].

3.5.4 Mesh processing

MeshLab 1.3.2 was used to modify the raw meshes generated by CTAnalyser. Meshes produced by marching cubes tend to have surfaces which may be described as "stepped", as shown in Figure 3.3 A. Poisson surface reconstruction [60] was therefore used to generate a smooth, uniformly sampled surface mesh that preserved the original surface toplogy (Figure 3.3 B). The next step was to remove internal structures (i.e. bone marrow) as only the outermost surface was to be included in the statistical shape model. This was achieved using ambient occlusion [61], which simulates the presence of light sources at various angles around objects in a scene (Figure 3.3 C). Each vertex in the mesh receives a value that is proportional to the number of light rays that reach the surface at that point. Internal structures receive low or zero-values (Figure 3.3 D) which can be isolated by thresholding over the ambient occlusion values, and removed (Figure 3.3 E).

In addition to the removal of unwanted structures, the total number of vertices (and

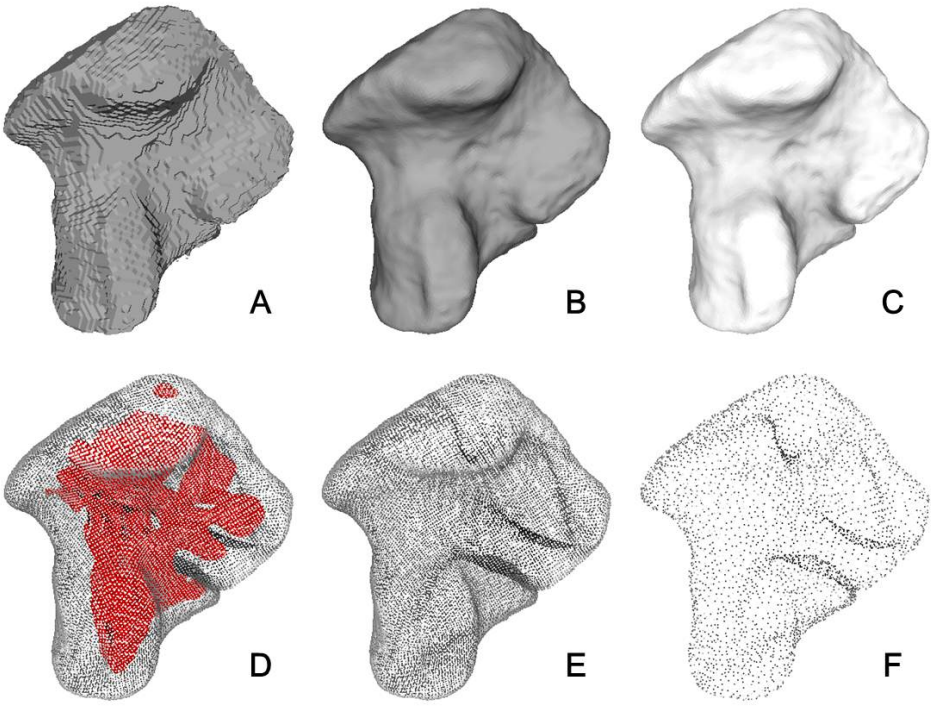


Figure 3.3: The mesh processing pipeline, as applied to a single bone. (A) Raw isosurface produced by marching cubes, with its characteristic "stepped" surface ($\sim 6.3 \times 10^4$ verts), (B) Smooth mesh achieved by Poisson surface reconstruction ($\sim 6.5 \times 10^4$ verts), (C) Poisson mesh with ambient occlusion values mapped onto the surface. Internal structures will receive a significantly lower value than points on the outer surface as they are not directly exposed to simulated ambient light, (D) Poisson point cloud with internal structures highlighted in red, achieved by thresholding over the ambient occlusion values, (E) Poisson point cloud with internal structures removed ($\sim 4 \times 10^4$ verts), (F) Simplified point cloud after quadric edge collapse decimation ($\sim 4 \times 10^3$ verts).

by extension, faces) was decimated. A hind-limb mesh generated from medium resolution image data is composed of approximately 2×10^6 vertices and 4×10^6 triangular faces. Mesh simplification was performed so that registration could be performed in a reasonable time frame, and the memory requirements for calculating covariance matrices and performing PCA calculations would be feasible. Quadric edge collapse decimation [62] was used to eliminate 90 % of the total vertices in the mesh, producing a simplified surface representation (Figure 3.3 F).

Chapter 4

Experimental implementation

4.1 The model construction framework

A pipeline was developed to generate articulated statistical shape models of the mouse hind paw from micro-CT image data. The proposed methodology is the subject of previously published work by the author of this thesis [63] (Appendix 1), and the details of its implementation are expanded upon in this section.

The process of creating a model consists of the following steps:

1. *Image pre-processing* (described in Section 3.5)
2. *Building an articulated model by manual annotation*
3. *Articulated registration of the model with samples*
4. *Statistical shape analysis.*

Steps 2 and 3 are inspired by previously reported methods for fully automated registration of a whole mouse skeleton atlas [29, 37] (discussed in Section 2.1.2). The methodology described here differs in several ways from the aforementioned, to account for the inherent differences in anatomy, joint complexity and proximity of individual bones:

- Additional coarse alignment of model and sample, to account for large variations in hind paw pose
- Rigid (rotation, translation) and affine (rigid + scaling) transformations are carried out in separate ICP registration steps to prevent falling into local minima
- Kinematic constraints using a viewing frustum, to avoid misregistration towards morphologically similar neighbouring bones.

4.1.1 Building an articulated model

4.1.1.1 Designing a mouse paw hierarchy

The mouse hind limb consists predominantly of the femur, tibia, fibula and hind paw (Figure 4.1). Much like the human foot, synovial joints can be found between the calcaneus, talus and distal tarsals (subtalar and transverse tarsal joints), the distal tarsals and metatarsals (tarsometatarsal joints), the metatarsals and phalanges (metatarsophalangeal joints) and between the phalanges (interphalangeal joints). This abundance of synovial joints make the hind paw a highly important anatomical structure in murine studies of inflammatory arthritis.

The hind paw is composed of 21 articulating bones [1], and a number of small non-articulating (sesamoid) bones beneath the metatarsals, which facilitate motion in the same way as the human kneecap (patella). Its articulated structure may be represented by a hierarchy or tree, where individual bones and joints are represented by nodes and connections respectively. Several possible hierarchies exist to model the articulation to varying levels of complexity. Although not a true reflection of the anatomy described in Figure 4.1, a simplified articulated model was deemed sufficient for the registration. An

[1]The exact number of bones varies depending on the gender, age and genotype of the mouse. In C57BL/6 mice for example, fusion of the 2nd distal tarsal and centrale/3rd distal tarsal varies between individual mice and is more prevalent in older females.

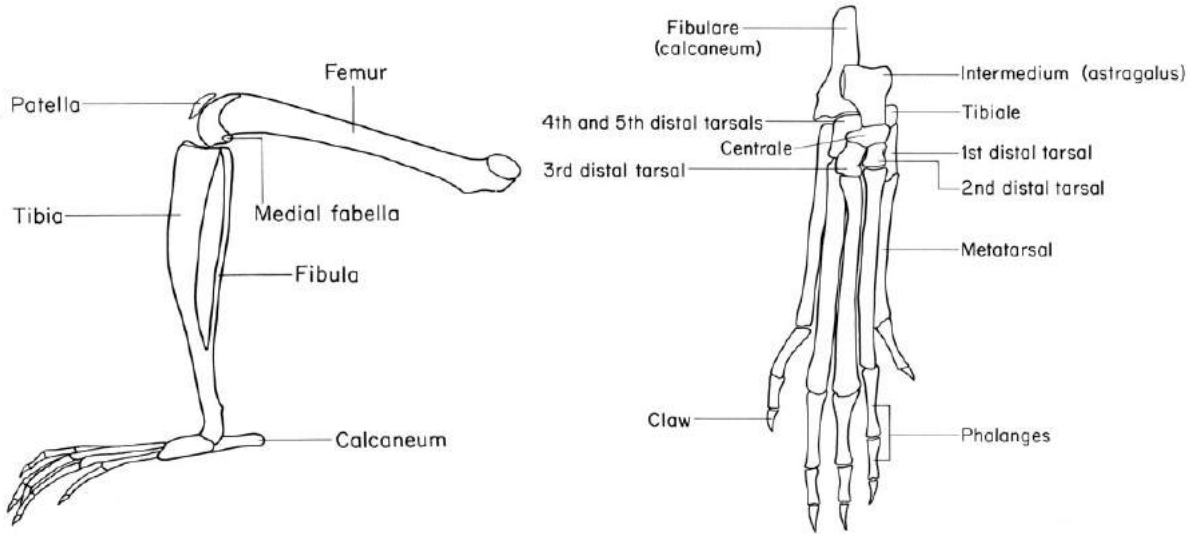


Figure 4.1: Skeletal anatomy of the mouse hind limb and paw, reproduced from *The Anatomy of the Laboratory Mouse* [64]. Left: the hind limb as dissected at the hip joint, showing the main components. Right: detailed diagram of the hind paw. In this particular diagram, the centrale and 3rd distal tarsal are drawn as separate bones, whereas they are completely fused in wild-type C57BL/6 mice. The articulated model incorporates only the bones shown on the right, as the upper limb is partially obscured by the limited field of view of the micro-CT instrument.

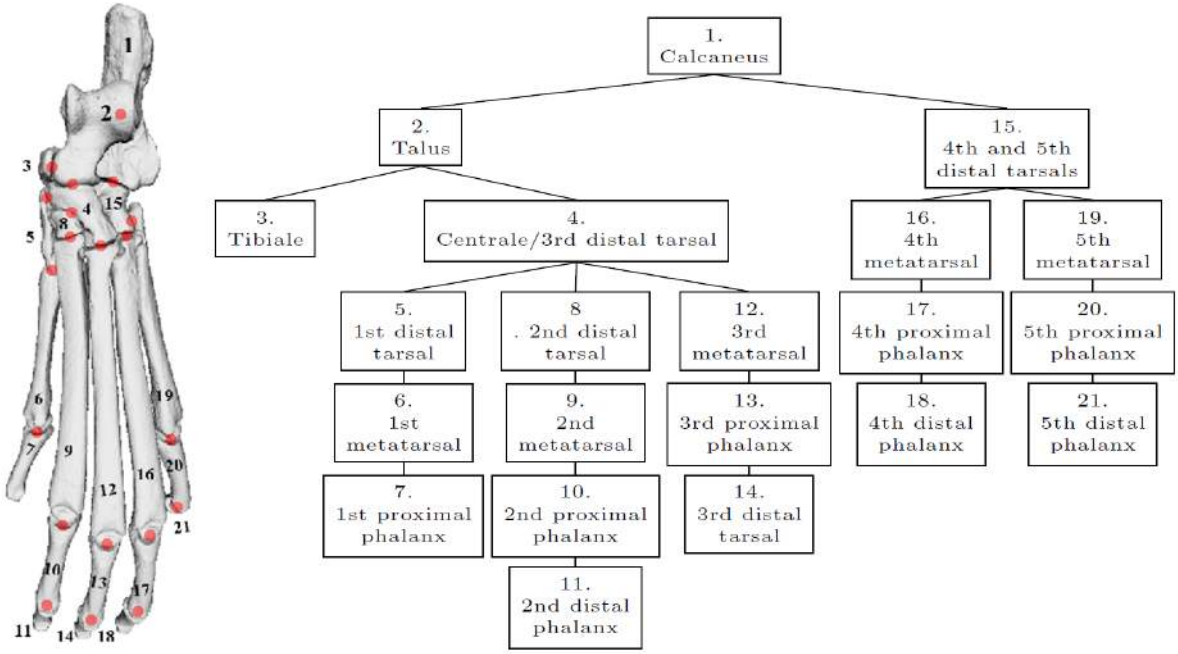


Figure 4.2: The hierarchy of the mouse hind paw, with corresponding labels. The connections between nodes represent joints, shown as red circles. The heel (calcaneus) bone is the root node of the hierarchy, and is therefore the first bone to undergo alignment. Transformations applied to any parent node are inherited by the entire subtree, allowing for the connected bones to be initialised and registered accordingly. This particular traversal order was selected from a number of less successful approaches which produced unsatisfactory registration results in initial experiments.

example hierarchy is shown in Figure 4.2. For any given pair of bones, the stationary (already registered) bone is described as the "parent", and the moving (to be registered) bone as its "child".

4.1.1.2 Manual segmentation and joint annotation

In order to initialise a statistical shape model of the mouse hind limb, a single sample was manually segmented to provide a labelled reference. A single C57BL/6 mouse paw sample was imaged by micro-CT and reconstructed according to the methods described in Chapter 3, Section 3.5. The sample was then segmented into its constituent bones using CTAnalyser (CTAn, Bruker) by manual contouring of individual slices. Having performed contouring on just a few slices, interpolation was used to generate contours

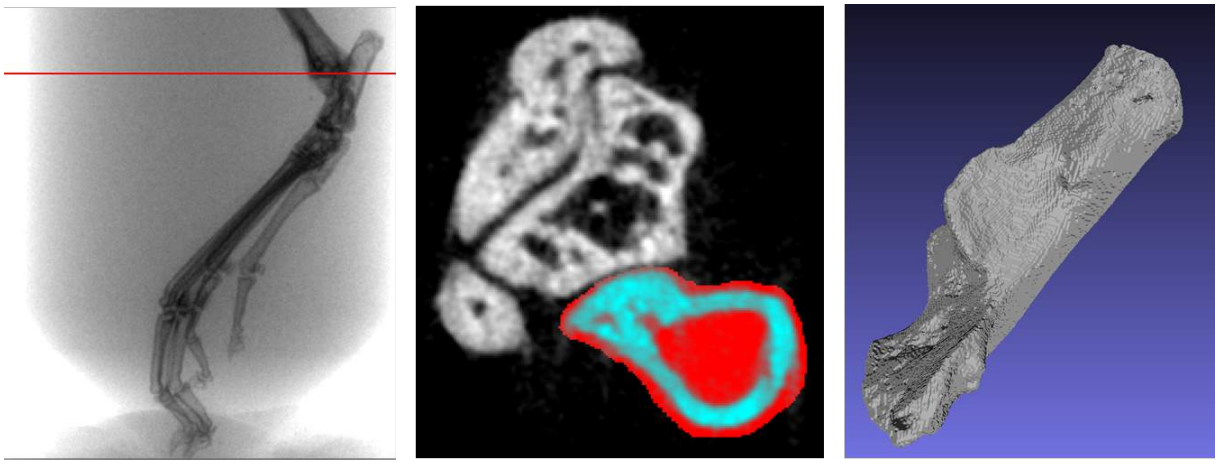


Figure 4.3: Manual contouring of individual bones using CTAnalyser. Left: individual projection of the mouse hind limb. The red line indicates the current slice being contoured. Centre: manually contoured cross-section of the calcaneus. Right: an isosurface mesh of the calcaneus after contouring and thresholding (rendered in MeshLab)

on the remaining slices. Interpolated contours were then manually edited to ensure that bones were segmented accurately. The entire micro-CT volume was then binarised by global thresholding and an isosurface mesh was generated for each bone using the marching cubes algorithm [35]. Figure 4.3 shows the result of manually contouring the calcaneus, alongside its corresponding isosurface mesh.

The locations of 20 joints were manually annotated as single point coordinates (shown as red circles in Figure 4.2) to define points of articulation. A set of motion constraints were then defined for each joint, the details of which are described in the next section. The articulated model incorporates only the kinematics (rigid transformations) of the hind paw; it does not model the physical forces and biomechanics that govern the motion of the bones.

4.1.2 Articulated registration of the model

4.1.2.1 Global alignment

The process of mounting specimens in the micro-CT instrument introduces variations in terms of their overall position and orientation. The aim of the global alignment process is to eliminate this variation, and provide an initialisation for the subsequent hierarchical registration. In order to prevent falling into local minima, the global alignment was designed to maximise the overlap between the model and sample by their root nodes, as specified by the model hierarchy. In cases where images of both left and right hind paws are acquired, the data is first manually mirrored to match the model.

The global alignment begins by coarsely aligning the model $\mathcal{M}$ and sample $\mathcal{S}$ based on their curvature. The curves were approximated from their reconstructed micro-CT volumes, having been binarised (see Section 3.5.3). For each sample, the volume was partitioned into n subvolumes along the longitudinal axis; the axis about which the sample is rotated in the micro-CT instrument. From each subvolume, the coordinates of non-background pixels are averaged

$$p_j = \frac{1}{b_j} \sum_{k=1}^{b_j} (x_k, y_k, z_k) \tag{1}$$

where b_j is the number of bone pixels in the j^{th} subvolume. The ordered set of points p_j $(j = 1, 2, \ldots, n)$ is referred to as a centreline, which approximates the overall curvature of the paw.

Using the known (albeit approximate) correspondence between model and sample curve points, a transformation is computed that minimises the root mean squared deviation (RMSD) between them. This is performed using the Kabsch algorithm [65], producing a rigid transformation composed of a translation and a rotation.

A centreline curve $\mathcal{G}$ is represented by an $n \times 3$ matrix of coordinates

$$\mathcal{G} = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix} \tag{2}$$

The centroid of each curve is calculated as the mean of the curve points, which is subtracted to centre the curve on the coordinate system origin $(0, 0, 0)$.

$$\mathcal{G}_{origin} = \begin{bmatrix} x_1 - \overline{x} & y_1 - \overline{y} & z_1 - \overline{z} \\ x_2 - \overline{x} & y_2 - \overline{y} & z_2 - \overline{z} \\ \vdots & \vdots & \vdots \\ x_n - \overline{x} & y_n - \overline{y} & z_n - \overline{z} \end{bmatrix} \tag{3}$$

This may also be defined as a translation matrix

$$T = \begin{bmatrix} 1 & 0 & 0 & -\overline{x} \\ 0 & 1 & 0 & -\overline{y} \\ 0 & 0 & 1 & -\overline{z} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Two translation matrices are in fact computed; one for the model and another for the sample ($T_{\mathcal{M}}$ and $T_{\mathcal{S}}$, respectively). A rotation matrix $R_{\mathcal{M}}$ is then computed which when applied to the model curve, rotates it to have the same orientation as the sample curve. The number of curvature points was chosen by increasing the value of n until no further reduction in RMSD was observed. The optimal value was found to be 20, which is approximately one point for every 60 micro-CT slices for an average volume.

Using the transformations $T_{\mathcal{M}}$, $T_{\mathcal{S}}$ and $R_{\mathcal{M}}$, the model mesh is coarsely aligned with the sample mesh. First, the model is translated to the origin by $T_{\mathcal{M}}$. It is then rotated

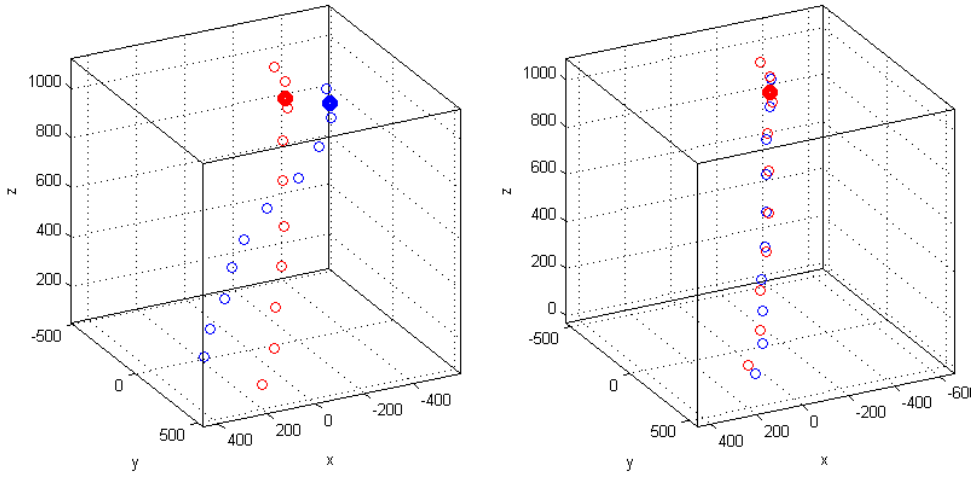
by $R_{\mathcal{M}}$, before being translated back by the inverse of $T_{\mathcal{S}}$ (i.e. back to where the sample mesh is located). Mathematically, this is defined by three matrix multiplications applied to the model points $\mathcal{M}$:

$$\mathcal{M}_{coarse} = T_{\mathcal{S}}^{-1} \cdot R_{\mathcal{M}} \cdot T_{\mathcal{M}} \cdot \mathcal{M} \tag{5}$$

Having brought the model and sample into approximate global alignment, an additional adjustment is performed to ensure that the topmost bones of the hierarchy are sufficiently well aligned prior to iterative closest point registration. This is achieved by approximating the location of the subtalar joint (where the tibia meets the ankle) as the point of highest curvature. For every point in the upper half of the curve p_j ($j = 2, 3, \cdots, \frac{n}{2}$), two vectors are computed: $v_1 = p_j - p_{j-1}$ and $v_2 = p_{j+1} - p_j$. The angle θ_j between the vectors is computed for every p_j, and the largest among them selected as the point of highest curvature. The model is then translated by the difference in the two points. The globally aligned model $\mathcal{M}_{global}$ is then provided as input to the hierarchical registration procedure.

4.1.2.2 Hierarchical registration

The iterative closest point (ICP) algorithm [31] is used to align the individual bones of the articulated model with the sample. The global alignment initialises the first bone in the hierarchy, with each successive registration initialising the next in line. The intermediate steps of the registration are shown in Figure 1. In early experiments, poor initialisations led to inadequate registration results. This was due to the inherent similarities between bones of the same type, such as the proximal phalanges. To counter this, model bones were grouped into three shape categories: small bones (tibiale, 1st and 2nd distal tarsals), irregular bones (talus, calcaneus, 3rd - 5th distal tarsals) and long bones (5 × metatarsals, 9 × phalanges). Small and irregular bones have little rotational freedom, allowing for both



(a) Before global alignment (b) After global alignment

Figure 4.4: Coarse alignment of a model (blue) and sample (red) centreline curve. The two filled markers are the approximate location of the subtalar joint. (a) The two curves in their original locations. (b) The result of applying the transformation from Equation 5 to the model curve (as opposed to the model itself), followed by adjustment according the points of highest curvature.

rigid and affine transformations to be solved for simultaneously without constraint. Long bones undergo larger rotations however, and require separate rigid and affine registration steps. For the small and irregular bones, the ICP algorithm is iterated until the difference in SSD between iterations falls beneath a specified threshold:

$$\frac{SSD_i}{SSD_{i-1}} < \mathcal{T} \tag{6}$$

A value of $\mathcal{T} = 10^{-6}$ was found to produce suitable alignments in initial experiments, based on visual inspection of the closeness of fit for each bone. For long bones, the rigid transformation constraints were achieved by reducing the search space of the ICP algorithm to a feasible subset of points. This process is known as *viewing frustum culling*; a technique used in computer graphics for rendering of 3D scenes [66].

To define a viewing frustum, a virtual camera system composed of three orthonormal vectors is initialised at a reference point VRP. The virtual camera axes are initialised

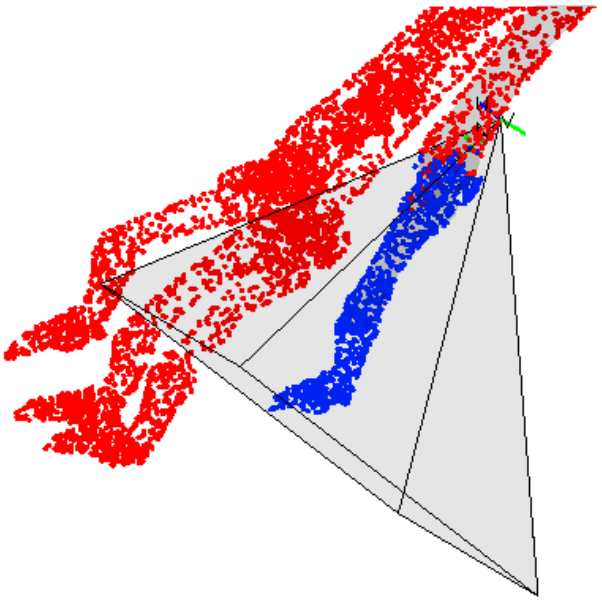


Figure 4.5: An example of viewing frustum culling applied to first proximal phalange. The virtual camera is defined at the terminal joint of the parent bone, from which the viewing frustum eminates. Feasible sample points (those within the frustum) are shown in blue, with culled points (those outside the frustum) are shown in red. Culled points are eliminated from the ICP search space, and are not used by the ICP algorithm during the rigid step. After rigid alignment, affine alignment is performed without these constraints in place.

for each model bone based on the manually annotated joint locations, and two planes defined at the near (proximal) and far (distal) ends of bone. The relative dimensions of the planes are defined by four angles with respect to the camera coordinates (up, down, left and right). The set of frustum angles used in experiments are given in Figure 4.6. These parameters were defined once only, based on the inherent variation observed among ten wild-type C57BL/6 samples. The near and far planes are then connected to produce six planes in total, forming a square frustum (the lower portion of a square pyramid) as shown in Figure 4.5. Finally, the viewing frustum is used to eliminate the sample points that lie outside its confines, through a series of normal vector comparisons. The model bone is registered towards the "feasible" sample points by ICP, solving for translation and rotation only. This initialises a second and final ICP step, solving for both rigid and affine

transformations (in the same fashion as small and irregular bones). Unlike the previous step, the culled points are included in the search space. Upon registering the final bone in the hierarchy, the final result is a complete piecewise alignment of the articulated model with the sample. An example result is shown in Figure 4.7. A diagram of the whole process is given in Appendix 2.

Bone name	Frustum angle (radians)			
	Up	Down	Right	Left
1st distal tarsal	π/8	π/8	π/32	π/16
1st metatarsal	π/8	π/8	π/16	π/8
1st proximal phalanx	π/32	π/8	π/32	π/8
2nd metatarsal	π/8	π/8	π/16	π/8
2nd proximal phalanx	π/8	π/8	0	π/8
2nd distal phalanx	π/8	π/8	π/8	π/8
3rd metatarsal	π/8	π/8	π/32	π/32
3rd proximal phalanx	π/8	π/8	π/128	π/8
3rd distal phalanx	π/4	π/4	π/128	π/8
4th metatarsal	π/8	π/8	π/32	π/32
4th proximal phalanx	π/8	π/8	π/8	π/8
4th distal phalanx	π/4	π/4	π/8	π/8
5th metatarsal	π/8	π/8	π/8	π/32
5th proximal phalanx	π/8	π/8	π/8	π/16
5th distal phalanx	π/8	π/8	π/8	π/8

Figure 4.6: Frustum parameters used for articulated registration.

4.1.3 Statistical shape modelling

In order to model bone shape variation across a set of samples, individual instances of the model are registered to each of the samples and the associated transformations stored. Statistical shape analysis requires that the individual training shapes be superimposed, and so it is necessary to perform an inversion step to bring the set of samples into a common coordinate frame.

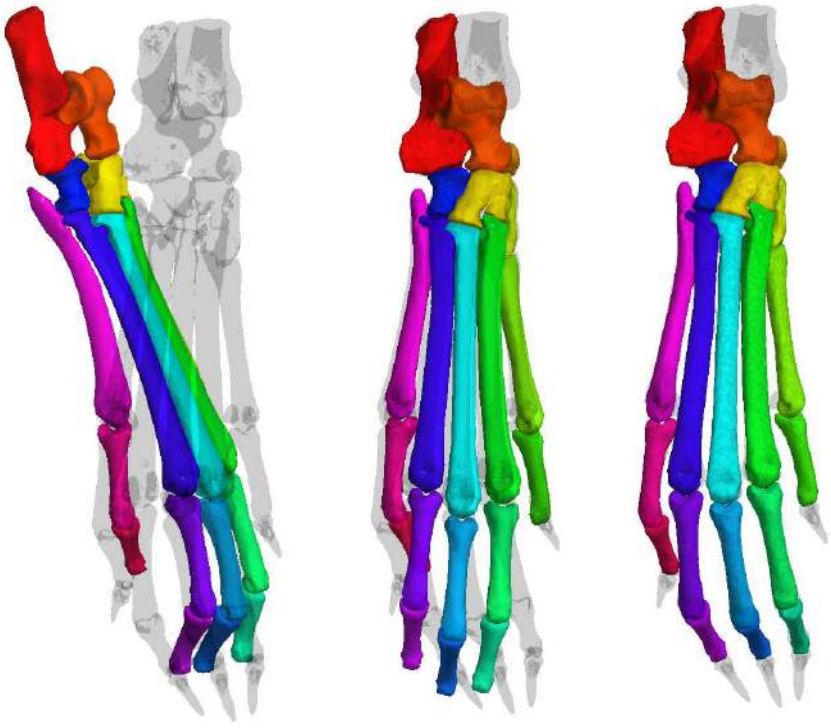


Figure 4.7: Model and sample shown before registration, after global alignment and after hierarchical registration. It should be noted that the most differences in pose are due to misalignment of the metatarsals and phalanges (with emphasis on the latter). This is particularly evident in the example above, where the uppermost bones are already in close alignment prior to piecewise registration.

4.1.3.1 Label propagation

The articulated registration procedure provides correspondence between the model and sample points. Using this information, labels may be propagated from model to sample, partitioning it into its as yet unlabelled bones. If a one-to-one mapping were to exist between the model and sample points, then a complete labelling of the sample mesh would be achieved. In all likelihood however, a number of sample points will receive a "null" label (Figure 4.8 (A). Furthermore, points located on the tibia, fibula, sesamoid bones and claws (which are not part of the model) may be incorrectly assigned labels (Figure 4.8 (B)). Such points are generally sparsely distributed over the sample mesh,

and can be corrected using a local "mode" filter (Figure 4.8 (C)). The filter considers the set of k-nearest neighbours (by Euclidean distance) around each unlabelled point, and the most common label among them is chosen to replace the non-labelled vertex. The value for k was chosen as the mean vertex connectivity over the whole mesh (the number of edges connected to a vertex on average). Optimally, the result is a labelled sample mesh, with only non-model regions having the "null" label. In practice however, some mislabelling does still occur and inevitably introduces some error in model construction.

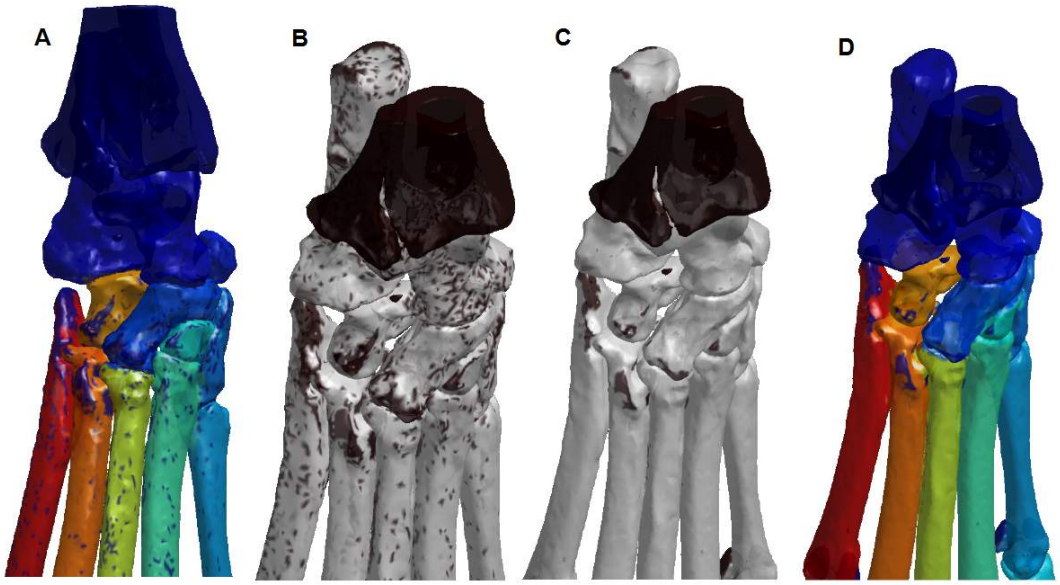


Figure 4.8: Label propagation from model to sample, showing intermediate steps. An alternative colour scheme is used here in order to distinguish it from the model colour scheme. (A) The initial labelling of the sample by nearest neighbour. (B) Binary labelling showing labelled regions in grey, and "null" labelled patches in black. (C) The result after applying the mode filter, correcting small mislabelled patches. (D) The final labelling of the sample after correction.

4.1.3.2 Inverse transformation

During articulated registration, individual transformations were stored in order for them to be applied to the labelled sample mesh. For each model bone m_i, its transformation T_i from the registration is inverted to produce an inverse transformation T_i^{-1}. All rigid

and affine transformations have a unique inverse, such that $TT^{-1} = T^{-1}T = I$ (where I is the identity matrix). That is to say, the forward transformation followed by its inverse (or even the other way around) has no practical effect.

It is important to note that in the context of articulated registration, the individual transformations are relative with respect to the previously registered bone. In order to bring a sample into correct alignment with the model, inverse transformations are performed in reverse hierarchical order (i.e. the last model bone registered is the first sample bone to be inversely transformed). The result of this process is complete alignment of the sample with the model in its original coordinate system. This was applied to all of the samples used in model construction, superimposing them with the model.

4.1.3.3 Principal component analysis

Using the model points as a reference and the known point correspondence across the training set, an average (mean) shape was calculated (Figure 5.4). In order to introduce shape variation into the articulated model, a statistical shape model was constructed for each bone separately using principal component analysis. The technical details of this process are described in Section 2.2.

4.2 Application of an ASSM

Using the framework described in the previous section, an ASSM may be constructed that is representative of a population of hind paw samples, controlling for factors such as age, gender and genetic background. The ASSM then may be used to analyse arbitrary hind paw samples, with the following steps:

1. *Image pre-processing* (described in Section 3.5)

2. *Articulated registration of the model* (in part, described in Section 4.1.2)

3. *Shape model fitting*

4. *Abnormality detection and analysis.*

4.2.1 Articulated registration of the model

This step is identical to the process used for model construction as described in Section 4.1.2. The only difference here is that rather than registering the annotated reference sample (as used in model construction), the average model shape is registered. This facilitates the next step of the process, where non-rigid deformation of the average shape is performed.

4.2.2 Shape model fitting

Section 2.2 presented how statistical shape models can be used to generate arbitrary instances of a class of shape, as defined by a training set. Under the proposed scheme, the ASSM is trained exclusively with non-pathological samples to produce a model of "normal" bone shape. Any given instance of the model should therefore only consist of bone shapes that could realistically be found among normal mice. The objective of the shape model fitting step is to find an instance of the model that closely approximates the sample being analysed. By constraining model instances to fall within three standard deviations of the mean shape, only non-pathological differences are "filtered out", leaving those that might have arisen due to pathology.

Prior to shape model fitting, the sample is labelled and inversely transformed into the model coordinate system as described in Sections 4.1.3.1 and 4.1.3.2. Shape model fitting can then be performed for each bone in the model independently. Initially, correspondence

between model and sample points is provided by the articulated registration. A non-rigid deformation is then found that minimises the distances between the model points and corresponding sample points. Point correspondence is then re-evaluated, and the process repeats until convergence. The algorithm described here is similar to that which is used in active shape model fitting [67]. The articulated registration effectively initialises the fitting process.

The shape of a sample bone may be approximated by a non-rigid deformation of the average bone shape. This is defined by a linear combination of the principal components, as computed by PCA during model construction:

$$x \approx \overline{x} + Pb \tag{7}$$

The vector b parametrises the deformation, which may calculated as

$$b = P^T(x - \overline{x}) \tag{8}$$

where $\overline{x}$ and x are the mean model points and corresponding sample points, respectively. Following deformation, the process is iterated for a fixed a number of iterations n. A value of $n = 100$ was used for this work, above which neglible deformation was observed for any of the shape model experiments.

4.2.3 Abnormality detection and analysis

4.2.3.1 Signed Euclidean distance and heatmap visualisation

The process of fitting the model onto a sample allows for identification of abnormalities in terms of fitting error. Regions where the model and sample show good agreement are assumed to be non-pathological. Conversely, regions where the sample deviates from the model statistics are regarded as abnormal. Numerically, the Euclidean distance between

corresponding points can be used as a measure of how closely the model matches the sample. Let $\vec{m} = \langle m_1, m_2, m_3 \rangle$ and $\vec{s} = \langle s_1, s_2, s_3 \rangle$ be a model point and its corresponding sample point, respectively. The unsigned Euclidean distance is calculated as:

$$E(\vec{m}, \vec{s}) = \sqrt{(m_1 - s_1)^2 + (m_2 - s_2)^2 + (m_3 - s_3)^2} \tag{9}$$

In order to account for differences in the nature of the abnormalities, the signed Euclidean distance (SE) can also be calculated. The sign is determined by the relative positions of model and sample points, or how their surfaces sit above or below one another. The model and sample meshes each consist of a set of points (vertices) and connections (edges) between them. The meshes were constructed to consist entirely of triangular faces (three vertices with three edges). For any given face, a vector that lies perpendicular to its surface can be computed as the cross-product of any two of its edges. This process is performed for every face in the mesh, yielding a complete set of "face normals". Following this, a set of inner vertex normals is computed. For each vertex in the mesh, the set of faces connected to it are found, and the average of their inner normal vectors computed to give an inner vertex normal $\vec{n}$. Using this knowledge, the signed Euclidean distance may be calculated:

$$SE(\vec{m}, \vec{s}) = E(\vec{m}, \vec{s})(\mathbf{sgn}(\vec{n} \cdot (\vec{m} - \vec{s}))) \tag{10}$$

where the function **sgn** outputs the numerical sign of a number (either -1, 0 or +1). Zero (or near zero) values are where the model and sample surfaces closely match one another. Positive values indicate the presence of bumps, whereas negative values are due to indentations. Physiologically, this translates into regions of bone resorption (loss) and deposition (gain), respectively. This can be visually represented as a heatmap, as shown in Figure 4.9.

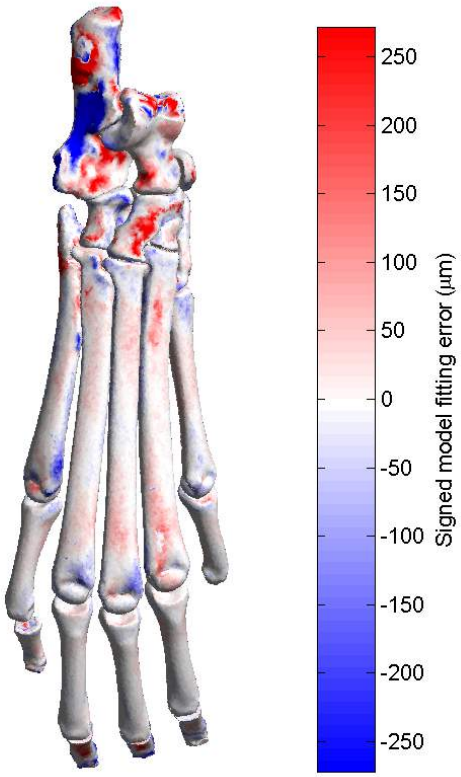


Figure 4.9: Heatmap visualisation of model error. In the example shown, the fitted model instance is coloured according to the signed error values (μm) at each point (voxel size is 13.59 μm). The red and blue regions of the colour spectrum correspond to bone formation and bone erosion, respectively.

4.2.3.2 Delineation of abnormal regions

In order to provide further meaningful numerical data about bone destruction, abnormal regions were delineated and measured in terms of their surface area. A tolerance level was first defined to account for imperfections in registration and shape model fitting. The threshold was determined by calculating the average Euclidean distance between model and sample points across the original training set. The statistics were computed for each bone independently, rather than for the hind limb as a whole, to account for the variability in registration accuracy between different bones. In the results presented

below, error values within three standard deviations of the mean error (per bone) were regarded as normal. Any points with error values above this threshold were therefore regarded as abnormal, and used in surface area calculation.

In order to separate abnormal patches into measurable areas of erosion and formation, individual faces were classified according to the distance values at vertices of neighbouring faces. A pre-computed library of rules were defined to assign each face one of the three labels; negative (bone erosion), zero (normal) or positive (bone formation). Each vertex is then visited in turn, and the surrounding faces are assigned labels based on the current vertex label, and the labels at neighbouring vertices. Given an input configuration of labels in clockwise order, the library provides the appropriate assignment of face labels. The trivial case is where an individual face consists of three normal vertices, and therefore receives a zero value itself. Non-trivial cases emerge when patches consist of a combination of negative, positive and zero values. An example is given in Figure 4.10. As each case emerges, the relevant rule is extracted from the library and the correct label assigned. The library was constructed to deal with between three and eight adjacent faces. This ternary labelling system provides a segmentation of the mesh surface, as shown in Figure 4.11a.

Bone surface areas were calculated by summing the areas of individual mesh triangles. For any given triangle $\Delta = \vec{v_1}\vec{v_2}\vec{v_3}$, its area is given by:

$$A(\Delta) = \frac{1}{2}|(\vec{v_2} - \vec{v_1}) \times (\vec{v_3} - \vec{v_1})| \tag{11}$$

For each of the three labels, the sum total surface area was computed. The degree of bone destruction measured for each bone as a whole can then be expressed as percentage coverage e.g. $x\%$ erosions, $y\%$ formation. In order to provide a more detailed characterisation of the abnormalities present, the surface areas of individual patches were computed. An queue-based algorithm for connected-component labelling was implemented for trian-

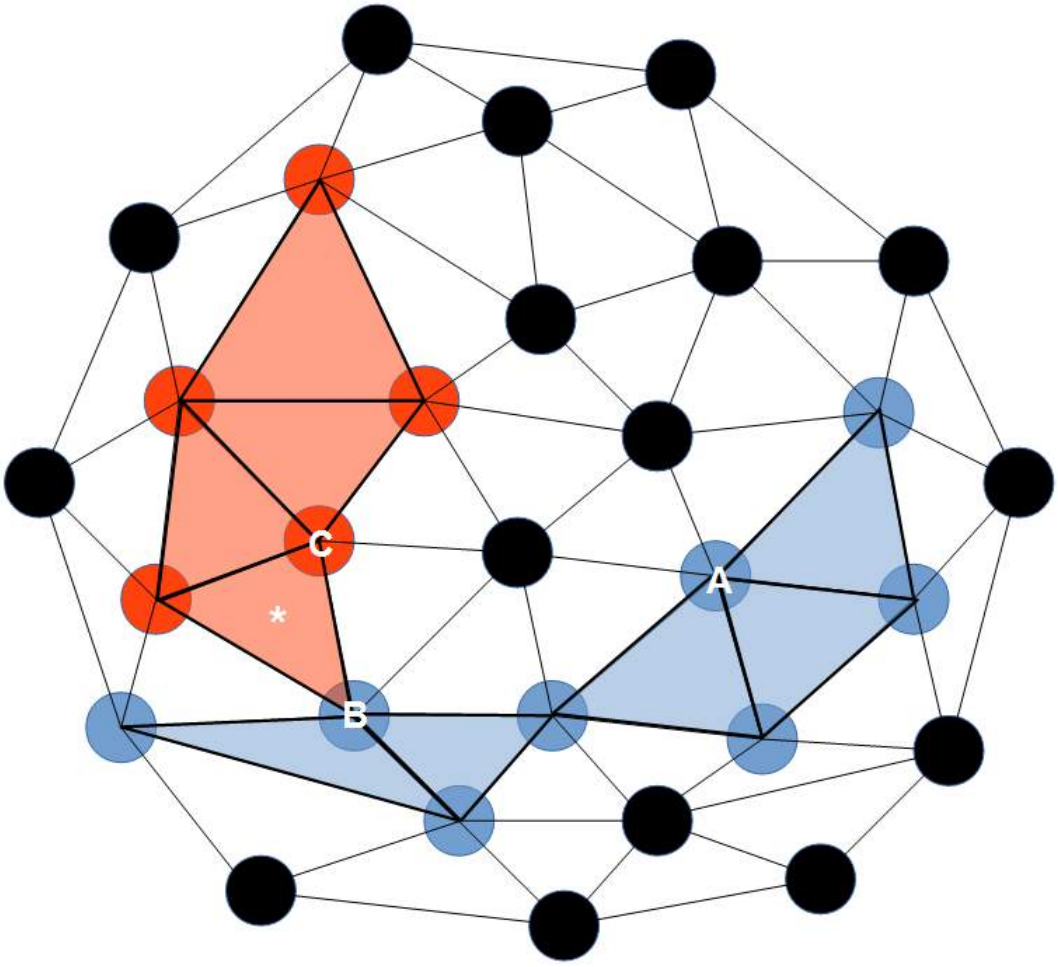
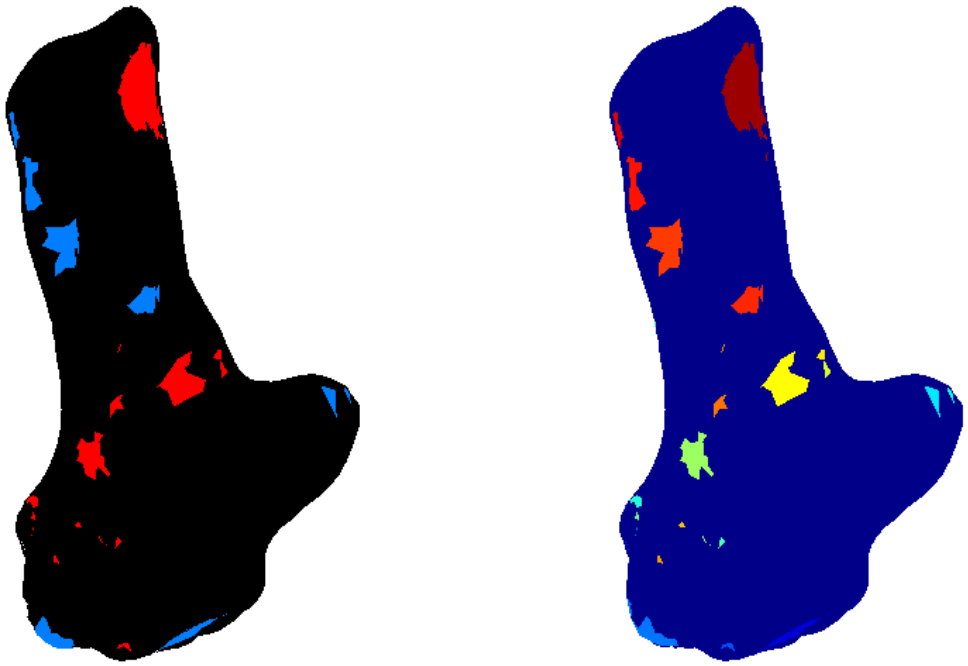


Figure 4.10: Labelling of faces as regions of bone erosion/formation. Each vertex is visited in turn, and the labels of surrounding faces assigned based on its own label. Vertex A shows a simple case, where its blue label is given to three of the surrounding faces, due to each of them having two blue vertices. At vertex B, despite there being a mixture of red and blue vertices, only the two blue faces are assigned. The face marked by an asterisk does not receive its red label until vertex C is visited.

gulated meshes (Algorithm 1), with an example result shown in Figure 4.11b. The surface areas of each abnormal patch may then be calculated, giving a distribution of sizes that may be visualised as a histogram.



(a) Ternary labelling (b) Connected component labelling

Figure 4.11: Delineation of abnormal regions on the calcaneus. (a) Ternary labelling of normal patches (black), erosions (blue) and formation (red). (b) Connected component labelling of patches, with normal bone shown in dark blue and separate abnormal patches shown in randomised colours.

Algorithm 1 A non-recursive algorithm for connected-component labelling of triangulated meshes

Initialisation:
$BW \leftarrow$ binary labelling of mesh faces
$N \leftarrow$ number of faces in the mesh
$visited \leftarrow$ zero array of length N
$CC \leftarrow$ zero array of length N
Q.**push**(0)
$currentLabel \leftarrow 1$

Outer loop:
while Q is *not empty* **do**

    $f \leftarrow Q[0]$

    if $BW[f] == 1$ **then**
        W.**push**(f)

        Inner loop:
        while W is *not empty* **do**:

            $g \leftarrow W[0]$
            $CC[g] \leftarrow currentLabel$
            $BW[g] \leftarrow 0$

            $abnormalNeighbours \leftarrow$ indices of faces adjacent to g (where $BW == 1$)
            $BW[abnormalNeighbours] \leftarrow 0$
            W.**pop**()
            W.**append**($abnormalNeighbours$)

        end while

        $currentLabel \leftarrow currentLabel + 1$

    end if

    Update:
    $visited[f] \leftarrow 1$
    $notVisited \leftarrow$ indices of faces adjacent to f (where $visited == 0$)
    Q.**pop**()
    Q.**append**($notVisited$)

end while

Chapter 5

A wild-type model of the mouse hind paw

5.1 Articulated registration of wild-type samples

The articulated registration framework was used to construct a "normal" model of the mouse hind paw from a set of wild-type samples. In accordance with most of the experimental designs described in Section 3.1, the mice were all 12 week old females on a C57BL/6 background. Ten hind paw samples were prepared and imaged according to the protocols outlined in Sections 3.4 and 3.5. One sample was selected at random to be the reference, and manually annotated to produce an articulated model consisting of labelled bones and joints between them. Separate instances of the model were then registered to each of the remaining samples, one by one. The results of the registration shown here are given as the mean Euclidean distance between corresponding model and sample points.

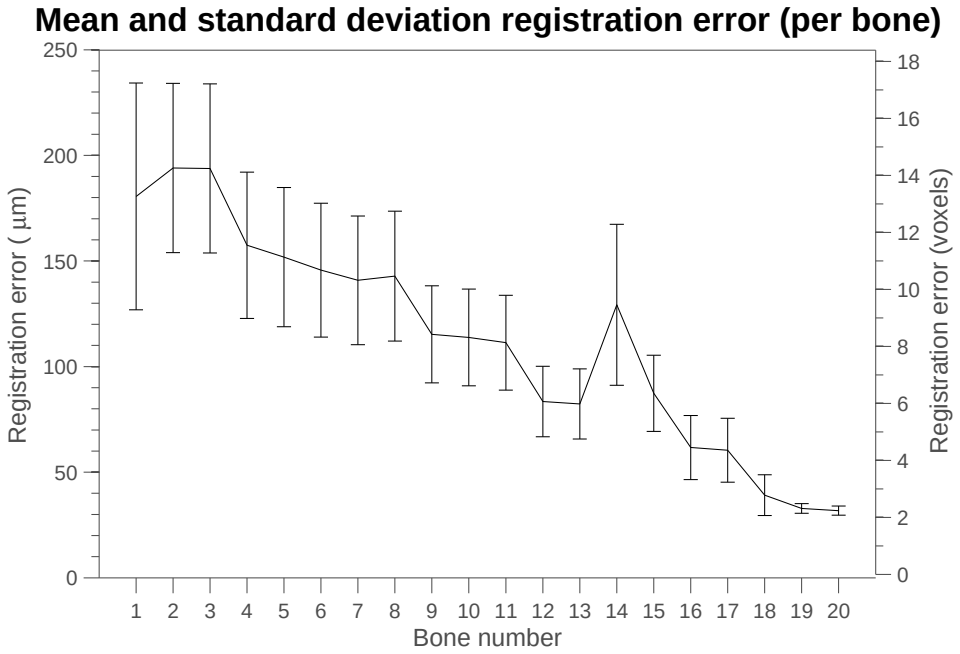


Figure 5.1: Mean registration error over all nine samples, with error bars at one standard deviation. Despite large variations among the samples in terms of initial pose, the registration algorithm performs consistently across all nine samples ($\sim 2.5 \pm 1.5$ voxels, after the final bone is registered). The non-monotonic decrease in registration error suggests that local alignment of bones does not necessary result in an overall decrease in error, due to the articulated nature of the samples.

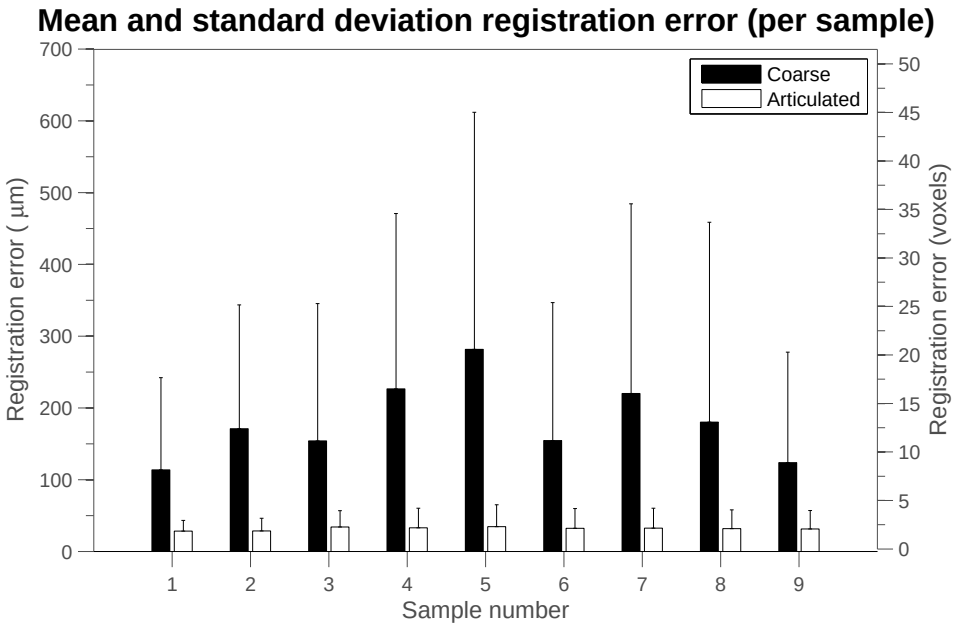


Figure 5.2: Mean registration error after coarse and articulated stages, shown for each sample individually. Error bars represent one standard deviation. Comparing the coarse and articulated steps together for each sample, it is clear that large variations in bone position are handled well by the registration algorithm.

Figure 5.1 shows the results of the registration at each intermediate step. One of the most striking behaviours of the articulated registration is the non-monotonic decrease in mean registration error as registration proceeds down the anatomical hierarchy. Although the registration error will always decrease locally after a successful registration step, there is no guarantee that there will be a decrease in the overall registration error. This is most prominent when registering the 4th/5th distal tarsal (#14), as indicated by the sharp increase in mean registration error. This is due to the fact that the unregistered child bones inherit the transformations of the parent, regardless of how close they already are to their intended destinations. Overall, the registration produces consistent results across the ten samples as shown in Figure 5.2. Visualisations of the results are shown in Figure 5.3.

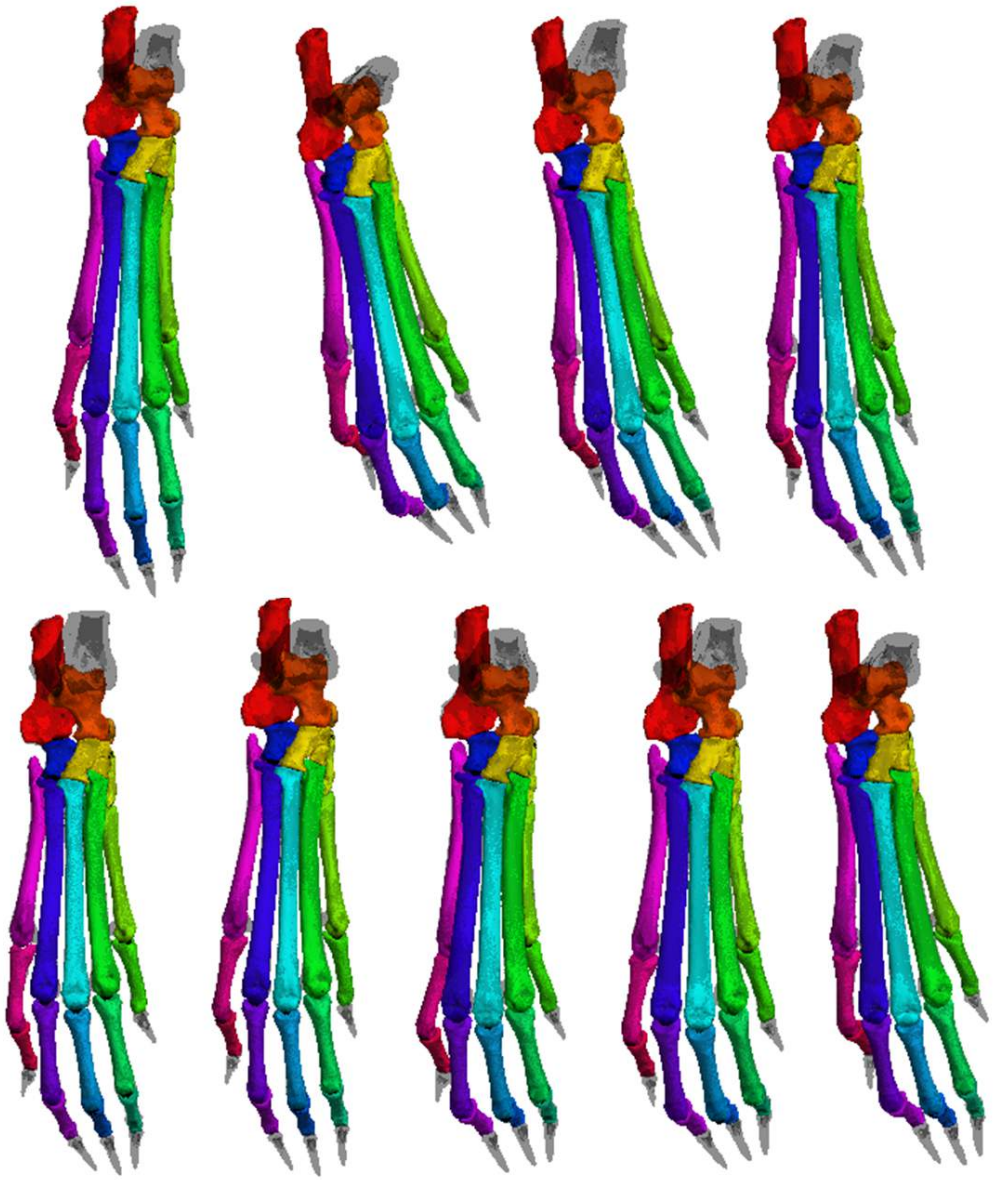


Figure 5.3: Results of registration for nine wild-type samples. Qualitatively, the registration framework produces consistent results across all samples, correctly registering all of the bones in the hierarchy. Variations in bone shape are not obvious at this scale, but can be seen most clearly in the calcaneus bone (red).

Figure 5.4: Superimposition and average of ten wild-type samples after registration. On the left, the superimposition of nine sample hind paws with the model is shown. Regions depicted in gray are either non-model parts (e.g. claws), variations in bone shape, or misalignment of individual bones. On the right, the average of the samples is shown based on the correspondence derived from the registration. Each bone can be considered in isolation, and used in the subsequent statistical shape analysis step.

5.1.1 Sensitivity to initialisation

The iterative closest point (ICP) algorithm is highly sensitive to its initialisation [31]. This becomes increasingly significant in the proposed framework, where multiple successive registrations are dependent on the success of their predecessors. Registration is deemed to have failed when severe misalignment of a parent causes the entire subtree to fail. When such cases emerged, the point of failure was found to be early on in the registration, highlighting the significance of a good initial alignment. This is supported by Figure 5.5, showing final registration error as a function of registration error after coarse alignment.

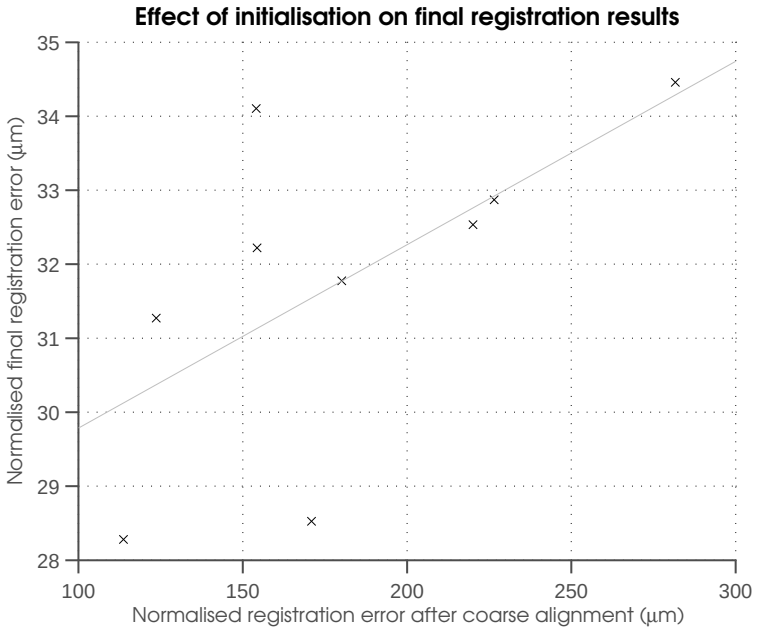


Figure 5.5: The effect of initialisation on final registration quality, showing the line of best fit in grey. The positive correlation suggests that in most cases, a poor initialisation will lead to a worse registration result overall. This is likely due to cumulative errors introduced during the course of the articulated registration.

5.2 Labelling of registered wild-types

Having registered the ASSM to each of the samples in turn, its labels were propagated based on their correspondence to provide labelled sample meshes. This allowed for each sample to be inversely transformed into a common coordinate system, as a precursor to statistical shape analysis. Due to the lack of manual segmentations for validation, the quality of the automated mesh labelling was assessed through visual inspection and identification of mislabelled regions. The correspondence provided by the registration determined the accuracy of the labelling, and so small deviations led to small patches of bones being mislabelled. This was observed in all of the samples registered, but were generally few in number and size. The local "mode" filter was effective in eradicating small isolated patches, replacing the labels with the most likely label based on the surrounding neighbourhood. Such patches could be observed between the metatarsals, where

adjoining bone surfaces come into close contact. Another example of mislabelling was observed around the joints, where labels had "bled" onto connected bones due to slight misalignment.

Labelling of the distal tarsals and metatarsals highlighted a limitation of the proposed method, most notably where two or more bones are fused. In such instances, the bones are composed of a single continuous surface which, upon separation by label, introduces holes in the bone surface. The observed fusion of some bones may be attributed to the resolving power of the micro-CT instrument, image artefacts, or the choice of threshold value for isosurface generation. An alternative, less avoidable explanation is that the bones are genuinely fused, which can be observed in other rheumatic diseases such as ankylosing spondylitis.

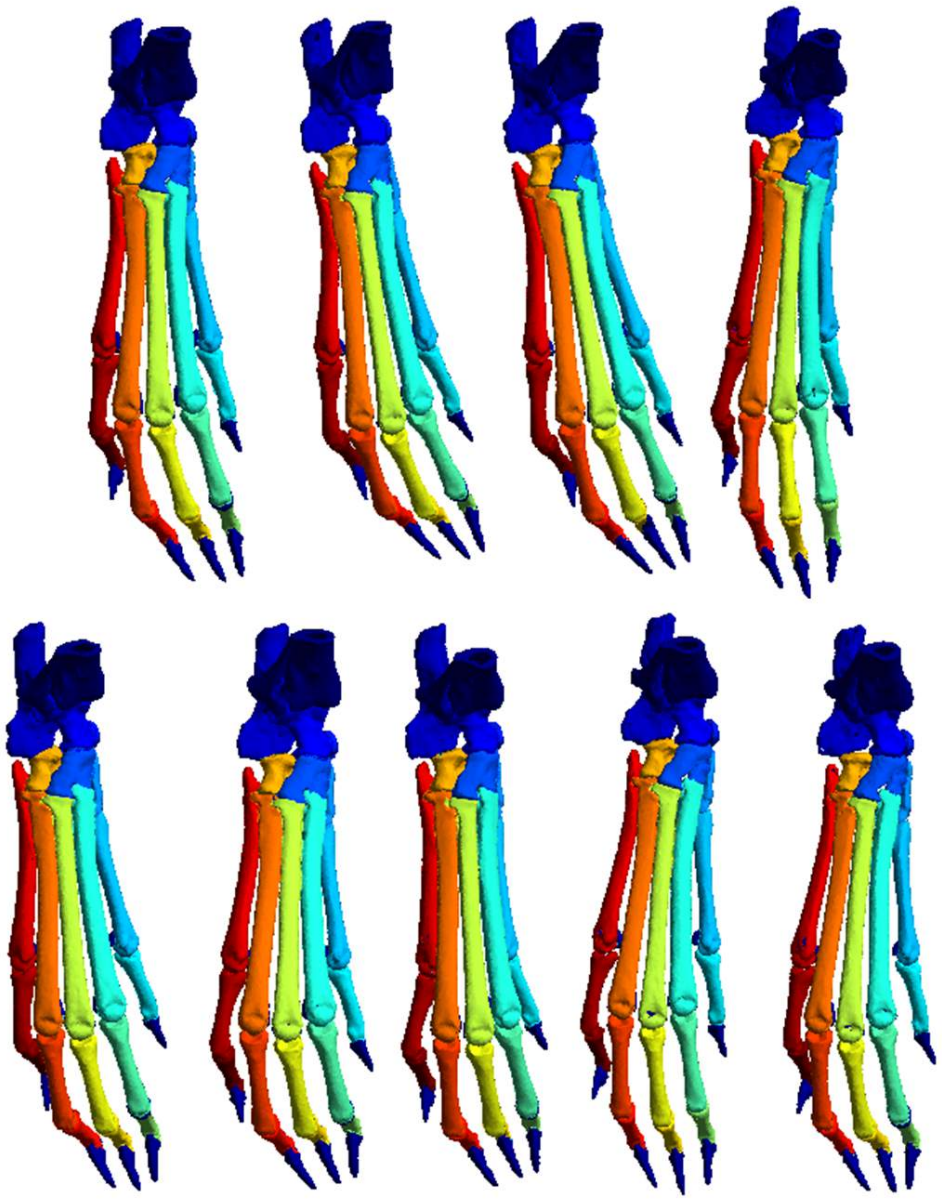


Figure 5.6: Results of segmentation for nine wild-type samples. Bones that lie in close proximity to one another can on occasion receive the incorrect label, which is evident in the upper regions of the metatarsals. Non-model "unlabelled" parts are shown in dark blue, which includes the tibia, fibula, claws and sesamoid bones (partially obscured).

5.3 Statistical shape modelling

A statistical shape model (SSM) was computed for each bone in the model hierarchy, based on the correspondence provided by the registration and subsequent label propagation. Deviations from the mean shape were produced to evaluate the models ability to produce biologically plausible shapes. Figures 5.7 - 5.9 show the resulting bone models from the hind limb as 0.0, $\pm$ 1.5, and $\pm$ 3.0 standard deviations from the mean shape along the first (most significant) principal component. In Figure 5.7, the most notable variation can be seen along the left side of the calcaneus, becoming flatter on the left side (from negative to positive σ). An indentation on the upper portion of the calcaneus can also be seen to depress, becoming flush at -3.0 σ. Differences in bone shape can also be observed along the lower portion of the talus, and along the indented articulating surface of the tibiale. Figure 5.8 shows the distal tarsal bones together.

The results of the statistical shape analysis demonstrate that bone shape variation can be captured using PCA. In most cases, these variations can be reasonably attributed to biological variation, such as increased/decreased curvature, elongation/shortening, narrowing/broadening, and bending. However, there are several examples of variation that are more likely to be due to poor correspondence, as indicated by regions with sharp corners or folding.

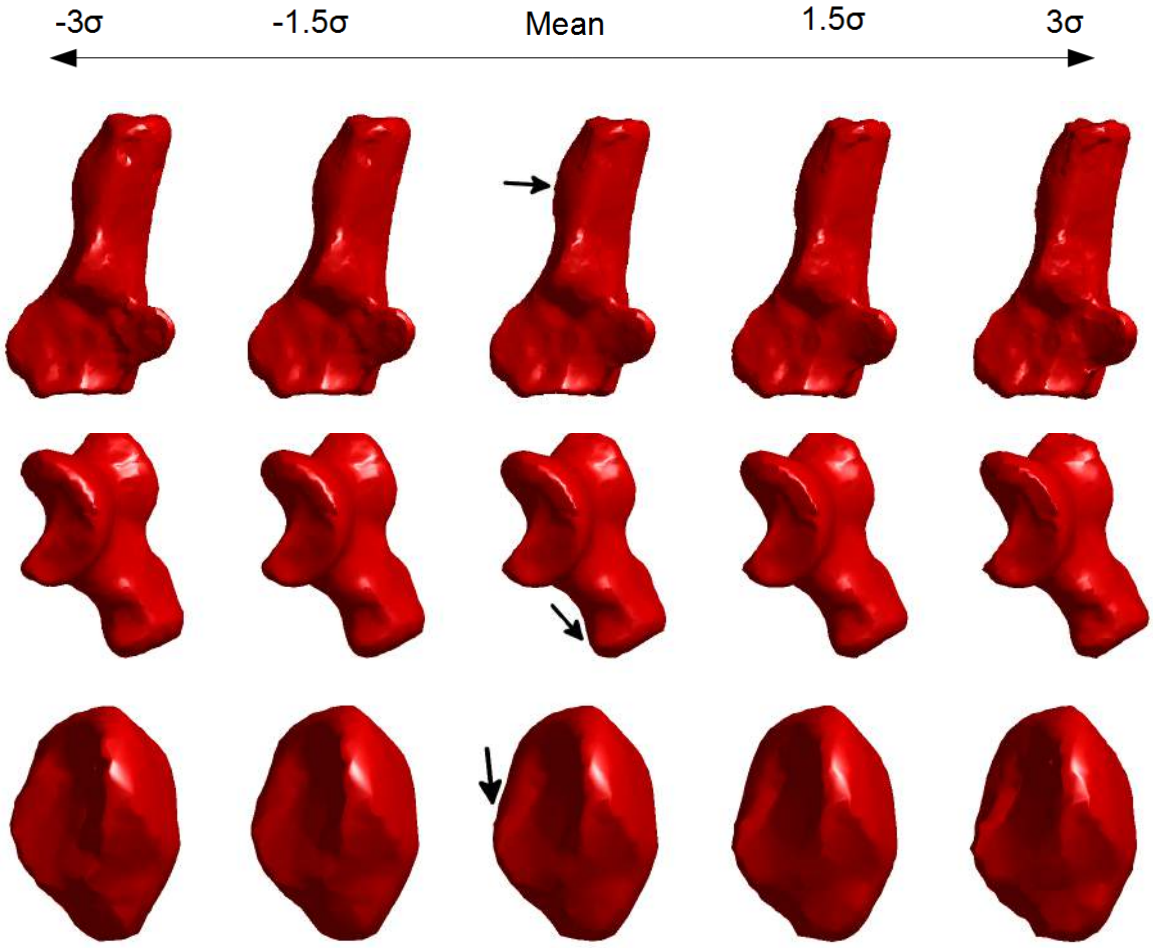


Figure 5.7: Shape models of the calcaneus (top), talus (middle) and tibiale (bottom). From left to right are shape instances -3, -1.5, 0, 1.5 and 3 standard deviations from the mean shape along the first principal component. Variations in bone shape are indicated by black arrows.

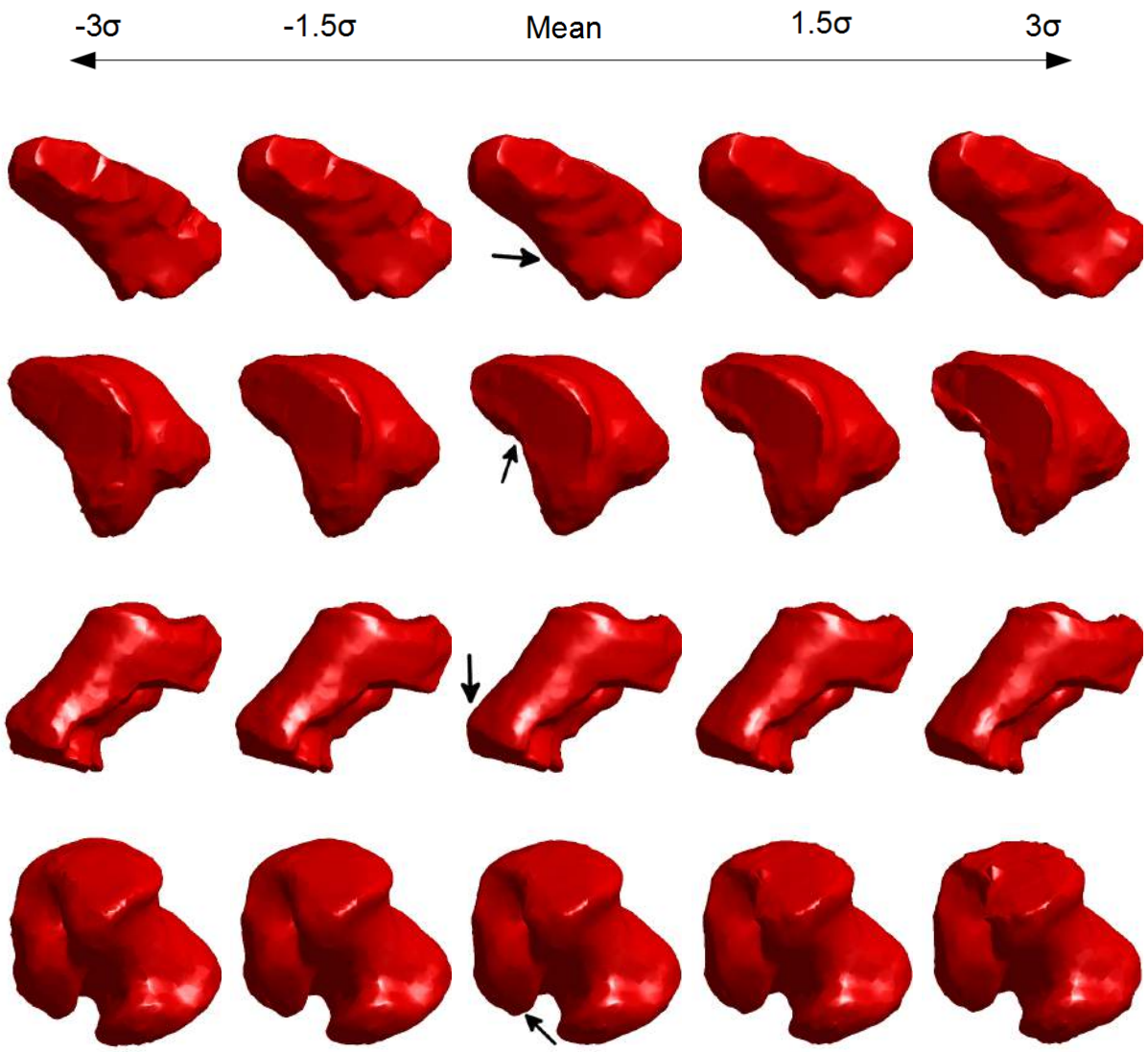


Figure 5.8: Shape models of the 1st (top), 2nd (upper middle), 3rd (lower middle) and fused 4th/5th (bottom) distal tarsals. From left to right are shape instances -3, -1.5, 0, 1.5 and 3 standard deviations from the mean shape along the first principal component. Variations in bone shape are indicated by black arrows.

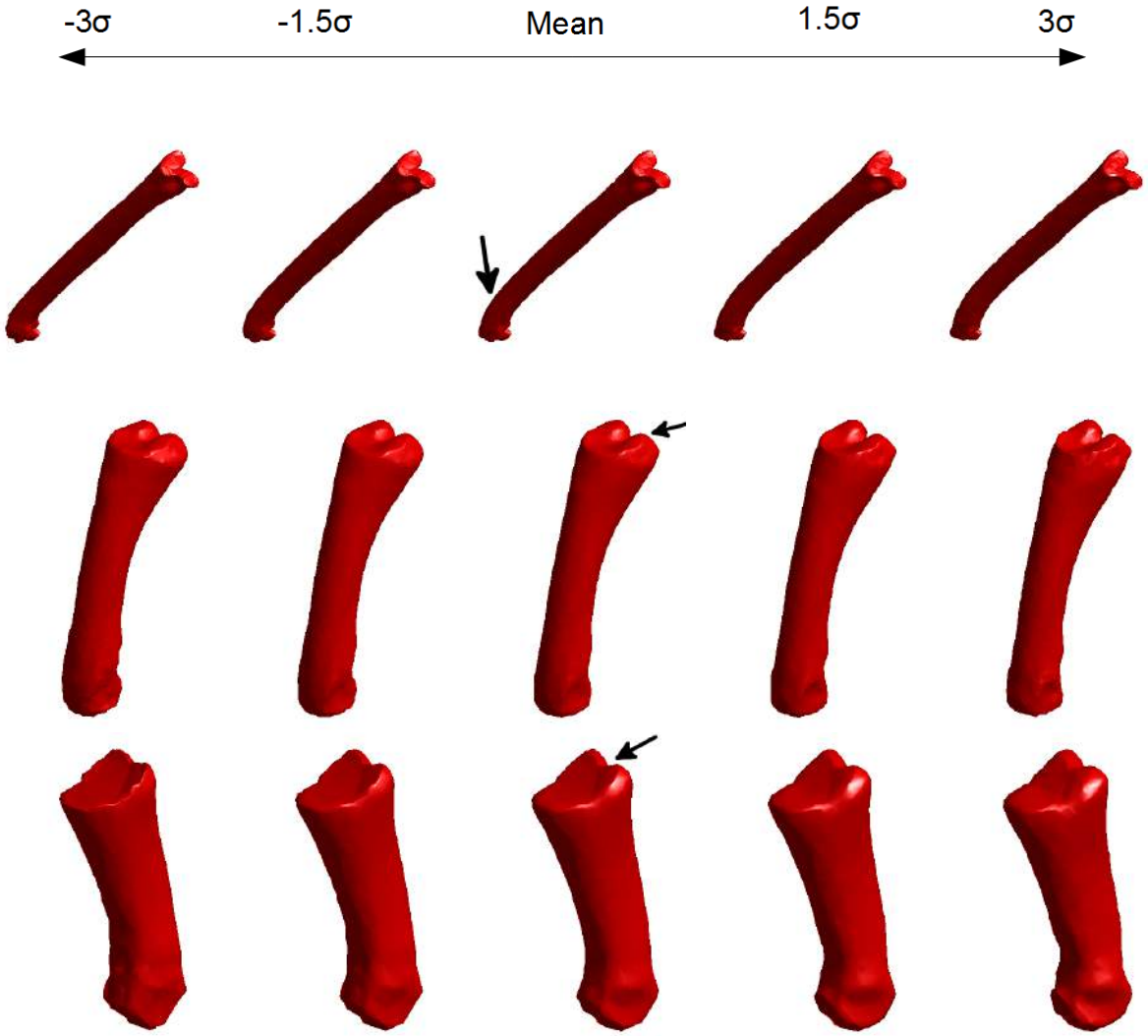


Figure 5.9: Shape models of the 3rd metatarsal (top), 3rd proximal (middle) and 3rd and distal (bottom) phalanges. From left to right are shape instances -3, -1.5, 0, 1.5 and 3 standard deviations from the mean shape along the first principal component. Variations in bone shape are indicated by black arrows.

Chapter 6

Validation and limitations of the model

The proposed method was validated in order to determine its sensitivity, specificity, accuracy and clinical applicability. Samples with artificially introduced deformations were used as the ground truth in a series of experiments to establish how well the model characterises abnormalities. Additionally, the model was applied to a set of arthritic hind paw samples, and the results compared with clinical scores obtained by three independent observers. In this chapter, the results of these experiments are presented and discussed. Based on these results, the limitations of the proposed method are outlined and suggestions made for its improvement. The results of applying the model to a number of different mouse models are presented and discussed in Chapter 7.

6.1 Comparison with ground truth

The model's ability to classify and measure bone abnormalities was validated using digital phantoms. Three leave-one-out experiments were conducted using wild-type samples that had been artificially deformed to simulate the presence of bone erosion and formation. This was achieved by elevating or depressing the normal bone surface at a number of

locations, with variations in their cross sectional area and height/depth. An example of a sample with artificial bone formations is shown in Figure 6.1. Knowledge of their location and surface area was used as the "ground truth", which was then compared to measurements generated by the ASSM. Is is important to note that generation of biologically feasible bone abnormalities is not straightforward, and the data generated for these experiments required a number of refinements to achieve a satisfactory ground truth.

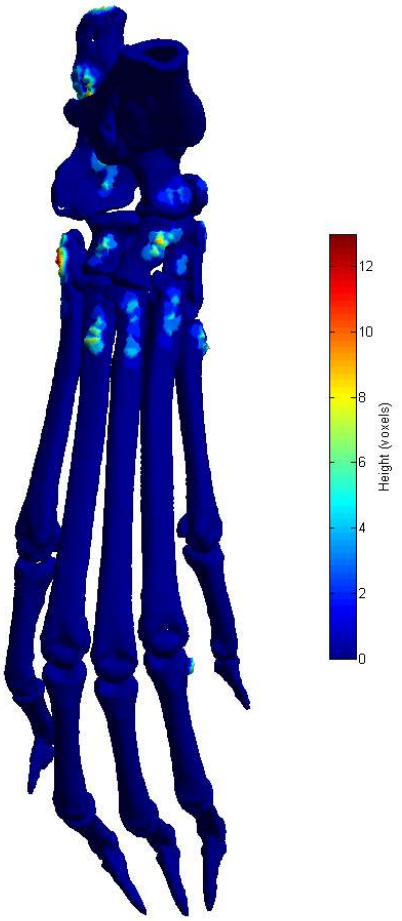


Figure 6.1: Sample with artificially deformed osteophytes and enthesophytes, visualised as a colourmap. The abnormalities were generated at locations which could realistically be observed in real data, and of a similar size and shape.

Figure 6.2 shows the results of the validation for each sample, grouped by bone destruc-

tion type. This data provides a broad overview of the affected regions, as the percentage of the total surface area affected. Figure 6.3 shows the same data grouped by anatomical region. The model was registered onto each of the samples, and the model fitting error used to calculate erosion and growth surface area. The measured values were then compared to the ground truth. In all three samples, the measured bone erosion was accurate to within 0.1% of total surface area (Figure 6.2). Furthermore, its accuracy was relatively consistent across all three anatomical regions (Figure 6.3, left column). For bone formation, the results were more variable (Figure 6.3, right column). The model was most accurate in the metatarsal region, with sample 3 showing a larger error than samples 1 and 2. In the heel region, the degree of bone formation was consistently underestimated. One possible explanation for this is a lack of variation in the model, preventing it from deforming inwards to fit the sample surface. This may have be caused by there being too few samples in the model, or poor point correspondence due to misregistration. In the phalanx region, bone formation was grossly overestimated. Bone formation was rarely found in the phalanx regions of the mouse models studied, hence the absence of ground truth values in that region. Furthermore, the lowest articulated registration accuracy was generally found in the phalanx region, providing a poor initialisation for shape model fitting.

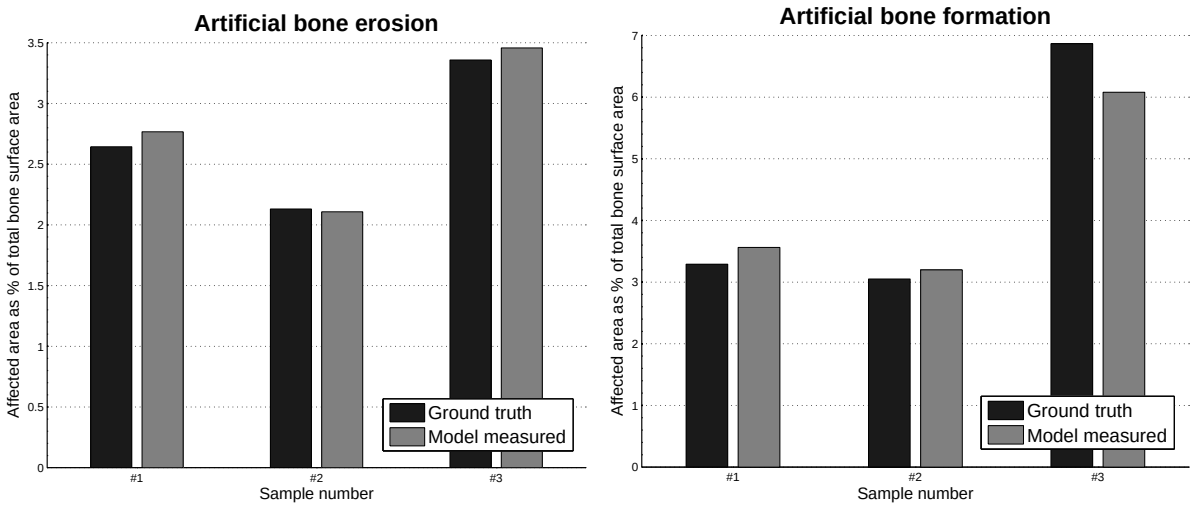


Figure 6.2: Results of three leave-one-out experiments using artificially deformed wild-type samples. The amount of overall bone erosion measured using the model is accurate to within 0.1 %. For bone formation, the model showed similar accuracy for samples 1 and 2 (~0.2 %) but produced a larger error for sample 3 (~1 %).

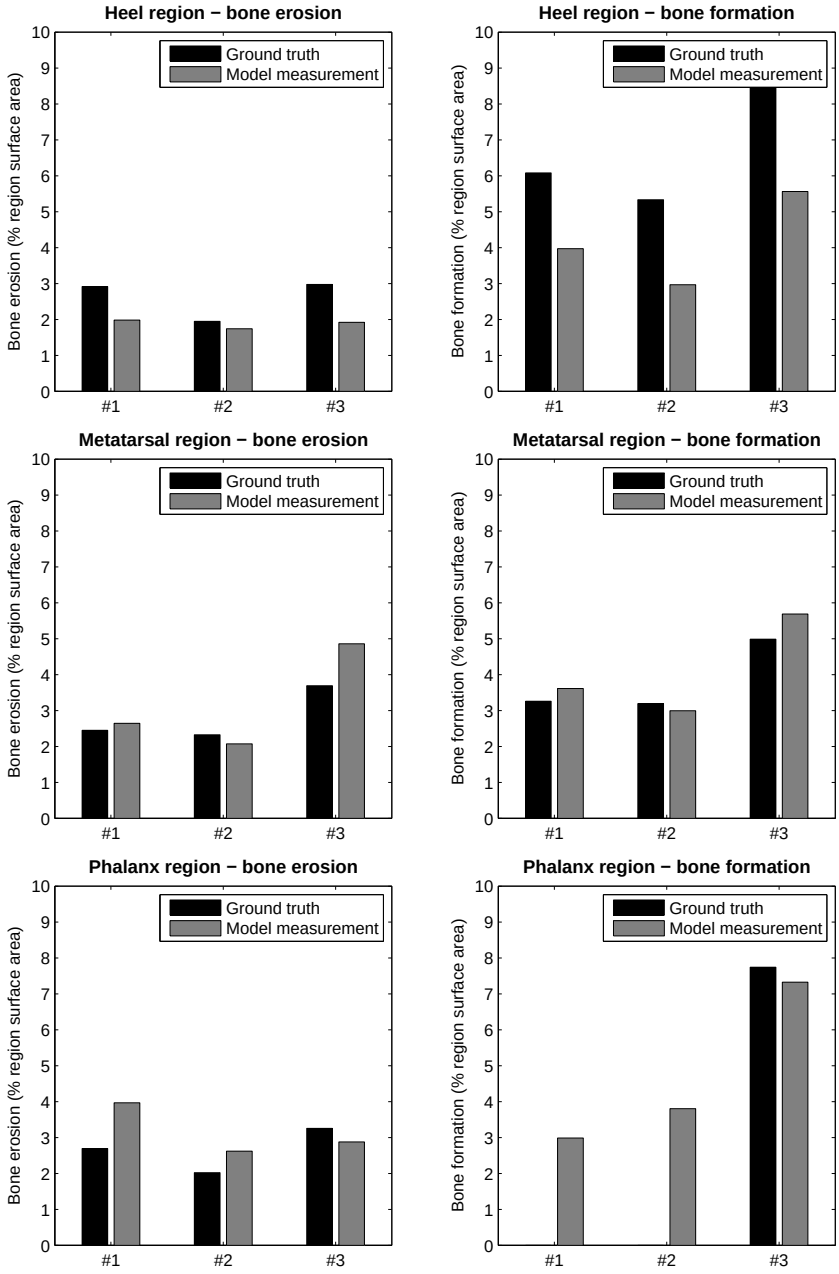


Figure 6.3: Results of three leave-one-out experiments, grouped by anatomical region. Model is most successful in measuring bone erosions in the heel region, and least successful at measuring bone formations in the phalanx region.

6.2 Comparison with clinical bone destruction scores

A clinical scoring system was devised for classifying individual samples imaged by micro-CT. Micro-CT data was scored according to the criteria shown in Table 6.4. Scores were assigned for each of the three main regions, as shown in Figure 6.5. Furthermore, scores were acquired from three independent observers and averaged to give a combined score for each hind paw sample. Model measured abnormalities are given as the percentage surface area affected, given separately for erosion and formation.

[12pt 8pt]

Type	Score	Observations
Bone erosion	0	Normal, no signs of erosion
	1	Roughness of bone surface
	2	Pitting/indentations
	3	Full thickness holes
Bone formation	0	Normal, no signs of formation/deformity
	1	Rough appearance, small osteophytes
	2	Spurs or signs of bone fusion
	3	Whole bone deformity/complete fusion

Figure 6.4: Micro-CT scoring system for bone destruction observed from micro-CT image data. Scores are assigned by independent observers.

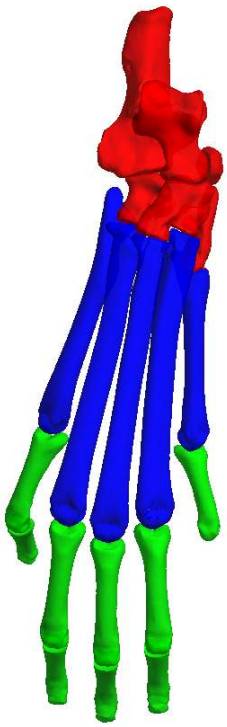
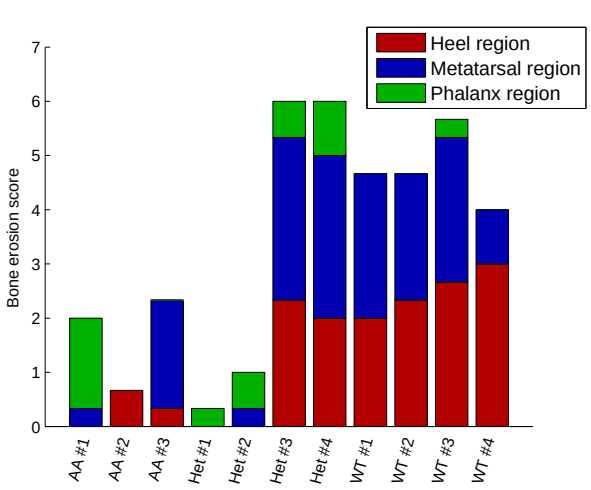


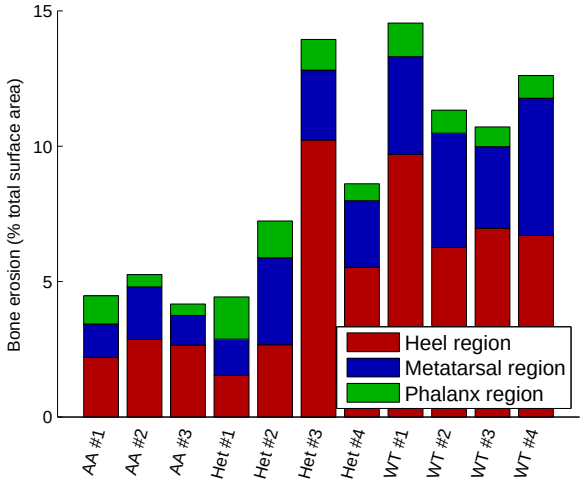
Figure 6.5: Colour coded anatomy of the mouse hind paw with region labels corresponding to the heel, metatarsal and phalange regions (red, blue and green respectively).

Although the data are not directly comparable, there are several commonalities that can be observed. Figure 6.6 shows the bone erosion scores alongside the percentage area affected by erosions, as measured by the model. The clinical scores reveal low levels or erosion in the phalanges as compared to the rest of the paw, which is recapitulated by the model measured results. Furthermore, the "AA" and two of the "Het" mice (1 and 2) all received lower erosion scores than the the "WT" mice, which can also be observed in the model measured erosions. It is important to note that the inter-operator variability was relatively high for erosions, highlighting the need for an unbiased approach when assessing bone destruction (see Appendix 3).

Figure 6.7 compares the bone formation scores with the amount of bone formation



(a) Clinical erosion scores



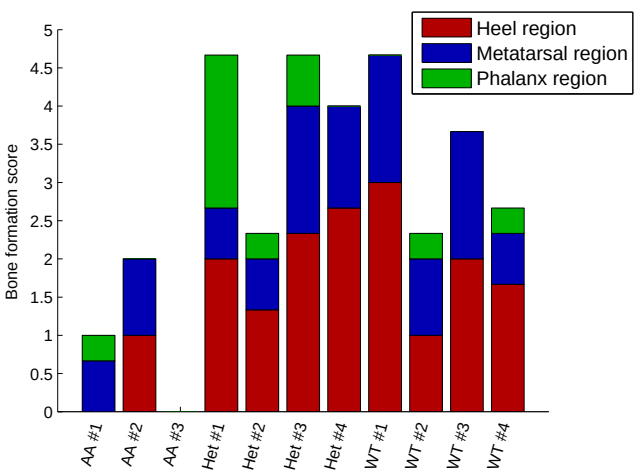
(b) Model measured bone erosion (% surface area)

Figure 6.6: Comparison between clinical erosion scores and model measured erosions. The mice used in the clinical validation (AA: $TTP^{AA/AA}$, Het: $TTP^{+/AA}$ and WT: $TTP^{+/+}$) are described in greater detail in Section 7.2.

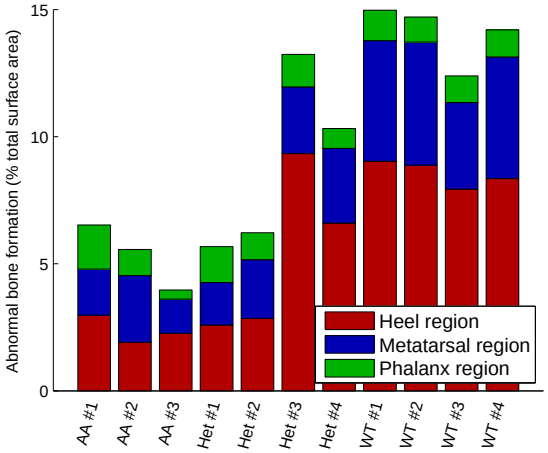
detected by the model (as percentage surface area). As with the erosion data, relatively low levels of bone formation are found in phalanges by clinical scoring. One sample, however, has significantly higher phalangeal bone formation scores than the other samples (Het #1). In examining the image data, narrow strips can be observed along the lengths of several of the phalanges. These are similar in appearance to image artefacts caused by incorrect misalignment compensation ("tails" - see Section 3, Figure 3.2). Their narrow width also explains the model's inability to detect and measure their surface area.

6.3 Limitations

The articulated statistical shape model has been tested and validated against a number of different mouse models in order establish its capabilities. Although the proposed method shows promising results for certain mouse models, others shed light on the limitations of the approach and suggest how they might be overcome.



(a) Clinical formation scores



(b) Model measured bone formation (% surface area)

Figure 6.7: Comparison between clinical bone formation scores and model measured formation. The mice used in the clinical validation (AA: $TTP^{AA/AA}$, Het: $TTP^{+/AA}$ and WT: $TTP^{+/+}$) are described in greater detail in Section 7.2.

6.3.1 Large variations in pose

In the human form of the disease, inflammatory arthritis can impair normal joint function due to swelling and destruction of the bone surface. This can lead to structural deformities such as swan-neck, Boutonniere, and ulnar deviation. Interestingly, similar structural deformities have been observed in mice during *in vivo* experiments in both the fore and hind paws. Variations in relative bone position are handled to some extent by the kinematic constraints imposed during registration. However, in cases where mice have developed extreme structural deformities, registration of the model can fail due to these unforeseen variations in pose. At present, such cases are dealt with as they arise by manual adjustment of the kinematic parameters previously described in Section 4.1.2.2. Approximately one in five cases require some form of manual adjustment, usually applied to a single bone. Another potential source of large pose variation is during sample preparation for micro-CT imaging, where fixation and storage of tissue samples can lead to formation of unnatural poses.

6.3.2 Large bone shape changes

Certain types of bone shape change can cause the articulated registration procedure to fail. The destruction of salient bone features can drive the ICP algorithm into local mimina, resulting in an incorrect alignment. This has an undesirable knock-on effect, as any bones within the same subtree will likely suffer the same fate. This is particularly prevalent in the TNF-DARE model, where extreme cases of bone destruction lead to failure very early on (Figure 6.8). This "error propagation" is an inherent limitation of articulated registration, and can only be mitigated by making the model as robust as possible to large bone shape changes.

When applying the model to diseased samples, the presence of abnormalities can also influence the articulated registration in a way that affects the shape model's ability to approximate non-pathological bone shapes. No adjustments are made to account for rigid or affine misalignment during shape model fitting, meaning that abnormalities may be under- or overestimated in terms of severity. This has been tackled in other studies by allowing for rigid transformations during shape model fitting, incorporating additional constraints based on the underlying image intensity [52].

6.3.3 Poor correspondence due to misalignment and bone fusion

In some hind paw samples, the distal tarsal bones appeared to be partially or completely fused together. This fusion, whether it be biological or due to digitisation, led to misalignment of the model and poorer point correspondence in those regions. An additional complication was observed when applying the shape model to meshes with holes. In some mouse models, full thickness bone erosions (complete loss of cortical bone) were observed as holes in the isosurface mesh. It is speculated that their presence misguides shape model fitting due to the lack of point correspondence in these regions.

Figure 6.8: Failure of the model when applied to a TNF-DARE sample.

6.3.4 Changes in bone mineral density

In constructing a model from micro-CT images, reconstruction and thresholding parameters are kept consistent to ensure that the same isosurfaces are being compared. This presents a problem in studies of mouse models where changes in bone mineral density occur. An example of such a model is the CD248 (endosialin) knockout (-/-) mouse which has increased bone mineral density [68]. Additionally, an ageing model of osteoarthritis was investigated in which abnormal fusion of bones and mineralisation of soft tissues led to poor registration results.

6.3.5 Classification of abnormalities

The purpose of the articulated statistical shape model is to quantify bone shape changes caused by pathology. One of the assumptions is that pathological shape changes can be identified as deviations from normal bone shape. It is evident that whilst this assumption holds true to an extent, some false negatives (failure to classify a genuine abnormality) and false positives (normal bone classified as being abnormal) may be observed.

Chapter 7

Application to mouse models of rheumatoid arthritis

Mouse models of rheumatoid arthritis work by targeting the genes and proteins that are implicated in its pathogenesis. The choice of mouse model is dependent on which aspects of the disease are under investigation, with factors such as the method of induction, incidence rate, duration and severity of response coming into play. In this chapter, the articulated statistical shape model (ASSM) is demonstrated as a useful tool for characterising the different ways in which bone destruction can take place in three commonly used mouse models. Following this, the ASSM is demonstrated in a clinical scenario that explores the role of macrophages in RA pathogenesis using a novel transgenic mouse. Results of its application to mice of three different genotypes mice are presented, and compared with previous findings.

7.1 Comparison of different mouse models

A comparison of three commonly used mouse models (K/BxN, CAIA and TNF dARE) was performed, giving insight into the potential mechanisms that give rise to bone destruction. In order to provide some context for this comparison, Table 7.1 briefly lists

some of the key differences between the models.

It is important to note that this study is not comprehensive, but instead lays down the groundwork for an in-depth investigation into the inherent differences between the models. In the few examples shown, differences in the nature of bone destruction are discussed and possible explanations for these differences explored. In addition, this study has also provided a valuable opportunity for some the practical limitations of the proposed method to be identified.

Model	**Type**	**Description**	**Resolves?**
K/BxN serum-transfer	Immune mediated	Administration of glucose-6-phosphate isomerase antibodies	Yes
CAIA	Immune mediated	Administration of type II collagen antibodies	Yes
TNF dARE	Cytokine mediated	Deletion of a regulatory sequence on the TNF gene	No

Table 7.1: Key differences between the three models used in this study.

7.1.1 The K/BxN serum-transfer model

Glucose-6-phosphate isomerase (G6PI) is an enzyme that is involved in the conversion of glucose into pyruvate via glycolysis, providing energy in the form of adenosine triphosphate (ATP). Increased expression of G6PI and its associated antibodies in RA patients has incited interest in its diagnostic value [69]. The K/BxN transgenic mouse produces antibodies to G6PI, due to the presence of the KRN T-cell receptor transgene and the I-A(g7) MHC allele (discussed in Section 1.3.1). Serum extracted from the K/BxN mouse may be used to induce a transient form arthritis in a number of different mouse strains, referred to as the K/BxN serum-transfer model [9]. Such mice develop severe inflammatory arthritis, with evidence of pannus formation, synovial hyperplasia and erosion of cartilage and bone that is reminiscent of human RA [70].

In order to characterise its phenotype in terms of bone destruction, four K/BxN serum-transfer mice were analysed and compared using the proposed method. Two of the mice were "non-responders" which had no signs of inflammation and received no clinical score. MicroCT imaging studies indicate that normal bone integrity is maintained in the non-responders, showing no evidence of localised erosions. Application of the wild-type ASSM yields results that are supportive of these observations, as shown in Figure 7.1. Both samples present no obvious signs of bone destruction, in both the microCT data and corresponding heatmaps generated by the model.

Figure 7.2 shows the two remaining samples that responded to the K/BxN serum, and received clinical scores. However, the microCT data does not reveal any obvious signs of bone destruction upon visual inspection. Application of the model reveals both erosion and formation of bone, predominantly around the synovial joints. Interestingly, the presence of bone formation along the metatarsals indicates that some abnormal thickening or perhaps bowing has taken place that is not accommodated by wild-type variation. It is also clear that whilst both mice were responsive to the serum, there is a clear difference in severity between the two samples shown. Furthermore, erosions can be found in common locations on both samples, such as the calcaneus, talus and metatarsals. More importantly, it is apparent that bone remodelling is not exclusively erosive. Evidence of bone formation can also be found on the calcaneus, distal tarsals and metatarsals of both samples.

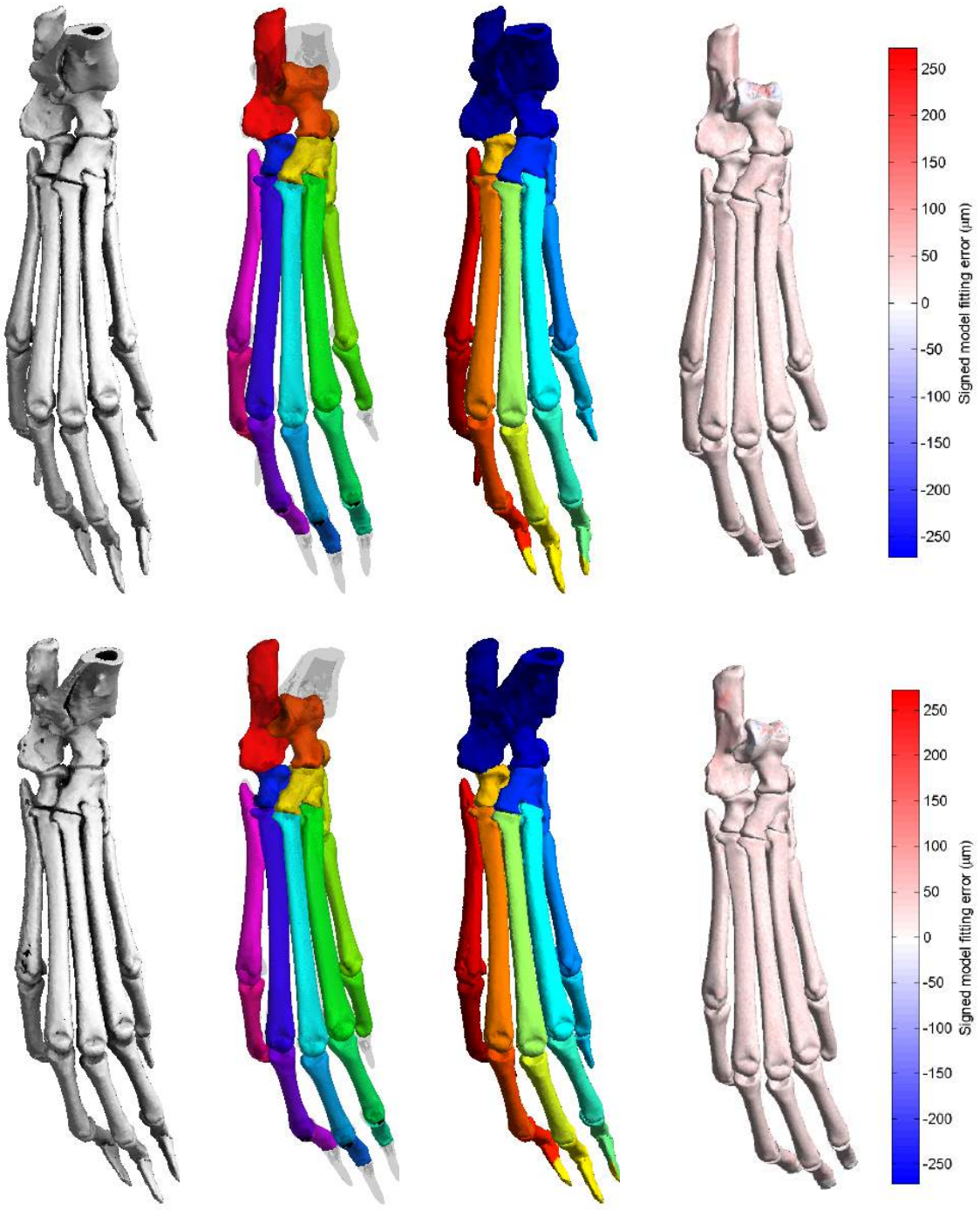


Figure 7.1: Non-responsive mice after K/BxN serum-transfer. The microCT data shown on the left is consistent with the lack of inflammation, with neither having any obvious signs of bone destruction. Registration and shape model fitting yields results that can be visualised as a heatmap of "model error", shown on the far right. With the exception of some uniform baseline error, there is no evidence of localised bone erosion having taken place.

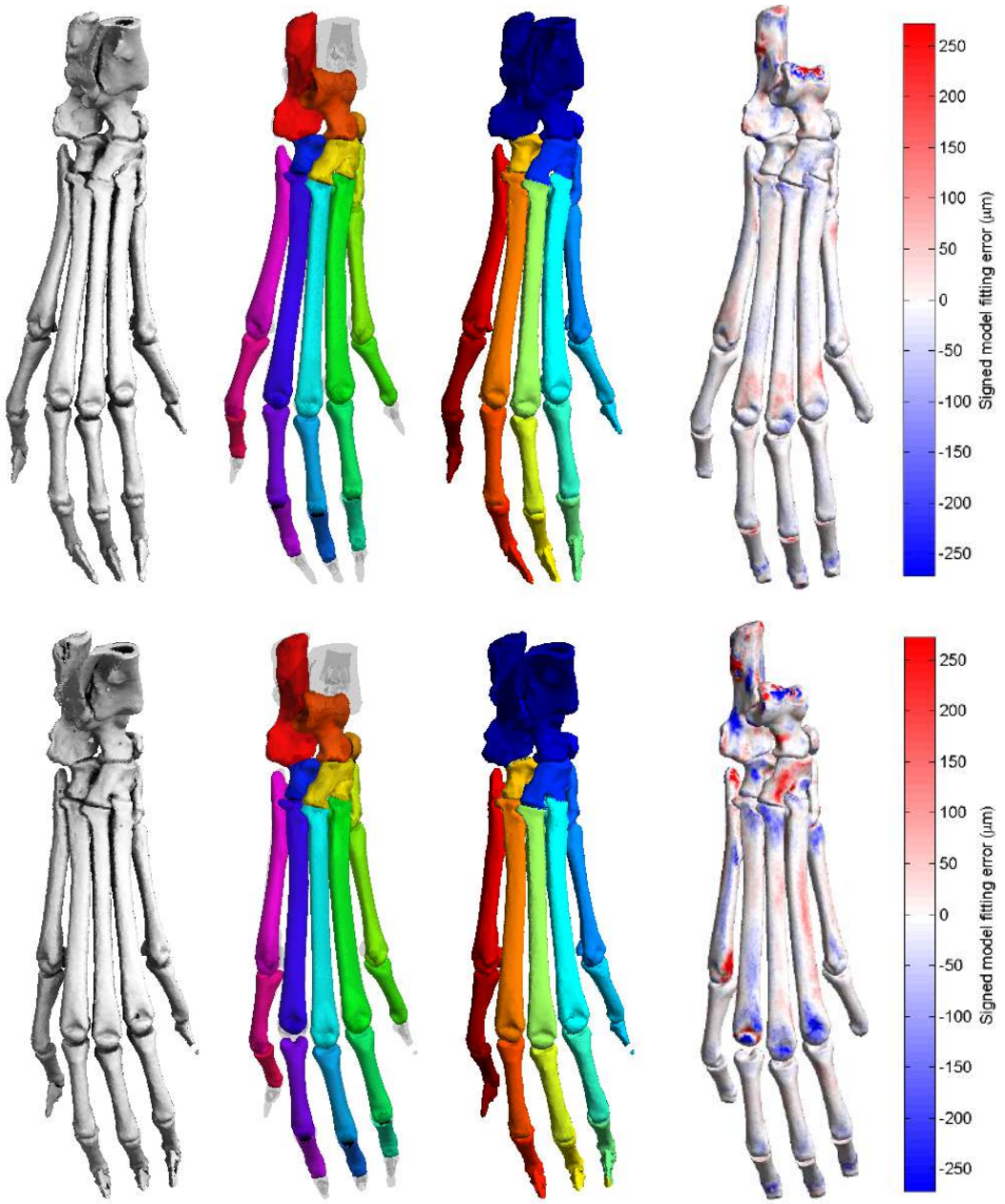


Figure 7.2: Responsive mice after K/BxN serum-transfer. In contrast to the non-responders, it is immediately apparent from the heatmaps that bone remodelling has occurred. The degree of bone destruction differs between the two samples, with the top sample showing relatively small changes when compared with the bottom sample.

7.1.2 Collagen antibody induced arthritis (CAIA) mice

Type II collagen (CII) is a structural protein that forms articular cartilage on the surface of bones [71]. In the collagen antibody induced arthritis (CAIA) model, CII antibodies are administered to the mouse, followed by immune stimulation with bacterial lipopolysaccharide (LPS), binding to immune cells and promoting the release of pro-inflammatory cytokines [72]. As in the K/BxN serum-transfer model, CAIA mice experience transient inflammatory arthritis, characterised by inflammation of the synovium and destruction of both cartilage and bone [73].

Two CAIA mice were analysed using the ASSM, as shown in Figure 7.3. It should be immediately evident that these mice differ greatly from the K/BxN mice. Furthermore, the two samples are very different from one another in terms of bone phenotype. In the first sample, a bulbous formation of bone on the distal region of the 4th metatarsal causes the registration to partially fail, resulting in subsequent misalignment of the connected phalanges. Elsewhere, the model reveals widespread bone erosion and formation of the calcaneus, talus, metatarsals and phalanges. The second sample offers a very different bone phenotype consisting of full thickness bone erosions, and small studded patches of newly formed bone. Application of the model reveals speckled hotspots (shown in red) that are consistent with the sparse nature of the bone formation located on the calcaneus, distal tarsals and metatarsals. Full thickness erosions present on the metatarsal are not well handled by the model, being detected either partially or not at all.

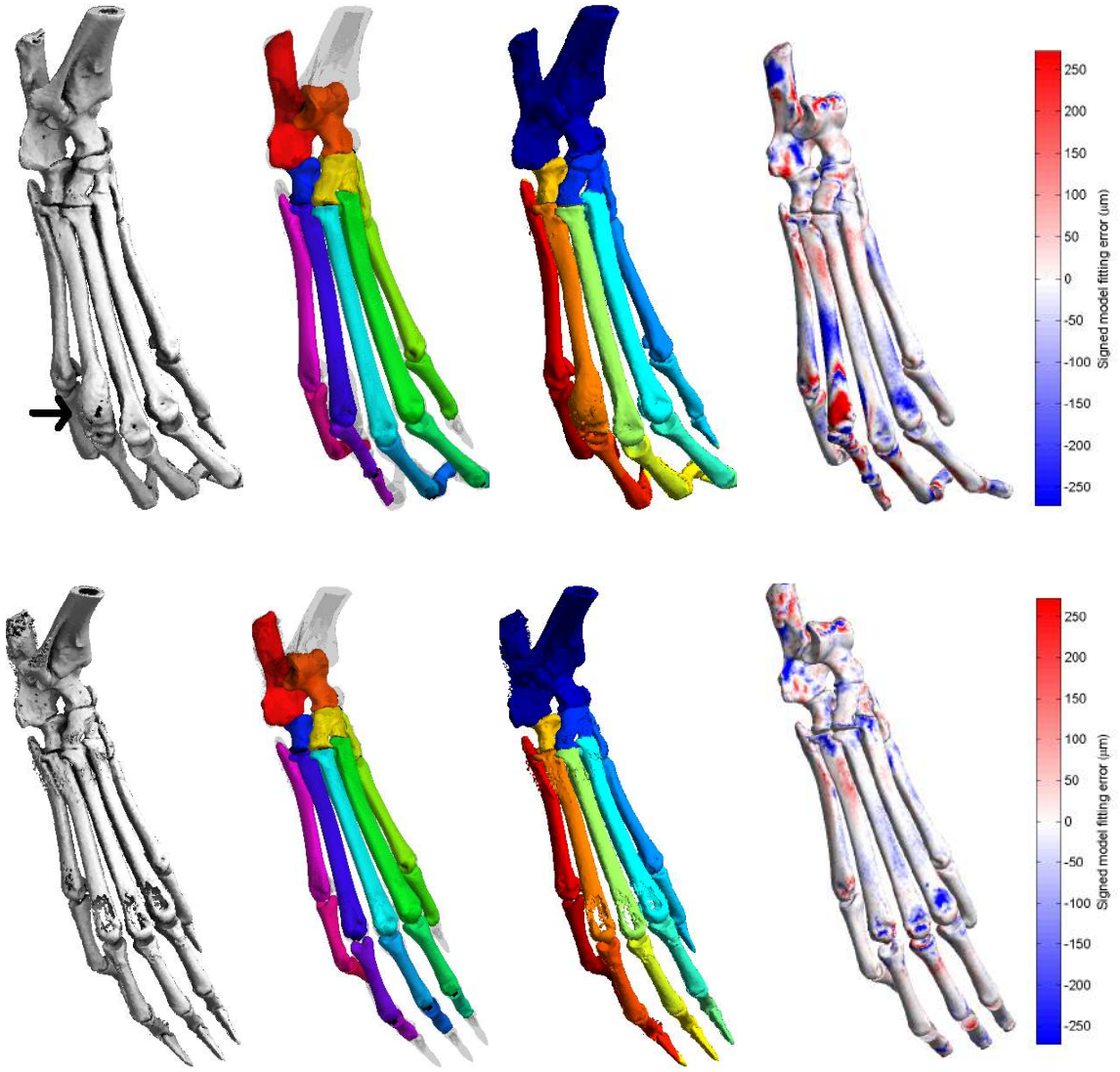


Figure 7.3: Application of the ASSM to collagen antibody induced arthritis (CAIA) mice. These samples present two very different bone phenotypes that reveal two important limitations of the proposed method; robustness to extreme bone formation, and sensitivity to full thickness erosions. The black arrow (top row) indicates a rarely observed, bulbous region of abnormal bone formation which caused the connected subtree to be misaligned. The resulting poor correspondence prevents the model from reliably identifying the abnormalities present. The lack of sensitivity to full thickness erosions is most evident on the metatarsals (bottom row), where blue patches indicate that they can only be identified when particulaly large.

7.1.3 TNF dARE transgenic mice

TNF is a pro-inflammatory cytokine that is strongly implicated in RA pathogenesis. The TNF dARE mouse model overexpresses TNF, due to deletion of a regulatory sequence on the TNF gene (AU rich element, or ARE). TNF dARE mice experience chronic polyarthritis (affecting five or more joints) that is non-resolving, unlike the K/BxN and CAIA mice. Despite this, TNF dARE mice often lack the severe swelling of the joints observed in the other models. Two TNF dARE mice were examined using the ASSM to determine whether they bore similarities to the K/BxN mice in terms of bone destruction. The results are shown in Figure 7.4.

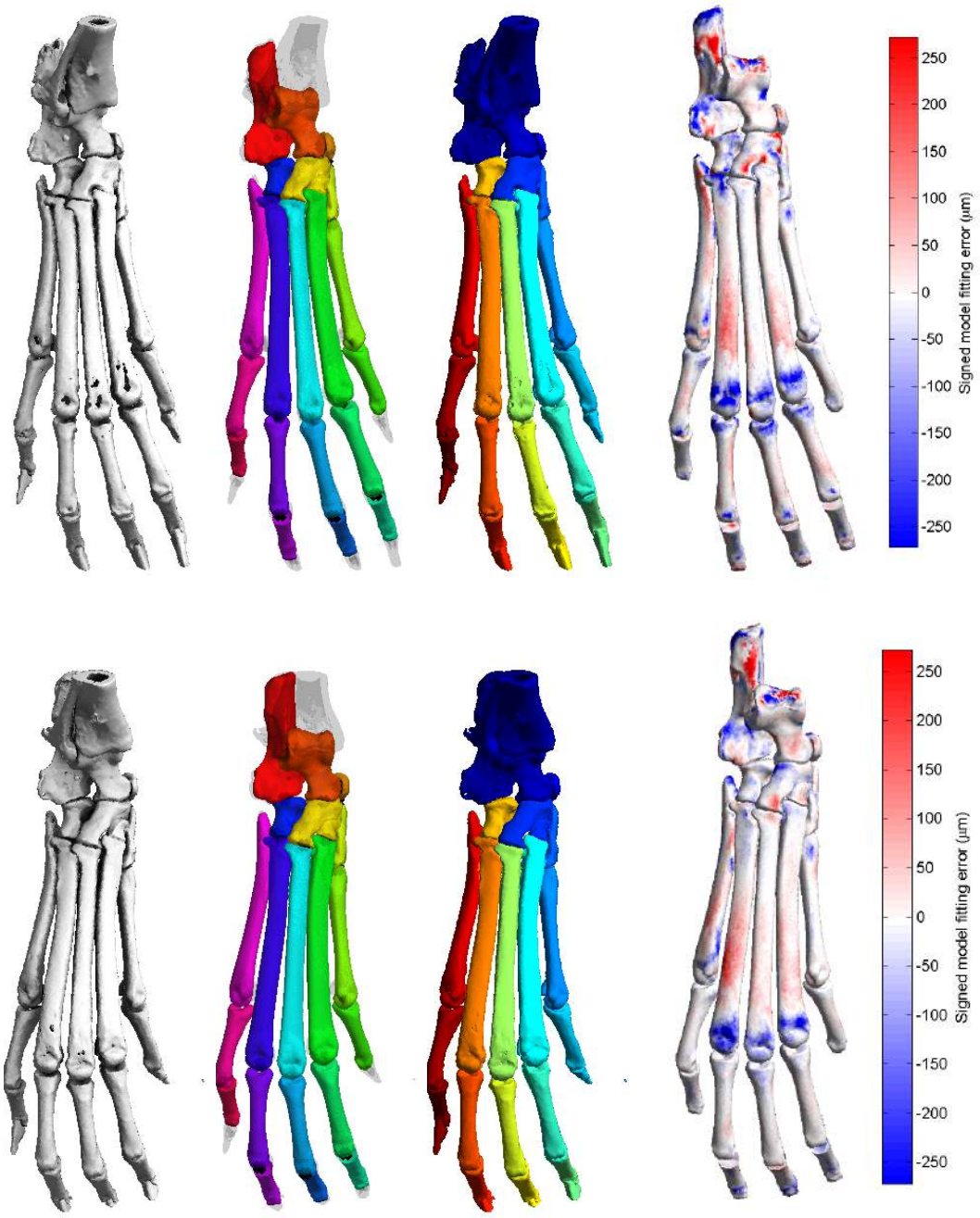


Figure 7.4: Application of the ASSM to TNF dARE transgenic mice. The two samples are very similar in terms of the location and severity of the bone destruction present. Notable erosions can be observed on the 2nd - 4th metatarsals, as well as the flat portion of calcaneus. Interestingly, a small degree of bone formation can be observed along the bodies of the metatarsals which is indicative of bowing. This can be verified upon inspection of opposite side of the paw, where the apparent "bone loss" is in fact due to curvature of the bones.

7.1.4 Summary

Although K/BxN and CAIA mice are very similar in terms of their transient antibody-induced nature, they each present very different bone phenotypes. Application of the ASSM reveals that whilst both models are predominantly erosive, both also bear clear signs of bone formation having taken place. This is supported by previous studies of both models [73, 74].

7.2 Application to a novel transgenic mouse

In order to assess the model's suitability in a real research scenario, the proposed method was used to assess bone destruction in a novel transgenic mouse. The gene of interest in this study - tristetraprolin - is predominantly expressed in macrophages. Macrophages are known to have a role in RA, serving as the primary source of TNF-α [75]; a cell signalling protein (cytokine) which regulates immune cell activity. Within the macrophage, production of a cytokine begins in the cell nucleus, where DNA is transcribed into a molecular template known as messenger RNA (mRNA). Upon leaving the nucleus, the mRNA is translated into a sequence of amino acids called a polypetide chain. This chain is then folded to form a complex three-dimensional structure, undergoing various modifications before being released by the cell to regulate immune activity.

Normally, cytokine expression itself is carefully regulated through interactions with genes and other proteins as part of a complex network of feedback loops. This process is called *post-transcriptional regulation*. Prior to translation into a protein, other regulatory proteins may bind to cytokine mRNA in order to promote or inhibit its expression. Tristetraprolin (TTP) is one such protein that is expressed by macrophages during inflammation, binding to a site on cytokine mRNA called the three prime untranslated region (3'-UTR). In its "inactive" phosphorylated form, the TTP protein stabilises the mRNA, allowing it to be translated into a working cytokine. When activated by a phosphatase

enzyme, cytokine mRNA is destabilised and more rapidly degraded.

Mice that are deficient in the TTP gene have been previously reported to exhibit inflammatory disease phenotypes, due to its role in regulating the expression of TNF-α [76]. Recent research into TTP has determined that modification of two amino acids causes permanent activation of the protein, resulting in enhanced mRNA degradation. Initial results in a transgenic mouse model show that this leads to decreased expression of pro-inflammatory cytokines such as TNF-α and CXCL1; a chemokine that is highly expressed in the joints of rheumatoid arthritis patients where it is known to recruit neutrophils [77]. Neutrophils are known to have a role in damaging synovial tissue and bone through secretion of protease enzymes [78]. The protective effects of modifying the TTP gene are therefore of great interest, and have been assessed in experimental mouse models by quantifying the bone destruction present in micro-CT image data.

7.2.1 Comparative study of wild-type and mutant mice

The K/BxN serum-transfer model was used on a C57BL/6 background to compare disease severity in wild-type (TTP $^{+/+}$) and transgenic (TTP $^{+/AA}$ and TTP $^{AA/AA}$) mice [1]. The K/BxN serum-transfer model is an immune mediated model, where glucose-6-phosphate isomerase (GPI) antibodies are administered to produce short-term inflammation [11]. Micro-CT were acquired according to protocols described in Section 3.4. Several of the distal phalanges were inadvertently cut off by the limited field of view of the micro-CT scanner, and so have been omitted from all of the graphs shown hereafter (specifically 10, 13 and 17).

[1] Wild-type mice have two normal copies of the TTP gene. The TTP $^{AA/AA}$ (homozygous) mouse has two mutant copies of the TTP gene, whereas the TTP $^{+/AA}$ (heterozygous) mouse has one normal and one mutant copy.

7.2.2 Correlation between paw thickness and bone destruction

Paw thickness is typically measured in mouse studies of rheumatoid arthritis to indicate the degree of inflammation. Although not generally considered a good predictor of bone destruction, a strong correlation between paw thickness and bone destruction can be observed in Figure 7.5. Interestingly, the homozygous TTP $^{AA/AA}$ mice and wild-types form two distinct clusters, with the heterozygotes spanning a wider range of severity.

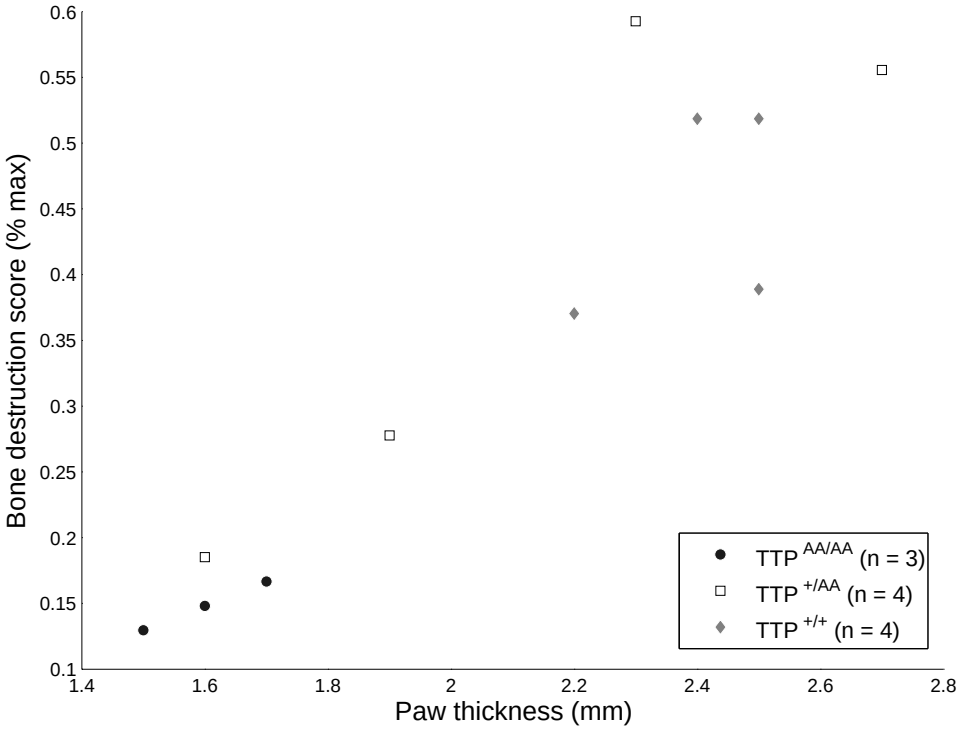


Figure 7.5: Correlation between paw thickness and clinical scores of bone destruction. *In vivo* paw thickness was measured using calipers as described in Section 3.1.6, where elevated thickness indicates the presence of inflammation. Clinical bone destruction scores were acquired for the same limb of each sample by three independent observers. The details of the scoring system are given in Section 6.2.

7.2.3 Heatmap visualisation of bone destruction

The ASSM was applied to the homozygous, heterozygous and wild-type hind limb samples and the signed model fitting error computed. The resulting heatmap visualisations are

given in Figures 7.6, 7.6 and 7.6, respectively. Colour scales were normalised across all samples to make them directly comparable. Across all three groups, the majority of bone destruction can be seen in the heel and metatarsal regions. The phalanges and interphalangeal joints exhibit little or no bone destruction in any of the heatmaps, which correlates well with the clinical scoring results (see Appendix 3).

The homozygous TTP $^{AA/AA}$ mice shown in Figure 7.6 exhibit lower levels of both bone erosion and formation than the other two groups. Abnormalities are highlighted on the heatmap around the lower metatarsals, calcaneus and distal tarsals, indicating that some bone remodelling has taken place. This is consistent with the non-zero clinical scores recorded for those samples, but showed no signs of inflammation based on paw thickness and histological findings (not shown). The TTP $^{+/AA}$ mice have a somewhat variable phenotype, with 50% being only slightly more severe than the TTP $^{AA/AA}$ mice. The remaining two samples are notably worse, with higher levels of erosion around the calcaneus, distal tarsals and upper metatarsals. This partial penetrance is also apparent in Figure 7.5. Bone formation is also evident at these locations, with red patches indicating osteophytosis (bone spur formation) or entheseal ossification (mineralisation of soft tissue around the tendons) having taken place. The same phenotype observed in these mice is also present among the TTP $^{+/+}$, except with increased converage and intensity. Upon examination of the micro-CT data, the calcaneus, distal tarsals and metatarsals are highly abnormal in terms of morphology, which is reflected in the higher clinical scores when compared to the other groups.

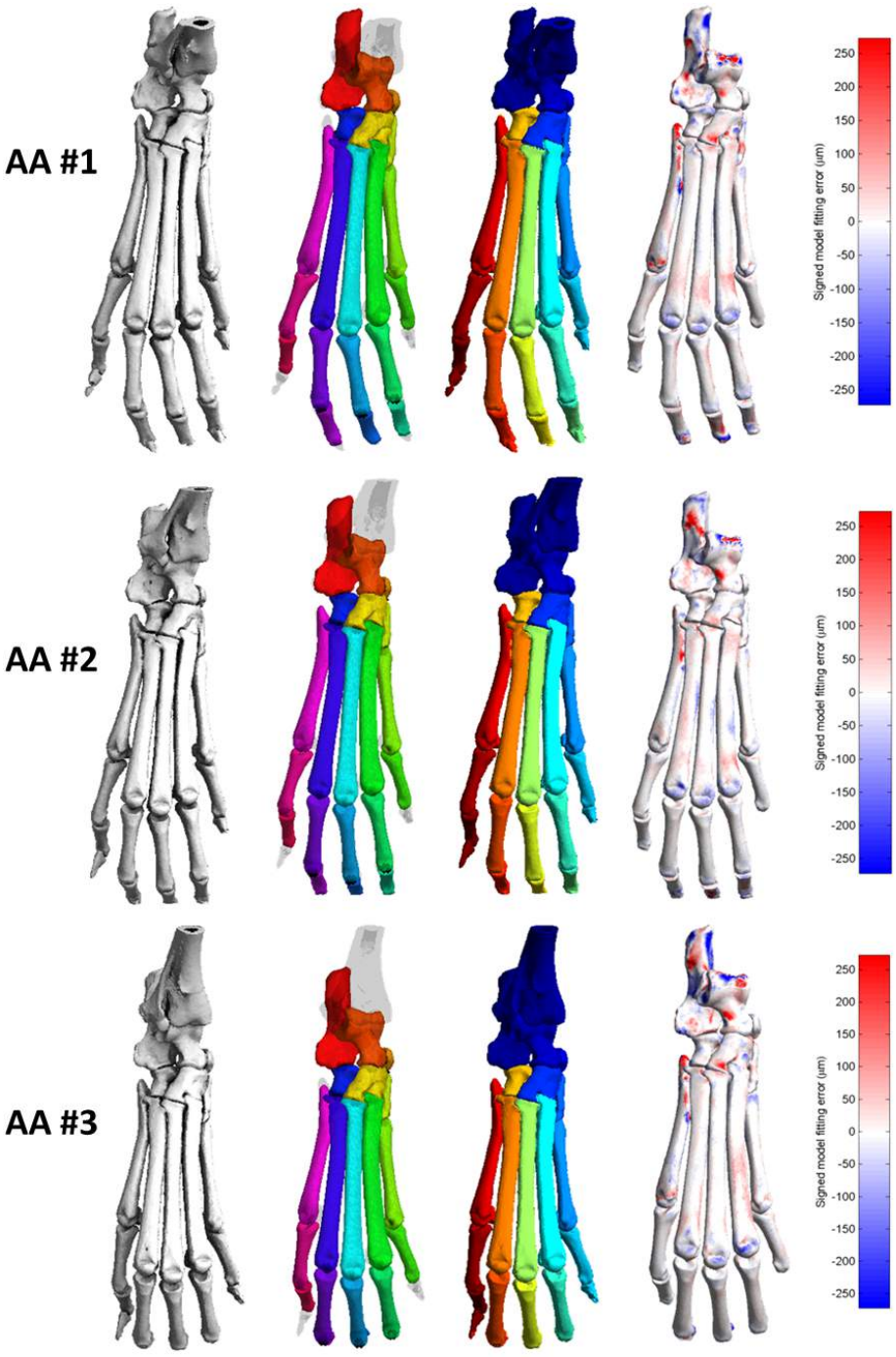
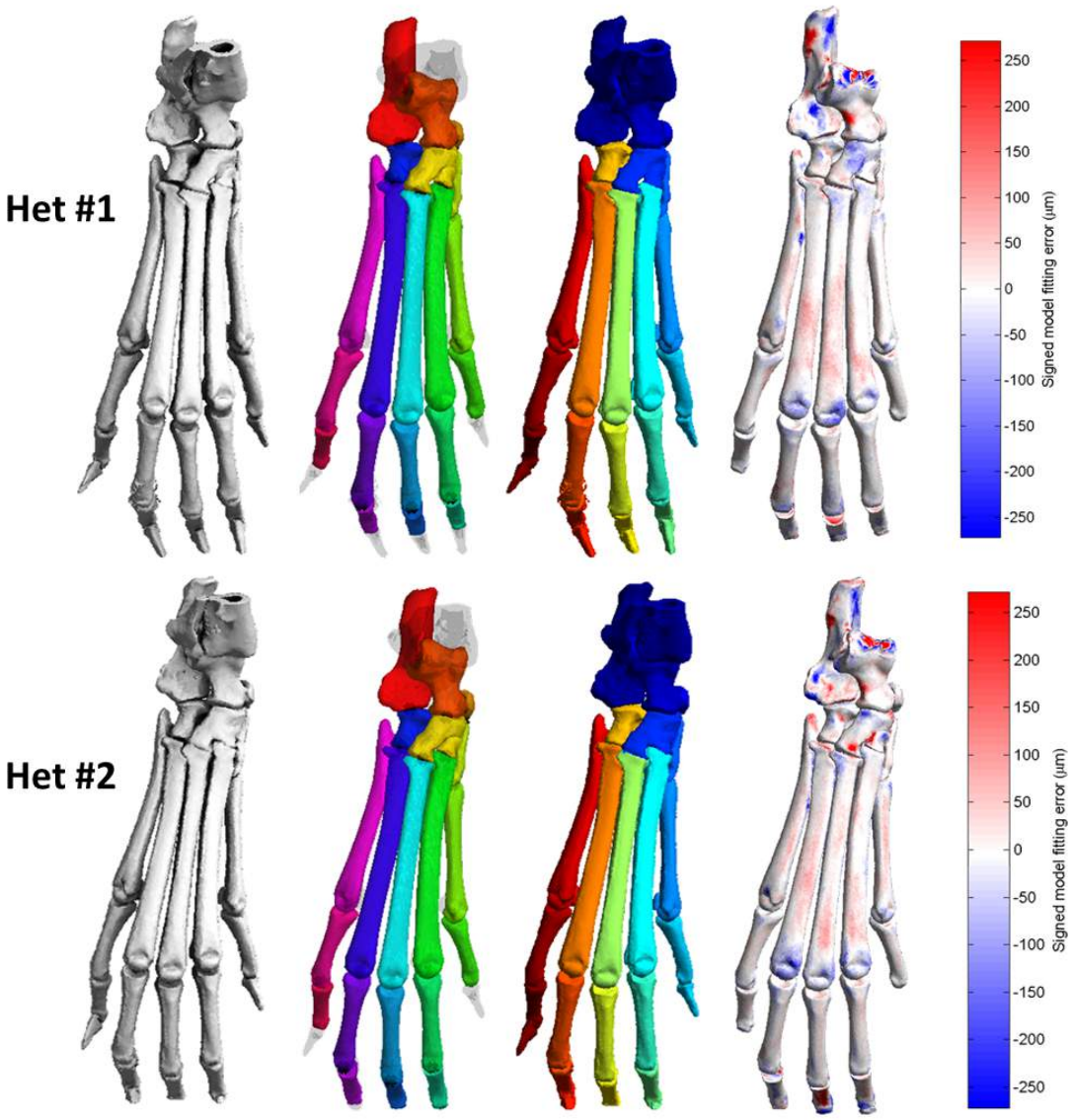


Figure 7.6: Application of the model to TTP $^{AA/AA}$ mice. From left to right: original sample mesh, registration result, segmentation result, signed error mapping. In all three samples, small abnormalities can be observed at numerous locations including the sides of the heel, distal tarsals, and metatarsophalangeal joints. Some larger, but shallow bone loss can be found around the calcaneus and talus.


Het #1
Het #2
250
200
150
100
50
0
-50
-100
-150
-200
-250
Signed model fitting error (µm)

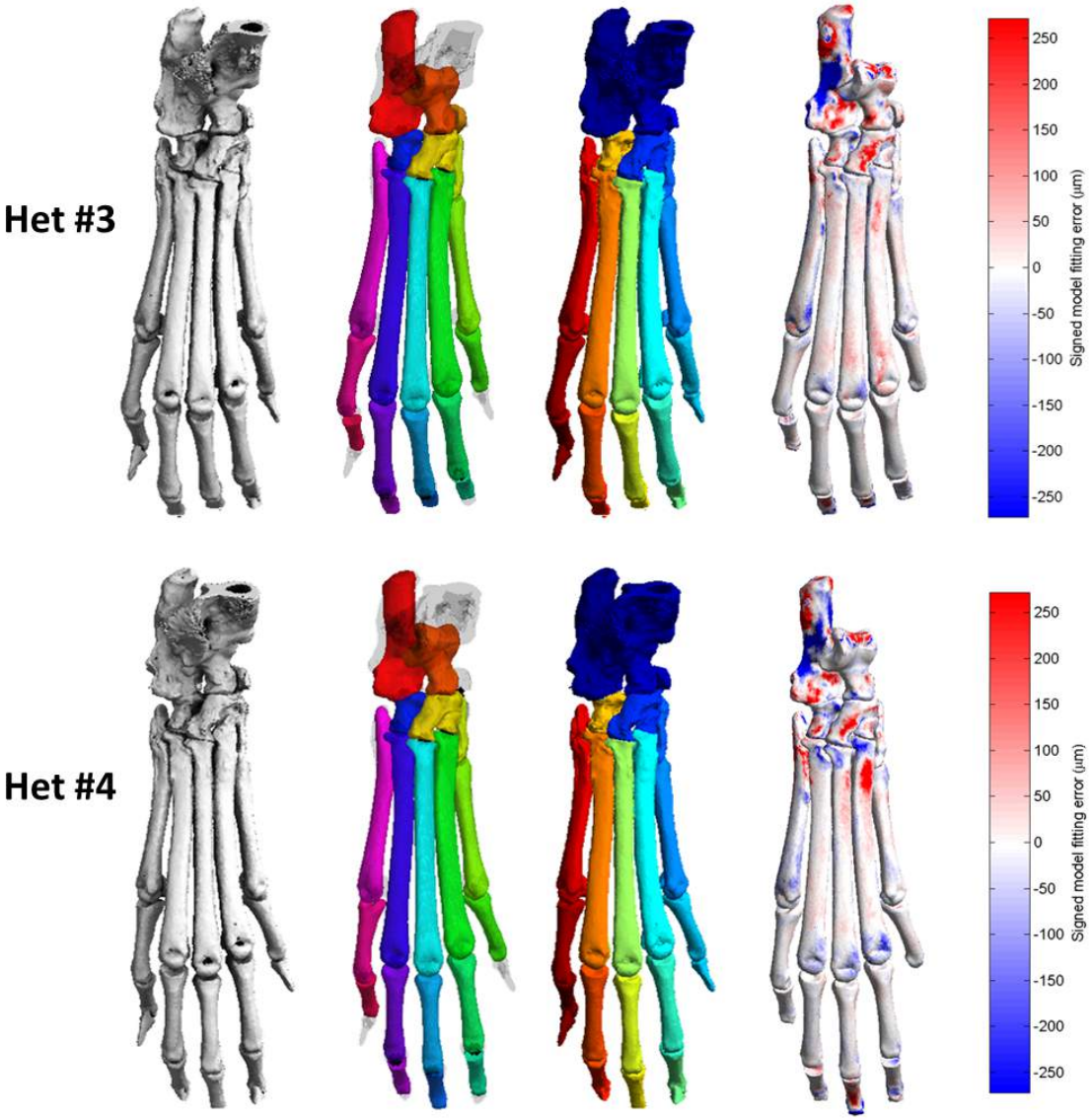
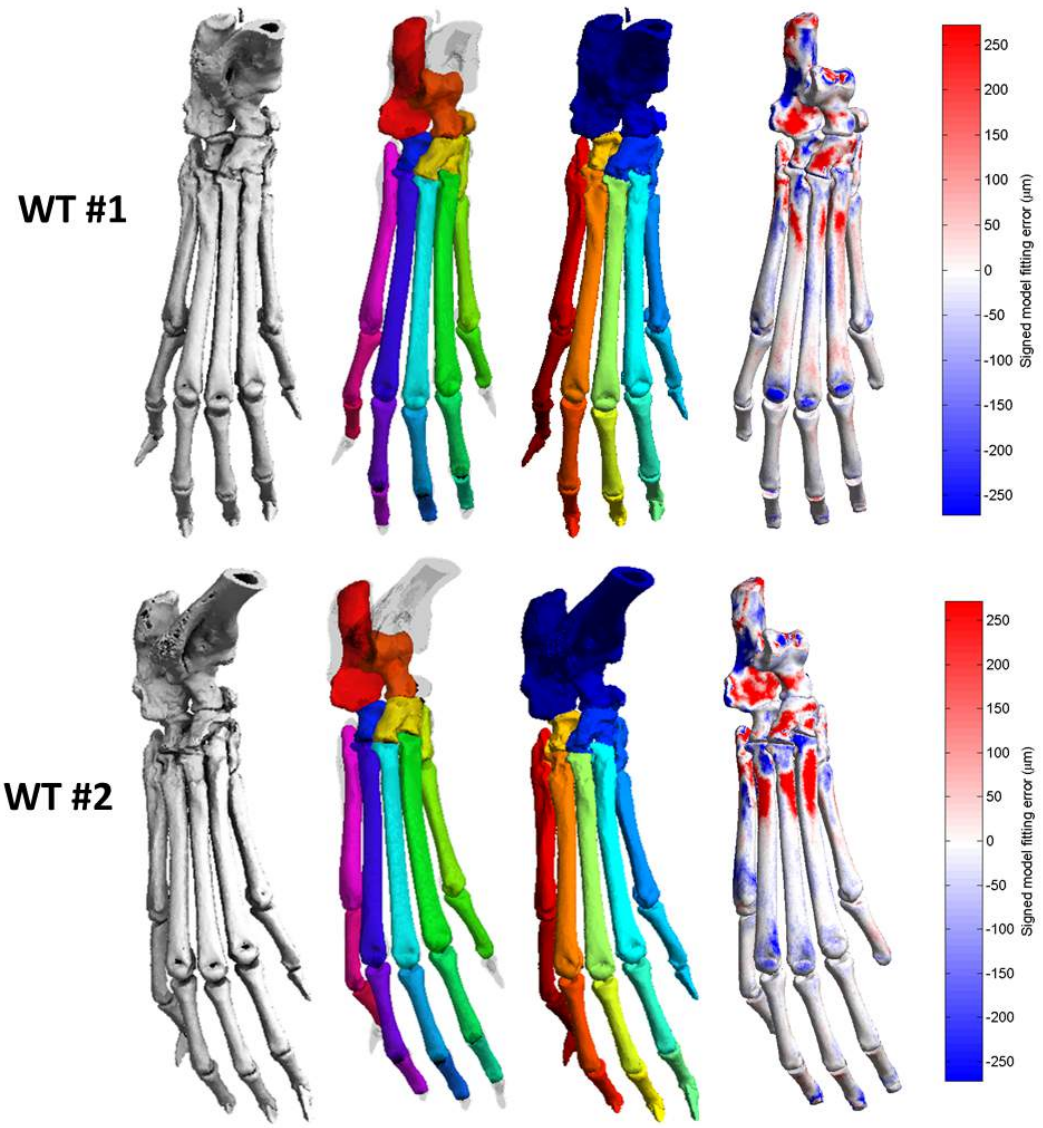


Figure 7.6: Application of the model to TTP $^{+/AA}$ mice. From left to right: original sample mesh, registration result, segmentation result, signed error mapping. The heterozygotes are somewhat variable in terms of severity, with #1 and #2 having relatively low levels of model fitting error compared to #3 and #4.


WT #1
WT #2
250
200
150
100
50
0
-50
-100
-150
-200
-250
Signed model fitting error (µm)

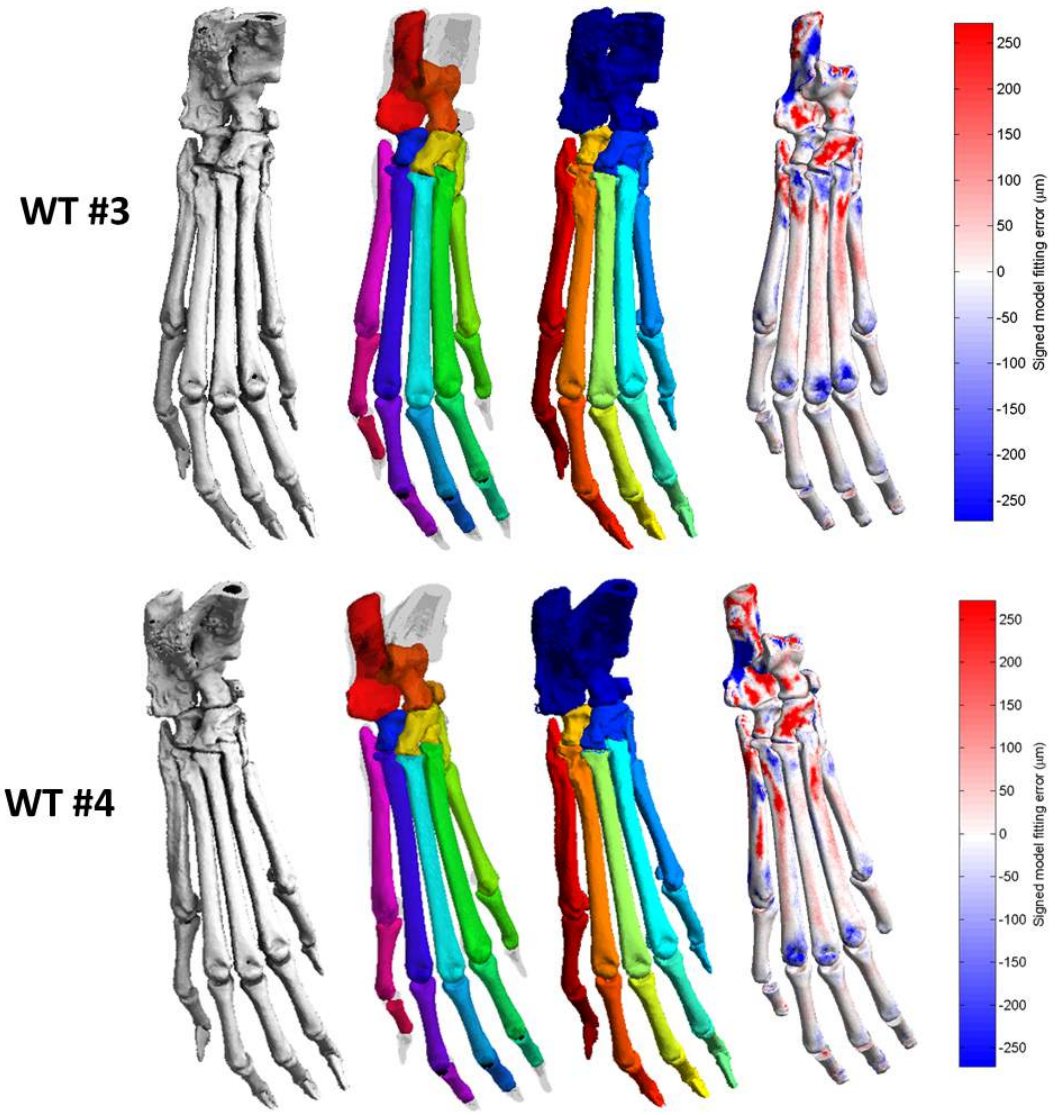


Figure 7.6: Application of the model to TTP $^{+/+}$ (wild-type) mice. From left to right: original sample mesh, registration result, segmentation result, signed error mapping. Examination of the micro-CT data reveals highly abnormal bone morphology of the calcaneus, distal tarsal and metatarsals. The heatmaps reveal large regions of abnormal bone growth on these bones, as well erosion around the metatarsophalangeal joints. Low levels of red spead along the length of the metatarsals suggested thickening of the bones, as opposed to localised bone remodelling.

7.2.4 Comparison between clinical and model scores

Clinical scores were acquired for the wild-type, heterozygous and homozygous mice as described in Section 6.2. Erosion and formation scores were acquired for each sample from three independent observers. The ASSM was then applied to each sample, and the surface areas of affected regions calculated as a percentage of total surface area.

Figure 7.7 compares the mean clinical scores with the mean ASSM-measured results, arranged by group and destruction type. Similar trends can be observed for both methods, with TTP $^{AA/AA}$ mice having notably lower scores/abnormal surface area than the wild-types. This is consistent with the predicted biological mechanism, as well as previous work (unpublished). The most notable difference between the two methods was the lower clinical score for TTP $^{AA/AA}$ bone formation, when compared to the model measured results.

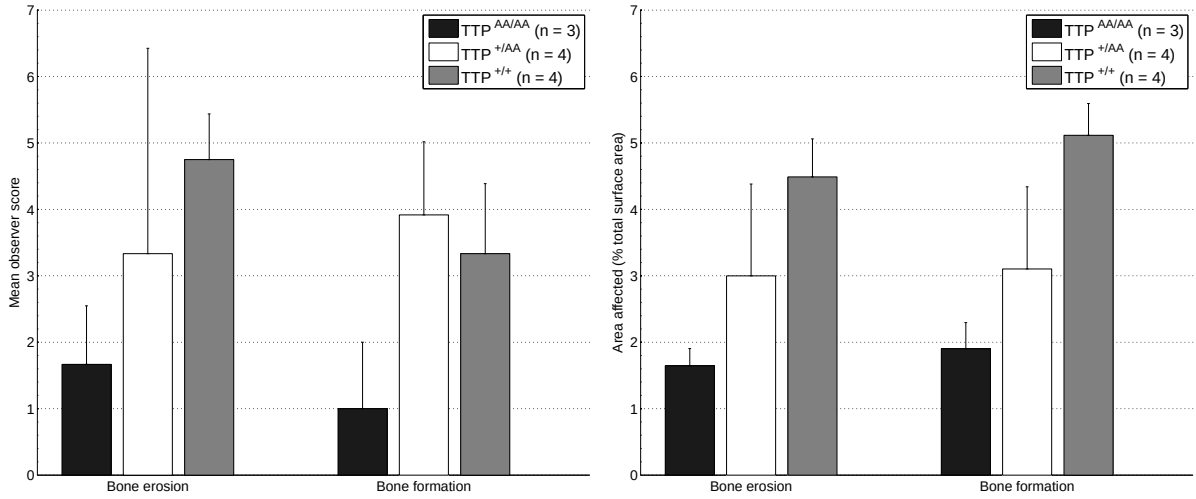


Figure 7.7: Comparison between mean clinical bone destruction scores (left) and the results of applying the ASSM as percentage surface area affected (right). Error bars represent one standard deviation. Both graphs exhibit similar trends, demonstrating the model's ability to reliably identify abnormalities.

7.2.5 Regional differences in bone destruction

Figure 7.7 compares the ASSM-measured surface areas, grouped by individual bone. With the exception of bones #9, #12 and #16 (all of which are metatarsals), the graph follows the same trend as observed for the combined scores in Figure 7.7. The low levels of destruction on the metatarsals suggest a detection limit of approximately 2% surface coverage, a value below which abnormalities cannot be reliably detected. The degree of erosion and formation is highly variable for the calcaneus and tibiale (#1 and #3), which was also observed for the clinical scores in that region.

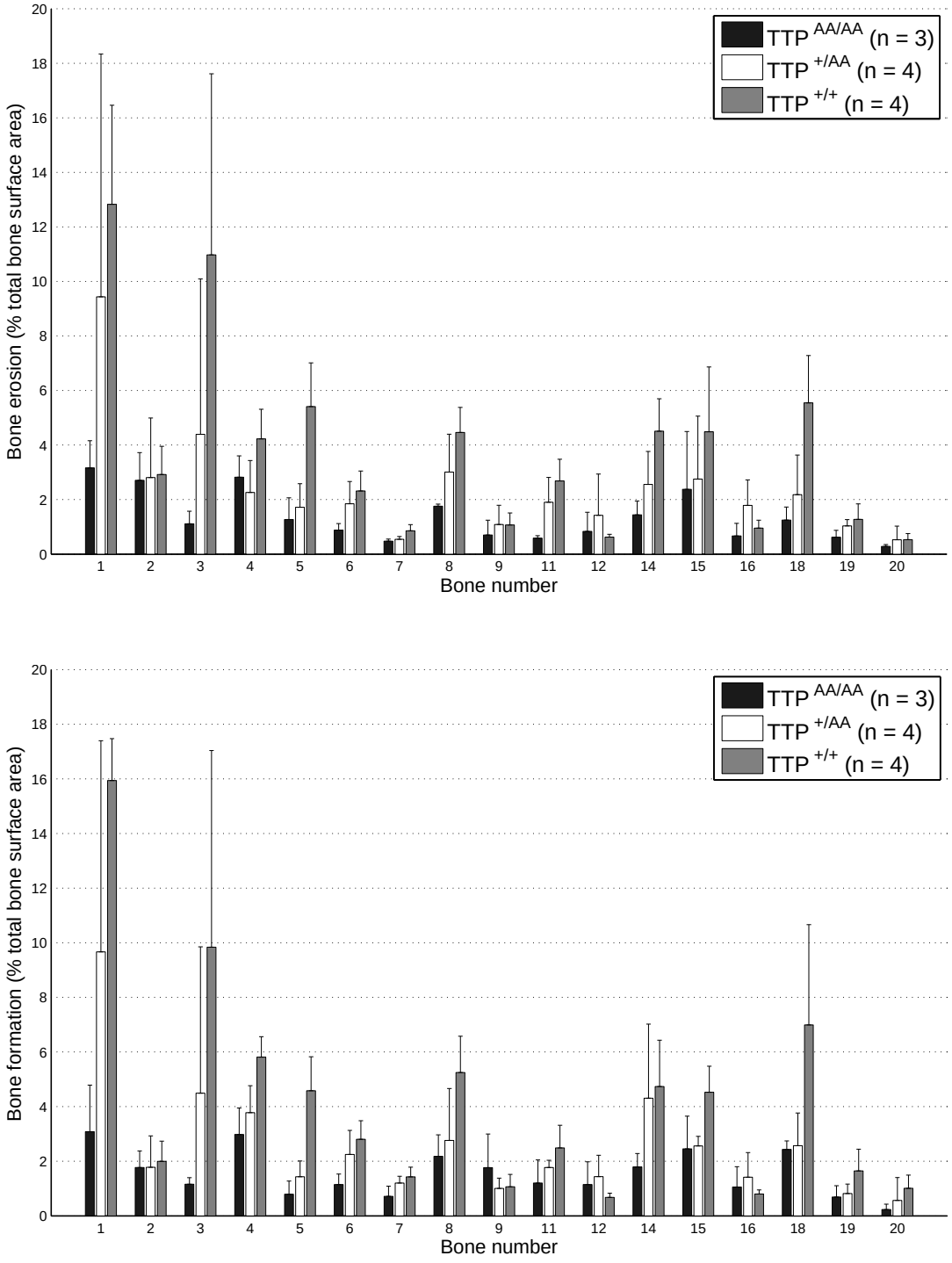


Figure 7.7: Bone erosion and formation as a percentage of total bone surface area. The values shown are averaged over the set of samples as shown in the legend. In general, TTP $^{+/+}$ (wild-type) and TTP $^{++/AA}$ have elevated levels of bone erosion and formation when compared to TTP $^{AA/AA}$ mice. This follows the trend that was also observed through physiological and micro-CT scoring, discussed in Section 7.2.4.

Chapter 8

Conclusions and future work

The adoption of mouse models in rheumatoid arthritis (RA) research is widespread, providing a powerful means to investigate the genes and processes that are implicated in disease. Bone destruction is an important hallmark of disease severity, which is typically erosive in nature and can affect multiple joints (polyarthritis). X-ray microtomography (micro-CT) is commonly used in mouse studies to visualise bone destruction in the fore and hind limbs, providing volumetric image data of bones at high resolution. The nature and severity of bone destruction in mouse models is of interest to members of the rheumatology research community, where different bone phenotypes may be indicative of certain modes of action or states of repair. Existing methods for the analysis of such destruction are often subjective, and provide only a partial assessment of the destruction that has occurred. In this thesis, an articulated statistical shape model (ASSM) was proposed as a viable approach for the analysis of bone destruction in mouse models of RA.

8.1 Method for model construction

Detection of abnormalities in medical images relies on eliminating differences that could be reasonably attributed to biological variability. The statistical shape model proposed by Cootes *et al.* [28] offers a solution to this problem, by capturing the shape variation that

exists among a set of training examples with known correspondence. A three-dimensional statstical model of non-pathological bone shape variation was proposed as a means to detect abnormalities in arthritic mice. Using high resolution micro-CT, images of non-pathological hind paw samples were acquired and processed to generate surface representations. A registration framework based on previous work by Baiker *et al.* [29] was then developed to align a single manually annotated reference sample onto the remaining samples. Global alignment of the reference with the samples based on their curvatures was appropriate for normalising their relative orientations, but in some cases failed to align them along their longitudinal axes. An additional alignment step based on their points of highest curvature ensured adequate initialisation for the subsequent articulated registration.

An anatomical hierarchy was designed for the mouse hind limb to facilitate the articulated registration. Bone transformations were constrained according to manually specified angular ranges, implemented as viewing frustums. This method was deemed appropriate for an object such as the mouse hind limb, due to it having several identically shaped bones in close proximity to one another. Additionally, the viewing frustums provided a visually intuitive way to adjust parameters experimentally. Having decided upon a set of "general case" frustum parameters, the articulated registration was applied to nine wild-type hind paw samples. Qualitatively, the registration framework was observed to produce accurate and consistent results across all of the samples used. Furthermore, each of the bones in the hierarchy were registered to a similar degree of accuracy. The registration error was consistent with these observations, demonstrating the method's robustness to differences in sample orientation and pose. Despite this, error propagation from parent bones to child bones was evident in some cases. Using the learned correspondence, labels were propagated from model to sample that would allow for individual bone shapes to be analysed. Labelling accuracy was not investigated quantatively, but deemed suitable for

shape analysis despite small labelling errors at joint boundaries.

Bone shape variation was analysed using principal component analysis (PCA). Having transformed the samples into a common coordinate system, differences in bone morphology were visualised as deviations from the average (mean) bone shape. In observing shape instances generated by the model, differences in bone morpology were identified that highlight the need for a model of wild-type variability.

8.2 Validation of the ASSM

Validation experiments were conducted to assess the model's accuracy when measuring abnormalities in different regions. Three leave-one-out experiments were conducted, having manually introduced abnormalities of a known surface area (ground truth). The results showed that in general, the model is more accurate at measuring bone erosion than bone formation. The reason for this is not entirely clear, but may be attributed to poor correspondence (and by extension, poor sensitivity) in regions where bone formation is common. This leads to false positives and false negatives, which translate into over- and underestimations in the area calculations, respectively. Additional experiments with a wider range of abnormality sizes are required to establish the limits of detection. Bone destruction scores were also acquired from three independent observers as a gold standard, and compared to the area calculations obtained from the model. Although not directly comparable, the model measured data showed reasonable agreement with bone erosion scores, but were less well correlated with bone formation scores. It is not entirely clear why this is the case, but likely to be due to systematic errors introduced in the model construction.

8.3 Analysis and comparison of mouse models

The K/BxN serum-transfer model is an immune mediate model that produces transient (resolving) inflammatory arthritis, but has a limited success rate. The ASSM was applied to four mice having been administered the K/BxN serum, of which two were "non-responders" based on clinical scores. Concordantly, the ASSM did not reveal any significant bone abnormalities in these mice. Following this, the ASSM was applied to the responsive mice to reveal notable both erosion of the metatarsals, as well as bone formation in heel region. The results of this analysis demonstrated the model's ability to discriminate between non-responsive and abnormal samples, and suggest that its limit of detection is within an acceptable range.

The CAIA mouse is similar to the K/BxN serum-transfer model in terms of the transient inflammation produced. The nature of the bone destruction is somewhat different however, presenting full thickness bone erosions that are only partially detected by the model. In addition, the presence of highly abnormal bone formation in one case caused the registration to fail. Despite this, the model was still able to highlight several interesting bone shape differences, including apparent "thinning" of cortical bone, as well as enthesophytes on the calcaneus. In contrast to the two K/BxN serum-transfer and CAIA models, the TNF dARE model is used in studies where chronic, non-resolving inflammation is required. The heatmap generated by the model indicated that TNF dARE mice have a primarily erosive phenotype, which is consistent with reported findings [18]. Interestingly, the erosions were more pronounced than those observed in the other models, which follows given that TNF dARE experience synovitis (joint inflammation) from birth.

An in-depth investigation into was also conducted for a novel mutant mouse, having been administered K/BxN serum. Homozygous, heterozygous and wild-type hind paw samples were analysed using the model, and the results compared with clinical scores.

Overall, the results showed very good agreement with clinical bone destruction scores, as well as the predicted trends in severity according to genotype. These results are highly encouraging, and suggest that the model is well suited to the task of detecting bone abnormalities in pre-clinical research scenarios.

8.4 Limitations and future work

Construction of an ASSM based on a single reference sample is inherently biased towards the geometry of the chosen sample. The effects of this bias were not investigated, but is an important area of investigation in future work. Different approaches to registration should also be investigated, with emphasis on utilising non-rigid correspondence in order to extract as much variation as possible from the training data. This would also improve the articulated registration accuracy at intermediate steps, minimising error propagation down the hierarchy.

Registration accuracy as assessed in terms of the error between corresponding points, which whilst useful does not inform of the "correctness" of the registration. Additional experiments should be undertaken to compare the results of label propagation with a gold standard, such as a manual segmentation. This would allow for problem areas to be highlighted, and help to explain the model's lack of sensitivity in certain regions. A more thorough validation is also required to determine the model's true capabilities and failures. Additional leave-one-out experiments spanning a wider range of abnormality sizes should be performed to determine the model's upper and lower limits of detection. Furthermore, colocalisation of abnormalities with annotated histological sections may be useful to further demonstrate its sensitivity.

8.5 Summary

In this thesis, an articulated statistical shape model has been demonstrated as a suitable approach for quantifying bone destruction in mouse models of RA. The proposed method for automatic construction of an ASSM has yielded promising results that demonstrate robustness to variation, and consistently low registration error across the training data. The results of validation experiments show that the model's accuracy varies by region, providing important insight into the limitations of the proposed method. Application of the model to real data acquired from a number of mouse models suggest that bone shape changes can be identified as deviations from the model statistics. The ASSM has also shown great promise in a more in-depth study of a novel transgenic mouse, with results that are supportive of previous findings.

The software that was developed during this research was written in Matlab, and has been applied to a number of mouse models. In order for this work to continue, an open-source Python implementation of the software is currently under development to provide a more user-friendly experience, and to eliminate the need for a Matlab license. The software is being designed to operate as a fully automated analysis pipeline that is capable of automatically generating graphs and figures. It is hoped that members of the biomedical community will take interest in this software, and use it to analyse their own mouse models in the foreseeable future.

Appendix 1

3D Articulated Registration of the Mouse Hind Limb for Bone Morphometric Analysis in Rheumatoid Arthritis

James M. Brown[1,2,3], Amy Naylor[3],
Chris Buckley[3], Andrew Filer[3], and Ela Claridge[1,2]

[1] PSIBS Doctoral Training Centre, University of Birmingham, UK
J.M.Brown.1@cs.bham.ac.uk
[2] School of Computer Science, University of Birmingham, UK
[3] Rheumatology Research Group, School of Immunity and Infection, University of Birmingham, UK

Abstract. We describe an automated method for building a statistical model of the mouse hind limb from micro-CT data, based on articulated registration. The model was initialised by hand-labelling the constituent bones and joints of a single sample. A coarse alignment of the entire model mesh to a sample mesh was followed by consecutive registration of individual bones and their descendants down a hierarchy. Transformation parameters for subsequent bones were constrained to a subset of vertices within a frustum projecting from a terminal joint of an already registered parent bone. Samples were segmented and transformed into a common coordinate frame, and a statistical shape model was constructed. The results of ten registered samples are presented, with a mean registration error of less than 40 μm ($\sim$ 3 voxels) for all samples. The shape variation amongst the samples was extracted by PCA to create a statistical shape model. Registration of the model to three unseen normal samples gives rise to a mean registration error of 5.84 μm, in contrast to 27.18 μm for three unseen arthritic samples. This may suggest that pathological bone shape changes in models of RA are detectable as departures from the model statistics.

1 Introduction

Rheumatoid arthritis (RA) is an autoimmune disease that affects approximately 1% of the world's population [1]. The autoimmune response mounted by the body gives rise to chronic inflammation of the synovial joints, which can cause active destruction of cartilage and bone. Although the exact cause of RA is unknown, new therapeutic targets may be discovered by investigating genes or processes that exacerbate or ameliorate disease progression. Animal models of inflammatory arthritis are frequently employed for this purpose, in conjunction with imaging techniques which provide data for deriving measures of disease severity [2, 3]. Histological scoring is commonly used to ascertain the amount of

bone destruction, whereas x-ray microtomography (micro-CT) provides qualitative assessments of the damage. These commonly used techniques are subjective. In response to the need of the biomedical community we are working towards developing objective and quantitative measures of bone destruction from micro-CT images of the mouse hind limb.

The hypothesis underpinning our work is that shapes of bones affected by a pathology depart from statically normal bone shape variations. When a diseased limb sample is registered with a statistical shape model of a normal limb, any diseased regions will show a gross departures from the model. Such regions can then be characterised as erosions or spurs, and have their morphology and volume assessed. Statistical shape models describe the variation that exists within a set of aligned training shapes described by points. The active shape model (ASM) is commonly used to identify shape instances in medical image data by utilising the variability extracted from the training set by principal component analysis (PCA) [4]. In building such a model, it necesary to establish point correspondences across the training set. This is often achieved by registering a single reference onto each sample, using algorithms such as iterative closet point (rigid) and B-spline free form deformation (non-rigid). This approach has been employed previously in building shape models of bones for the assessment of morphological variations in the primate humerus and scapula [5].

In our research, registration plays a vital role in both model construction and in abnormality detection. For the construction of a statistical shape model the individual samples must be co-registered in order to remove any differences that are not attributable to shape, such as their position, orientation and size. As the mouse hind limb is composed of multiple bones of various shapes and sizes, registration of a complete sample requires a 3D anatomical model that describes both structure and articulation. Having registered this model onto a series of samples, the pose-normalised bone shapes may be compared. The resulting model is similar to the hip joint model detailed in [6] in which only bone shape variations are modelled statistically, having previously aligned the samples based on known kinematic constraints. Finally, for abnormality detection, a sample in question must be co-registered with the model before establishing whether its shape deformations fall within the bounds defined by the model statistics. The closest work related to the bone pathology detection via model registration detailed the development of a statistical model of the rabbit femur, which was used to segment osteophytes (bone spurs) present in osteoarthritic femurs imaged by micro-CT [7]. Research described in this paper explores the possibility of identifying bone shape changes over the whole mouse hind limb in models of rheumatoid arthritis, such as periarticular bone loss and full thickness cortical bone damage. Although bone damage observed in RA is generally confined to the joints, we consider the entire hind limb in order to examine a variety of mouse models that may develop bone abnormalities elsewhere (e.g. spondyloarthropathy).

2 Method

This section first describes the methodology used to acquire and process the necessary image data used to construct an articulated model of the mouse hind limb. The framework for model-based registration and segmentation of a training set is then described, followed by the construction of a statistical model of non-pathological bone shapes and model validation. Analysis of an abnormal sample is performed by registering the articulated model, and iteratively deforming the mean bone shapes to produce the closest biologically feasible fit, and then assessing departures from the model as a measure of disease severity.

2.1 Image Acquisition and Processing

All experiments were carried out at the University of Birmingham following strict guidelines governed by the "Animal (Scientific Procedures) Act 1986" and approved by the local ethics committee. Female C57Bl/6 mice (Harlan, UK) were housed in individually ventilated cages in groups of 3-6 individuals on a 12 hour light-dark cycle with ad lib access to standard laboratory mouse chow diet and water. For arthritis experiments 200 µl KBxN serum was injected intraperitoneally into 10 week old mice, details of which can be found in [8]. All mice were sacrificed at 12 weeks of age. Both hind limbs were dissected and fixed in formalin over 24 hours in preparation for imaging.

Samples were imaged using a Skyscan1172 micro-CT scanner (Bruker), at a source voltage of 60 kV and source current of 167 µA, with a 0.5mm aluminium filter. Projections were taken every 0.45 degrees at 1000 ms exposure, with an image pixel size of 13.59 µm. Flat field corrections were performed to remove any effects caused by varied pixel sensitivity. Image slices (2000 x 2000 px) were reconstructed using NRecon 1.6.1.5 (Bruker), and beam hardening correction was applied to reduce cupping artefacts. Bone regions were segmented from soft tissue by global thresholding and a 3D surface mesh was computed using the marching cubes algorithm (CTAn 1.12, Bruker). The global threshold value was chosen manually, and kept consistent for all samples. Meshes were resampled using Poisson surface reconstruction to produce a smooth uniformly sampled mesh, and simplified using quadric edge collapse decimation [9]. Any mesh structure due to marrow space is of no interest in itself, and may misguide registration due to its highly variable morphology. Therefore, internal structures were isolated by ambient occlusion, and removed to give a completely hollow surface mesh (MeshLab 1.3.2, open-source).

2.2 Construction of an Articulated Model

To bootstrap the model construction, a single micro-CT scan of a wild-type mouse hind limb was first manually segmented into the constituent bones of interest by outlining individual slices (CTAn 1.12, Bruker). Global thresholding and mesh processing was then performed as described in Section 2.1. Joint

positions were approximated by isolating the articulating bone surface, and calculating the mean vertex position. In constructing the model, the leg bones (tibia and fibula) were ignored due to the limited field of view in the micro-CT instrument. Sesamoid bones and claws were ignored as they are irrelevant to pathology detection.

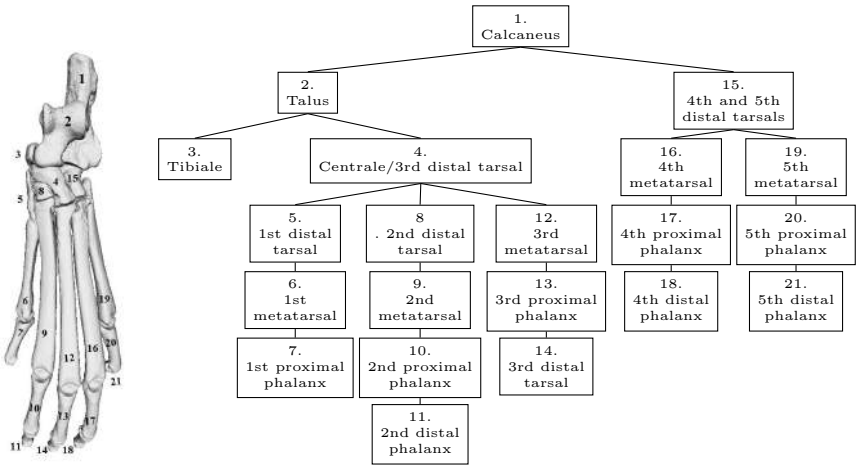


Fig. 1: (a) Micro-CT reconstruction of mouse hind limb with bones labelled. (b) Hierarchical representation of bones and their joints in the mouse hind limb. This representation provides the order in which model bones are registered to the sample mesh, which can be modified with ease for experimental purposes.

The topology of the mouse hind limb was represented as a tree (or hierarchy) where nodes and connections correspond to bones and joints, respectively (Figure 1). All nodes have exactly one parent node (except for the root node) with any geometric transformations applied to a parent bone being inherited by its children (i.e. if the 2nd metatarsal is rotated, then so are the 2nd proximal and distal phalanges). This hierarchical model was represented and stored as an eXstensible Markup Language (XML) document, allowing for construction of models with arbitrary hierarchies and traversal sequences.

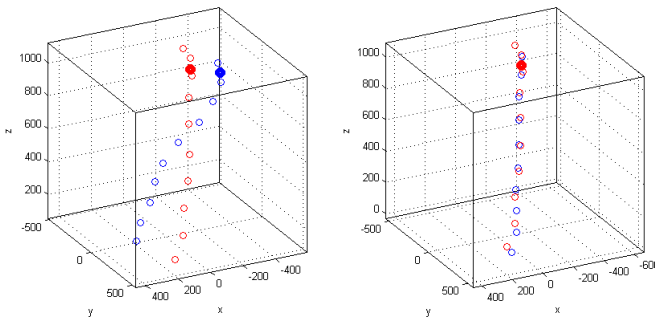
2.3 Articulated Registration

The articulated registration algorithm follows the scheme outlined in [10], where an initial coarse alignment of a whole mouse atlas with the sample is followed by consecutive registration of individual bones, initialising subsequent registration steps. The method described in this paper differs in several ways to account for

the differences in bone anatomy, joint complexity and proximity of parts. In particular (1) there is an additional coarse alignment of model and sample based on their "centres of mass"; (2) rigid (rotation, translation) and affine (rigid + scaling) transformations are carried out in separate ICP registration steps; and (3) motion constraint uses a viewing frustum, to account for the proximity and similarity of neighbouring components. Without incorporating these modifications registration yields unsatisfactory results in the form of misaligned bones. During this process, all transformations are applied to separate instances of the model, leaving individual samples stationary. In order to perform statistical shape analysis, sample bones are segmented (using the registration correspondence) and then inversely transformed into a common coordinate system as shown in Figure 4.

Coarse alignment

The coarse alignment process globally aligns the model and sample meshes, providing an initialisation for the subsequent articulated registration. The curvature of the model and sample limbs is first approximated by equally subdividing the image volume along the longitudinal axis (the axis about which the specimen is rotated in the micro-CT instrument) and computing the centroid for each subvolume. This gives rise to two corresponding "centre of gravity" curves, from which a rotation matrix can be computed by solving a system of linear equations in a least-squares fashion.



(a) (b)

Fig. 2: Coarse alignment of model (blue) and sample (red) "centre of gravity curves" shown (a) before coarse alignment, and (b) after coarse alignment. The two larger points shown are the respective centres of mass. The alignment of these points aims to filter out rotational and translational effects associated with mounting the sample in the micro-CT scanner, prior to performing articulated registration.

The rotation matrix serves to align the meshes such that they both face the same direction. This does not however guarantee that the two meshes are aligned along the longitudinal axis. This is achieved by approximating their centres of mass, located near the ankle joint. The whole image volume is first projected onto its xy and yx planes, and the brightest pixels in each projection are then located (corresponding to the thickest regions of the specimen.) Of these pixels the topmost (nearest the leg) is chosen, and the centroid of the slice in which it resides is computed. The difference between the two centres of mass is used to determine the translational offset. The rotation and translation are then applied to the model, aligning it coarsely with the sample mesh. Figure 2 shows the result of applying the two transformations to two example curves.

Motion Constraints

Having coarsely aligned the model and sample mesh, the individual bones are registered by ICP consecutively down the hierarchy, with connected sub-trees inheriting the transformations computed at each step. The bones that comprise the mouse hind limb can be grouped into three shape categories; long bones (metatarsals, phalanges), small bones (tibiale, 1st and 2nd distal tarsal) and irregular bones (talus, calcaneus, 3rd - 5th distal tarsals). Small and irregular bones have limited natural motion, and so the entire search space is made available to the iterative closest point (ICP) algorithm when approximating point correspondence. By contrast, long bones have a greater range of motion which can yield an incorrect registration result if rotation is not constrained. This problem has been solved using a *field of view* approach.

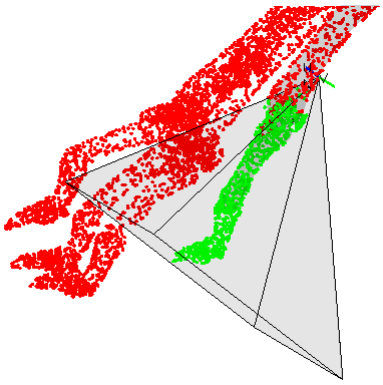


Fig. 3: Motion constraint as applied to a proximal phalanx bone by viewing frustum culling. The set of legal points within the viewing frustum are shown in green, and the set of illegal (culled) points are shown in red.

Having already registered its parent, registration of a child bone begins by finding an initial set of corresponding points by nearest-neighbour criteria. Rather than testing against all of the available sample points, the set is reduced to a set of feasible points that fall within a viewing frustum, parametrised by four angles (up, down, left and right). The viewing frustum is projected from the end of the parent bone along its principal axis, and the vertices that fall outside the frustum are eliminated (Figure 3). Correspondence is then approximated from the remaining points, and the optimal transformation found by ICP. The parametrisation of the viewing frustums for each bone was determined manually, and found to be consistent for all of samples used in the results presented. Having calculated the optimal rigid transformation, the entire search space is opened up once again for an additional ICP step that solves for differences in scaling. The iterative closest point (ICP) algorithm used in this work is a freely available MATLAB implementation [4].

Segmentation and Shape Modelling

The result of registering the articulated model to a set of n training samples is a set of n transformed model instances. The point correspondence that the registration yields is used to segment the individual samples, by propagating labels between model and sample vertices. Each of the meshes is composed of several thousand points, and in all likelihood will not have the same exact amount in each. As a result, not every point will receive a label, and so unlabelled mesh patches are assigned a label based on neighbouring vertices (those that share an edge). Anatomical structures not represented in the model (e.g. claws, small non-articulating bones) are left unlabelled.

The labelled samples are inversely transformed using the learned registration parameters so that the whole training set adopts a common coordinate frame. The coregistered samples can now be integrated into a multi-part statistical model, where shape variation of each individual bone is modelled separately. For each bone, an $n \times 3m$ system matrix is formed (where m is the number of points) from which the mean shape is subtracted from each row and a covariance matrix P is computed. Principal component analysis (PCA) is performed to compute the eigenvectors of P, which correspond to the modes of shape variation, in order of decreasing variance. New shape instances may be generated from the model as weighted deviations from the mean shape along the first p modes:

$$X = \bar{X} + Pb \tag{1}$$

where b is a vector of weights, constrained to fall within three standard deviations of the mean. Normal bone shapes are approximated by finding a vector b for each bone that minimises the least-squared distance (using the Levenberg-Marquardt algorithm) between the model and sample points. By constraining

[4] Finite Iterative Closest Point: http://www.mathworks.co.uk/matlabcentral/fileexchange/24301-finite-iterative-closest-point)

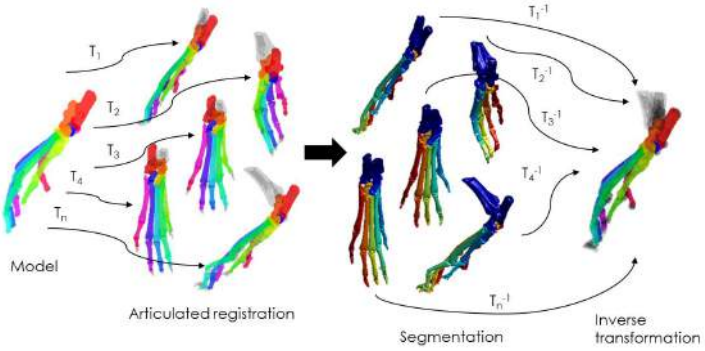


Fig. 4: The registration and segmentation workflow for statistical shape model construction. The process begins with the articulated model being registered onto each of the n samples, which remain stationary. After registration, each sample is segmented according to the learned correspondence, and inversely transformed into the common (model) coordinate system according to the transformations gathered during registration.

the deformations to biologically feasible limits, the differences between the model and sample may be attributed to pathological shape changes.

3 Results and Validation

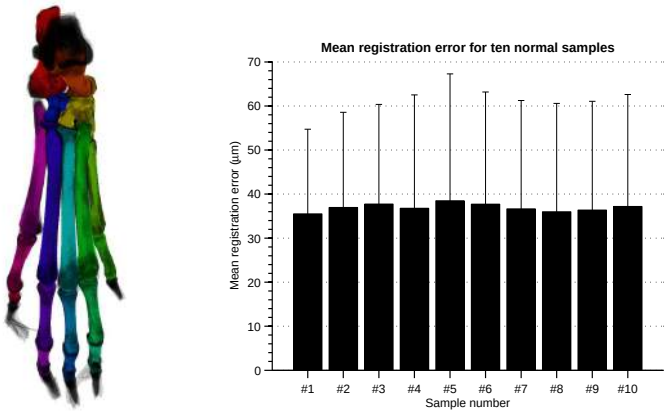
Ten wild-type (normal) mouse hind limb samples (5 females, C57Bl/6, 12 weeks old) were acquired and imaged by micro-CT, and processed according to protocols outlined in Section 2.1. Articulated registration was performed on the ten samples which are shown overlaid with the original model in Figure 5 alongside the mean registration error for each sample. Registration error is defined as the mean Euclidean distance in voxels between model (M) and corresponding sample (S) points:

$$E(M,S) = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(M(x_i,y_i,z_i) - S(x_i,y_i,z_i))^2} \tag{2}$$

To test the model's ability to approximate samples from outside the training set, articulated registration and shape model fitting was applied to three normal samples (from outside the training set) and three arthritic samples. The results are shown in Figure 6 as mean error histograms.

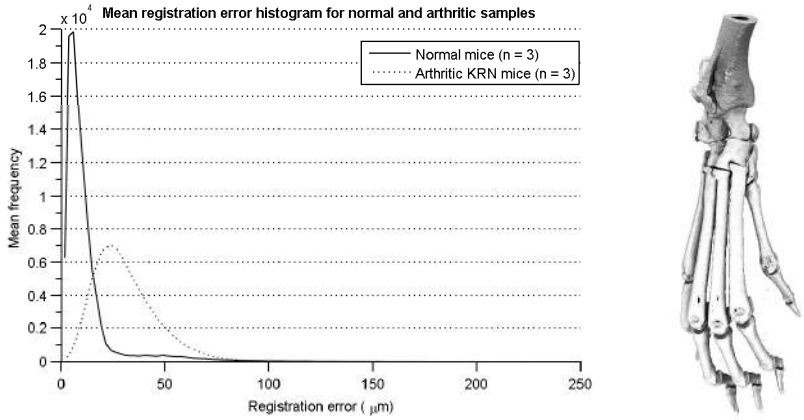
4 Conclusion

In this paper, we demonstrated a method for constructing a statistical model of the mouse hind limb. Manual segmentation and labelling of a single sample



(a) (b)

Fig. 5: Results of articulated registration of ten samples: (a) the original model overlaid with the registered samples and (b) the mean registration error for the ten samples. The results demonstrate a registration accuracy of less than 40 μm (∼3 voxels) over the whole hind limb (error bars correspond to one standard deviation).



(a) (b)

Fig. 6: (a) Result of articulated registration with shape model fitting to three unseen normal samples and three unseen arthritic samples. On average, the normal samples are more accurately approximated by the model that the arthritic samples (with mean errors of 5.84 μm and 27.18 μm respectively) indicating the presence of statistically abnormal shape features. (b) An example of a KRN mouse hind limb, with evidence of bone destruction around the metatarsals.

provides an articulated model which may be used to extract bone shape variation from a set of unlabelled samples. Registration of the model onto a set of ten samples achieved a mean registration error of less than 40 μm ($\sim$ 3 voxels). Whilst some errors are due to natural shape variation, others may attributed to cumulative misregistration of connected parts. The effect of coarse aligment on the final results will therefore be investigated in future work. After registration, labels were transferred onto the samples which were then inversely transformed into a common (model) coordinate system. The co-registered samples were used to build a bone-by-bone statistical model of shape variations via PCA. The ability of the model to represent unseen normal shapes was successfully validated by registering a small set of normal samples and obtaining a mean registration error of 5.84 μm voxels. A mean registration error for unseen abnormal samples was larger, at 27.18 μm, and its mode was shifted towards larger displacements. This suggests that the erosions and spurs present in the abnormal samples depart from the model of normal bone shape. Further analysis will be aimed at demonstrating that the latter results are mainly due to large errors at the locations of the arthritic abnormalities, which may be mapped onto the meshes to determine their height/depth or volume. Computing the differences between an abnormal sample and a normal model at these locations will aid the quantification of pathologies.

References

1. Buckley, C.D.: Science, medicine, and the future: Treatment of rheumatoid arthritis. Brit. Med. J. **315** 236–238 (1997)
2. Asquith, D.L., Miller, A.M., and McInnes, I.B. and Liew, F.Y.: Animal models of rheumatoid arthritis. Eur J Immunol. **39** 2040–2044 (2009)
3. Diarra, D., Stolina, M., Polzer, K., Zwerina, J *et al.*: Dickkopf-1 is a master regulator of joint remodeling. Nat. Med. **13** 156–163 (2007)
4. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J. and others.: Active shape models - their training and application. Comput. Vis. Image. Und. **6** 38–59 (1995)
5. Yang, Y.: Shape modelling of bones: application to the primate shoulder. Ph.D thesis, Imperial College London, UK (2008)
6. Kainmueller, D. and Lamecker, H. and Zachow, S. and Hege, H.C.: An articulated statistical shape model for accurate hip joint segmentation. IEEE EMB-M. 6345–6351 (2009)
7. Saha, P.K., Liang, G., Elkins, J.M., Coimbra, A., Duong, L.T., Williams, D.S. and Sonka, M.: A New Osteophyte Segmentation Algorithm using Partial Shape Model and its Applications to Rabbit Femur Anterior Cruciate Ligament Transection via Micro-CT Imaging. IEEE T. Bio-Med. Eng. **55** 2212–2227 (2011)
8. Monach, P. A., Mathis, D. and Benoist, C.: The K/BxN Arthritis Model. Curr. Protoc. Immunol. **81:15:22** 1–12 (2008)
9. Garland, M and Heckbert, P.S.: Surface simplification using quadric error metrics. SIGGRAPH. 209–216 (1997)
10. Baiker, M., Milles, J., Vossepoel, A.M., Que, I., Kaijzel, E.L., Lowik, C., Reiber, J.H.C., Dijkstra, J. and Lelieveldt, B.P.F.: Fully automated whole-body registration in mice using an articulated skeleton atlas. IEEE ISBI: From Nano to Macro. 728–731 (2007)

Appendix 2

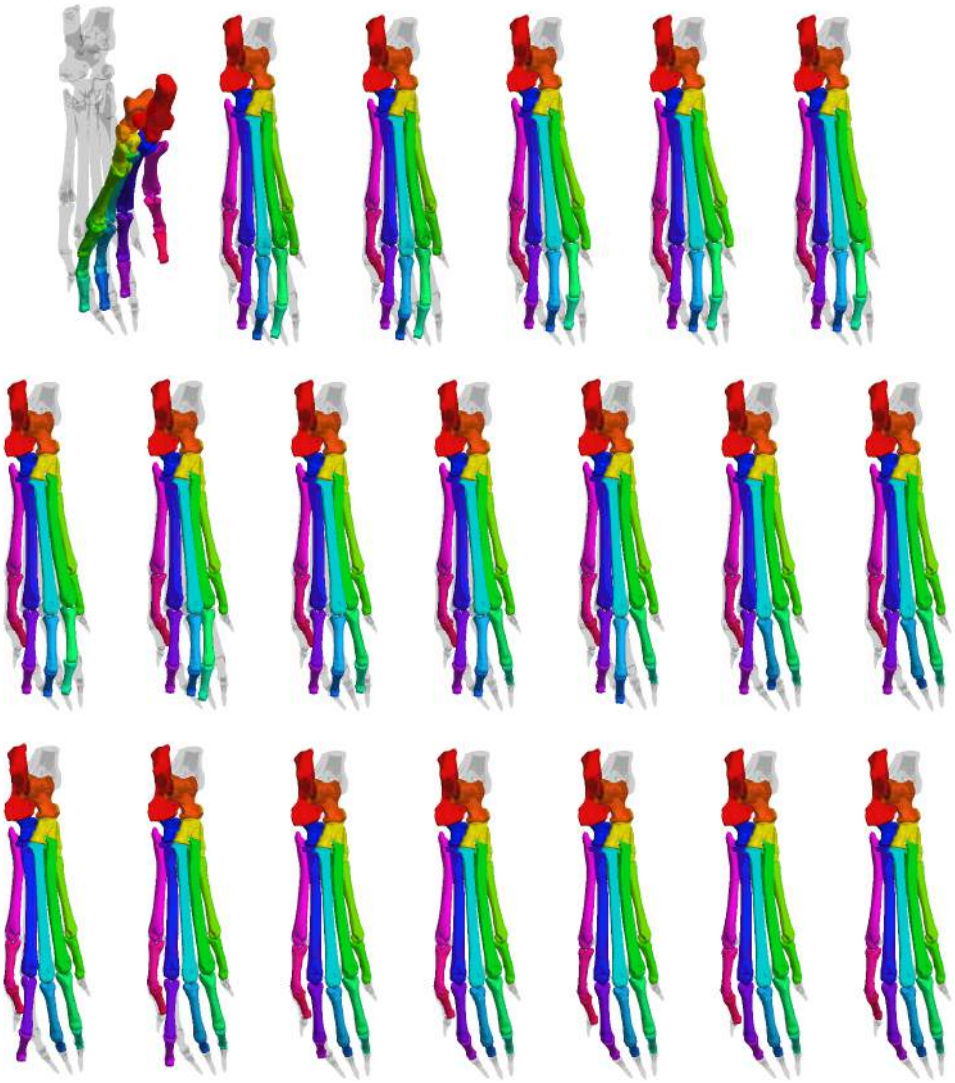


Figure 1: The hierarchical registration process. The first step (top left) shows the model and sample prior to global alignment. Going from left to right, the intermediate hierarchical registration steps are shown.

Appendix 3

Wild-types	Operator #1		Operator #2		Operator #3		Average		Standard deviation	
	Erosion	Formation	Erosion	Formation	Erosion	Formation	Erosion	Formation	Erosion	Formation
Heel region	2	0	2	0	3	3	2.33	1.00	0.47	1.41
Metatarsals	3	0	1	0	3	3	2.33	1.00	0.94	1.41
Phalanges	0	0	0	0	0	1	0.00	0.33	0.00	0.47
Totals	**5**	**0**	**3**	**0**	**6**	**7**	**4.67**	**2.33**	**1.25**	**3.30**
Heel region	3	0	3	3	2	3	2.67	2.00	0.47	1.41
Metatarsals	3	1	3	2	2	2	2.67	1.67	0.47	0.47
Phalanges	0	0	1	0	0	0	0.33	0.00	0.47	0.00
Totals	**6**	**1**	**7**	**5**	**4**	**5**	**5.67**	**3.67**	**1.25**	**1.89**
Heel region	3	1	3	3	3	1	3.00	1.67	0.00	0.94
Metatarsals	1	0	1	1	1	1	1.00	0.67	0.00	0.47
Phalanges	0	0	0	1	0	0	0.00	0.33	0.00	0.47
Totals	**4**	**1**	**4**	**5**	**4**	**2**	**4.00**	**2.67**	**0.00**	**1.70**
Heel region	1	3	2	3	3	3	2.00	3.00	0.82	0.00
Metatarsals	3	2	2	1	3	2	2.67	1.67	0.47	0.47
Phalanges	0	0	0	0	0	0	0.00	0.00	0.00	0.00
Totals	**4**	**5**	**4**	**4**	**6**	**5**	**4.67**	**4.67**	**0.94**	**0.47**
					Average		**4.75**	**3.33**		
					Standard deviation		**0.69**	**1.05**		

Hets	Operator #1		Operator #2		Operator #3		Average		Standard deviation	
	Erosion	Formation	Erosion	Formation	Erosion	Formation	Erosion	Formation	Erosion	Formation
Heel region	0	0	0	0	0	1	0.00	0.33	0.00	0.47
Metatarsals	0	0	0	2	0	0	0.00	0.67	0.00	0.94
Phalanges	1	0	0	2	0	1	0.33	1.00	0.47	0.82
Totals	**1**	**0**	**0**	**4**	**0**	**2**	**0.33**	**2.00**	**0.47**	**1.63**
Heel region	2	2	2	3	2	3	2.00	2.67	0.00	0.47
Metatarsals	3	0	3	2	3	2	3.00	1.33	0.00	0.94
Phalanges	0	0	3	0	0	0	1.00	0.00	1.41	0.00
Totals	**5**	**2**	**8**	**5**	**5**	**5**	**6.00**	**4.00**	**1.41**	**1.41**
Heel region	0	0	2	0	0	0	0.67	0.00	0.94	0.00
Metatarsals	0	0	0	0	0	0	0.00	0.00	0.00	0.00
Phalanges	0	0	0	2	0	0	0.00	0.67	0.00	0.94
Totals	**0**	**0**	**2**	**2**	**0**	**0**	**0.67**	**0.67**	**0.94**	**0.94**
Heel region	0	0	0	3	0	3	0.00	2.00	0.00	1.41
Metatarsals	0	0	0	2	0	0	0.00	0.67	0.00	0.94
Phalanges	0	1	1	2	0	3	0.33	2.00	0.47	0.82
Totals	**0**	**1**	**1**	**7**	**0**	**6**	**0.33**	**4.67**	**0.47**	**2.62**
Heel region	2	1	3	3	2	3	2.33	2.33	0.47	0.94
Metatarsals	3	0	3	3	3	2	3.00	1.67	0.00	1.25
Phalanges	0	0	2	2	0	0	0.67	0.67	0.94	0.94
Totals	**5**	**1**	**8**	**8**	**5**	**5**	**6.00**	**4.67**	**1.41**	**2.87**
Heel region	0	0	0	2	0	2	0.00	1.33	0.00	0.94
Metatarsals	1	0	0	2	0	0	0.33	0.67	0.47	0.94
Phalanges	1	0	0	1	1	0	0.67	0.33	0.47	0.47
Totals	**2**	**0**	**0**	**5**	**1**	**2**	**1.00**	**2.33**	**0.82**	**2.05**
					Average		**3.33**	**3.92**		
					Standard deviation		**3.09**	**1.10**		

AA	Operator #1		Operator #2		Operator #3		Average		Standard deviation	
	Erosion	Formation	Erosion	Formation	Erosion	Formation	Erosion	Formation	Erosion	Formation
Heel region	0	0	0	0	0	0	0.00	0.00	0.00	0.00
Metatarsals	1	0	0	2	0	0	0.33	0.67	0.47	0.94
Phalanges	1	0	3	1	1	0	1.67	0.33	0.94	0.47
Totals	**2**	**0**	**3**	**3**	**1**	**0**	**2.00**	**1.00**	**0.82**	**1.41**
Heel region	0	0	0	0	1	0	0.33	0.00	0.47	0.00
Metatarsals	3	0	0	0	3	0	2.00	0.00	1.41	0.00
Phalanges	0	0	0	0	0	0	0.00	0.00	0.00	0.00
Totals	**3**	**0**	**0**	**0**	**4**	**0**	**2.33**	**0.00**	**1.70**	**0.00**
Heel region	0	0	1	2	1	1	0.67	1.00	0.47	0.82
Metatarsals	0	0	0	2	0	1	0.00	1.00	0.00	0.82
Phalanges	0	0	0	0	0	0	0.00	0.00	0.00	0.00
Totals	**0**	**0**	**1**	**4**	**1**	**2**	**0.67**	**2.00**	**0.47**	**1.63**
					Average		**1.67**	**1.00**		
					Standard deviation		**0.88**	**1.00**		

Figure 2: Raw clinical scoring data collected by three independent observers. Samples highlighted in green were used in Chapter 7.2, whereas those highlighted in red were not used in analysis.

List of References

[1] C. Buckley, "Science, medicine, and the future: treatment of rheumatoid arthritis," *BMJ*, vol. 315, no. 7102, pp. 236–238, 1997.

[2] S. R. Goldring, "Bone and joint destruction in rheumatoid arthritis: what is really happening?," *The Journal of Rheumatology*, vol. 65, pp. 44–48, 2002.

[3] K. Sato and H. Takayanagi, "Osteoclasts, rheumatoid arthritis, and osteoimmunology," *Current Opinion in Rheumatology*, vol. 18, no. 4, pp. 419–426, 2006.

[4] M. Feldmann, "Development of anti-TNF therapy for rheumatoid arthritis," *Nature Reviews Immunology*, vol. 2, no. 5, pp. 364–371, 2002.

[5] P. Li and E. M. Schwarz, "The TNF-α transgenic mouse model of inflammatory arthritis," in *Springer Seminars in Immunopathology*, vol. 25, pp. 19–33, Springer, 2003.

[6] A. Pettit, H. Ji, D. Von Stechow, R. Müller, S. Goldring, Y. Choi, C. Benoist, and E. Gravallese, "TRANCE/RANKL knockout mice are protected from bone erosion in a serum transfer model of arthritis," *The American Journal of Pathology*, vol. 159, no. 5, pp. 1689–1699, 2001.

[7] K. Aya, M. Alhawagri, A. Hagen-Stapleton, H. Kitaura, O. Kanagawa, D. V. Novack, *et al.*, "NF-κB–inducing kinase controls lymphocyte and osteoclast activities

in inflammatory arthritis," *The Journal of clinical investigation*, vol. 115, no. 7, pp. 1848–1854, 2005.

[8] T. L. Vincent, R. O. Williams, R. Maciewicz, A. Silman, P. Garside, and Arthritis Research UK animal models working group, "Mapping pathogenesis of arthritis through small animal models," *Rheumatology*, vol. 51, no. 11, pp. 1931–1941, 2012.

[9] P. A. Monach, D. Mathis, and C. Benoist, "The K/BxN arthritis model," *Current protocols in immunology*, pp. 15–22, 2008.

[10] D. Kyburz and M. Corr, "The KRN mouse model of inflammatory arthritis," in *Springer seminars in immunopathology*, vol. 25, pp. 79–90, Springer, 2003.

[11] B. A. Binstadt, P. R. Patel, H. Alencar, P. A. Nigrovic, D. M. Lee, U. Mahmood, R. Weissleder, D. Mathis, and C. Benoist, "Particularities of the vasculature can promote the organ specificity of autoimmune attack," *Nature Immunology*, vol. 7, no. 3, pp. 284–292, 2006.

[12] P. S. Burrage, K. S. Mix, C. E. Brinckerhoff, *et al.*, "Matrix metalloproteinases: role in arthritis," *Front Biosci*, vol. 11, no. 1, pp. 529–543, 2006.

[13] G. Kollias, P. Papadaki, F. Apparailly, M. Vervoordeldonk, R. Holmdahl, V. Baumans, C. Desaintes, J. Di Santo, J. Distler, P. Garside, *et al.*, "Animal models for arthritis: innovative tools for prevention and treatment," *Annals of the Rheumatic Diseases*, vol. 70, no. 8, pp. 1357–1362, 2011.

[14] K. S. Nandakumar and R. Holmdahl, "Antibody-induced arthritis: disease mechanisms and genes involved at the effector phase of arthritis," *Arthritis research & therapy*, vol. 8, no. 6, p. 223, 2007.

[15] D. L. Asquith, A. M. Miller, I. B. McInnes, and F. Y. Liew, "Animal models of rheumatoid arthritis," *European Journal of Immunology*, vol. 39, no. 8, pp. 2040–2044, 2009.

[16] G. Shaw and R. Kamen, "A conserved AU sequence from the 3 untranslated region of GM-CSF mRNA mediates selective mRNA degradation," *Cell*, vol. 46, no. 5, pp. 659–667, 1986.

[17] C. Barreau, L. Paillard, and H. B. Osborne, "AU-rich elements and associated factors: are there unifying principles?," *Nucleic acids research*, vol. 33, no. 22, pp. 7138–7150, 2005.

[18] D. Kontoyiannis, M. Pasparakis, T. T. Pizarro, F. Cominelli, and G. Kollias, "Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies," *Immunity*, vol. 10, no. 3, pp. 387–398, 1999.

[19] E. A. Ross, T. Smallie, Q. Ding, J. D. O'Neil, H. E. Cunliffe, T. Tang, D. R. Rosner, I. Klevernic, N. A. Morrice, C. Monaco, A. F. Cunningham, C. D. Buckley, J. Saklatvala, J. L. Dean, and A. R. Clark, "Dominant Suppression of Inflammation via Targeted Mutation of the mRNA Destabilizing Protein Tristetraprolin," *Journal of Immunology*, vol. 195, no. 1, pp. 265–276, 2015.

[20] D. Diarra, M. Stolina, K. Polzer, J. Zwerina, M. Ominsky, D. Dwyer, A. Korb, J. Smolen, M. Hoffmann, C. Scheinecker, *et al.*, "Dickkopf-1 is a master regulator of joint remodeling," *Nature medicine*, vol. 13, no. 2, pp. 156–163, 2007.

[21] S. Schambach, S. Bag, L. Schilling, C. Groden, and M. Brockmann, "Application of micro-CT in small animal imaging," *Methods*, vol. 50, no. 1, pp. 2–13, 2010.

[22] L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *JOSA A*, vol. 1, no. 6, pp. 612–619, 1984.

[23] P. Pine, B. Chang, N. Schoettler, M. Banquerigo, S. Wang, A. Lau, F. Zhao, E. Grossbard, D. Payan, and E. Brahn, "Inflammation and bone erosion are suppressed in models of rheumatoid arthritis following treatment with a novel Syk inhibitor," *Clinical Immunology*, vol. 124, no. 3, pp. 244–257, 2007.

[24] A. Corthay, A.-S. Hansson, and R. Holmdahl, "T lymphocytes are not required for the spontaneous development of entheseal ossification leading to marginal ankylosis in the DBA/1 mouse," *Arthritis & Rheumatism*, vol. 43, no. 4, pp. 844–851, 2000.

[25] A. Bakker, F. van de Loo, H. Van Beuningen, P. Sime, P. van Lent, P. Van der Kraan, C. Richards, and W. Van Den Berg, "Overexpression of active TGF-beta-1 in the murine knee joint: evidence for synovial-layer-dependent chondro-osteophyte formation," *Osteoarthritis and Cartilage*, vol. 9, no. 2, pp. 128–136, 2001.

[26] S. Yang, A. M. Hollister, E. A. Orchard, S. I. Chaudhery, D. V. Ostanin, S. J. Lokitz, and J. M. Mathis, "Quantification of bone changes in a collagen induced arthritis mouse model by reconstructed three dimensional micro-CT," *Biological procedures online*, vol. 15, no. 1, pp. 1–8, 2013.

[27] A. A. Joshi, R. M. Leahy, R. D. Badawi, and A. J. Chaudhari, "Morphometry for early monitoring of treatment response in rheumatoid arthritis," in *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pp. 121–124, IEEE, 2013.

[28] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[29] M. Baiker, J. Milles, A. Vossepoel, I. Que, E. Kaijzel, C. Lowik, J. Reiber, J. Dijkstra, and B. Lelieveldt, "Fully automated whole-body registration in mice using an articulated skeleton atlas," in *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pp. 728–731, IEEE, 2007.

[30] G. K. Tam, Z.-Q. Cheng, Y.-K. Lai, F. C. Langbein, Y. Liu, D. Marshall, R. R. Martin, X.-F. Sun, and P. L. Rosin, "Registration of 3d point clouds and meshes: a survey from rigid to nonrigid," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 19, no. 7, pp. 1199–1217, 2013.

[31] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Robotics-DL tentative*, pp. 586–606, International Society for Optics and Photonics, 1992.

[32] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Communications of the ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[33] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding*, vol. 89, no. 2, pp. 114–141, 2003.

[34] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 12, pp. 2262–2275, 2010.

[35] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," in *ACM Siggraph Computer Graphics*, vol. 21, pp. 163–169, ACM, 1987.

[36] M. Baiker, J. Dijkstra, I. Que, C. Lowik, J. Reiber, and B. Lelieveldt, "Organ approximation in μct data with low soft tissue contrast using an articulated whole-body

atlas," in *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pp. 1267–1270, IEEE, 2008.

[37] A. Khmelinskii, M. Baiker, E. L. Kaijzel, J. Chen, J. H. Reiber, and B. P. Lelieveldt, "Articulated whole-body atlases for small animal image analysis: construction and applications," *Molecular Imaging and Biology*, vol. 13, no. 5, pp. 898–910, 2011.

[38] P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy," *Neuroimage*, vol. 46, no. 3, pp. 726–738, 2009.

[39] J. A. Dowling, J. Fripp, S. Chandra, J. P. W. Pluim, J. Lambert, J. Parker, J. Denham, P. B. Greer, and O. Salvado, "Fast automatic multi-atlas segmentation of the prostate from 3D MR images," in *Prostate Cancer Imaging. Image Analysis and Image-Guided Interventions*, pp. 10–21, Springer, 2011.

[40] H. Wang, D. B. Stout, and A. F. Chatziioannou, "A method of 2D/3D registration of a statistical mouse atlas with a planar X-ray projection and an optical photo," *Medical image analysis*, vol. 17, no. 4, pp. 401–416, 2013.

[41] J. Ashburner and K. J. Friston, "Voxel-based morphometrythe methods," *Neuroimage*, vol. 11, no. 6, pp. 805–821, 2000.

[42] M. D. Wong, Y. Maezawa, J. P. Lerch, and R. M. Henkelman, "Automated pipeline for anatomical phenotyping of mouse embryos using micro-CT," *Development*, vol. 141, no. 12, pp. 2533–2541, 2014.

[43] M. D. Wong, A. E. Dorr, J. R. Walls, J. P. Lerch, and R. M. Henkelman, "A novel 3D mouse embryo atlas based on micro-CT," *Development*, vol. 139, no. 17, pp. 3248–3256, 2012.

[44] I. L. Dryden and K. V. Mardia, *Statistical shape analysis*, vol. 4. John Wiley & Sons New York, 1998.

[45] K. Pearson, "LIII. On lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901.

[46] A. D. Brett and C. J. Taylor, "A method of automated landmark generation for automated 3d pdm construction," *Image and Vision Computing*, vol. 18, no. 9, pp. 739–748, 2000.

[47] A. Frangi, D. Rueckert, J. Schnabel, and W. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 9, pp. 1151–1166, 2002.

[48] M. Van de Giessen, M. Foumani, G. Streekstra, S. Strackee, M. Maas, L. Van Vliet, K. Grimbergen, and F. Vos, "Statistical descriptions of scaphoid and lunate bone shapes," *Journal of Biomechanics*, vol. 43, no. 8, pp. 1463–1469, 2010.

[49] S. Osher and N. Paragios, *Geometric level set methods in imaging, vision, and graphics*. Springer, 2003.

[50] 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.

[51] E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*. New York, NY, USA: John Wiley & Sons, Inc., 1989.

[52] P. K. Saha, G. Liang, J. M. Elkins, A. Coimbra, L. Duong, D. S. Williams, and M. Sonka, "A new osteophyte segmentation algorithm using the partial shape model and its applications to rabbit femur anterior cruciate ligament transection via micro-ct imaging," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 8, pp. 2212–2227, 2011.

[53] M. J. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *The computer journal*, vol. 7, no. 2, pp. 155–162, 1964.

[54] R. Bowden, *Learning non-linear models of shape and motion*. PhD thesis, Citeseer, 2000.

[55] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, and N. Ayache, "Geometric variability of the scoliotic spine using statistics on articulated shape models," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 4, pp. 557–568, 2008.

[56] D. Kainmueller, H. Lamecker, S. Zachow, and H.-C. Hege, "An articulated statistical shape model for accurate hip joint segmentation," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp. 6345–6351, IEEE, 2009.

[57] L. M. Khachigian, "Collagen antibody-induced arthritis," *Nature protocols*, vol. 1, no. 5, pp. 2512–2516, 2006.

[58] A. Nanda, B. Karim, Z. Peng, G. Liu, W. Qiu, C. Gan, B. Vogelstein, B. S. Croix, K. W. Kinzler, and D. L. Huso, "Tumor endothelial marker 1 (Tem1) functions in the growth and progression of abdominal tumors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 9, pp. 3351–3356, 2006.

[59] P. Bourke, "Ply-polygon file format (online)," 2009.

[60] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, 2006.

[61] G. Miller, "Efficient algorithms for local and global accessibility shading," in *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*, pp. 319–326, ACM, 1994.

[62] M. Garland and P. S. Heckbert, "Surface simplification using quadric error metrics," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pp. 209–216, ACM Press/Addison-Wesley Publishing Co., 1997.

[63] J. Brown, A. Naylor, C. Buckley, A. Filer, and E. Claridge, "3D articulated registration of the mouse hind limb for bone morphometric analysis in rheumatoid arthritis," in *Biomedical Image Registration*, pp. 41–50, Springer, 2014.

[64] M. J. Cook, *The anatomy of the laboratory mouse.* London (& New York): Academic Press., 1965.

[65] W. Kabsch, "A solution for the best rotation of two sets of vectors," *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, vol. 32, no. 5, pp. 922–923, 1976.

[66] E. Haines and T. Akenine-Moller, *Real-time rendering.* AK Peters, Ltd., 2002.

[67] T. Cootes, E. Baldock, and J. Graham, "An introduction to active shape models," *Image Processing and Analysis*, pp. 223–248, 2000.

[68] A. J. Naylor, E. Azzam, S. Smith, A. Croft, C. Poyser, J. S. Duffield, D. L. Huso, S. Gay, C. Ospelt, M. S. Cooper, *et al.*, "The mesenchymal stem cell marker CD248 (endosialin) is a negative regulator of bone formation in mice," *Arthritis & Rheumatism*, vol. 64, no. 10, pp. 3334–3343, 2012.

[69] N. Umeda, I. Matsumoto, I. Ito, A. Kawasaki, Y. Tanaka, A. Inoue, H. Tsuboi, T. Suzuki, T. Hayashi, S. Ito, *et al.*, "Anti-citrullinated glucose-6-phosphate isomerase peptide antibodies in patients with rheumatoid arthritis are associated with HLA-DRB1 shared epitope alleles and disease activity," *Clinical & Experimental Immunology*, vol. 172, no. 1, pp. 44–53, 2013.

[70] H. J. Ditzel, "The K/BxN mouse: a model of human inflammatory arthritis," *Trends in molecular medicine*, vol. 10, no. 1, pp. 40–45, 2004.

[71] Y.-G. Cho, M.-L. Cho, S.-Y. Min, and H.-Y. Kim, "Type II collagen autoimmunity in a mouse model of human rheumatoid arthritis," *Autoimmunity reviews*, vol. 7, no. 1, pp. 65–70, 2007.

[72] A. K. Abbas, A. H. Lichtman, and S. Pillai, *Basic immunology: functions and disorders of the immune system*. Elsevier Health Sciences, 2012.

[73] S. Oestergaard, K. Rasmussen, N. Doyle, A. Varela, L. Chouinard, S. Smith, P. Qvist, and M. Karsdal, "Evaluation of cartilage and bone degradation in a murine collagen antibody-induced arthritis model," *Scandinavian journal of immunology*, vol. 67, no. 3, pp. 304–312, 2008.

[74] G. Ruiz-Heiland, A. Horn, P. Zerr, W. Hofstetter, W. Baum, M. Stock, J. H. Distler, F. Nimmerjahn, G. Schett, and J. Zwerina, "Blockade of the hedgehog pathway inhibits osteophyte formation in arthritis," *Annals of the rheumatic diseases*, vol. 71, no. 3, pp. 400–407, 2012.

[75] N. Parameswaran and S. Patial, "Tumor necrosis factor- signaling in macrophages," *Critical reviews in eukaryotic gene expression*, vol. 20, no. 2, pp. 87–103, 2010.

[76] E. Carballo, W. S. Lai, and P. J. Blackshear, "Feedback inhibition of macrophage tumor necrosis factor-α production by tristetraprolin," *Science*, vol. 281, no. 5379, pp. 1001–1005, 1998.

[77] T. Iwamoto, H. Okamoto, Y. Toyama, and S. Momohara, "Molecular aspects of rheumatoid arthritis: chemokines in the joints of patients," *FEBS Journal*, vol. 275, no. 18, pp. 4448–4455, 2008.

[78] H. L. Wright, R. J. Moots, R. C. Bucknall, and S. W. Edwards, "Neutrophil function in inflammation and inflammatory diseases," *Rheumatology*, vol. 49, no. 9, pp. 1618–1631, 2010.