UNIVERSITY OF SOUTHAMPTON

Large Scale, Multi Femur Computational Stress Analysis Using a Statistical Shape and Intensity Model

by

Rebecca Bryan

A thesis submitted in partial fulfilment of the

degree of Doctor of Philosophy

in the

Faculty of Engineering Science and Mathematics

School of Engineering Sciences

June 2010

UNIVERSITY OF SOUTHAMPTON

Abstract

SCHOOL OF ENGINEERING SCIENCES
Bioengineering Science Research Group

Doctor of Philosophy

Large Scale, Multi Femur Computational Stress Analysis Using a Statistical Shape and Intensity Model

By Rebecca Bryan

The vast majority of orthopaedic computational studies only use a single bone model and attempt to extrapolate results to the population as a whole. This overlooks the large interpatient variability in bone geometry and quality which is inherently present and could explain the differing outcomes seen with otherwise comparable scenarios. A major barrier to correcting this is the substantial challenge involved in sourcing and generating large numbers of bone models. In order to address this situation a statistical model of the femur was developed which incorporated both geometry and material properties. The model generated 3D finite element models of the whole femur using principal component analysis. An elastic surface matching registration scheme and a mesh morphing algorithm were developed and applied to a training set of 46 femurs. Reconstruction tests showed that accurate reproduction of both geometry and material characteristics could be achieved with the first 35 modes. Sampling the statistical model was able to produce unique, anatomically realistic femur models with a high quality mesh. This led to a potential tool for the generation of femur models incorporating material properties and geometry for large scale multi femur finite element studies.

To test the ability of the model to generate realistic, unique, finite element femurs it was used as a source of bone models to drive a study on femoral neck fracture risk. Comparison to previous computational, clinical and experimental work investigating fracture risk factors revealed that the generated models showed similar characteristics in fracture location and type. Some geometric and bone quality traits suggested to increase risk were also supported. This case study developed a fully automated methodology for running an analysis, from model generation to FE post processing, on multiple bones with widely varying geometries and material properties.

The potential of the technique was further tested by using it in a study incorporating a femoral resurfacing implant, requiring the solution of a range of problems to enable correctly sized femoral components to be positioned appropriately in any given femur geometry. A fully automated methodology was established to implant any generated femurs and run an FE analysis. Comparison of the resulting strain changes through the bone to a previous study where implantation was manual showed similar trends.

The work described in this thesis shows how statistical modelling has been used to capture the variation in shape and material present within a set of femurs and use this to generate a large number of unique, realistic models. The case studies have described how large scale, fully automated analyses are possible using these techniques. A method has also been demonstrated to show how orthopaedic implant testing can be run with these generated bone models, opening up the prospect of far more extensive computational analyses than have been feasible in the past.

Contents

List of Figures

List of Tables

DECLARATION OF AUTHORSHIP

I, Rebecca Bryan, declare that the thesis entitled 'Large Scale, Multi Femur Computational Stress Analysis Using a Statistical Shape and Intensity Model' and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:

  Bryan, R., Mohan, S., Hopkins, A., Galloway, F., Taylor, M. and Nair,P.B. Statistical modelling of the whole human femur incorporating geometric and material properties, Medical Engineering and Physics 32(1), 57-65

  Bryan, R., Nair, P.B., Taylor, M., 2009. Use of a statistical model of the whole femur in a large scale, multi-model study of femoral neck fracture. Journal of Biomechanics 42(13), 2171-2176

Signed:.....

Date:.....

Acknowledgements

I owe a great deal of thanks to my supervisors Professor Mark Taylor and Doctor Prasanth Nair for their seemingly endless patience and enthusiasm towards this project over the last three years. Despite coming across some significant hurdles they were both always keen to listen and advise, never failing to encourage me to achieve as much as was possible at every stage.

I am extremely grateful to my colleagues who have given up their time to contribute to my thesis. Without the inputs made by Surya Mohan, Andrew Hopkins and Francis Galloway I would not have been able to reach the difficult goals of the project. By tapping into their skills, outside my areas of training, I was able to progress far more quickly than if I had been left to work alone. I also acknowledge the funding I have received from the Technology Strategy Board UK and DePuy Ltd, which have ultimately allowed me to complete this work but have also given me the opportunities to travel to conferences around the world.

My final thanks go to my family and friends, you have kept me sane and not allowed me to give up at any point. Whether colleagues who followed my PhD based rantings or those who politely nodded and sympathised, you have all had to put up with far too many conversations about femurs in the last three years - I apologise.

Chapter 1

Introduction

1.1 Introduction

The vast majority of orthopaedic computational studies are performed using a single bone model and then attempt to extrapolate their results to the population as a whole. This overlooks the large interpatient variability in bone geometry and bone quality which is inherently present between people, and could explain the differing outcomes seen with otherwise comparable scenarios. For example, why when two people have a comparable fall, one may suffer a bone fracture and the other not or why a joint replacement can be markedly more successful in one patient than another. For both cases, bone structure and quality have been shown to be highly significant in determining success or failure [1, 2, 3], along with surgical skill in the later case [4]. Currently there are no commonly used methods for running an analysis which incorporates naturally occurring variations without large computational cost and significant manual input, despite it having been recognised that by omission quantitative accuracy is sacrificed [5].

Of the few papers which have looked at intersubject variability two have incorporated this in attempts to predict implant stability. Pancanti et al [6] used data recorded by instrumented femoral prostheses to apply individualised loading conditions to a cemented hip model. Simulations of various activities, such as walking, fast walking, stair climbing/descent and bending, were then carried out. The model investigated the micro-motions produced at the bone-implant interface, an early indicator of failure if too great, and showed that inter patient variability was a more significant factor than the type of activity. However the study still assumed an optimal surgical outcome, resulting in a perfectly fitted implant, and used a generic bone

model with no incorporation of differences in bone quality between subjects. Wong et al [3] deterministically varied bone density within a femur model to simulate different grades of bone quality, then analysed the resulting micro-motion and strain produced at the bone-implant interface under normal walking loads. The study showed that these factors are greatly influenced by the overall stiffness of the femur and concluded that in order for FE testing to be any more than comparative, a wide range of patient models must be used.

The need for multiple models in computational testing has also recently been supported by Radcliffe et al [7], whose work investigated the number of femur models required to represent the variation present within a set of patient data. This study concluded that the time consuming and laborious task of creating multiple models from sources such as Computer Tomography (CT) scans was a significant limitation, making multi subject based finite element (FE) studies rare. It has been estimated that an optimal time scale for producing and solving a single proximal femur FE model from the original CT scan is $\sim$8 hours [7]. Without robust and reliable automated model generation techniques the time consuming and laborious task of creating multiple models from sources such as Computer Tomography (CT) scans is a significant limitation preventing multi subject based Finite Element (FE) studies becoming commonplace. In addition there is often a limited availability of CT data as they are not routinely taken before or after joint replacement operations. When carried out they are localised to reduce the radiation exposure to the patient.

This work investigates the potential of developing a statistical model to use as a source of FE bone models as a possible solution to the problem of model generation and limited data availability. The basis of the technique lies in statistical deformation models which have been widely used and developed in computer vision to capture the variations possible within a class of shapes. Commonly methods are based on Principal Component Analysis (PCA). These include active shape models (ASM) [8], where the shape is represented by manually or semi-automatically placing landmark points along boundaries, and active appearance models (AAM) [9] which incorporate the texture of the whole image into the model. The resulting statistical model can then be used to identify the shape in a new image or to generate a unique instance of the shape. However, because of the field in which these and other similar techniques have been developed they are often only suited to two dimensional images. Extending them to three dimensional shapes creates a real challenge in establishing accurate correspondence between any landmarked points and in the case of AAM leads to a dramatic increase in complexity. Registration of each of the training images must therefore be carried out in order for the model to be built. For many applications rigid registration techniques have proved sufficient, such as the Iterative Closest Point

(ICP) algorithm [10], used by Vos et al [11] to create a statistical model of bones in the hand, which has the advantage of not requiring a predefined relationship between points on the objects being registered. This is not true for more complex biological applications such as comparing breast and brain MRI images [12, 13] where the shapes are highly deformable. In these cases non-rigid registration techniques have been required based on computationally demanding free-form deformation models. An alternative mesh matching approach was developed by Couteau et al [14] to automatically generate 3D, patient specific, Finite Element (FE) meshes of the proximal femur.

The majority of statistical models focus on shape. An exception is Querol et al. [15], where non-rigid registration in the image space was employed in conjunction with principal component analysis to construct a statistical shape and intensity model of the proximal femur based on only 11 subjects. In theory image space registration schemes which make use of intensity based similarity metrics can lead to a more realistic statistical shape and intensity model. However, this approach is computationally expensive for high resolution images (such as good quality CT scans) and the statistical model output is not directly usable in FE, each instance would require meshing after generation.

1.2 Objectives

The broad aims of this project are to produce a three dimensional statistical model of the human femur, incorporating both geometry and material property distribution, and go on to use this in large scale computational studies. The model will be created from real patient data and hopes to capture as wide a range of inter subject variability as available data will allow, thus it will be able to reproduce the variation seen in a patient population and allow the affect of natural variability to be incorporated into computational analysis.

The model generation stage will require the creation of a whole process to turn CT data sets into a statistical representation which can then be sampled to output realistic femurs. Once constructed it will be used to generate large numbers of models by sampling through the variability it captures. It is hoped that this will allow many times more femur models to be produced than was needed to create it, each of which will be unique. In order to make these models as useful as possible the model will ideally produce FEA ready meshes. This will require the mesh structure to be fine enough to accurately capture changes in geometry and material property, and for the

mesh quality to be sufficiently robust to not require remeshing before analysis can be done. It is critical that the model only produces femurs which have plausible shape and bone density distributions, and that it has the ability to generate models which represent the whole range of variables present in the real femurs used to train it.

To test how successful the statistical model managed to be, it was used to drive a study on femoral neck fracture. Hip fracture is a potentially devastating injury to suffer, particularly for the elderly, leading many people to have investigated the problem, often trying to identify risk factors related to femoral bone quality and geometry [1]. By performing a multi femur study using the statistical model as a source of bones it was possible to compare any risk factors this identifies to the wealth of clinical and computational data available. It was important that it was possible to set up a high level of automation in performing this study, if manual intervention was needed then the methodology would have become impractical to use when large numbers of femurs were tested.

The ultimate aim of this work is to be able to use a statistical model of the intact femur to test orthopaedic implants by developing an automated methodology for implanting components into any given model. The main additional challenge of this target, as compared to the previous study, is to be able to align and fit a correctly sized component accurately into each bone, as poor positioning will bias any findings. As with the simulations on the intact femurs it will be critical that the processes of model generation and FE simulation are fully automated. By replicating the loading and analyses of a previous study, where implantation was manually performed, it is hoped that the process developed in this work will be shown to produce realistic results and so open up the prospect of population based analysis to wider orthopaedic testing. Further to this, the techniques potential to analyse large numbers of models will enable meaningful statistical analysis to be carried out to gain a better understanding of the complex, multifactorial influences behind implant performance.

1.3 Publications

The work described in this thesis has led to the following publications:

- Chapter 3: Generation of a Statistical Model of the Whole Femur

  Bryan, R., Mohan, S., Hopkins, A., Galloway, F., Taylor, M. and Nair,P.B., 2010. Statistical modelling of the whole human femur incorporating geometric and material properties, Medical Engineering and Physics 32(1), 57-65

  Bryan, R., Mohan, S., Taylor, M., Nair, P.B., 2008. Generation of a Statistical Model of the Whole Femur Incorporating Shape and Material Property Distribution. 54th Orthopaedic Research Society, San Francisco.

- Chapter 4: Femoral Neck Fracture Risk Study

  Bryan, R., Nair, P.B., Taylor, M., 2009. Use of a statistical model of the whole femur in a large scale, multi-model study of femoral neck fracture. Journal of Biomechanics 42(13), 2171-2176

  Bryan, R., Nair, P.B., Taylor, M., 2009. Use of a Statistical Model of Geometry and Material Properties in a 1000 Femur Study of Femoral Neck Fracture Risk. 55th Orthopaedic Research Society, Las Vegas.

- Chapter 5: Automated Implantation of Femoral Resurfacing Implant

  Bryan, R., Nair, P.B., Taylor, M. An Automated, Large Scale Finite Element Study into the Influence of Femoral Head Resurfacing on Load Transfer in the Proximal Femur, submitted to Journal of Biomechanics

  Bryan, R., Nair, P.B., Taylor, M., 2010. Influence of Femoral Head Resurfacing on Load Transfer in the Proximal Femur - An Automated FE Study of 260 Femurs. 56th Orthopaedic Research Society, New Orleans.

  Bryan, R., Nair, P.B., Taylor, M., 2010. Influence of Head Diameter on Load Transfer in the Resurfaced Femoral Head - A Finite Element Study. 56th Orthopaedic Research Society, New Orleans

Chapter 2

Orthopaedic Biomechanics Literature Review

2.1 Anatomy of the Hip

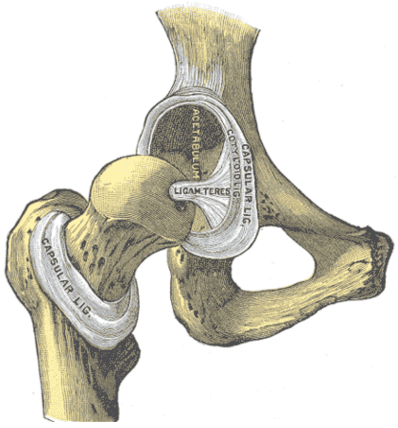


Figure 2.1: The Hip Joint [16]

The hip is a ball and socket joint, where the femoral head, the 'ball', rotates within the acetabulum, the 'socket' (Fig. 2.1). The joint is inherently stable due to this design and its reinforcement by the strong ligaments and muscles which surround it. The most significant ligament structures in maintaining joint stability are the ligamentum teres and the transverse acetabular ligaments which hold the femoral head in the socket.

The hip is a highly versatile joint, capable of a wide range of motions. The maximum range of passive motions, measured with the knee extended to 120°, are 120° flexion mirrored by 10-30° extension (Fig. 2.2b) and 45-50° abduction mirrored

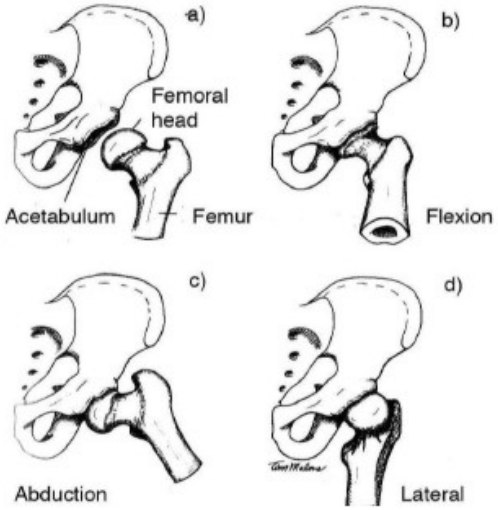


Figure 2.2: Hip motions. a) neutral, b) flexion, c) abduction and d) lateral rotation [17]

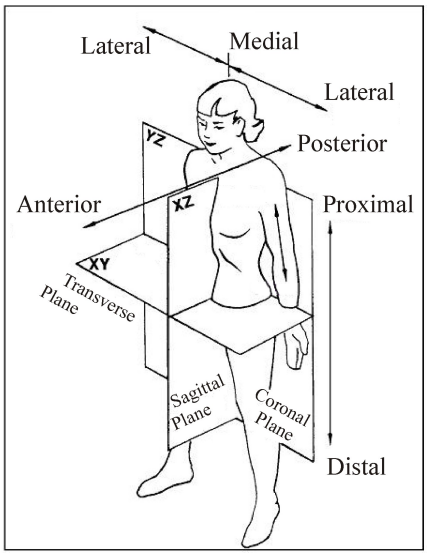


Figure 2.3: Illustration of anatomical planes of motion and terms indicating anatomical location [18]

by 20-30° adduction (Fig. 2.2c). Medial-lateral rotation (Fig. 2.2d) ranges from 42-50°, measured with the knee at 90° of flexion.

The muscles responsible for active control of hip motions are illustrated below (Fig. 2.4). There are a large number of muscles which act on the hip, each of whom are often primarily associated with one motion but may have a secondary role in another. In addition, no single muscle is responsible for any movement. Table 2.1 reflects this by grouping muscles by the movement they influence and indicating whether this is their primary or secondary function.

2.2 Anatomy of the Femur

The femur is the longest bone in the human body. The morphology of the femur has developed in response to the demands placed upon it by its function. Primarily the femur must transmit load from the acetabulum to the tibia but it also provides insertions points for the major muscles which control and stabilise the motions of the hip and knee joints. These requirements result in forces begin unevenly applied through the femur, which is reflected in both the internal microstructure and external geometry of the bone.

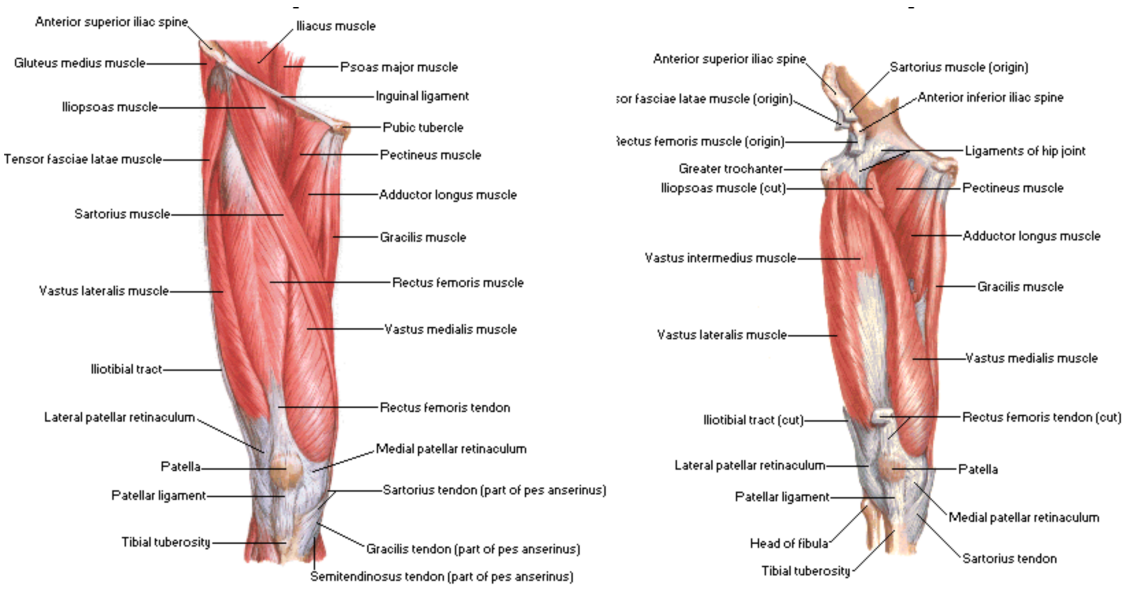


Figure 2.4: Anterior view of the superficial (left) and deep (right) muscles of the hip [19]

Function	Primary Muscles	Secondary Muscles
Flexion	Ilipsous Rectus Femoris Tensor Fascia Lata Sartoris	Pectineus Adductor Longis Adductor Magnus Gracilis
Extension	Gluteus Maximus Biceps Femoris Semitendinosus Semimembranosus	Gluteus Medius Adductor Magnus Piriformis
Adduction	Pectineus Adductor Brevis Adductor longus Adductor Magnus Gracilis	
Abduction	Gluteus Medius Gluteus Minimus	Gluteus Maximus Satorius Tensor Fascia
Lateral Rotation	Oburator Internus + Externus Gemellus Superior + Inferior Quadratus Femoris Piniformis	Gluteus Medius Gluteus Minimus Gluteus Maximus
Medial Rotation		Gluteus Medius Tensor Fascia Lata

Table 2.1: Table of primary and secondary muscles responsible for different hip motions [20]

2.2.1 Form of the Femur

The femur can be described in three sections; the proximal femur, the femoral shaft and the distal femur. The femur is bowed in the sagittal plane and, when standing,

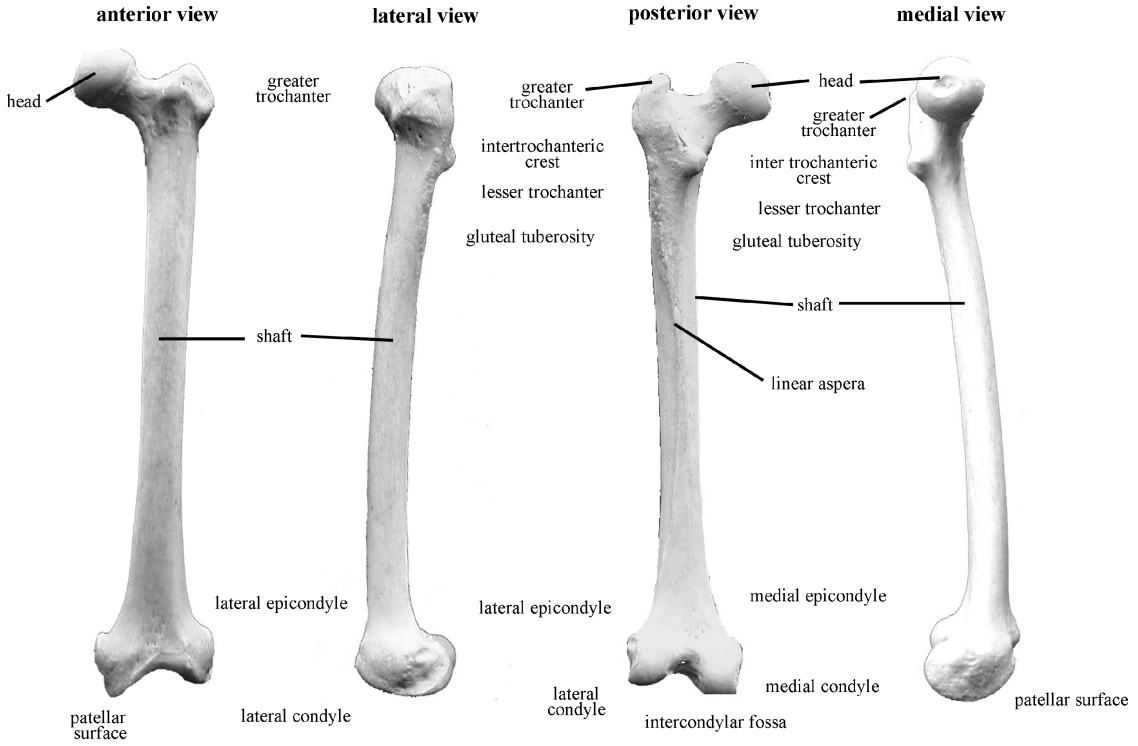


Figure 2.5: Photographs of the human femur in different orientations (Sawchuck and Padiack 2003)

inclined laterally in the coronal plane (Fig. 2.5). This results in the shaft of the femur being inclined by approximately 5-7° from the vertical. The femoral neck axis is rotated in relation to both the coronal plane, indicated by the anteversion angle, and the shaft axis, indicated by the neck shaft angle (Fig. 2.6).

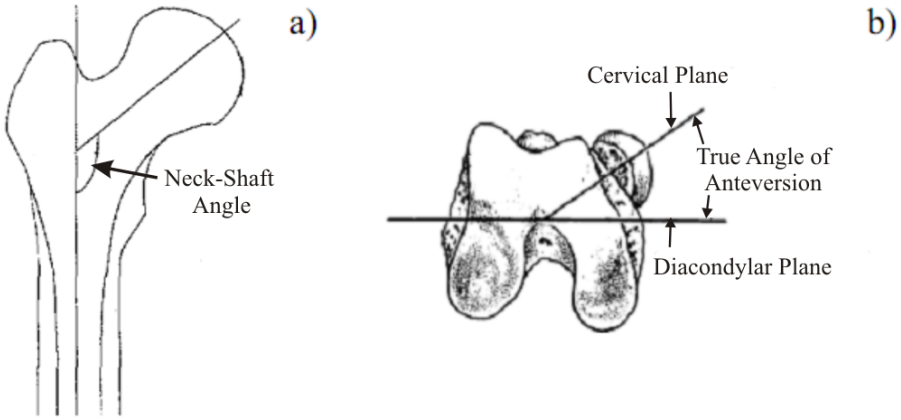


Figure 2.6: Diagrams of the femur illustrating (a) neck shaft angle [21] and (b) anteversion angle [22]

The prominent features of the proximal femur are the femoral head, femoral neck, greater and lesser trochanters (Fig. 2.5). The head is two thirds of a sphere with an indent, the fovea capitis femoris, just below and behind its center where the ligamentum teres attaches. It is covered with smooth articular cartilage which

provides the femoral side of the bearing surface of the hip joint. The neck joins the head to the body of the femur, merging with the lesser trochanter at its inferior limit and with the base of the greater trochanter at its lateral limit. The trochanters are growths which support the muscles controlling rotation of the thigh, they are irregularly shaped with rough surfaces where muscles attach to the bone and vary greatly in form from person to person. The greater trochanter is located laterally and backward of the head and, in an adult, approximately 1cm lower at its superior point. The lesser trochanter is located at the lower and posterior part of the base of the neck [16].

The shaft of the femur is close to cylindrical, marginally broader at the top than the center and with a subtly arched external profile on the posterior side. The deviation from cylindrical is caused by a ridge called the linea aspera, which runs longitudinally down the middle third of the posterior surface of the femur [16].

At the distal end the femur diverges into the medial and lateral condyles which form the superior portion of the knee joint. Anteriorly the condyles protrude slightly and are separated by a shallow depression called the patella surface, which is the articulating surface between the femur and patella. Posteriorly the condyles protrude significantly and are separated by the intercondyloid fossa which is a deep, rough surface notch providing attachment points for the cruciate ligaments. The lower and posterior surfaces of the condyles articulate as part of the knee joint and thus are covered with articular cartilage [16].

2.2.2 Structure of Bone

The femur is constructed primarily of a tube of cortical bone with the epiphyses being a shell of cortical bone filled with cancellous bone. The macrostructure of these two materials is very different. Cortical bone is transversely isotropic in nature, with greater strength in compression than tension (Table 2.2). It is a compact structure forming a hard casing which protects the interior and provides rigidity and strength against applied stresses. The exterior surface is made up of highly organised rings of lamellar bone.

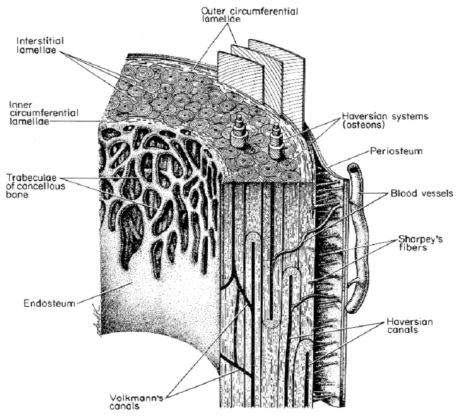


Figure 2.7: Section through cortical bone [23]

The bulk of it is made of compact osteons with the spaces between them filled by interstitial lamellae. Osteons contain nerve fibres and blood supply for the surrounding tissue, they are aligned along lines of high stress in the bone (Fig. 2.7).

Direction	Young's Modulus (GPa)
Longitudinal	17.4
Transverse	9.6
Bending	14.8
Shear	3.51

Table 2.2: Table of human cortical bone elastic modulus [24]. Illustrating the transversely isotropic nature of cortical bone.

Cancellous bone is an irregular lattice construction of trabeculae which is nourished by the red bone marrow which fills the cavities within it, making it far less dense than cortical bone. It is generally found in the epiphyseal regions of long bones, where load is transferred across a joint from one bone to another. Mechanical testing of cancellous bone has produced a wide range of results, illustrating that the properties of the bone are most dependant on its density, which varies with location in the body. The stiffest trabecular structures being in the highly loaded, densest packed sections of cancellous bone which are found in the femoral head and neck [25] (Table 2.3).

Site	Loading Direction	Young's Modulus (MPa)	Yield Strain	Yield Stress (MPa)
Vertebra	Compression	344 ± 148	0.0077 ± 0.06	2.02 ± 0.92
	Tension	349 ± 133	0.0070 ± 0.05	1.72 ± 0.64
Prox. Tibia	Compression	1091 ± 634	0.0073 ± 0.06	5.83 ± 3.42
	Tension	1068 ± 840	0.0065 ± 0.05	4.50 ± 3.14
Greater Trochanter	Compression	622 ± 302	0.0070 ± 0.05	3.21 ± 1.83
	Tension	597 ± 330	0.0061 ± 0.05	2.44 ± 1.26
Femoral Neck	Compression	3220 ± 936	0.0085 ± 0.10	17.45 ± 6.15
	Tension	2700 ± 772	0.0061 ± 0.03	10.93 ± 3.08

Table 2.3: Table of cancellous bone modulus, measured in different locations in the body [25].

2.2.3 Structure of the Femur

The internal structure of the femur is efficiently designed to transfer load through the bone with the minimum of material. The orientation, position and size of the

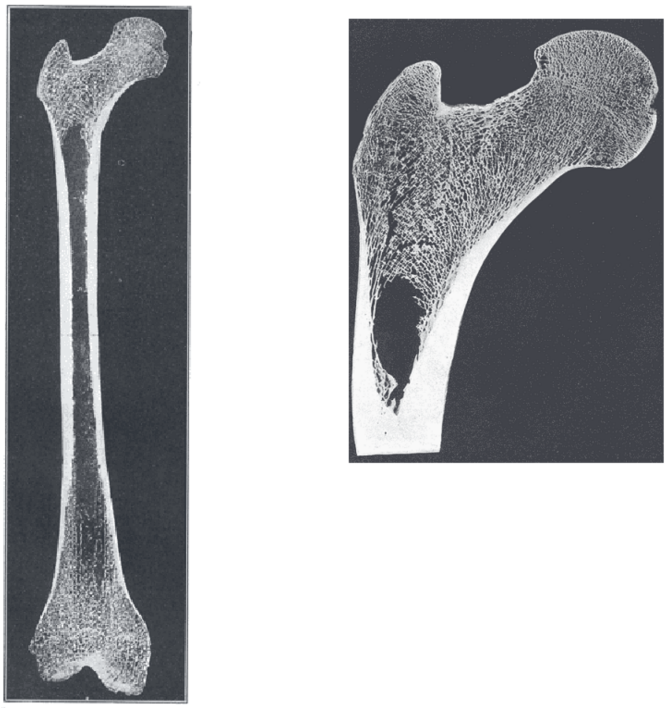


Figure 2.8: Frontal longitudinal midsection of the femur [21]

trabeculae relate to the paths and magnitude of principal stresses caused by load transfer through the hip and knee joints.

From the level of the head to the lesser trochanter, the femoral structure is predominantly a thin layer of cortical or compact bone, encasing cancellous or spongy bone. The cancellous structure consists of two distinct groups of trabeculae which intersect perpendicularly. The medial-lateral system fans outwards and upwards from the medial side and resists compressive stresses induced by loads on the femoral head. The lateral-medial system runs inwards and upwards from the lateral aspect and resists tensile stresses. The absolute tensile stress seen in the proximal femur is less than the absolute compressive stress therefore the trabeculae of the lateral-medial system are thinner than their opposing group (Fig. 2.8).

The architecture of the shaft or diaphysis of the femur exhibits maximum resistance to bending, shear, torsional and axial stresses. The section consists of a roughly cylindrical, hollow tube of thick cortical bone. The central area is the intermedullary canal, filled with bone marrow which serves no significant structural function but has an important metabolic role in the body.

The distal femur is structured similarly to the proximal end with a core of cancellous bone covered in a thin layer of cortical bone, except for the intercondyloid fossa where the cortical bone is notably thicker. The cancellous bone in this section is constructed of two main groups. The longitudinal system runs parallel to

the shaft and curves slightly through the condyles to meet the articulating surfaces perpendicularly. The transverse system braces the longitudinal system.

2.3 Hip Contact and Muscle Forces

The contact forces seen at the hip joint vary with the type of activity which the person is doing. It is logical that running will put higher loads through the joints than walking, however the pattern of loading will also change depending on the type of movement being made. Typically most analyses use the gait cycle as a standard loading condition, whether the dynamic cycle is being considered or simply the peak force seen which can be used in static FE simulations. The normal gait cycle consists of two main parts; the stance phase, where the leg is question is in contact with the ground, and the swing phase (Fig. 2.9). The peak forces seen at the hip occur just after 'heal strike' and are followed by a secondary peak coinciding with 'toe off'. The magnitude and direction of these forces has been investigated by both indirect and direct techniques.

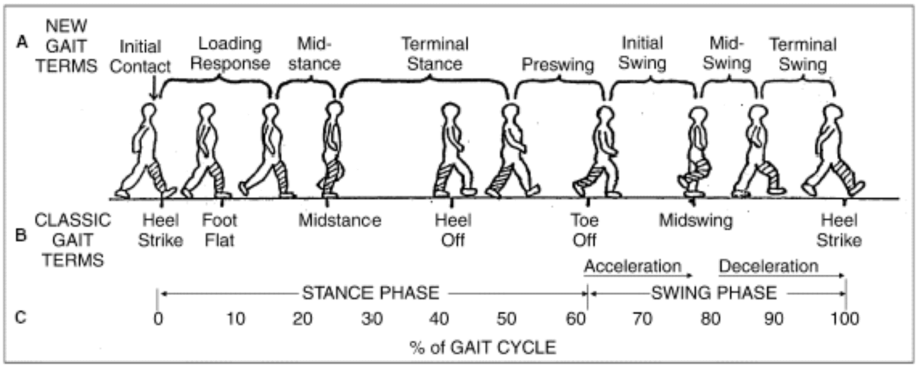


Figure 2.9: Diagram of the gait cycle [26]

Indirect prediction of joint, and muscle forces, have been carried out using a combination of mathematics and experimental data [27]. Data is captured by recording the motion of a subject walking; the limb displacement recorded and the reaction force on the ground measured using a force plate. This information is used to feed mathematical models which calculate lower limb kinetics from these given kinematics, providing information about the ankle and knee as well as the hip. These models are necessarily simplified representations as the number of unknown terms, multiple and complex muscle contributions as well as joint forces, vastly exceeds the number of known inputs. Some studies have indicated which muscles or groups of muscles are active at various stages of gait, however the direction and magnitude of their

individual contributions has not been established in vivo. The models use optimisation techniques to predict muscle forces, with results varying from model to model depending on initial conditions and muscle groupings [27, 28, 29].

Direct measurement of hip contact force magnitude and direction has been carried out using specialised femoral implants, instrumented with stain gauges, and inverse dynamic calculations. The forces going through these prostheses have been recorded while the patients performed different activities. The loads were presented in terms of percentage body weight, as the patient's weight clearly has a direct relationship to the magnitude of the forces going through the joint. Studies by Bergmann et al [30, 31] recorded the peak forces associated with walking, jogging and climbing. A patient was accidentally recorded stumbling, with surprisingly high forces being generated (Table 2.4).

Unlike the resultant contact force on the femoral head, muscle forces can not be directly measured using the techniques previously described. Instead they must be estimated using a mathematical model of the joint. The inverse dynamic approaches do this by assuming that at any point the system is in equilibrium and solving it through optimisation [27, 32, 28]. This has led to an understanding of muscle forces accross the gait cycle [28]. Heller et al. [33] generated patient specific models using CT scans and combined this with data from an instrumented prosthesis. Tasks were performed by the patients and the kinematic and kinetic results recorded and then applied to the model.

Once again it is worth noting that the use of data from an instrumented prosthesis introduces possible abnormalities as the subject's natural anatomy has been altered and soft tissues potentially weakened by being cut during surgery. All musculoskeletal models are necessarily simplified. This includes grouping sets of muscles together into one force vector, splitting large muscles with large attachment sites into several lines of action which together represent its effect and modelling lines of action as being straight, simplifying the true three dimensional volumetric structure with associate curved pathway.

These results give an indication of the magnitude of the forces seen regularly by the hip and such findings are the basis of the loadings applied to finite element models. It must be remembered that this data comes from patients with hip prostheses, not natural joints, with the added complication of the unknown effect that such traumatic and invasive surgery may have had on the surrounding soft tissue. However the

Activity	Peak Joint Contact Force (% BW)	Torsional Moment (%BWm)
1 Legged Stance	250-359	-
Walking	211-285	1.2-1.9
Jogging	550	5.3
Stair Climbing	227-314	1.8-3.0
Stumbling	870	5.4

Table 2.4: Hip joint reaction force and torsional moment for different activities, showing the range in peak forces between subject [31], jogging(1 subject) [30], single limb stance [34, 35].

findings are significant and as close as it is possible to get to directly measuring internal forces. They have shown that the peak force seen in-vivo may well be much higher than that predicted by normal gait, for example jogging doubles the magnitude. This may not be a concern for elderly patients but it is a consideration for younger, more active patients. The unpredictable event of stumbling showed a force around four times greater than walking, it is therefore the type of event which is most likely to result in a failure and could be a concern to any patient. The study also showed a large intersubject variability for each activity, and even intrasubject variability between left and right hips.

2.4 Finite Element Modelling

Finite Element Modelling was reportedly first used in orthopaedics in 1972, although had been commonly used across the wider engineering community before being adopted by biomedical engineers [36]. Since then rapid improvements in computational technology and processing power have allowed FE to become commonplace, capable of solving complex three dimensional, multi-body problems. FE in orthopaedics has been applied to problems of varying complexity, giving information on the stress and strain resulting from load cases for models representing individual bones up to the whole musculoskeletal system, with dynamic and adaptive FE also used to assess the affect of kinematics and bone adaption [5].

Extensive use of FE is evident in implant design and testing. Preclinical testing of orthopaedic implants aims to be as comprehensive as possible, with every effort made to avoid joint replacement failure. Failure being a highly traumatic experience resulting in revision surgery, a major operation, which is unpleasant for the patient and expensive for the health authority. FE allows designers to investigate ways of improving implants in three main areas; improving joint function, improving

implant fixation and improving wear characteristics. The latter two directly relating to avoiding and delaying revision surgery where possible.

FE in orthopaedics has commonly been carried out on single bone or small sets of bone models, taking a single available data set or a generic/average geometry and performing computational analyses. The lack of incorporation of interpatient variability is dangerous, as natural differences in bone geometry and quality has been shown to produce different stress/strain results for identical computational restraints and loads [37]. This has led to recent work in developing patient specific finite element models, where individuals exact bone geometry and bone material properties are reproduced as accurately as possible [38, 39, 40, 41]. Such models are usually generated from CT scans through manual segmentation, which is time consuming and susceptible to human error in identifying structures. Material property information can also be extracted from CT scans, see Section 3.3.3. Semi-automated and automated segmentation techniques have begun to be developed, but are not currently widely available as tools to apply to the problem and are often based on statistical models, which require their training data to be segmented in the first place [42, 43, 44].

Orthopaedic implant testing, of any form, does not tend to be performed on a patient by patient basis. Experimental and computational analyses are run to assess designs using available subject data, be that a number of cadaveric or synthetic bones or CT/MRI data sets. In either case the subjects used for testing are rarely chosen to represent the entire range of people who may later have the implant, yet that is exactly what the results they yield are expected to do [45, 46, 47, 48, 49, 50].

Attempts have been made to try to address this problem by incorporating intersubject variability in analysis, however often only taking one factor into account at a time. Good examples of this are studies which have used a generic model of bone geometry and superimposed other parameters onto it to investigate implant stability. Pancanti et al. [6] applied subject specific joint loads to such a model, reflecting four patients performing simple motor tasks (walking, sit to stand, climbing/descending stairs and so forth) recorded using instrumented prostheses. Wong et al. [3] created a single femur model from an available CT data set and systematically varied the material properties of the bone in proportion to the apparent bone density interpreted from the images, reducing modulus from its true healthy value down to 60% of its peak to represent the changes expected with age or disease. Both studies saw intersubject variability having a dramatic effect on the predicted results and concluded that it was important to incorporate this if any more than a comparative result was to be gained.

Truly patient specific, multi model studies are rare. There are many reasons for this but the most notable are; the scarce availability of high quality image data (suitable to produce accurate models with material properties) and the time investment required to generate models. Model generation not only includes the segmentation stages previously mentioned, but also meshing and application of model specific loading/boundary conditions, as well as actually running an analysis and processing the results. The time estimated by Radcliffe to complete this for each model was around 8 hours [7], even once the methodology was perfected and ran smoothly. Keyak et al. [51] experimentally and computationally analysed fracture load in a set of eighteen cadaveric femurs, Lengsfeld et al. [52] computationally assessed femoral strain change in eleven femurs twelve year post hip replacement surgery and Radcliffe and Taylor [7, 37] modelled femoral resurfacing in sixteen subjects, predicting the strain changes produced in the proximal femur. All of these studies generated individual FE models from subject CT data sets and applied subject specific loads (proportional to body weight). Their results showed that very different results were produced with different models, again emphasising the need for more comprehensive testing methods if population wide assessment is to be achieved. The limitations discussed have restricted those who have attempted this type of analysis to quite small data sets, not yielding sufficient results to perform meaningful statistical analysis.

More recent approaches have tried to avoid the 'brute force' method of manually generating a large number of models by incorporating statistical and probabilistic techniques. The majority still only use a very small initial set of models but their strength is in intelligently varying parameters known to affect the analysis, generating vast numbers of runs to examine their interactions in multi factorial problems. In biomechanics these techniques have been used to identify the parameters to which a scenario is most sensitive and examine the extent of their potential effect [53]. Studies have incorporated uncertainty in material properties [54] and geometry [55] through to more complete systems such as a hip implant [56, 57]. Viceconti et al. [58] simulated more than 1000 scenarios to investigate primary total cementless hip implant stability using a parametric analysis, four key features were identified as key to the problem (bone material properties, presence/extent of gaps at bone-implant interface, patient body weight and scaled femur size), these were varied between defined limits while assuming perfect implant alignment. Dopico-González et al. [57] also modelled a cementless hip prosthesis examining the influence of surgical positioning variation on strain within the bone. Both of these examples show the potential of examining combinations of variables together in a large, Monte Carlo simulation making it possible to observe interactions between them. However both made major simplifications to the system to allow this and preselected the variables which were

to be examined. Bone geometry variation was either ignored by using a single bone model [57] or dramatically reduced by using a parametrically scaled femur model [58]. Similarly, despite bone modulus being a variable factor in both studies, its description was made homogeneous across the entire bone or homogeneous for cortical and cancellous regions respectively.

Despite the extensive use and development of computational analysis, as well as experimental assessment, to test orthopaedic implants over the last few decades it has still not proved possible to eliminate failure. Current estimates show that the 10 year survivorship rates for hip replacements, the most successful of joint replacement designs, lies at 90-95% [59, 60, 61]. With ∼40,000 patients undergoing such an operation in the UK each year [62], and rising, this percentage represents a significant number of people and a large financial burden to the health service. The figure may even be suggesting better results than exist as it reflects only those implants which have required revision, it does not include those patients who are unhappy due to pain or poor mobility. It is therefore hypothesised that current assessment techniques may be missing these 5-10% of failures due to the lack of incorporation of interpatient and surgical variability, both of which have been shown to have significant influence on implant performance. The work discussed has demonstrated that techniques exist for incorporating surgical variation and subject specific modelling, however they have severe limitations and extending the techniques to population based approaches has not been established.

Chapter 3

Generation of a Statistical Model of the Whole Femur [1]

3.1 Statistical Modelling (SM)

3.1.1 Development of Statistical Models

Deformable shape models were first used in the fields of computer vision and computer graphics in the mid 1980's, using splines to describe malleable curves, surfaces and solid shapes [63]. Deformable models maintain the characteristics of the shape, or class of object, they represent while also being able to generate or deform to fit any other example of that shape which is legal or realistic [9]. This is achieved as the model is constructed from a training set of images, thus incorporating a priori knowledge of the form, while also being able to extract information from a new image being examined.

The potential of these models has since been seen by researchers in medical imaging to solve the challenges which have arisen from the rapid developments in medical imaging technology. As high-level imaging modalities such at CT, Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) scans have become widely available, these non-invasive techniques are increasingly being used for a variety of tasks from preoperative planning and intraoperative navigation to disease

[1] Bryan, Mohan, Hopkins, Galloway, Taylor and Nair, 2010. Statistical modelling of the whole human femur incorporating geometric and material properties. Medical Engineering and Physics 32(1), 57-65 - based on the work in this chapter

tracking and radiotherapy planning [64]. Effective clinical use of these images requires computational analysis in order to accurately extract and quantify the useful data. Models have achieved this by developing automated and semiautomated processes for segmentation of anatomical structures from images [65], patient specific model representations [66], tracking anatomical movements such as cardiac function [67] and matching intrapatient images to monitor disease progression [13].

A deformable model aims to conform to the following objectives; generality, specificity and compactness. *General* meaning the model is able to represent any instance within the class of shape and *specific* meaning that only legal or realistic instances can be produced by the model. The system aims to achieve this with as few parameters as possible, so being *compact* [68].

3.1.2 Existing Statistical Modelling Techniques

These techniques have mainly evolved within the two dimensional world of computer vision, and so the longest standing medical applications involve extracting shapes from 2D images. The most widely used examples are deformable contour models or 'snakes' [69]. These are malleable curves defined by control points which are attracted to features, such as lines and edges, within an image. Thus, the position and shape of an object can be identified by iteratively aligning the curves with the boundary (Fig. 3.1).

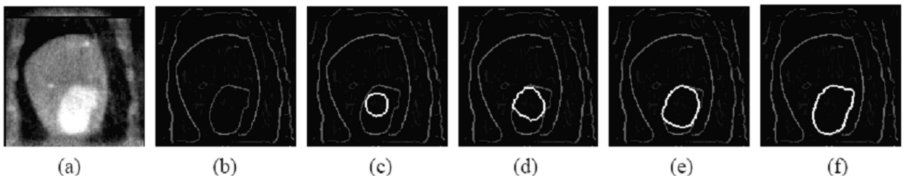


Figure 3.1: Illustration of iterative fitting of a snake contour onto an image of a heart ventricle. (a) intensity CT image, (b) Edges detected, (c-f) iterations of contour from initial shape guess [64].

A more advanced method, although still two dimensional, was developed for the creation of a statistical model of shape by Cootes et al [8], Active Shape Modelling (ASM) which is based around a training set of images. The idea is to capture the variation possible in an object using a range of examples collected in a training set (Fig. 3.2a). Each example used to train the model is described as a set of points, called landmarks, which are manually placed at the same relative locations, allowing direct comparison between them. For example, if landmarking a hand the first point

may always be placed at the tip of the first finger, the second at the tip of the second finger and so on (Fig. 3.2b). Once aligned to the same set of axes, the variations between each example at each point are used to create a Point Distribution Model (PDM). This model is then subjected to Principal Component Analysis (PCA) in order to reduce the large volume of information into its 'principal' parts, identifying the main modes of variation seen in the training data (Fig. 3.2c). Once the model has been created it can be used in two ways; firstly to identify new instances of the shape in an image and secondly to generate unique but legal instances of the shape. It is this later application which is of most interest to this project.

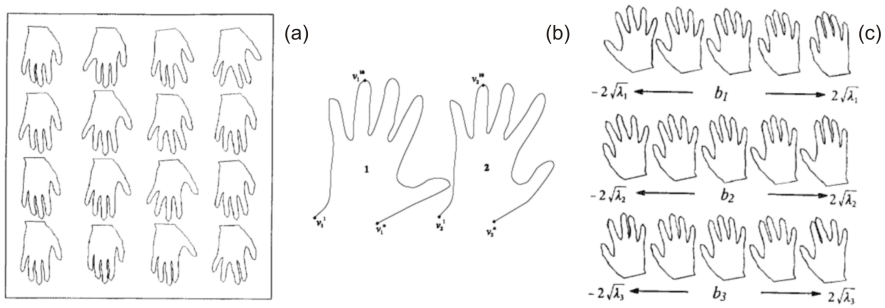


Figure 3.2: Example of ASM of a hand. (a) training set of example shapes, (b) landmarking of examples and (c) illustration of the effect of the first three principal modes of variation [8].

Cootes and Taylor [9] went on to further develop the ASM to create the Active Appearance Model (AAM). AAMs are constructed in the same way as ASMs but represent the whole image rather than just the shape boundary by incorporating texture into the model. They require the whole image to be sampled rather than just landmarked points. This extra information makes the AAM more computationally demanding but it has been shown that they are able to converge to match model-to-image with fewer landmark points and are more robust than ASMs. As landmarks are manually placed, the reduction in number of these required is an advantage in terms of time and reducing potential error in misplacement on complex shapes. In theory it is possible to extend these methods to three dimensional models, although in practice it is challenging. ASMs are relatively straightforward to compute but providing a suitable set of consistently placed, densely packed landmarks can be hard to achieve. Conventional approaches to AAM construction are impractical in three dimensions as the amount of data involved is just too large. Developments are being made to solve this problem, Babalola et al. [70] recently proposed a method to achieve 3D segmentation of the brain using AAM based on further work by the Cootes and Taylor group. Levels of AAMs are used, first identify large, global features before AAMs of small individual structures are used to segment areas of the brain.

Ballester et al. [71] proposed a statistical modelling technique which incorporated positional and intensity information across a three dimensional shape called a Generalised Image Model (GIM). Similarly to the work of Cootes et al., GIM requires the construction of a training set of examples, incorporating the variation seen between them in terms of position and intensity. As these factors are modelled together their relationship can also be described. This study created a registration technique for aligning the model to an image using an Iterative Closest Point (ICP) algorithm. This removed the requirement for explicit correspondence between points in the model and image. The method aims to be used for segmentation and detection of abnormalities, not for generating viable instances of the model.

The ICP method was used by Vos et al. [11] to map sets of point clouds, representing the carpal bones of the hand to each other, without landmarks. With correspondence established, a statistical model describing the carpal bones' combined shape as a surface mesh was created using PCA. The model was able to illustrate the ways in which these bones vary between the left and right hand and between the subjects used to build the model. PCA analysis was also used by Yang et al. [72] to investigate the variations in shape characteristics of long bones in primates, and was able to use the results to discriminate directly between great apes and monkeys. Again, a triangulated surface mesh was generated for each example bone but the registration technique was a little more complex to allow for the increased variability in shape (although any size differences were removed by scaling). ICP was used to allow the initial rigid alignment of bones to be automated, but free-form deformation using B-splines was used to establish shape correspondence [13], deforming a reference mesh onto each other example.

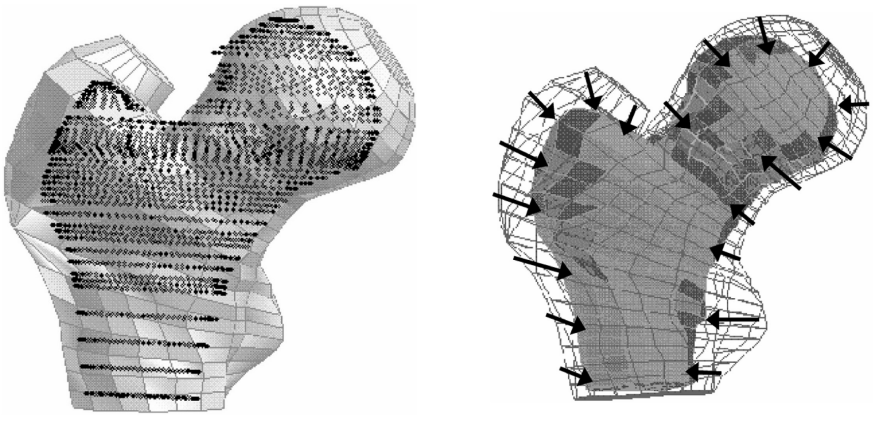


Figure 3.3: Illustration of the method defined by Couteau et al. [14]. Showing superimposition of a grey 3D reference mesh with a new target femur defined by a cloud of points (left), and the final mesh generated for the target femur from the reference (right).

A further methodology to allow a baseline or reference mesh to be morphed to fit another instance of that shape is the elastic based registration scheme. As with many ideas in this field, the principles originated in computer vision and graphics [63]. Couteau et al. [14] applied and developed these ideas into a 'mesh-matching' algorithm which deformed a reference proximal femur surface mesh onto a new femur shape, described by a cloud of points (Fig. 3.3). The method used a relatively course mesh but was able to retain mesh quality and to match the shapes of a large range of femur geometries. The work's aim was to allow 3D surface meshes to be generated automatically from segmented volumetric structures. This was also the aim of Kaus et al. [73], who again developed a unique elastic based registration scheme which initially aligned the baseline mesh to a new instance described as a triangulated mesh, before applying localised adaptations to the mesh to match the shapes precisely. A statistical model was then built using PCA, allowing the robustness of the model to be interrogated with leave-one-out testing. The model was used to achieve automated segmentation, by identifying the model shape within an image (Fig. 3.4). Statistical Shape Models (SSM) were successfully built for both the proximal femur and vertebra, again at quite a coarse mesh density.

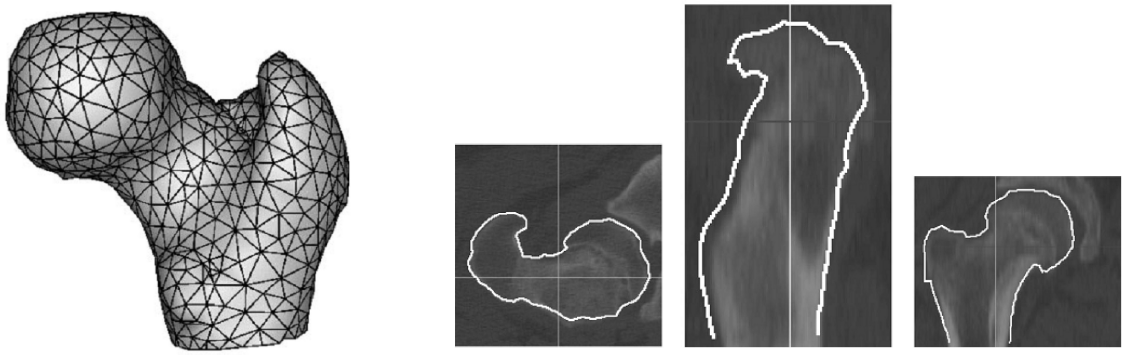


Figure 3.4: Illustrations of the work of Kaus et al. [73], showing the proximal femur surface model (left) and the identification of an 'unseen' femur instance from a CT image using the model (right).

Rajamani et al. [74, 75] also developed a statistical model of the proximal femur, however their intended application was quite different to the previously discussed papers. The statistical model was developed to generate patient specific models from sparse data, with the idea that a few key landmark points could be quickly taken from sources such as ultrasound, and the model used to generate a femur which fits these measurements (Fig. 3.5). The application was specifically aimed at inter and intra operative visualisation, where 3D scanning techniques would be impractical or undesirable to apply.

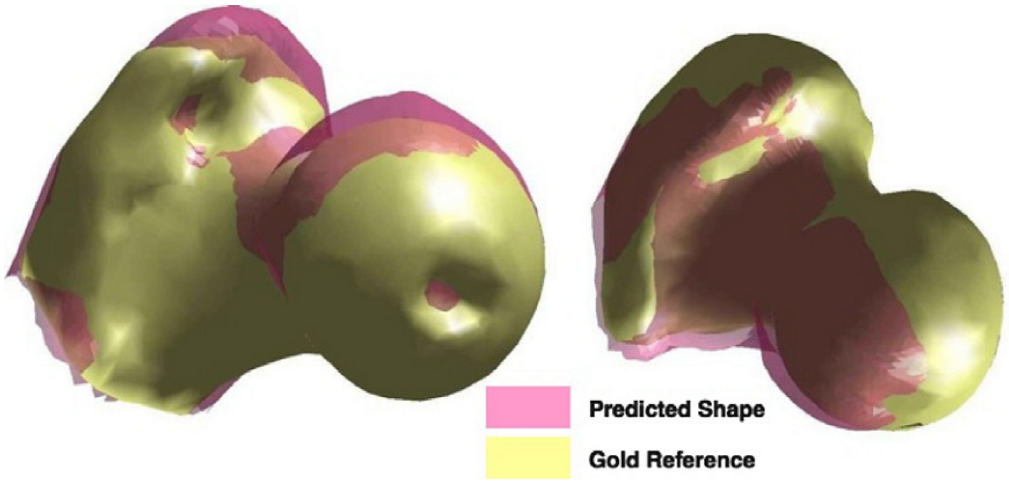


Figure 3.5: Illustration of the work of Rajamani et al. [75], showing two attempts to match their proximal femur model to real femurs described by points gathered from ultrasound data. Where the predicted shape (red) were compared to registered surface models, segmented from CT scans of the femur (gold).

All the three dimensional studies discussed thus far have involved surface shape models. A key aim of this work is to create a model which incorporates material properties as well as geometric information, therefore not only will this information need to be incorporated but a solid model will be required. A statistical tetrahedral model of the proximal femur has been developed by Querol et al. [15] which incorporates elemental material modulus values. The aim here was to develop a model which could be used in finite element analysis studies and so a very simple static FE analysis was run using the model to investigate the different stress distributions produced in femurs generated to be at the extremes of the variation seen in the model. These were created by taking just the first eigenmode produced by PCA and exaggerating the influence of this by +/-2 standard deviations from the mean geometry (Fig. 3.6). The statistical model itself was relatively simple, built from only 11 femur CT scans and containing approximately 3500 elements. A reason for the coarse mesh may have been the FFD registration scheme applied, which is computationally demanding and may have restricted the model resolution. This study provides a proof of

concept, showing that generating a statistical model of shape and material property is achievable and this model can then be used for FEA.

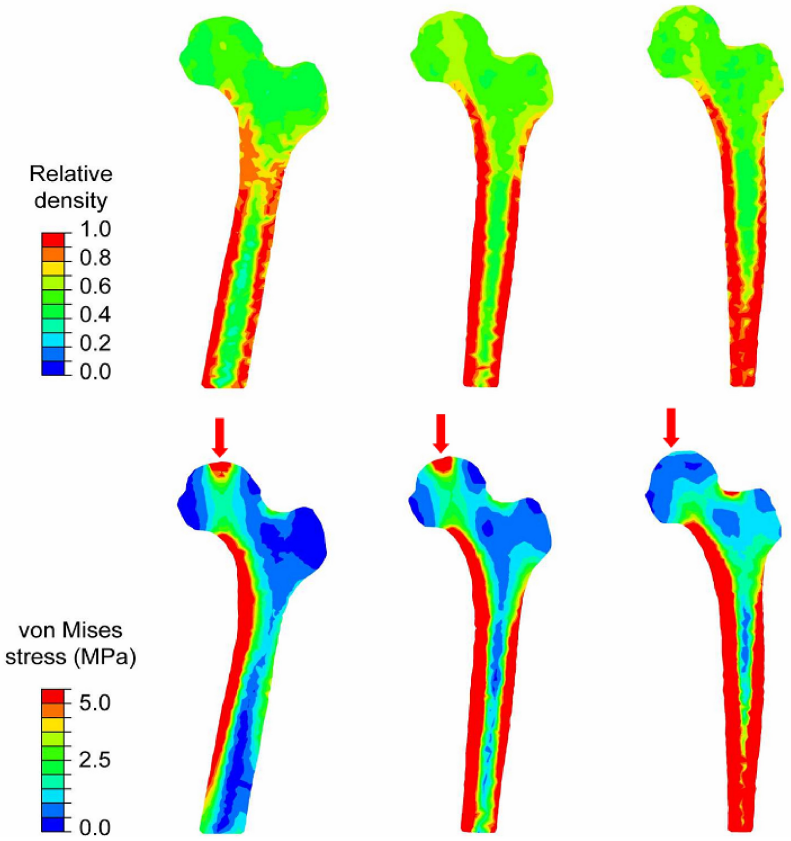
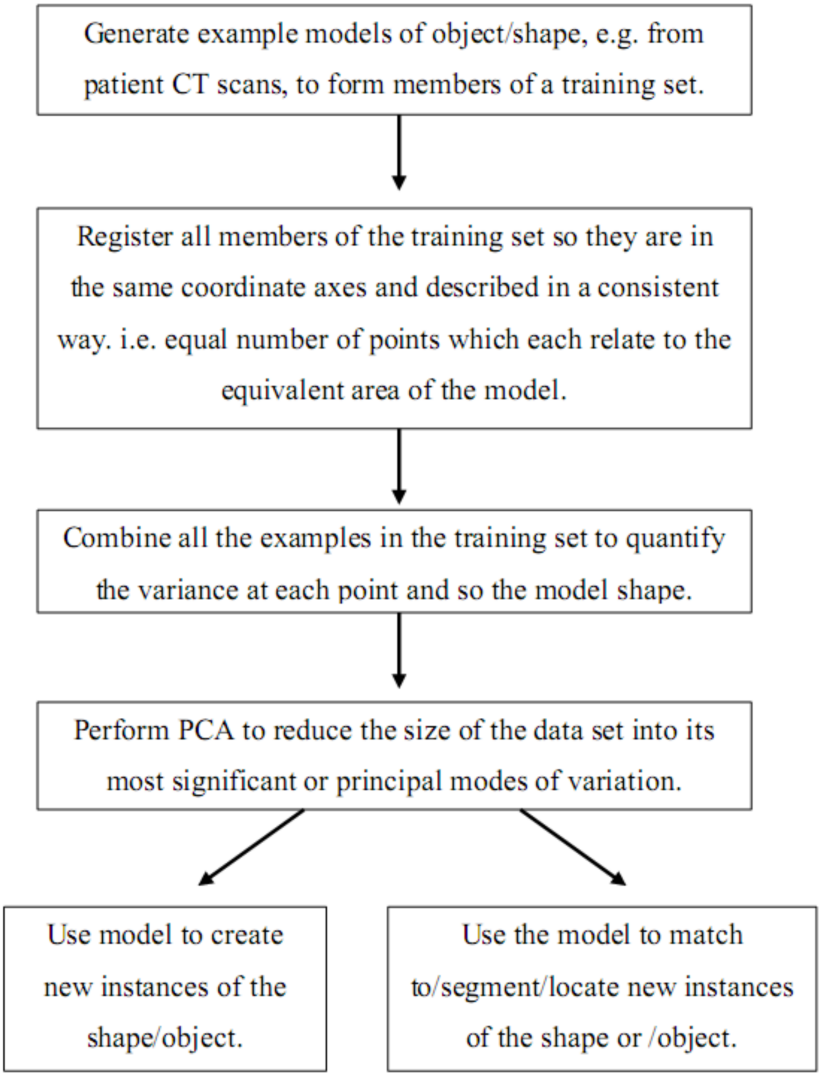


Figure 3.6: Results of the FE simulation run by Querol et al. [15] on a femur generated from their statistical model by varying, from left to right -2 s.d. to +2 s.d. from the average. The figure shows the relative bone density (top) and the von misses stress (MPa) produced by a 1600N vertical load (bottom).

3.1.3 Process for the Construction of a Statistical Model

There are several key stages involved in the creation and use of a statistical model. A clear way to describe these is in the flow diagram below, which breaks down these steps into the main sections of work required.



3.2 Data Registration

3.2.1 Introduction to Registration - Why is it Needed?

Registration is required to accurately relate one image to another. Medical image registration is a huge field which is developing rapidly because the ability to compare images between subjects or to compare images from the same patient, has useful applications for diagnostics, operative planning and radiotherapy treatment planning. Registration techniques range in complexity and flexibility, the most simple account for only translational differences while the complex can correct global and local deformations. However, the fundamental aim of them all is to establish correspondence between two shapes and determine the transformation required to achieved this.

The field has applied techniques to solve the relatively straightforward problem of intrasubject registration of rigid structures, such as bones, most commonly required for monitoring disease progression [76]. There has also been work to try to register such rigid objects from different subjects, intersubject variability requiring a greater level of deformation [72, 15]. More complicated still is the problem of registering deformable anatomical structures such as comparing breast MRI [13], or using registration to compensate for motions such as breathing to allow cardiac function to be tracked [77]. Registration is vital in creating statistical models from medical images in order to find correspondence between members of a training set [9, 71, 72]. If an accurate correspondence between shapes is not found then illegal shapes can be generated by the model [63].

3.2.2 Classifications of Registration

Registration techniques can fundamentally be classed in one of two categories; rigid or non-rigid. These terms describing the types of transformation which that method allows. Rigid registration relies on the principle that the objects being matched are rigid. Therefore, at its most basic it only allows them to be translated and rotated, giving 6 degrees of freedom. An extension of the rigid technique is affine transformation, where anisotropic scaling and skews are allowed, giving the shape up to twelve degrees of freedom in which to deform and align to match the other shape [78]. Non rigid registration is a more complex process which allows for global changes explained by rigid transformations but also allows localised deformations to occur [10].

Medical image registration can be approached in two ways; by geometry based similarity metrics and by intensity matching. Geometric approaches build explicit models of anatomy by elements such as points, curves and surfaces, while intensity based methods match patterns over the images without incorporating anatomical knowledge. Hybrid methods using both geometric and intensity information to establish correspondence also exist and have been applied to more complex problems like brain images [12]. The other important feature which separates registration schemes from one another is the amount of user intervention required. Some demand significant manual input [8] while others aim for an automated approach [71].

3.2.3 Rigid Registration Techniques

A good example of a manual rigid registration technique is the commonly used landmarking approach employed by Cootes et al [8, 9] in building their statistical shape models. Landmarks were placed manually at key points on each example in the training set, most densely over the more complex details. The hand model, see Fig. 3.2, was represented by 72 points around the boundary, more densely packed at the ends of the fingers and in the joints. Although this is only required for the construction of the model and is not needed for later image interpretation, so only done once per training example, landmarking is still an arduous task. There is also room for error arising from personal judgement as to the exact position of a feature which may not always be clear. The biggest restriction to using landmarking, or any other manual technique, in this study is the complexity of the three dimensional problem. It is relatively simple to place points at equivalent positions on a 2D image but much harder to reliably identify points on a three dimensional feature between objects, especially when surfaces are smooth.

One of the most commonly used automatic methods is the Iterative Closest Point (ICP) algorithm [10], an optimisation technique which fits a cloud of points to a model with the advantage of not requiring a predefined relationship between points. Proposed by Besl and McKay in the early 90's as a simple and computationally efficient method for matching surfaces [79], it has become widely used in medical imaging. The technique is guaranteed to converge to a minimum, is fast to compute and relatively simple to implement. Its main flaw is that it can find a local rather than the global minimum, this being highly dependant on the initial alignment of the shapes being matched. The robustness of the technique is therefore poor where large changes in scale, rotation or position are not pre-corrected.

The popularity of ICP has led to the development of many improved algorithms such as that described by Chui et al [79]. This involves a system for the automatic removal of outliers, which can otherwise affect the effectiveness of the method, making convergence to a local minimum more likely. The method becomes a hybrid rigid/non rigid approach using thin plate splines. Initially transformation of the points are uniform, then as the points come closer to correspondence the algorithm allows localised deformations.

3.2.4 Non Rigid Registration Techniques

Spline based techniques are popular in non rigid registration; these are built on control points with a known, defined relationship. Thin plate splines (TPS) are widely used [80]. They were developed for computer vision applications in order to interpolate a surface over a fixed number of nodes. TPS relate two sets of landmark points, modelling a global affine transformation and then superimposing 'principal warps' which incorporate local deformation patterns over the shape. The potential for describing local deformations has long been seen as capable of modelling the complex deformation possible in biological image analysis. The method requires corresponding landmarks on both images which, as previously explained, is not ideal. The spline warp approach is computationally expensive, limiting the number of degrees of freedom which can be incorporated into the model [13]. B-Splines have been used to reduce this computational cost, achieved thanks to the local relationship between control points. This means that moving a control point in a spline only has a localised affect, whereas TPS require global recalculation [10].

B-Splines have been used as part of the Free Form Deformation (FFD) models developed by Rueckert et al. for modelling anatomical structures [13]. FFD models are an advanced, automated method for modelling object deformation in three dimensions. The object is placed in an underlying mesh of control points and is deformed indirectly through manipulation of this mesh, producing a smooth and continuous transformation. This again is a computationally demanding and complex technique which has been used to model high level deformation required for registration of breast MRI (correcting for motion) and brain images [12, 13]. These schemes assume that everything within the image is equally nonrigidly deformable, in reality the the human body is composed of many different structures with different rigidities. Staring et al. [81] incorporated filtering into a nonrigid registration scheme to allow different tissues to deform differently, in relation to their stiffness, an approach which would be important for modelling systems and tracking disease progressions such as tumours.

Yao and Taylor [82] utilise a range of techniques including snakes, b-splines and analysis of image intensity changes to automatically mesh CT images and identify bone contours. The bones are represented as solid tetrahedral meshes where the intensity-density information is known on an element by element basis. This density atlas has potential application for radiograph reproduction (2D from 3D) and for automated 3D intensity based registration. Davies et al. [83] offer an alternative approach, parametrising the training models and finding an optimal set of variables by which to deform this model to define new instances. Described as the Minimum Description Length (MDL) optimisation has the potential to offer a much more compact set of model parameters than landmarking approaches. However it requires a known shape function similar to the target (such as a sphere) to fit to a model, which is not appropriate for all applications and in its current form is computationally slower than alternative methods.

3.3 Generation of Training Data

3.3.1 Sources of Data

The creation of a representative three dimensional model required accurate three dimensional data to be recorded. One of the most commonly used and available sources of three dimensional information for orthopaedic data for human subjects is Computer Tomography scans. CT scans produce multiple X-Ray slices through the subject which can then be reconstructed digitally into a volume. The output is divided up into small volumes called voxels, each containing a grey level value which is derived from X-ray attenuation at that point and so relates directly to the density of the material. In medical images, voxels are typically 0.5-1mm square in slice with a slice distance anywhere between 1-10mm, depending on the resolution required over that area. A lower resolution will be used where possible to reduce the radiation dose to the patient. Minimising the exposure, cost and imaging time are key reasons why planar X-rays are routinely used in preference to CT.

The raw voxalised output from CT would clearly produce a jagged, blocky image so a variety of computational software is available to extract the voxels of interest from the scans and produce smooth, realistic looking models. This is known as segmentation. Studies routinely use a CT generated model as the base bone model for FE experiments [84]. Patient specific models, for example of the femur [38], have

been created for finite element studies, capable of predicting individualised stress fields.

Other medical imaging modalities have been investigated to produce 3D models, these include Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Single Photon Emission Tomography (SPECT), Ultrasound and X-rays [78]. MRI, PET and SPECT scans are not regularly used for orthopaedic applications, they are more commonly used for imaging soft tissue or dynamic systems such as the brain [12], or the heart [85]. Ultrasound has been proposed as an attractive alternative to the radiation based modalities, as there is no radiation dose associated with it and it is cheap. A tracked ultrasound, allowing the positions of the slices of data to be recorded, can be reconstructed into a 3D model incorporating geometry and density information [78, 86]. Although the resolution produced is not as good as CT, X-ray data, despite being two dimensional in nature, has been used to create three dimensional models by combining them with statistical shape models [87, 66].

3.3.2 Methodology for Creation of 3D Models from CT scans

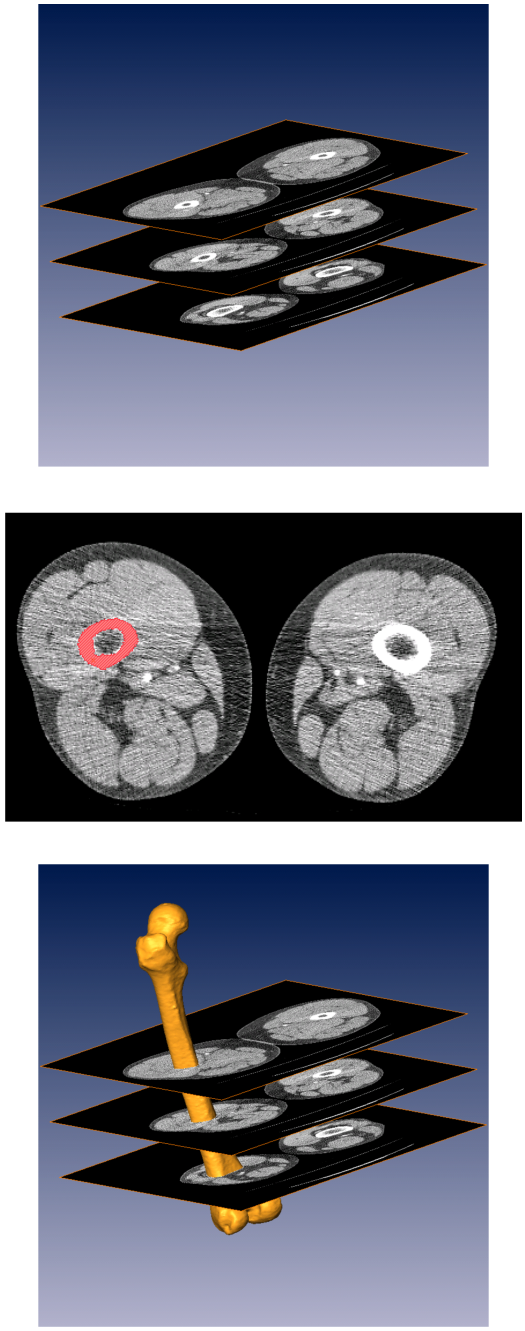


Figure 3.7: Illustration of the stages required to segment a femur from CT scans using Avizo^{TM}

This study requires the generation of three dimensional femur models from CT data. The created representations should accurately reproduce each subject's femur, requiring the extraction of geometric data and grey level values from the scans. Many software packages are available, fundamentally achieving this through the same process. The package chosen for use in this project is Avizo® (Visualization Sciences Group,USA/France formerly VSG of Mercury Computer Systems, USA).

The package reads in the CT slices in order and stacks them as shown (Fig. 3.7). A segmentation algorithm was then implemented to highlight the areas of interest on each slice, discriminating by voxel greylevel value. Defining a high value will mean only dense material, bone, will be selected. The algorithm is not intelligent to shape so often areas which are not required are selected, for example the pelvis along with the femur. These features were removed manually or by other selection tools within the program. The selected areas were then used to create a model of the surface of the segmented shape from which surface and volumetric meshes could be produced.

3.3.3 Methodology for Extraction of Material Property Data

A significant feature of the model developed in this study is the incorporation of material property variation within the femur as well as external differences in geometry. This information therefore also needs to be extracted from CT scan data along with geometry. This is possible because, as previously discussed, the grey level of each voxel within a CT scan directly relates to the density of the material it represents.

To assign material properties to a solid finite element mesh this study used a program called BioMesh (A. Hopkins, Imperial College London), which was designed to work specifically with Avizo generated geometries and further developed to meet the requirements of this work. A commonly used alternative, which works in a similar way, is the free software BoneMat (Rizzoli Institute, Bologna) [88]. Both programs define a number of 'sampling points' within each element and interpolate the greylevel relating to their coordinates from the original CT. The element greylevel is assigned from the sampling point values. BioMesh has the added advantage to this study of also assigning greylevel values to each node, which is particularly useful when building the data into a matrix for PCA. This allows the defining data for each training member to be significantly shorter than if elemental values were extracted.

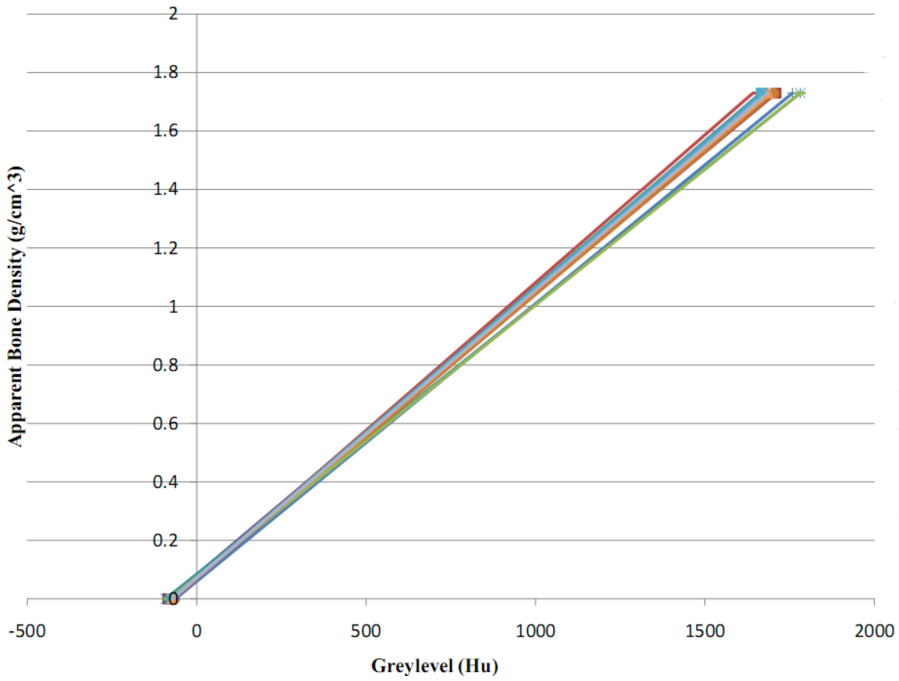


Figure 3.8: Graph showing a range of calculated calibration lines for different CT files

Greylevel is proportional to apparent bone density, however to accurately convert between the two the CT scans must be calibrated to correct for any slight offset due to

machine set up. Ideally this is calculated using calibration phantoms within the CT scan, usually blocks of materials included in the image which have a known density. Unfortunately, calibration phantoms are not routinely used in clinical CT scans and were not present in many of the data sets available in this study. The proportional relationship was therefore defined by identifying the specific greylevels in each CT set for two materials with 'known' density values. Bone marrow, the lowest modulus readings found in the medullary canal, was taken as having a density equivalent to water, $0g/cm^3$, and the highest value cortical bone identifiable was considered to be $1.73g/cm^3$ [89] (These density values relate to the density of mineralised tissue). Although this could be a source of error in the final model as it assumes that every subject has this maximum density of bone. The relationships for a range of the data sets used are shown above and define the calibration values used by BioMesh (Fig. 3.8).

Young's Modulus is one of the properties required by finite element analysis to define a material. Schileo et al [41] investigated mathematical relationships between apparent bone density, ρ, and Young's Modulus, E, through experimental testing of bone. The equation found to best correlate with the experimental findings was that established by Morgan et al. [90], despite all their test specimens lying in the lower density band of bone ($< 0.8g/cm^3$).

$$E = 6850\rho^{1.49} \tag{3.1}$$

This relationship was therefore the one used to convert nodal greylevel into nodal modulus, which could then be incorporated into the PCA model at a later stage. FE usually requires element modulus, which could be recalculated by averaging the modulus values of the four nodes making up each element.

As previously mentioned, the lack of calibration phantoms is not ideal and the methodology described was established to try to calibrate the data as well as possible. The extracted material properties generated for the training set femurs were all examined and found to lie within realistic bounds and have realistic distributions. For the approach required in this study these results were suitable, however if more precise correction was required further investigation of density-modulus relationships should be pursued. In particular, different relationships can be used to assign material properties to cancellous and cortical bone, however the initial definition of a divide between cortical and cancellous in uncalibrated CT scans would be difficult.

3.4 Methodolgy Developed and Applied for Construction of SM of the Femur

3.4.1 Overview of Methodology for SM Construction

This section aims to provide a step by step overview of the stages used in this study to construct a statistical model of the femur from real subject CT scans. The significant challenge in doing this is establishing correspondence between each training example in such a way that the location and material properties at any given point in one model can be directly related to an equivalent point in another model. Although the previous sections have described how this has been achieved by previous work, the novel difficulties in this case lie in achieving high accuracy in registration, at a fine mesh quality suitable for FE studies, over the complex morphology and relatively large volume of the whole femur.

1. Segment out the region of interest, i.e. the femur, from each CT data file and describe its shape as a dense cloud of surface points.

2. Select one femur instance as the baseline or reference to which each other model will be matched and convert this model into a high quality solid tetrahedral mesh.

3. Register the baseline femur to each training example using a scheme based on an elastic matching algorithm[2] and volume mesh deformation strategy[3].

4. Assign every point in the morphed meshes a grey level value from their original CT scan files using a material property extraction program BioMesh.

5. Construct a Point Distribution Model using the mesh based femur representation complete with material properties, and create the statistical model using principal component analysis. The statistical model produces a volume mesh along with spatially varying material properties that can be directly used by a finite element solver.

[2]Software developed by Dr Prasanth Nair, Computational Engineering Design Center, University of Southampton

[3]Software based on open source MatLab FEA code developed by Surya Mohan, Computational Engineering Design Center, University of Southampton

Figure 3.9: Illustration of the target mesh (black) with morphed baseline mesh (grey) superimposed, showing the accuracy achievable by the registration scheme.

The registration stages of this process are the most complex and the most crucial to the success of the final model. The two stages, surface registration and volumetric morphing, result in three dimensional correspondence between a baseline tetrahedral mesh and each target surface mesh in the training set. The surface matching step iteratively deforms the baseline surface vertices to match the given target surface and is able to achieve a smooth, accurate final mesh through user defined inputs controlling the magnitude and speed of the deformation and a two stage smoothing process. These parameters specify the area affected by moving a single vertex and the size of each iterative step, both decreasing over a specified number of iterations, and prevent mesh distortion. The registration scheme initially allows almost global deformations to act which roughly align the surfaces before narrowing the regions of influence to produce fine, local manipulations of the mesh until a stopping criteria is reached. The accuracy of the surface registration achievable using this algorithm is illustrated in Fig. 3.9, showing the morphed and target surfaces superimposed. The volumetric registration step used the registered surface points as a scaffold and positioned the internal mesh points based on the surface node displacement vectors by solving decoupled Laplacian equations with appropriate essential boundary conditions.

3.4.2 Baseline Femur Model

Establishing correspondence between each member of the training set depends upon the the accurate registration of a common baseline mesh to each femur surface model. Significant efforts were made in developing the baseline mesh as any femur representation produced by the statistical model would be described in this form. The baseline mesh quality therefore had to be as high as possible to allow generated models to be used in finite element studies, however the number of elements had to be constrained to reduce computational cost.

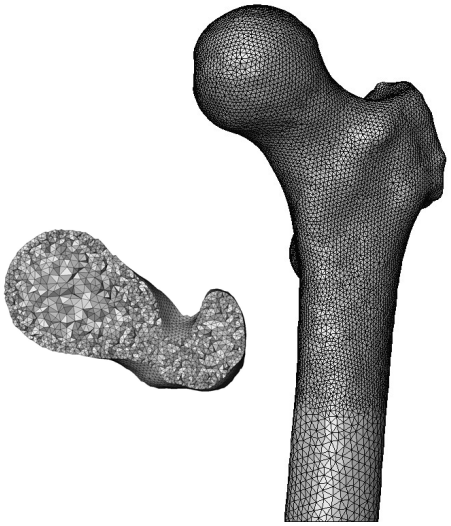


Figure 3.10: The baseline meshed geometry. Shows a section of mesh through the femoral head and the transition in mesh size between the proximal femur and femoral shaft.

To maintain mesh quality over all examples, the median length femur was chosen to be the reference, with the rationale that this would lead to the minimum element distortion when stretching or compressing the mesh to fit another instance. The chosen femur surface mesh was imported into meshing software, ANSYS© ICEM CFD™ (ANSYS. Inc, Canonsbury PA), and converted into a high quality, 4-noded solid tetrahedral element mesh with a global size of 3mm. To achieve a balance between model definition and computational cost the model was then split into three regions, with the upper and lower thirds' mesh size refined to 1-1.5mm at the surface (Fig. 3.10).

This was justified for three reasons; 1) these areas are of more clinical interest so a fine mesh will be important for future use of the model, 2) these areas contain the most rapidly changing geometry and material properties, hence require a higher density of surface points to achieve accurate registration and 3) Perillo-Marcone et al. (2003) recommended choosing a element size equivalent to the CT slice distance in order to achieve convergence of material property distribution. The baseline tetrahedral mesh, and therefore any subsequent mesh produced by the model, consisted of 615,523 elements and 117,225 nodes, of which 27,171 were on the surface.

3.4.3 Registration Strategy

The aim of the registration scheme was to manipulate the baseline tetrahedral mesh to achieve three dimensional correspondence to the geometry of each of the other femurs, where each target femur is represented as a surface mesh. The developed process had two stages; 1) surface registration and 2) volumetric morphing.

Surface Registration[4]

A surface registration scheme was originally developed based on an algorithm proposed by Moshfeghi et al. [91]. Two key modifications were made to allow registration of the high density meshes in this work. Firstly k-d trees [92] were used for nearest neighbour searching, this is a computationally efficient method of accelerating nearest neighbour searching in a large data set. Secondly Laplacian smoothing [93] was incorporated at each iteration to try to prevent any significant degradation of mesh quality through element distortion.

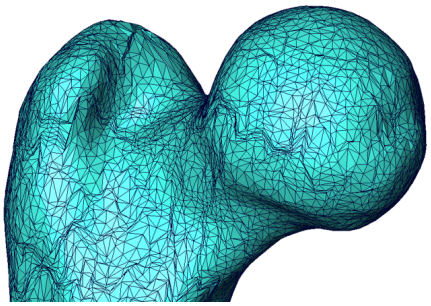


Figure 3.11: Evidence of mesh folding and degredation.

The addition of k-d trees was to speed up computation, without this modification the registration scheme would take many, many hours to solve. The final code was able to register the >27,000 surface point meshes in approximately 40 minutes. This time could have been less but a conservative set of parameters where chosen to control the registration speed, ensuring the final accuracy and mesh quality. The need for additional smoothing became evident from initial studies where there was evidence of element bunching, element distortion and mesh folding (Fig. 3.11), mainly observed in areas of concavity on the target surface. This was a known issue with the algorithm of Moshfeghi et al. The Gaussian smoothing term present in the scheme was deemed unable to maintain mesh quality alone, so a Laplacian smoothing step was added at each registration step. This two stage smoothing approach did not add significant computational time but showed a marked improvement in the mesh quality of the registered models (Fig. 3.12).

[4]Software developed by Dr Prasanth Nair, Computational Engineering Design Center, University of Southampton

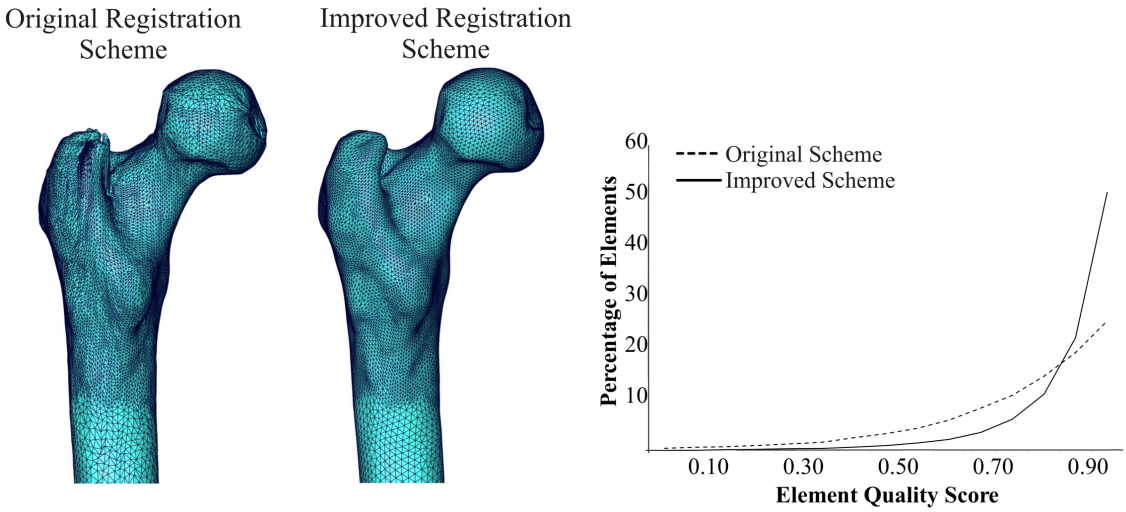


Figure 3.12: Illustrations of the registered mesh quality achieved with the original algorithm (left) and the improved algorithm (center). The surface mesh quality distribution of the two schemes is plotted on the right, as measured by triangle aspect ratio.

The algorithm required user defined inputs specifying the magnitude and speed of surface matching and parameters of the smoothing to be applied. This ensured the baseline surface vertices were iteratively deformed onto a given target surface smoothly and accurately. Once defined, these were applied to all registration operations for this statistical model. The parameters controlled the area affected by moving a single vertex and the magnitude of each iterative step, both decreasing over the specified number of iterations. This initially allowed almost global deformations to act which roughly aligned the surfaces before narrowing the regions of influence to produce fine, local manipulations of the mesh until the required number of iterations were completed.

Volume Morphing[5]

The problem of deforming a volume mesh to track moving/perturbed surfaces has been extensively studied in several areas including computational geometry, moving surface fluid flow simulation and engineering design optimisation [94, 95, 96]. The central idea underpinning many existing approaches is to solve partial differential equations (for example the elasticity or diffusion equations), for the coordinates of the deformed mesh with the known surface displacements imposed as essential boundary conditions.

[5]Software based on open source MatLab FEA code developed by Surya Mohan, Computational Engineering Design Center, University of Southampton

The diffusion based mesh deformation strategy was adopted for this work. It is computationally efficient and has been shown to work well for a variety of applications [97, 94]. This approach solved the following decoupled three-dimensional Laplace equations to deform the baseline volume mesh

$$\nabla^2 \delta x = 0,\ \nabla^2 \delta y = 0,\ \nabla^2 \delta z = 0, \tag{3.2}$$

where δx, δy and δz were the displacement fields applied to the baseline mesh coordinates in the $x-$, $y-$ and $z-$ directions, respectively, and ∇^2 was the standard Laplacian operator defined on the baseline volume mesh.

The x, y, z components of the surface registration deformations vectors were used to specify appropriate Dirichlet boundary conditions for equation 3.2. The result was a morphing of the internal points of the baseline volume mesh onto the target volume. Once the whole registration process was applied to all members of the training set, each was described by a solid tetrahedral mesh, with direct correspondence established between each point and element.

An issue with this approach is that it can allow element reversal, which is possible if the magnitude of the surface point displacement vectors become large. Two strategies were considered to address this; firstly by integrating volume morphing into the registration strategy, making the volume morphing process incremental and secondly by repairing the mesh. The incremental morphing strategy did not suffer element reversal issues and also provided high quality meshes, agreeing with observations made by Shontz and Vavasis [96], but this was at the expense of a dramatic increase in computational cost. Therefore, due to computational constraints the second approach was adopted. Fully automated, efficient mesh checking and repair was coded into the registration strategy following the completion of volumetric meshing.

At this stage each of the training set geometries is represented by a comparable solid, tetrahedral mesh with nodal and elemental correspondence. Using the methodology described in Section 3.3.3 each registered mesh was assigned nodal Young's Modulus from their original CT scans. The training set now contains model specific geometric and material variation for each example femur.

3.4.4 Mesh Quality Checks

Mesh quality is a key concern when automating the generation of finite element models. As the statistical model is created by stretching and morphing a baseline

mesh some level of mesh distortion is inevitable. It is crucial that despite this, mesh quality is maintained such that meshes can be used in finite element simulation without causing errors or producing inaccurate results. Initially the mesh needed to be scrutinised to ensure the methodology was sound, but in the longer term mesh checks are also necessary as an automated step in creating large numbers of models.

As mentioned in section 3.4.3, some mesh repair is required to fix element reversal in this method. The problem is possible to solve within the morphing process, but would be highly computationally expensive whereas a simple code to check and reverse any 'inside-out' elements is quick and simple to apply.

Metric	Warning	Error
Maximum Angle (°)	165	179.9
Aspect Ratio	20	100 000

Table 3.1: Table defining mesh quality criteria and related level for warning and failure as defined by ANSYS©.

Specific mesh quality limits are difficult to define. The approach taken in this work was to look at the levels of element distortion which would result in a warning or failure within ANSYS©, Table 3.1. The tetrahedral elements produced as a result of a maximum face angle or aspect ratio defined by these rules were then graded by two mesh quality metrics which could be coded into model generation. Both metrics interrogate a mesh on an element by element basis, giving a score between 0 and 1, where 1 indicates a perfectly regular tetrahedron. The score produced by the warning and error shaped elements were then used to define a pass-fail criteria so that if a mesh was generated with poor mesh quality then it would be automatically discarded (in practice this did not occur).

The normalised shape ratio (NSR) [98] is used as an indicator of aspect ratio and is calculated as

$$NSR = \frac{3r}{R} \tag{3.3}$$

where r is the radius of the element insphere and R is the radius of the elements circumsphere. The second metric used is a shape distortion ration (SDR) [99]

$$SDR = \frac{12(3V)^{2/3}}{\sum_{i=1}^{6} l_i^2} \tag{3.4}$$

where V is the element volume and l the element edge length.

3.4.5 Principal Component Analysis - Covariance and Correlation Approaches

Principal component analysis is a statistical technique which allows high dimensional data sets to be decomposed into their significant components, creating an accurate, reduced order model of the original data. PCA calculates the eigenvectors and eigenvalues of a data set which represent the principal modes of variation, with the later indicating their relative significance in capturing the variation between members of the training data set matrix, $\mathbf{X}$. Following surface registration and volumetric mesh morphing all femur models were described in the same coordinate system by the same number of corresponding nodes and elements. Each femur, $\boldsymbol{x}_i$, defined by a single vector containing nodal coordinates, x, y, z, and modulus values at each node, E, can be compactly written as;

$$\boldsymbol{x}_i = [x_{1i}, y_{1i}, z_{1i}, E_{1i}, ..., x_{ni}, y_{ni}, z_{ni}, E_{ni}]^T, \quad 1 \leqslant i \leqslant N. \tag{3.5}$$

Where N is the number of geometries in the training set and n is number of nodes in each mesh. Each training geometry could then be combined to create the training matrix $\mathbf{X}$, as shown below.

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]^T \in \mathbb{R}^{N \times 4n} \tag{3.6}$$

As each node in the femur was described by 4 components, x, y, z, E, each training vector was $4n$ in length, making the overall training matrix extremely large. The scale of the data set made explicit calculation of the covariance matrix and eigenvectors impossible (the matrix is simply too large, $486,900 \times 486,900$), so Singular Value Decomposition (SVD) was used as a more suitable method for this application.

Two methods of PCA were investigated to establish which best decomposed the training data to produce as rapid a decay of eigenvalues as possible. The data was analysed using the commonly used covariance method and then using the correlation method, which is more suitable for data sets containing mixed units [100], in this case position and modulus. The differences between the methods behind the two approaches follow.

In order to calculate the basis vectors, or eigenvectors, both schemes required modifications to the training matrix. The covariance approach removed the training data mean from each data set member

$$\tilde{\boldsymbol{x}}_i = \boldsymbol{x}_i - \bar{\boldsymbol{x}}, \tag{3.7}$$

where $\bar{\boldsymbol{x}}$ is the mean of the training geometries contained in $\mathbf{X}$. The modified training matrix was defined as $\mathbf{A}$, where $\mathbf{A} = [\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, ...\tilde{\boldsymbol{x}}_N]^T$ and the SVD of the data set written in the conventional form was

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T. \tag{3.8}$$

where $\mathbf{U} \in \mathbb{R}^{N\times\ N}$ and $\mathbf{V} \in \mathbb{R}^{4n\times\ 4n}$ are orthogonal matrices, or left and right singular vectors. $\boldsymbol{\Sigma} \in \mathbb{R}^{N\times 4n}$ is a diagonal matrix with diagonal elements $\boldsymbol{\Sigma}_{ii}$ consisting of $q = min(N, n)$ nonnegative numbers $\boldsymbol{\sigma}_i$, arranged in decreasing order. $\mathbf{V}$ is the eigenvector matrix, the columns contain the proper orthonormal modes of the system, so the basis vectors $\boldsymbol{\psi}_i = \mathbf{V}(:, i)$ can be used in equation 3.9 to approximate $\boldsymbol{x}$.

The covariance approach approximated each femur vector as follows

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \sum_{i=1}^{m} b_i \boldsymbol{\psi_i}. \tag{3.9}$$

where $\boldsymbol{\psi}_i \in \mathbb{R}^{4n}$, $i = 1, 2, ..., m$ define the set of m basis vectors and b_i, $i = 1, 2, ..., m$ are a set of coefficients controlling the relative influence of each basis vector on the final output $\boldsymbol{x}$. $\mathbf{b}$ can be described as a shape parameter, in this application it describes the shape and modulus of the femur model when combined with $\boldsymbol{\psi}$. It is unique for each training set member and is the key to generating new instances, possible by altering the value of this vector.

The modifications to the training data for the correlation method consisted of two steps. Firstly each member was divided by the standard deviation of the data set $\boldsymbol{\alpha} \in \mathbb{R}^{4n}$, where α_i is the standard deviation across every i^{th} term in the training matrix. Secondly, the mean of the newly calculated data set was removed from each set member. Equation 3.8 was then applied with the modified training matrix $\mathbf{A}$ constructed from the normalised set members, as with the covariance approach. The output of the statistical model generated through this approach will generate normalised femur vectors. Equation 3.9 is still applicable to approximate each femur geometry but only after the standard deviations, $\boldsymbol{\alpha}$, are multiplied back in.

Calculation of the dominant modes was illustrated by the ability of each method to capture the variation in the data set, showing how great a percentage can be described by any number of included modes. Interpretation of this result allowed the appropriate number of dominant basis vectors, m, to be established where $m \ll N$. The energy within the first m modes was found from

$$E(m) = \frac{\sum_{i=1}^{m} \sigma_i^2}{\sum_{i=1}^{N} \sigma_i^2}. \tag{3.10}$$

3.5 Results of SM Construction

3.5.1 Training Data

This study used 46 subjects, with CT voxel resolutions ranging from 0.488×0.488×1.5mm to 0.781×0.781×2mm. The subjects were aged between 43 and 91, 15 female and 31 male. Semi-automated segmentation of bone from surrounding tissue was achieved with grey level thresholding tools and manual slice by slice corrections using Avizo (Visualization Sciences Group, USA/France) formerly VSG of Mercury Computer Systems, USA. By preference the left femur was segmented from each patient but in eleven cases this was not possible so the right femur was modelled instead and subsequently mirrored through the mid sagittal plane. This resulted in a training set of 46 left femurs, each represented by a triangulated finite element surface mesh with a resolution of a similar order to the baseline surface.

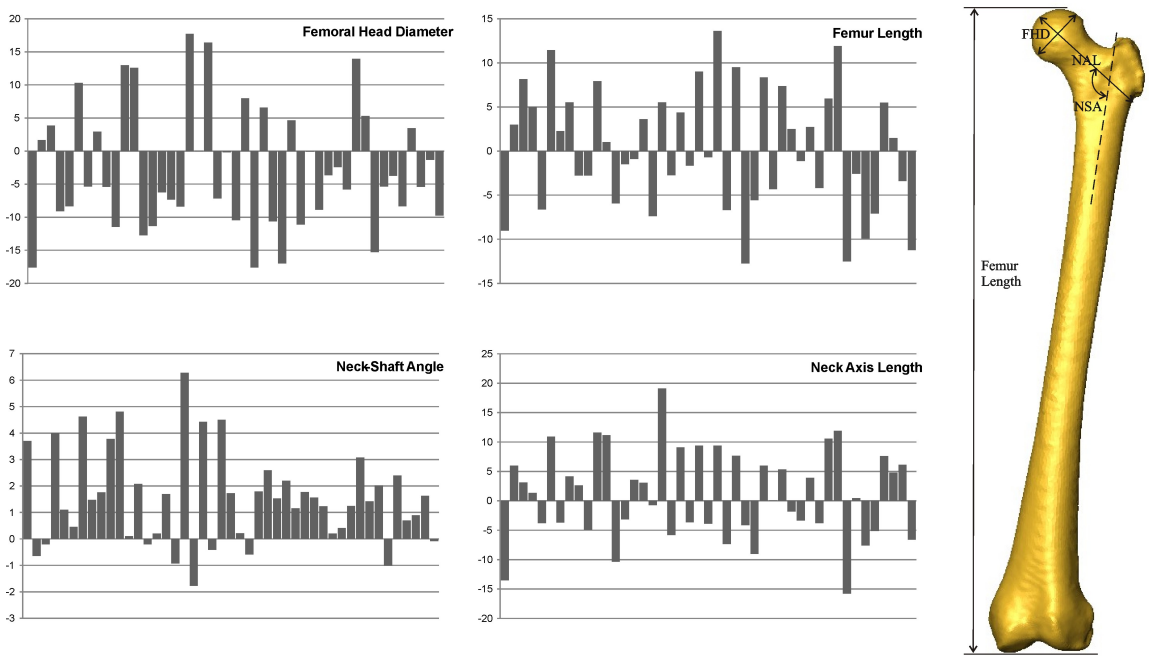


Figure 3.13: Graphs showing the variation in femur geometries present in the training data set as a percentage difference to the baseline femur of four geometry metrics; Neck Axis Length (NAL), Femoral Head Diameter (FHD), Neck Shaft Angle (NSA) and Femur Length.

The degree of variation present within the training set was relatively large. To illustrate this some key anatomical measurements relating to the major dimensions of the femur were taken from each of the training femur models. These included femur length, femoral head diameters, neck shaft angle and neck axis length. The measurements were then described as a percentage difference from the reference femur

geometry (Fig. 3.5.1). The figure shows the dramatic range of geometric variations present, also highlighting the fact that the training set contains more than simple scaling as an increase in one measurement does not always coincide with an increase in another.

3.5.2 Assesment of Registration Strategy

Assessment of Surface Registration Scheme

The success of the first stage of the registration scheme, surface matching, was examined from two perspectives; 1) how closely the registered surface fitted the target surface and 2) the quality of the registered surface mesh. Registration accuracy was assessed by calculating the shortest distance between each registered node and the target surface. This found the mean surface registration error over all 45 femurs to be 0.598mm (45 as the 46th femur was the baseline model). On visual inspection the meshes appear smoothly matched (Fig. 3.9), so it is likely that if the registered nodes are offset from their target then they will still be on the target surface. A boxplot illustrating the distribution of error across the training set is shown in figure 3.14. This plot contains a wealth of information about the registration error. The vast majority of nodes are within $1mm$ of their target with no error greater than $3mm$.

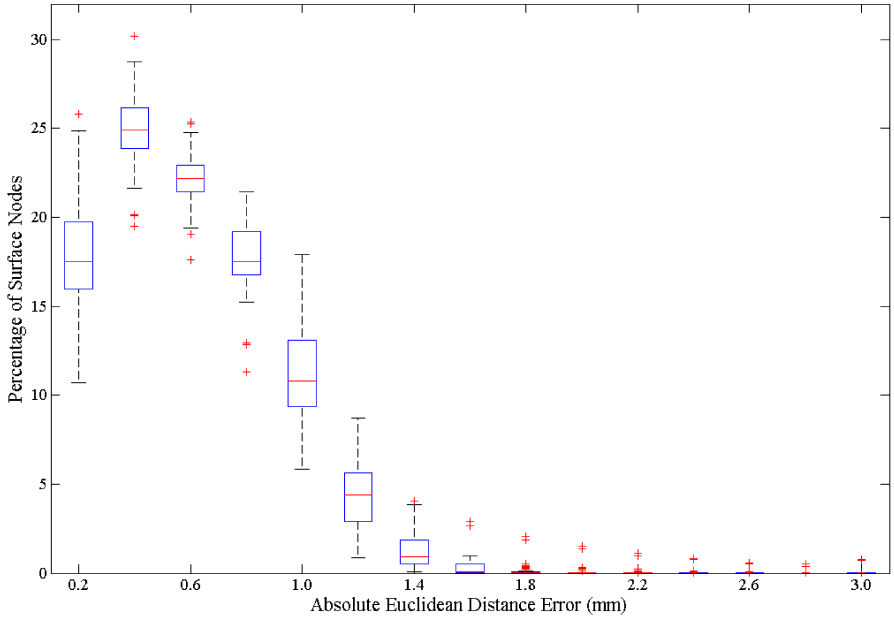


Figure 3.14: Boxplot of final surface registration errors over all training set members, calculated as the shortest euclidean distance between each registered surface node and the target surface.

The surface mesh quality was calculated for each registered femur as well as for the initial baseline mesh using a Normalised Shape Ratio metric (NSR) [98] which indicated element aspect ratio. This gave a score between 0 and 1 for each element, where 1 indicated perfect regularity. The mean, minimum and maximum distribution of element scores for the registered data set, alongside the initial baseline quality are shown in Figure 3.15. In addition to the distribution of mesh quality an overall quality score was assigned to each mesh by binning the elements by NSR score and multiplying the value of the bin by the percentage of elements it contained. Therefore, a perfect mesh would have 100% of elements in the top bin and so score 100 and vice versa if all elements were in the lowest bin the mesh would score 1. This provided a measure of mesh degradation due to the registration process. The mean score of the registered femurs was 93.0 with the minimum seen being 89.0 and the maximum 95.6, as compared to the baseline mesh score of 98.2.

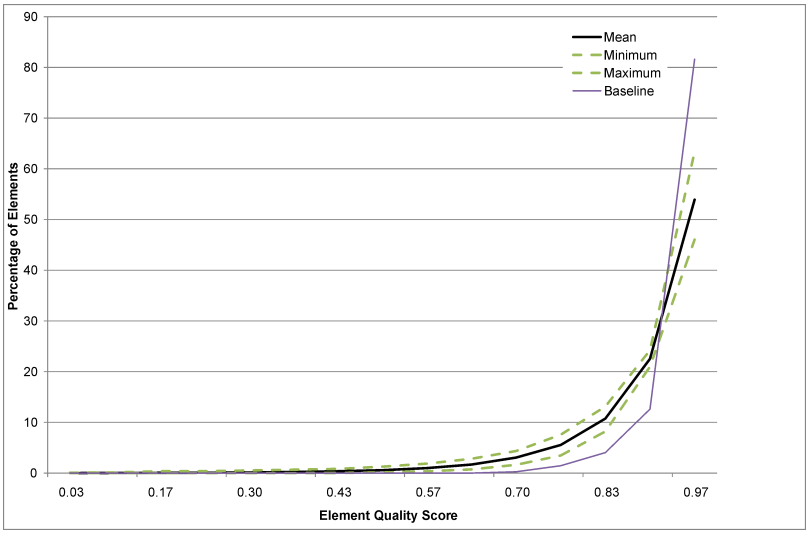


Figure 3.15: Plot of the surface element aspect ratio score of the baseline mesh compared to the mean, minimum and maximum scores of the registered training geometry surface meshes.

Assessment of Volume Morphing

To assess the success of the volume morphing step of the registration scheme it was again appropriate to examine the mesh quality produced. This was done using the same method as for the surface mesh, analysing the NSR quality distribution of the morphed meshes in comparison to the original baseline. The calculated mesh scores showed that the morphed meshes had a mean quality of 78.9, a minimum of 77.49 and a maximum of 81.13, as compared to the baseline score of 79.53. As with the surface mesh there was a slight decrease in mesh quality but no significant increase in low quality elements (Fig. 3.16).

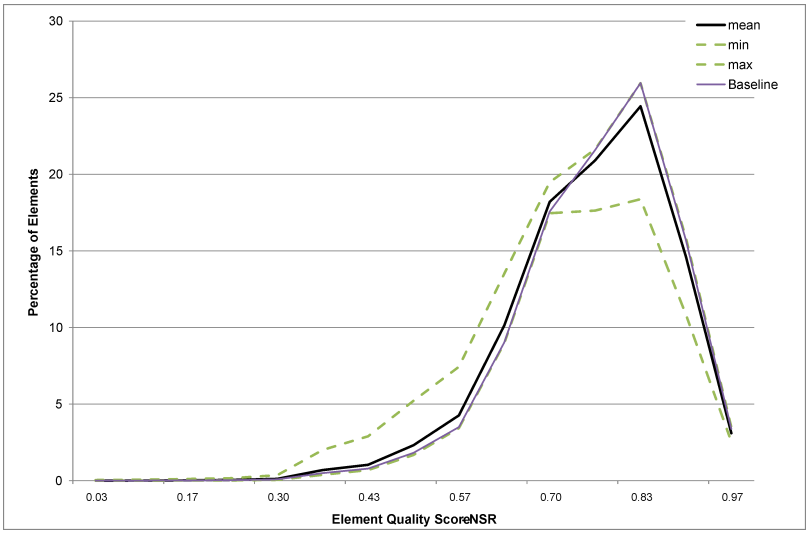


Figure 3.16: Plot of the solid element aspect ratio score of the baseline mesh compared to the mean, minimum and maximum scores of the morphed training geometry meshes.

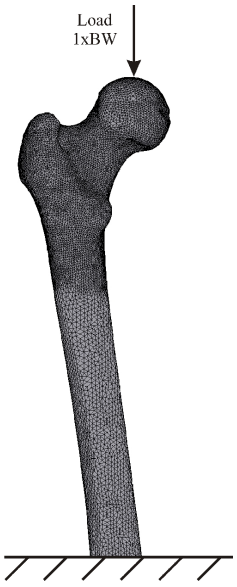


Figure 3.17: Loading condition applied simulating stance.

However, mesh quality alone was only an indicator of how successful volume morphing had been. As the intention was to use the model for FE analysis it seemed appropriate to investigate how a model meshed through the morphing approach compared to a model of the same instance meshed manually. To test this a simple static, linear elastic FE simulation was run to look at predicted strain distribution. A stance loading condition was chosen, involving a vertical force equivalent to 1x body weight being applied vertically over a $\sim \phi$1cm area of the femoral head, and fixing of the distal portion of the femur (Fig. 3.17). If the registration scheme was robust then each automatically generated mesh should produce the same strain field as its counterpart meshed from scratch to the same element criteria as the baseline mesh using ANSYS ICEM CFD^{TM}(ANSYS. Inc., Canonsburg PA). Figure 3.18 shows plots for three instances, each meshed by both methods, illustrating modulus and strain through a cross section and showing comparative strain distribution within the proximal femur. These plots showed that the models generated through registration and morphing produce almost identical results to those of their uniquely meshed counterparts.

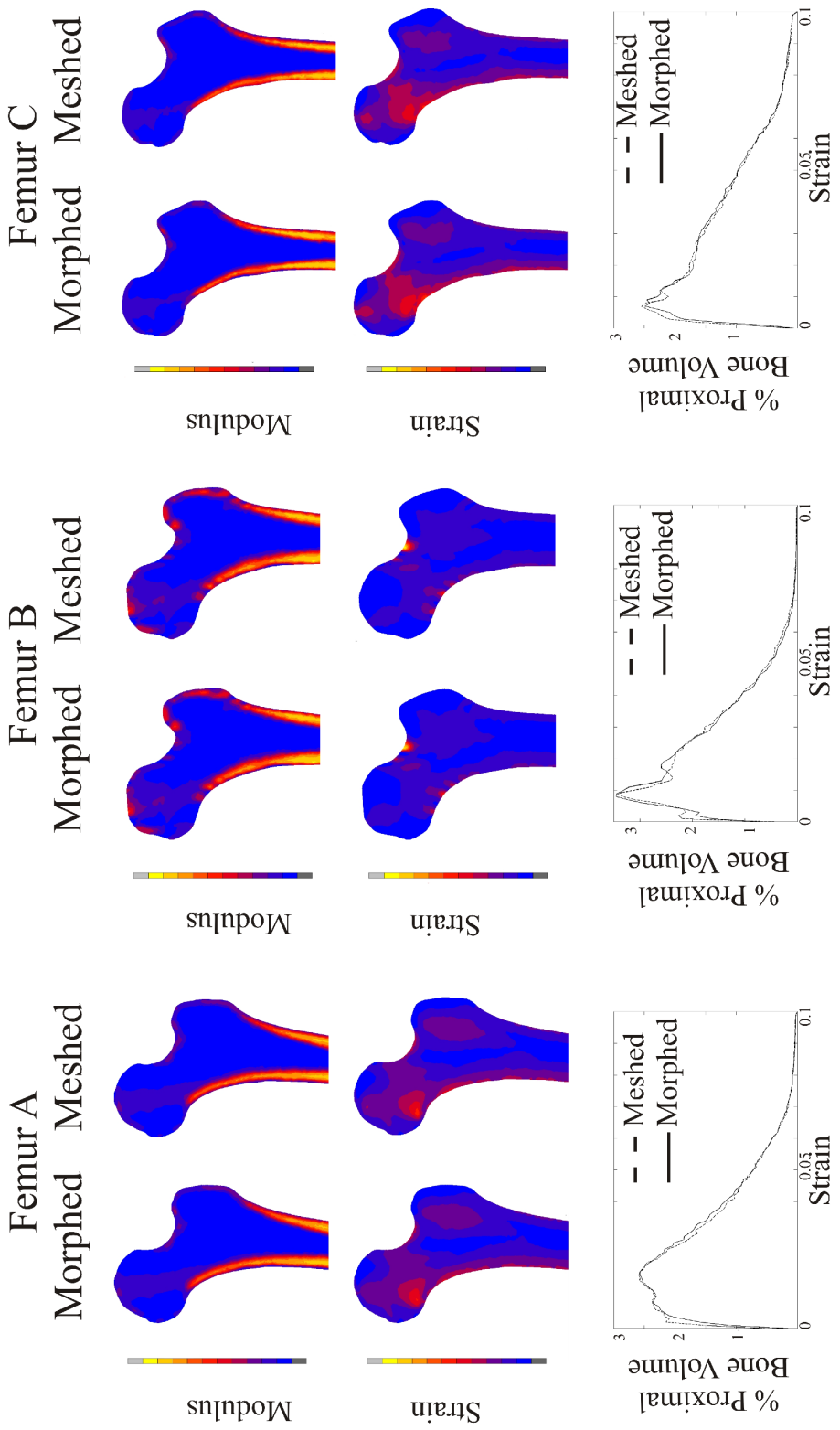


Figure 3.18: Plot of morphed and manually meshed instances of the same geometry, comparing material modulus representation and strain distribution resulting from a 1x body weight load applied vertically to femoral head, simulating one-legged stance.

	Covariance Method		Correlation Method	
Mode	Variance Captured	Total Variance	Variance Captured	Total Variance
1	20.74	20.74	45.59	45.59
2	11.01	31.75	7.74	53.32
3	6.85	38.59	6.25	59.56
4	5.59	44.19	4.80	64.37
5	4.33	48.52	3.50	67.86
6	4.18	52.70	2.56	71.30
7	2.96	55.65	2.29	73.56
8	2.73	58.39	2.00	76.15
9	2.40	60.79	1.62	78.15
10	2.28	63.07	1.55	81.72

Table 3.2: Percentage variation captured by first 10 eigenmodes computed using the covariance and correlation methods.

3.5.3 Interpretation of Eigenmodes

Comparing the eigenvalues produced from PCA of the data set using the covariance and correlation approaches can not be done directly as the correlation method works on a normalised version of the data whereas the covariance method does not. Table 3.2 details the eigenmode decay by each technique, suggesting that the correlation approach's decay is more rapid. Reconstruction error gives a better indication of this decay than the raw numbers, and thus the success of the method (Fig. 3.21). Investigations using a smaller training set used reconstruction error to compare the covariance and correlation approaches, examining geometry and modulus separately (Fig. B.3). The geometry showed a markedly improved result with the correlation approach, this was not matched in the modulus reconstruction but the results were no worse than the covariance alternative. The analyses indicated that the correlation method was better suited to the analysis of data with mixed units and that although there is some correlation among the training set it is not strong.

The physical effect on shape and material characteristics of each eigenmode was investigated by manipulating each mode in isolation and visualising the femur produced (Fig. 3.19). The first mode was dominated by scaling effects alongside an increase in anteversion angle due to rotation of the femoral head. The modulus values of the shaft cortex were seen to increase significantly as the femur size decreased. By the second mode little change in femoral length was seen, however a thickening of the bone was observed alongside a reduction of the highest modulus (cortical) bone and an increase in medulary cavity volume. These characteristics are very similar to those reported by Ruff and Hayes [101] when describing the effects of ageing on the

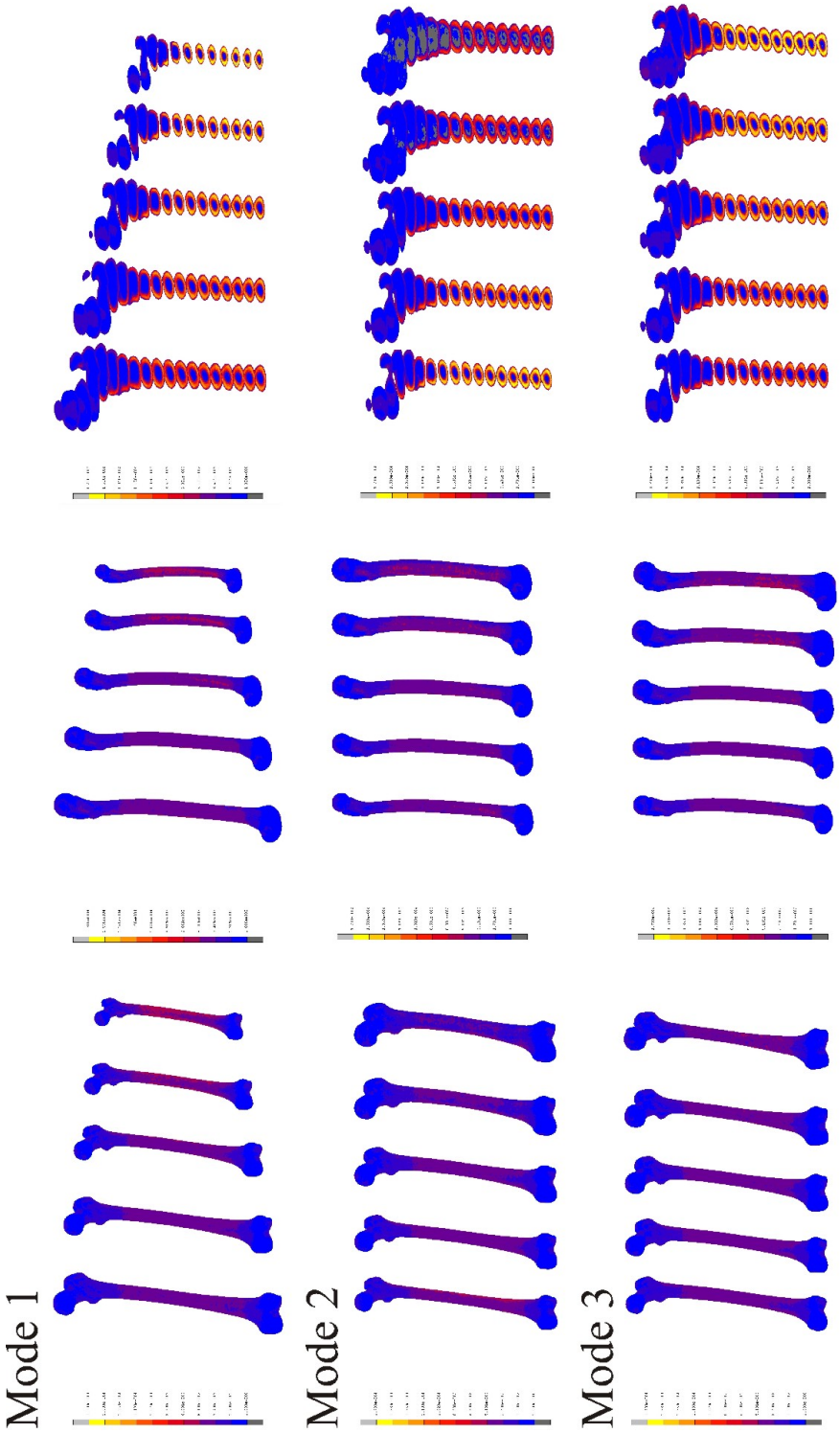


Figure 3.19: Geometry and material property changes with first 3 modes, varied between ± 3 standard deviations, acting in isolation.

femur. There was a slight decrease in neck-shaft angle and a noticeable change in anteversion angle. Mode three again showed an alteration in anteversion, this time decreasing through the mode. It mainly indicated an increase in average bone modulus with some further subtle geometric traits such as an increase in femoral head

diameter, condylar size, lateral inclination of the femoral shaft and a reduction of bowing in the shaft in the sagittal plane.

The nature of the changes observed are interesting since measures based on metrics such as cortical thickness, average bone density and neck-shaft angle have been shown to be risk factors for clinical problems such as proximal femoral fracture [102, 103, 104]. The influence of the modes provides an insight into the significant ways in which this set of femurs vary, however it must be noted that these modes will never occur in isolation. In reality any femur will be the product of the combined effect of a number of modes, which may result in the features observed being cancelled out or exaggerated.

3.6 Construction of New Femur Models and Assessing Model Robustness

3.6.1 Methodology for Creation of New Femur Models

A key aim of this study is to use the statistical model to generate new femur instances. To become a truly useful tool in the future, the created models must be realistic femurs and ideally be able to be used directly in FE so must have good mesh quality and require no manual correction or remeshing.

The equations defining the statistical model for both the covariance and correlation approaches are described in detail in section 3.4.5. Equation 3.9 describes how a femur instance can be constructed by the covariance approach, from the mean femur geometry plus the sum of a number of eigenvectors whose specific influence is controlled by a shape and intensity parameter vector b. The correlation approach reconstructs a geometry in a similar way although the standard deviations of the training set must by multiplied back in as an additional, final step. New, unique femurs can be generated from the model through the control of the two following factors; firstly by randomly varying the shape parameter, which is assumed to be randomly distributed, between upper and lower bounds b_u and b_l and secondly by including the m dominant eigenmodes. The value of b_u, b_l and m must be set specifically to the optimum values for the model. Investigations were required to find the number of modes needed, m, and the extent of the bound over which the model is sampled Y, such that the limits become $\pm Y\sigma_{b_m}$, where σ_{b_m} is the standard deviation of the individual shape parameters across the training set. If possible the model

should be sampled through ±3 standard deviations (as common in the literature and describing 99.7% of the distribution assuming it is Gaussian), unless the resulting shapes become unrealistic or element quality degrades. This would be a risk if the training set was not large enough to describe the variation satisfactorily.

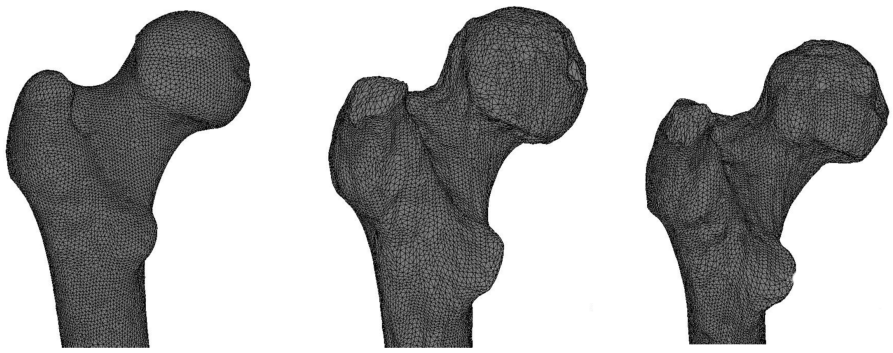


Figure 3.20: Figure shows the proximal portion of three generated femurs. A good quality mesh (left) is produced when the model bounds are set correctly, however mesh degeneration can occur (right) if this does not happen.

The importance of finding suitable bounds within which to sample the statistical model relates to the potential quality of the femur models it produces. The shape parameter bounds define how far from the mean the generated femurs will be, if stretched too far this could result in very unlikely configurations. Establishing the optimum number of modes is a balance of including all relevant information from the model without taking too much low level fluctuation, which will act as noise. The main result of this will be local distortions and poor mesh quality (Fig. 3.20).

Two approaches were taken to assess the requirements of the model; firstly by calculating the error in reconstructing the training examples with different numbers of eigenmodes and secondly looking at the variation seen in the models produced by changing the number of modes and sampling bounds.

3.6.2 Reconstruction Error

Reconstruction tests were performed to interrogate the model, where each of the training examples was generated from an increasing number of included eigenmodes by applying its known shape parameter weightings (Eq. 3.9). The tests investigated the robustness of the model in capturing the variability within the training set and provided a way of assessing the number of modes required to describe the majority of it. When sufficient modes are used the error in reproducing a given instance will fall to a low (or in this case where a small training set is used, acceptable)

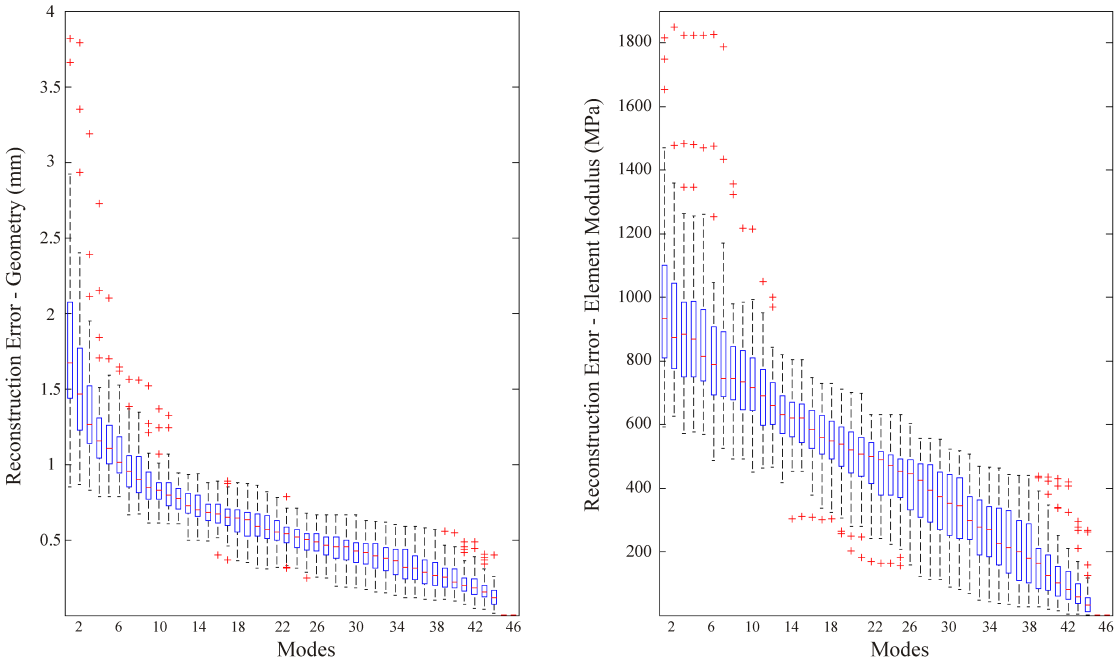


Figure 3.21: Reconstruction error of geometry (left) and elemental modulus (right) with increasing eigenmodes.

level. With this method of analysis the use of all modes will result in the target model being replicated exactly, however the highest modes are only contributing minor perturbations. For further applications of the model, such as reconstructing an unseen instance or generating new models, these higher order, low energy modes will simply add noise and distortion to the shape. The aim of the test is therefore to identify the least number of modes required to describe the highest proportion of variation, resulting in the lowest reconstruction error.

To fairly examine the results of the tests the geometric and modulus reconstruction data were analysed separately as they relate to different units of measurement. When purely geometric error was considered by comparing corresponding nodal positions throughout the volume of the mesh, a mean Euclidian distance error of $< 1mm$ was possible for geometry using the first 7 modes, which fell to $< 0.5mm$ by 26 modes. This was 50% of the total eigenmodes (Fig. 3.21). When elemental modulus reconstruction error was calculated using the same method the error decay was not as rapid as geometry. When 50% of the eigenmodes were incorporated the mean error fell below 500MPa (Fig. 3.21), yet when femur instances were visualised the material distribution and modulus values looked realistic. Therefore a further investigation was carried out to see whether the reconstruction error suggested was present or this method of calculation was unfairly harsh.

An analysis akin to that used to asses the volume morphing accuracy was run, examining the effect of an increasing number of included modes on the model's ability to predict strain distributions as compared to its original instance (i.e. the training set member being reconstructed). The modulus and strain distribution through the proximal femur were compared to the target femur being rebuilt to assess the ability of the model to capture the material variation with increasing modes. The results for a typical femur tested are shown in figure 3.22. The analysis concluded that if less than 35 modes were used then some features were lost, however the material property distribution remained realistic at all levels.

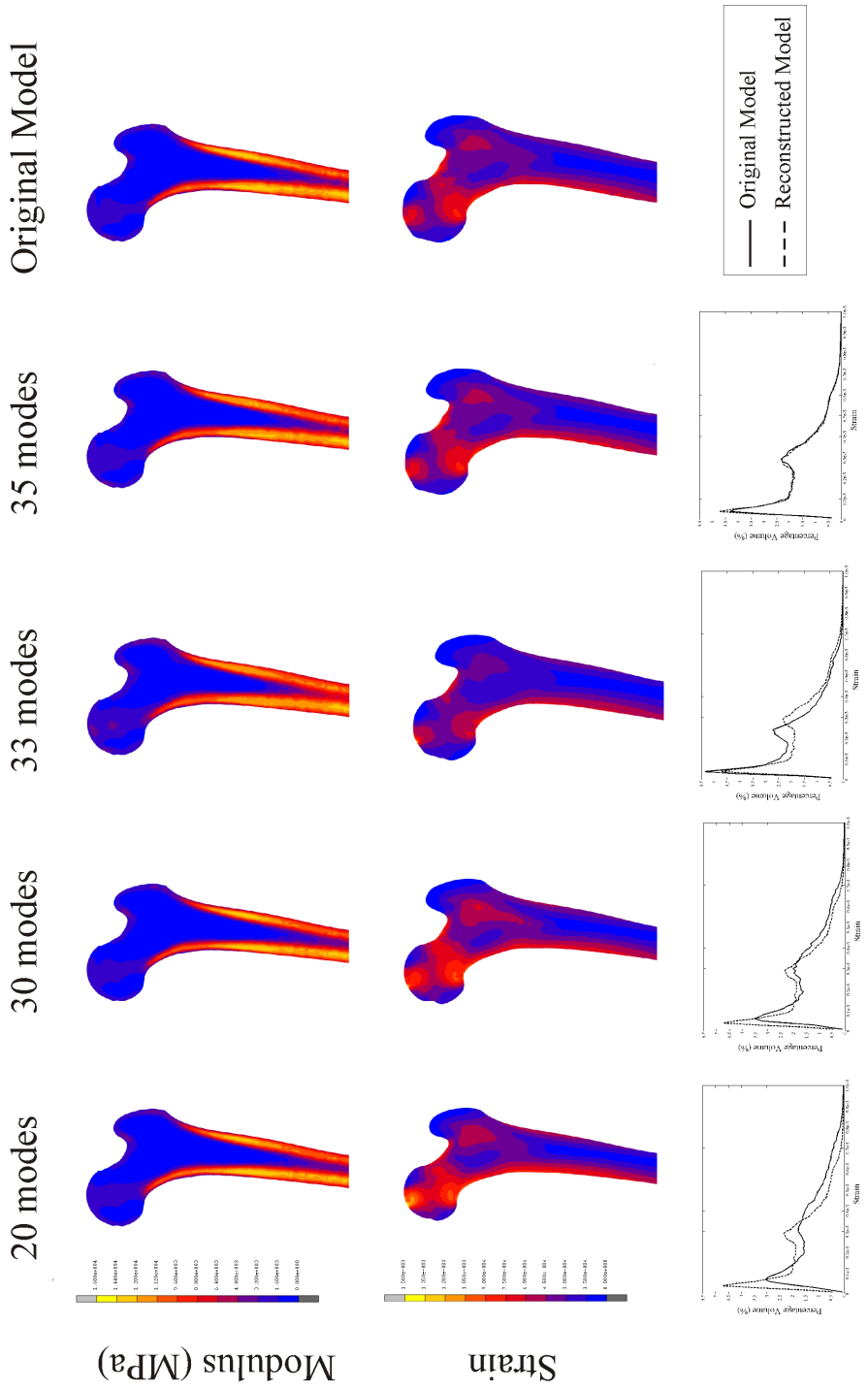


Figure 3.22: Plot of modulus (top) and strain (middle) distribution in a section through the proximal femur with 20, 30, 33 and 35 modes included and in the original femur being reconstructed (far right), following a stance loading FE simulation. All plots were made on a generic femur model so the geometric changes with modes were not included. Plots of percentage bone strain volume distribution through the whole femur are also shown (bottom).

3.6.3 Sampling the Model

The previous section has established an optimal number of modes to be included for this study (based on a 45 model training set). The remaining parameter to define is the sampling range along with the sampling technique. As previously discussed the sampling range is $\pm Y\sigma_{b_m}$. If this is chosen correctly then the generated models will exhibit a wide range of variation spanning that seen in the training set and possibly beyond. If too large a range is chosen then the new models could exhibit unrealistic/distorted features (a greater risk with a small training set) and mesh quality degradation, whereas too conservative an option will clip the variations possible. The standard sampling range used is ± 3 standard deviations of the training set mean, with the theory that if the data distribution is Gaussian then this will cover 99.7% of the variation in the data. This is an assumption, so to ensure that the majority of the data is represented a plot of the shape parameter values for each mode was created with the bounds of $\pm$3s.d. (Fig. 3.23). The figure shows that almost all the training data lies within this boundary so the sampling limit is not artificially cropping the information.

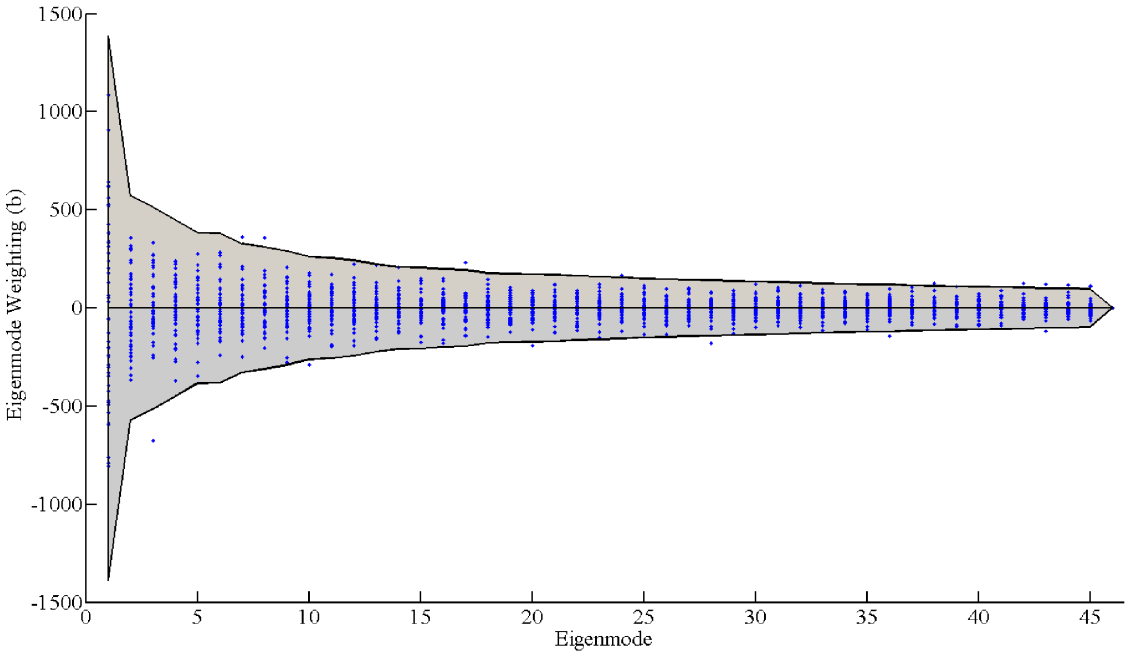


Figure 3.23: Figure showing the eigenvalue weightings of the training set (blue dots) superimposed onto the sampling space created by a $\pm$3s.d. range (grey area).

The next concern with the sampling limit is that the mesh quality does not become degraded. As discussed in a previous section it should be remembered that mesh degradation will also be affected by the number of included modes. To examine the influence of this, NSR and SDR scores of femurs reconstructed using an increasing number of modes with shape parameters 'cropped' to ±3s.d. were calculated (Fig. 3.24). The previous figure illustrates that few points fall outside this range, but those that do will be pulled back to the ±3 s.d limits. The results show that at the highest modes the mesh quality of the outlying results decreases progressively, however the mean qualities at the selected 35 modes are high for both scores.

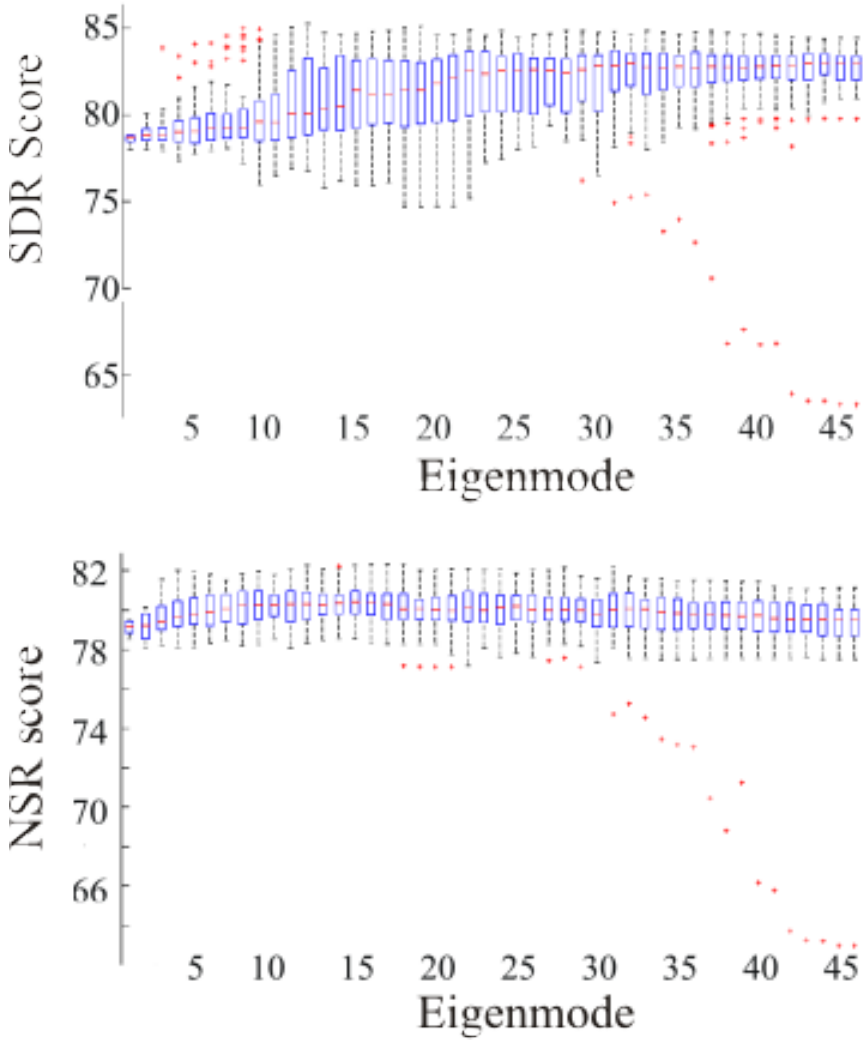


Figure 3.24: Boxplots of the NSR and SDR scores for reconstructed training set femurs with increasing numbers of eigenmodes, using a set of shape parameters 'cropped' at ±3s.d.

The choice of sampling distribution is key to the range of models which are generated. A starting point is examining the distributions existing in the training data set, namely the distributions of the shape parameters or eigenvalue weightings, b. The

first five, which have the most significant influence and together account for nearly 70% of the variation in the model, are plotted in figure 3.25. These show that the distributions are all different. Modes 1 and 2 (blue and red) appear to be bimodal but are relatively flat, whereas modes 3 and 5 show singular peaks but these are offset (as an aside, the bimodal peaks in mode 1, which represents scaling, could relate to gender differences).

In order to truly represent the training data each of these distributions would need to be modelled and sampled according to their own pattern, advanced techniques are available to do this if it is desired [105]. However in this work the aim is not to reproduce a specific data set but to generate a wide range of varied femur examples, ensuring that they remain realistic. Therefore two simple sampling distributions approaches were explored, firstly a Gaussian distribution, which is a reasonably approximation of many modes, and secondly a uniform distribution, which best matches the first, most dominant mode. The most appropriate way to examine the suitability of these approaches was to look at anatomically meaningful measurements, this is the focus of the next section.

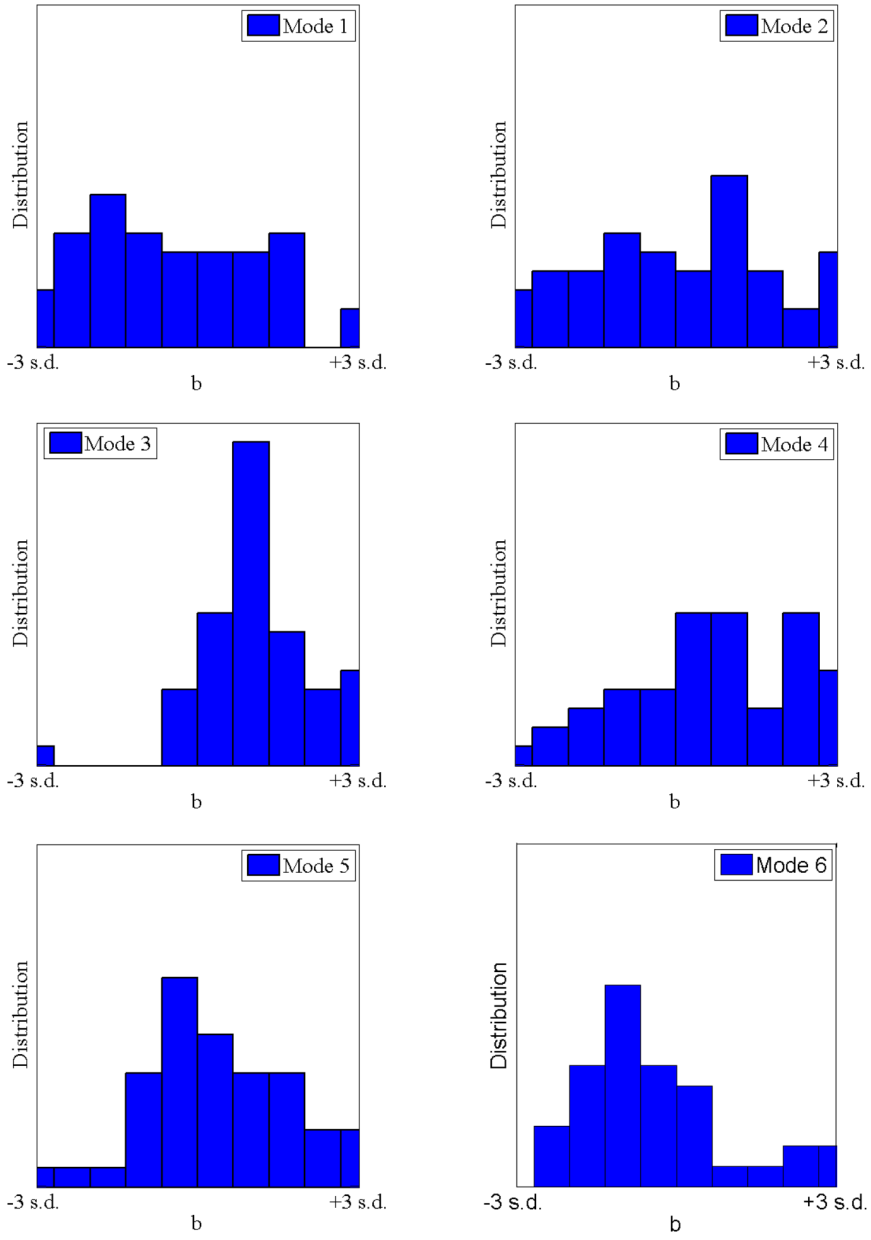


Figure 3.25: Histogram illustration of shape of the distributions of the first six eigenmodes, which together account for over 70% of the training set variation.

3.6.4 Characteristics of Generated, 'Synthetic' Femurs

Reconstruction tests are an excellent mathematical test of a statistical model, providing details of how accurate the model is in reproducing a specific instance. The prospective application of the model in this project is not, however, to regenerate specific geometries but to generate new instances with realistic characteristics. Future studies will aim to exploit the ability of the statistical model to generate new femur instances. This will only be a useful tool if the femurs produced are unique and realistic. Whether these criteria are met is dependent on the method by which the

model is sampled. If these criteria are wrongly set then the generated femurs could end up stretched too far and become unrealistic. Alternatively the created models could vary very little from instance to instance therefore only represent a small subset of a population. In view of this, it seems logical to investigate anatomically meaningful characteristics of the models which can be generated and compare these to both the training set data and the wider population.

The statistical model contains both geometric and material information therefore it was important to examine both of these forms of variation. This was performed using semi-automated and automated checks to ensure the validity of the models. The semi-automated assessments involved the generation of screen shots of all the generated models, followed by manual inspection of the images. These showed that non anomalous instances were produced. The registration method allowed a for an automated method of comparing models as each node and element lie in the same relative position in the bone in every instance - whether from a generated or training set. A range of geometric and material metrics were defined on the baseline model by identifying key nodes and element volumes which related to standard femur measurements and areas of clinical interest [103, 106].

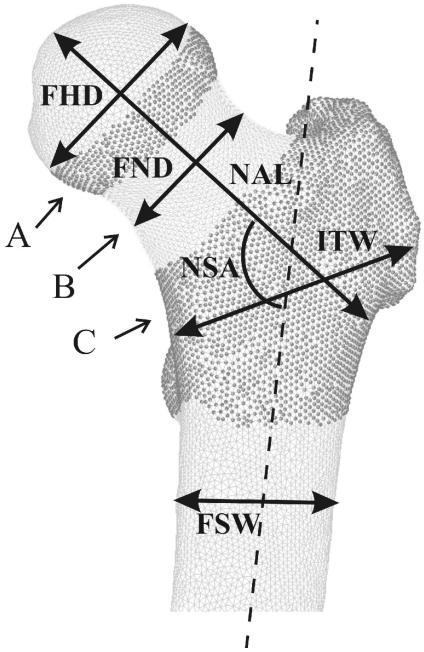


Figure 3.26: Illustration of the key proximal geometric measurements extracted from each model and the three proximal volumes examined for material property characteristics.

The following geometric measurements were automatically extracted and calculated, based on known node positions: neck axis length (NAL), neck shaft angle (NSA), femoral head diameter (FHD), femoral neck diameter (in both anterior-posterior, FND-AP, and proximal-distal, FND-PD, directions), anteversion angle (AA), femoral shaft width (FSW, measured ∼3cm below the lesser trochanter), intertrochantic width (ITW) and femur length (FL). To examine bone quality three key proximal sections were defined: lower femoral head (A), femoral neck (B) and intertrochanteric area (C) (Fig. 4.11). Comparative bone quality was assessed by the cortical element percentage (where cortical bone was defined as bone >3000MPa) and by average element modulus in each section.

To compare the alternative sampling techniques both the Gaussian and uniform distributions were used to generate 1000 femur instances, the results of which are shown in Table 3.3 alongside the data for the original training set. Both methods match the mean geometric parameters very well, but there is an underestimation in material modulus by the uniform approach. The range or spread of outputs, calculated as the maximum measurement minus the minimum, was less well matched to the training data. The Gaussian method consistently extends beyond the original data however the uniform method consistently underestimates the variation. The range in shapes present in the base 46 models is not completely reflected in the 1000 models generated by the uniform approach, for this reason the Gaussian approach seems better suited to this particular set of shape parameters in this model.

A key observation of the models generated is that they all appear realistic on visual inspection and none failed the automated pass-fail element quality checks, therefore all are potentially able to be used directly in FE. However, limited observations can be made by comparing generated examples to the original set when, as discussed, the distribution of each shape parameter would need to be modelled to reproduce the set closely. A more suitable examination was to compare the data to a true population,

Metric	Training Data		Gaussian Sampling		Uniform Sampling	
	mean	spread	mean	spread	mean	spread
Neck Axis Length (mm)	101.3	35.7	101.3	43.2	100.6	11.2
Neck Shaft Angle (deg)	127.0	10.4	127.2	13.3	127.5	3.8
Femoral Head Dia. (mm)	48.0	16.5	48.0	24.9	46.9	6.8
Femoral Neck Dia. A-P (mm)	34.7	17.2	34.7	24.4	35.2	6.9
Femoral Neck Dia. P-D (mm)	33.8	16.5	33.8	26.6	34.1	6.8
Anteversion Angle (deg)	18.7	37.7	18.4	38.4	20.2	10.0
Femoral Shaft Width (mm)	16.4	6.0	16.3	8.9	16.4	2.1
Intertrochantic Width (mm)	56.0	18.7	56.0	28.7	56.3	7.9
Femur Length (mm)	436.3	115.3	462.9	211.9	452.6	38.4
Stature (mm)	1630.9	431.0	1730.3	792.0	1691.9	143.5
Fem. Head Cort Element (%)	3.2	30.2	1.8	23.5	0.1	1.2
Fem. Head Mean Mod (MPa)	1317.5	2005.8	1318.5	2142.8	1295.0	649.5
Fem. Neck Cort Element (%)	28.7	59.3	29.7	62.1	19.1	21.5
Fem. Neck Mean Mod (MPa)	2566.6	3230.7	2579.6	3932.7	2108.7	1232.9
Intertroc. Cort Element (%)	33.2	49.0	34.7	53.7	28.2	17.4
Intertroc. Mean Mod (MPa)	2877.9	3276.4	2879.6	3889.6	2439.8	1045.7

Table 3.3: Table of geometric and material metrics for 1000 femurs generated by Gaussian sampling, 1000 by uniform sampling and for the original 46 training femurs. Both sampling methods used the first 35 modes and a range of ± 3 standard deviations of the mean. The table shows the mean and spread (max-min) results for each of these groups.

the U.S. National Health and Nutrition Examination Survey (NHANES) was used to do this [107]. The survey published hip geometry and anthropometric data for 13,615 subject X-Rays, covering a population of 6942 women and 6673 men ranging from 20 to 90 years old. Several reported parameters were directly comparable to the measurements extracted from the generated models, namely FND-PD, ITW and FSW. These were suitable as they have only a limited three dimensional component unlike neck shaft angle, for example. As the training data was in the older half of the age group an age-matched subset of the data was used, scatter plots of these variables from with the Gaussian models superimposed are shown in figure 3.27. The figure also shows a comparison of standing height in the NHANES population and the generated set. The generated data can be seen to sit over the majority of the survey data, although it does appear slightly larger in general. This size difference is evidence by the slight mismatch in the the neck v's shaft diameter and the the height distribution plots. This is likely related to the ethnic discrepancies between the training and NHANES data sets.

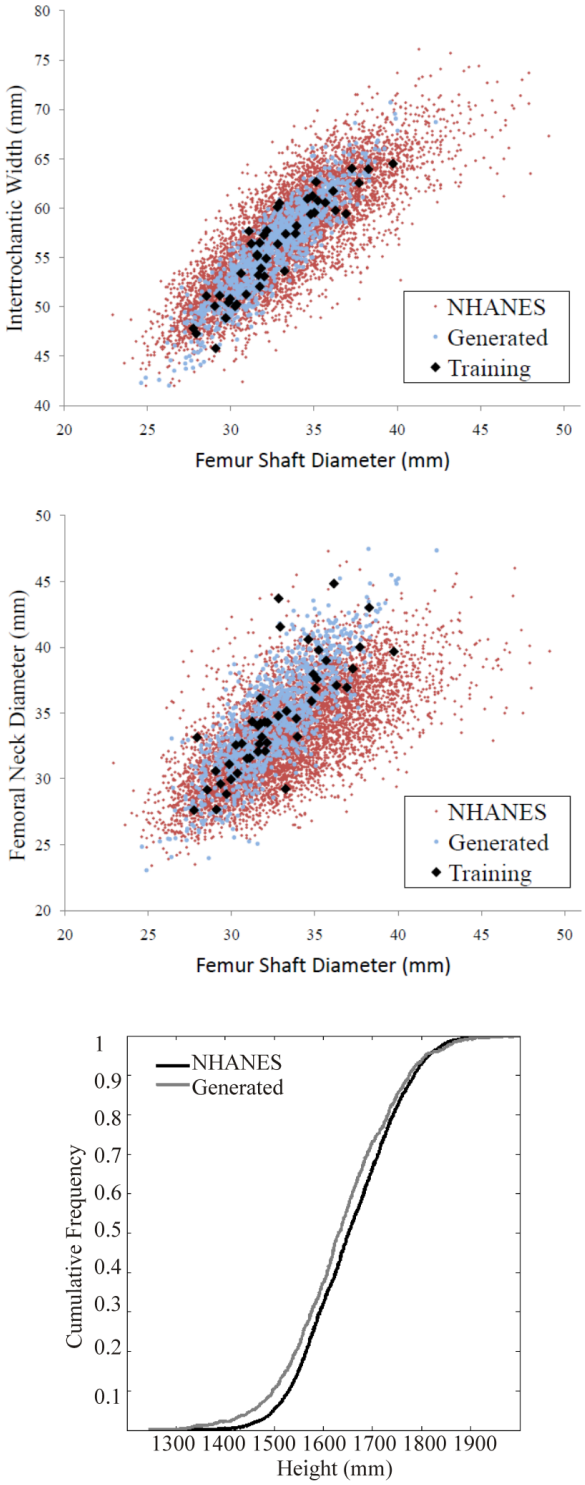


Figure 3.27: Plots comparing statistics taken from the statistical model and training set to data from the National Health and Nutrition Examination Survey (NHANES) findings 13,615 subjects [107]. Plots compare an age matched subgroup of the NHANES population containing 5803 subjects and a generated femur population of 1000 models .Top: scatter of femoral shaft diameter against intertrochantic width. Center: scatter of femoral shaft diameter against femoral neck diameter. Lower: Plot of standing height cumulative distribution between (heights in mm).

Another important question which can be answered by comparing the generated models to the original training set is whether the statistical modelling approach is generating new, unique instances. i.e. providing combinations of geometric and/or material properties which are not present in the training set but are still realistic. Figure 3.28 shows scatter plots of some of the exacted metrics, the combinations of plots were chosen specifically to interrogate the performance of the statistical model in capturing and reproducing the femur. The top row contains plots where the metrics should logically have some correlation to each other, the first concerns geometric features and the second material metrics. The geometric features of neck diameter and head diameter should be related, i.e. one should not change dramatically without the other also being similarly affected. The same logic applies to the material modulus in the neck and intertrochantic regions. The lower row of plots again contain both geometric and material data, however metrics which are not directly related to each other were chosen.

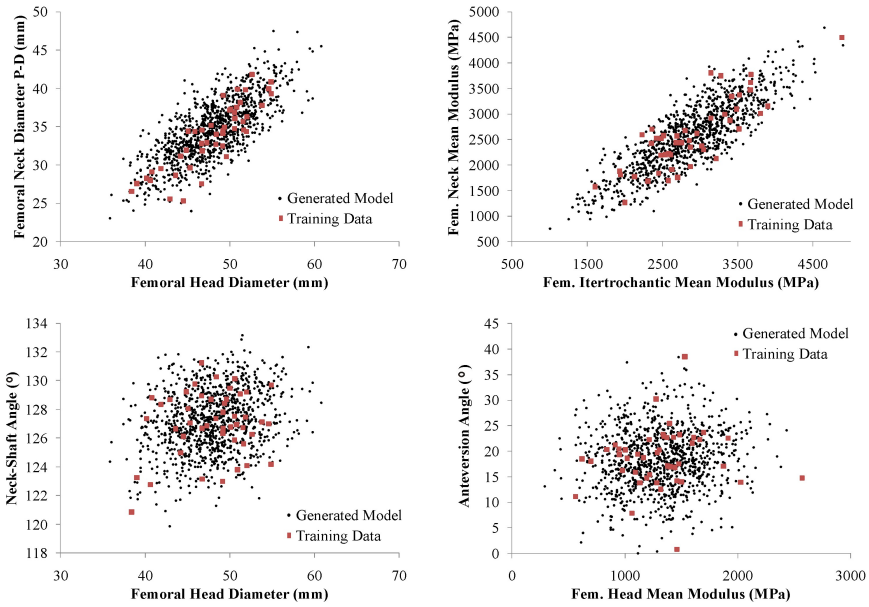


Figure 3.28: Scatter graphs comparing the ranges of various geometric and material metrics seen in the 1000 Gaussian generated data set (black dots) in relation to the original 46 training femurs (red squares).

The plots show that the model generation approach adopted samples the range of parameters in the training set and interpolates between them, allowing combinations of geometries/material qualities which are not seen in the original data. There is also evidence that the methods allows some extrapolation beyond the original data limits, which is controlled by the sampling parameters chosen. Reassuringly, the interpolation/extrapolation retains the relationships defined within the training set, e.g. measurements which should be proportional remain related (Figure 3.28 top),

whereas when such a relationship is not present the entire space is sampled (Figure 3.28 lower).

In 2D plots it is easy to consider the observed metrics in isolation or in pairs. It may be logical to think that a parametric model could be used to produce the same sort of relational changes in shape or material, for example changing femoral head diameter size in isolation to examine the effect this has. However, the femur is a complex geometry and the addition of material information adds to the problem, therefore what is not well expressed by the previous plots is that changing one parameter will have a knock on effect to several others. Using changes in FHD as an illustration, figure 3.28 shows that for a given head diameter there is a band of allowable neck diameters. Extending this to other parameters, figure 3.29 shows that with increasing FHD there are related allowable ranges of NAL and ITW which would also have to be adjusted to make the generated geometries realistic. Plausible material properties, on the other hand, are not so narrowly correlated and so a more random method of parametrically altering these between limits would need to be devised. The problem of representing a realistic femur is therefore not straightforward, incorporating many different changes which need to be considered together. From the distributions of data shown in these scatter plots it can be seen that the statistical modelling approach used is able to account for this and allow for both correlated and uncorrelated variations.

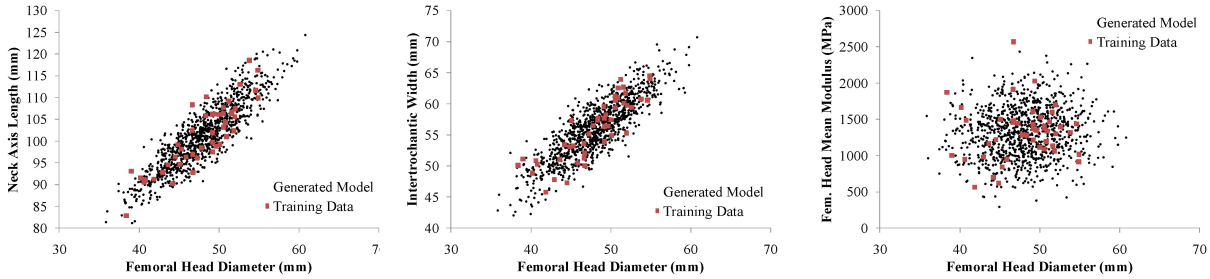


Figure 3.29: Scatter graphs comparing the ranges of various geometric and material metrics seen in the 1000 Gaussian generated data set (black dots) in relation to the original 46 training femurs (red squares), examining the changes in NAL, ITW and mean head modulus with changes in FHD.

3.7 Discussion

The target future use of this technique was to be able to generate large numbers of new and unique femur models in an FE analysis ready format. This would enable larger combinations of realistic variations in femoral anatomy and material characteristics to be used in computational studies than is currently feasible. To examine this possibility a statistical model, trained on 46 subjects, was generated using PCA and sampled to generate synthetic femurs. The first 35 eigenvectors were used for this with random, Gaussian perturbations between ± 3 standard deviations of the mean eigenvalue weightings indicated by the training set. To test whether the generated femurs related to a real population, thus could be useful in a later large scale analysis, comparisons were made between several metrics taken from the NHANES survey of 13,615 Americans and 1000 generated models. This indicated that the generated population did show similar trends and limits. Cumulative distribution of standing heights indicated that the generated set was slightly taller than the U.S. population. Some differences were to be expected and reflect the range of femurs in the training data and the sampling method. Plots of the distribution of the eigen weightings in the training data indicate that the trend is not Gaussian at 46 examples, this is particularly true of the first mode which is the most significant and seems to have the most influence on femoral length. The correlation between the generated and real populations could be made closer if each weighting distribution were modelled individually and these used to sample the model. Such techniques are available and should be considered for future studies if deemed necessary[105].

A main feature of the statistical model discussed in this chapter when compared to existing methods is its use of a fine mesh density to allow capture of geometric and material property distributions, and produce an 'analysis ready' FE mesh [15]. The subsequent increase in model complexity led to two major developments being made to the elastic surface matching registration scheme of Moshfeghi et al. [91], the incorporation of k-d trees and Laplacian smoothing. The use of k-d trees [92] accelerated nearest neighbour searches during each iteration of the elastic surface matching algorithm. This reduced the complexity of the original elastic surface registration scheme significantly at each iteration, resulting in calculation speed increasing by orders of magnitude. It was verified that this improvement was achieved without any adverse impact on the accuracy of the registration algorithm. Laplacian smoothing was also introduced at each iteration to try to prevent mesh quality degradation through the matching process and reduce the risk of mesh folding and bunching which is known to potentially occur in areas of concavity [91]. This dense mesh could be considered

a limitation for some future applications, however for this work it was easily manageable and a future detailed study of mesh density could be done to investigate whether the mesh could be coarsened.

Surface registration was performed within 40 minutes on an Intel Xeon X5365 processor. By the scoring system defined in Section 3.5.2, the mean surface mesh quality score of the registered models dropped by 5% from the baseline, however the quality remained over 89/100 in all cases with a mean of just <0.2% of elements in the lowest quality quartile. The inclusion of Laplacian smoothing had the potential to reduce registration accuracy, so an investigation of the mean distance between each registered point and target surface was run. This showed a mean error of <0.6mm, with >99% of nodes under 1.5mm from the target. This level of error was within the resolution of the CT's used, 0.488x0.488x1.5mm to 0.781x0.781x2mm, and comparable to the 1-2 voxel error associated with manual segmentation [73].

Volume morphing was assessed by comparing registered-morphed meshes to individually meshed instances of the same initial geometries, with comparable element sizes. Both the morphed and ideally meshed instances were then identically loaded to simulate single legged stance and the material and strain field distributions through them compared. Figure 3.18 shows very little difference between the two examples. This, together with the morphed model quality metrics remaining high, added confidence in the combined success of the surface registration and volume morphing methods as well as the final ability of the statistical model to generate meshes of sufficient mesh quality to allow direct use in FEA.

The training data was dominated by men and the older population, with a mean age of 70 years. A large degree of variation was present within the training set. No pre-selection was applied as whole femur CT data was hard to source. Ideally, in future, sufficient scans would be available so that subdivision could be applied on the basis of age, sex and ethnicity, which all have affects on femur geometry and material properties. As the femur has a complex form and distribution of material variation, the training set size will always be a limitation. This is emphasised by the relatively large proportion of eigenmodes required to reproduce shape and modulus (Figure 3.22), showing meaningful variations are contained in up to 35 of the 46 modes. At 35 modes >95% of the variation in the training set is captured, the remaining 5% therefore likely contains noise and minor fluctuations. These could have been introduced due to segmentation errors or low level mesh degradation which, by using only 35 modes, has been filtered out.

Reconstruction tests are a conventional method for assessing a statistical model by examining its ability to reproduce each member of the training set, the fewer modes

required to do so accurately the more correlated the data. As two different qualities were being captured by the model, geometry and modulus, their reconstruction errors were calculated independently (Figure 3.21). This revealed a much higher correlation within geometry where mean error falls below 1mm after just 7 modes, and becomes equivalent to registration error at 20 modes. Modulus error was reported to be poor by this test, with errors >500MPa at 20 modes. However from the plots of the influence of the first 3 eigenmodes on shape and modulus, figure 3.19, the overall distribution of material property seemed realistic. This theory was supported by comparing the strain results, following loading, of original femurs to instances of that femur constructed with increasing numbers of modes (Figure 3.22). At 35 modes the strain fields produced were comparable, whereas reconstruction error alone indicated an error between 37-464 MPa.

The reasons for noise in the modulus data which resulted in high reconstruction error and the need for a large proportion of modes are likely due to the nature of the data and the methods used to extract it. CT data is inherently noisy, with voxalisation causing partial volume effects where materials with dissimilar densities are close together. This would occur on the femur surface between cortical bone and soft tissue and within the bone where there are areas of rapid transition between cortical and cancellous bone. The CT data used in this study were from clinical sources and without calibration phantoms. Calibrations for the proportional relationship between greylevel and apparent bone density was defined individually for each set by identifying the greylevel of marrow in the medullary canal and the highest greylevel visible in the femur itself. Marrow was considered to be equivalent to water and given a density of $0g/cm^3$ and the maximum greylevel was assumed to be compact cortical bone at $1.73g/cm^3$ (these values relate the the density of mineralised tissue) [108]. This calibration technique is a potential source of error in modulus data although every effort was made to reduce possible human error. The use of quantitative-CT data in future would be a solution if it were available.

3.8 Conclusions

This chapter has discussed a technique for the construction of a highly detailed statistical model, which captures geometry and material property variations within a training set of data. The elastic registration and mesh morphing schemes developed enabled a solid tetrahedral model to be constructed from a large number of points, resulting in highly accurate registration of the training data and the direct production of a mesh suitable for finite element analysis. A larger number of training examples would help to improve the model's ability to describe a wider range of the population and perform patient specific reconstruction, although this is not what the model is intended to be used for and alternative approaches such as intensity based registration may prove more appropriate [12, 109]. The results presented indicate that the model is capable of capturing key anatomical features, with distributions of results comparable to the training data and a true population set even with simplified sampling methods. The significance of this is that it provides a method of generating a large number of femurs for analysis, allowing the identification and investigation of the ways in which they vary, making running large scale, multi model FE analyses feasible. This may prove that more comprehensive preclinical computational testing is possible and could, in the long term, lead to the generation of a tool for clinical assessment with patient specific potential applications [66, 52, 110].

Chapter 4

Femoral Neck Fracture Risk Study [1]

4.1 Femoral Neck Fracture Risk

Hip fracture is one of the major health problems facing our increasingly ageing population at the current time. This is not only due to the traumatic and acutely painful nature of the injury but also its related effects of reduced mobility, long term disability, reduced capacity to live independently and morbidity [1]. From a socio-economic perspective, hip fracture has a substantial final cost thanks to the common requirement for long term care, continuing treatment and rehabilitation [111]. Evidence suggests that the number of age-specific fractures is increasing [112], combining this with globally increasing life expectancy and associated ageing populations [113], it is clear that hip fracture has the potential to provide a huge social and economic burden.

[1] Bryan, Nair and Taylor, 2009. Use of a statistical model of the whole femur in a large scale, multi-model study of femoral neck fracture. Journal of Biomechanics 42(13), 2171-2176 - based on the work in this chapter

4.1.1 Aim of Femoral Neck Fracture Risk Study

The aim of the study described in this section is to apply the statistical model of the femur developed in Chapter 3 to the problem of proximal femoral fracture. Due to the limitations of the model it will not necessarily be possible to draw conclusions which relate to the entire population. However, it is hoped that by following the examples of previous analyses done in this area, the model will be able to produce meaningful results. *Meaningful*, being that the results will show comparable trends to previous studies, particularly when suggesting characteristics of femoral geometry or bone quality which may indicate an increased risk of suffering femoral fracture. The study described in this chapter hopes to illustrate the potential of using statistical modelling techniques in analysis and to demonstrate how this could be achieved.

4.1.2 Description of Statistical Model Used in Study

The study described in this chapter was performed using a statistical model of the whole femur built using the methodology described in the previous chapter. However, the work was done before the final iteration of the model was complete therefore it was built from a training set of 21 subjects (as opposed to the final 46). As a result of this the characteristics of the model differ from those previously described, as do the sampling methods and parameters applied. Characteristics of the model used in the femoral neck fracture risk study are available in Appendix B.

4.1.3 Description of Hip Fracture

Proximal femoral fractures are most basically described as either intra- or extra- capsular, referring to whether the fracture occurs within or outside the hip's joint capsule. This capsule is a synovial membrane which encases the articulating joint, constructed from two sets of fibres arranged in circular and longitudinal directions, it connects to the acetabulum and to the proximal femur just above the intertrochanteric line (Fig. 4.1) [16].

Intracapsular fractures are defined as either femoral neck or femoral head fractures. Femoral neck fractures are further subcategorised as subcapital, cervical or basicervical (Fig. 4.1). Femoral head fractures are rare and usually only the result of high energy trauma. Extracapsular fractures are broadly defined as subtrochantic or

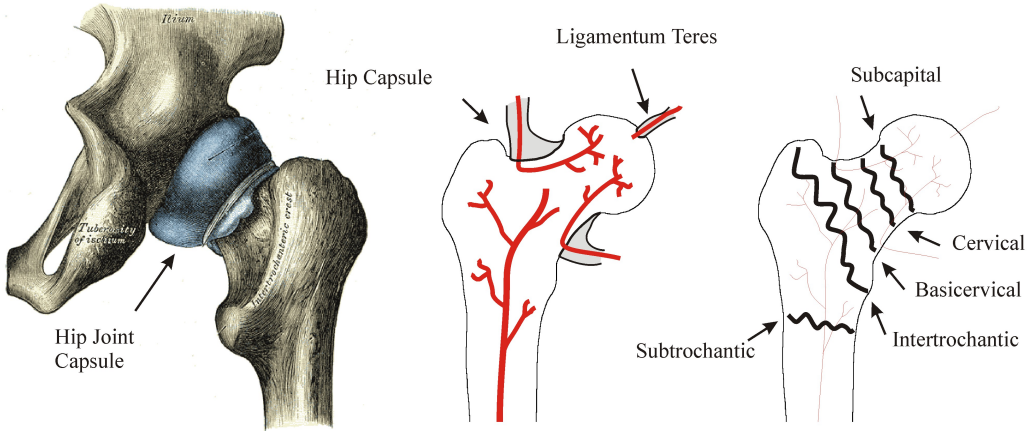


Figure 4.1: Illustration of hip joint capsule (left) [16], major blood supply paths in the proximal femur (center) and the main fracture types seen in the proximal femur (right).

Location	Percentage
Intertrochantic	49
Intracapsular	37
Subtrochanctic	14

Table 4.1: Table illustrating the clinically seen proportion of fracture locations [115].

intertrochanteric, the later being the most commonly seen location, accounting for about half of the fractures seen (Table 4.1) [114, 115].

Fracture location is significant as it defines the treatment course taken and seriously influences the chances of recovery. Figure 4.1 illustrates the major blood supply paths to the proximal femur, if the blood supply is disrupted of cut off as a result of fracture then areas of bone can suffer avascular necrosis and die. A good blood supply is also essential for bone to heal. The diagram shows how intracapsular fractures, particularly subcapital, can be highly disruptive to blood supply to the femoral head, if the fracture is severe and displaced then there is a low chance of successful healing. This scenario is likely to result in a hemiarthroplasty or full joint replacement. Intertrochanteric fractures are less disruptive to blood supply so there is more chance of natural healing and typical treatment would be internal fixation with screws [1]. However, patients suffering intertrochantic fractures have been reported to have a poorer short term outlook, having a higher rate of mortality at 2-6 months and are slower to regain pre-fracture activity levels than patients with femoral neck fractures [116]. This may be partly explained by the demographic who seem more prone to intertrochantic fracture, being older and in poorer health.

4.1.4 Risk Factors for Hip Fracture

Hip fractures most frequently occur in the elderly, and it is reported that 90% of the time they follow a fall [46, 117, 118]. The most dangerous type of fall has been one which results in an impact to the lateral aspect of the proximal femur, i.e. the greater trochanter, or the side of the leg [119, 120, 121]. Logically the velocity of the fall and the amount of 'cushioning', either by soft tissue, the surface landed on or by protective measures such as hip protectors, will affect the impact energy and therefore the likelihood of bone fracture [122]. There is some evidence to suggest that fracture may result from sudden, powerful muscle contraction immediately before impact or even as a reaction to beginning to fall [123]. However Hayes et al. [119] saw that even falls from a standing height or less are no minor trauma to elderly patients. Through cadaveric testing, it was shown that the impact energy relating to a fall was more than an order of magnitude greater than that required to fracture elderly bone. Worryingly, it is estimated that up to 80% of all elderly people over the age of 80 suffer at least one fall per year [124].

Muscle strength and body size have been suggested as possible indicators of fracture risk. Their effect seems, however, to most likely be significant only in their interaction with the risk of falling. The energy of a fall will be greater for a taller and/or heavier person, simply because there is further to fall and a greater mass [125]. In relation to body mass, both high and low body masses can be problematic. A high body mass may be a problem as reduced mobility, proportionally low muscle mass and strength and the significant load applied by their size may be sufficient to exceed the femur's fracture load, despite the positive effects of body fat dissipating the impact force [1]. In contrast, underweight elderly people may have insufficient soft tissue to absorb the energy of a fall, and associated muscle mass loss may make falling more likely [125]. Similarly, other related factors which have the affect of either making a fall more likely or heavier are: decreased muscle strength, inactivity (affecting muscle strength and coordination), environmental hazards (such as ice or slippery floors), impaired cognition or perception, poor vision and neurological disorders [1].

With time, bone mass is lost and the collagen within bone degrades, reducing the bones elasticity, and increasing bone brittleness. Low bone mineral density and bone mass, often the result of osteoporosis and natural age related bone loss, have long been logically related to increased fracture risk as bone mass and mineral content is directly related to bone modulus [108]. However in hip fracture, despite sufferers having below average bone mass for middle aged subjects, control studies have shown

that there is significant overlap between non-fractured and fractured subjects of the same age and gender groups [126]. Due the greater natural rate of bone mass loss and increased propensity to osteoporosis, women are at significantly higher risk of hip fracture (Fig. 4.2). It has been estimated that women over the age of 50 are 2-4 times more likely to suffer hip fracture than men [111, 127, 128, 129]. The risk continues to increase with age; at 50 the percentage of women at risk is 17.5% compared to 6% for men [129], by the age of 90 this has increased to 25% for women and 16% for men [124].

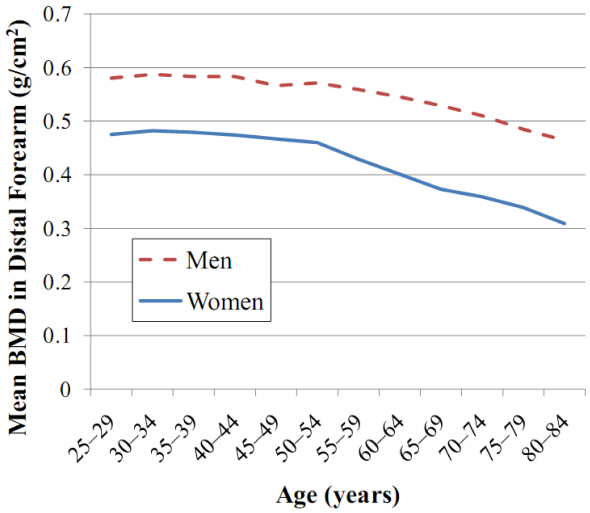


Figure 4.2: Graph illustrating the difference in both overall bone mass and rate of bone mass loss in men and women. The data is taken from a study of 3062 men and 4558 women, measuring BMD in the distal forearm using x-rays [130].

4.1.5 Outcomes of Hip Fracture

The immediate reaction to hip fracture is extreme pain, leaving the victim unable to walk or stand. As discussed above, depending on the fracture, the blood supply to areas of bone can be cut off which will lead to avascular necrosis if left untreated. Hip fracture will require hospital attention for surgical treatment ranging from reduction and internal fixation through to hemiarthroplasty and full joint replacement, although for non-elderly, non-displaced intracapsular fracture a long period of bed rest may be sufficient [1, 131, 132].

The immediate outlook for a hip fracture patient over 50 is uncertain (Fig. 4.3a). Statistics range between 15-25% mortality within the first six months after the injury, rising to as high as 30-40% after 1 year, depending on the location and severity of the fracture and the age and pre-fracture health of the patient [111, 124, 132, 133].

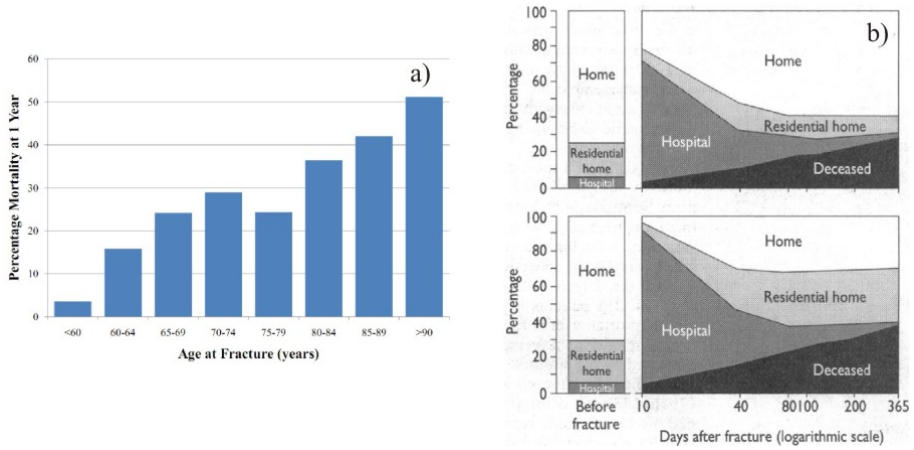


Figure 4.3: (a) Graph of the mortality rate for different age groups following hip fracture of 1000 patients admitted to hospital. (b) Figures illustrating the residency of patients who suffered from intracapsular (top) and extracapsular fracture (bottom) in the year following fracture [132].

The patient's risk of dying within the first 6 months is seen to increase significantly if they are male and/or over 75 [134, 135]. Having survived the first year, a significant number of patients, 20-30%, will have been moved to long term care facilities for the first time and only around 45% are expected to be discharged home from hospital [132, 111, 135](Fig. 4.3b). Recovery rates are also reported to be far from ideal. After 1 year, Cooper et al. [135] describe 80% of patients as unable to perform activities that were previously possible, with 60% finding at least one essential daily task difficult unaided. Most worryingly, they claim that 40% are still incapable of walking unaided, a figure supported by Keene et al. [132]. Not only can recovery prospects be poor and slow, the effect of this vastly raises the likelihood of a second or even third hip fracture [1]. At 1 year, Karlsson et al. [136] showed patients to have a 7% reduction in BMD in the fractured hip, 5% reduction in lean body mass and an 11% increase in body fat, all undoubtedly not helped by enforced inactivity.

4.1.6 Occurrence Rate, Cost and Projections

It was estimated that the number of hip fractures happening globally per year was around 1.7 million in the mid 1990's, with around 250-300,000 per year in the U.S. and over 50,000 per year in the UK [1, 111, 127]. As previously discussed, hip fractures are most likely in the elderly who have weakened bone strength and increased propensity to fall. These numbers are projected to rise substantially. The injury is only really seen in younger patients following serious trauma, such as a car crash [134]. A major factor behind this is that the proportion of old and elderly in the population is increasing, particularly in the developing world. Globally the percentage of people

over 60 has increased from 8% in 1950 to 10% in 2000, and is predicted to hit 21% by 2050, with the fastest growing age group being the 'oldest old', over 80 years, who will be majority female and are set to make up 1/5th of the over 60's by 2050 [113]. The increasing elderly population, with the 'oldest old' making up a significant proportion, and the growing number of very elderly women is very worrying as these are the highest risk groups for suffering hip fracture.

Year	A		B		C		D	
	Men	Women	Men	Women	Men	Women	Men	Women
1990	338	917	338	917	338	917	338	917
2000	417	1086	460	1200	483	1242	520	1332
2010	553	1321	615	1611	727	1760	854	2049
2025	783	1821	1109	2580	1386	3102	1976	4145
2050	1381	3112	2509	5653	3905	8430	6794	14516

Table 4.2: Table showing projected numbers of hip fractures globally (000) per year. A - Assumes unchanging age and sex specific incidence B - Assumes a 1% increase in age and sex specific incidence worldwide, C - Assumes no secular change in the US and N. Europe but a 2% increase in age and sex specific incidence elsewhere and D - Assumes no secular changes in US and N. Europe but 3% increase in age and sex specific incidence elsewhere [137].

Gullenburg et al. [137] investigated the incidence of hip fracture in each region of the world and combined this with global population increase projections. Several calculations were performed; firstly simply scaling by predicted population growth and then by several other approaches where the increasing proportion of elderly and females was included by varying amounts (Table 4.2). The results showed that by 2050 the number of hip fractures will reach 4.5 million purely by population growth and there could be many as 21.3 million globally, if only a conservative 3% increase in age and gender specific rates is seen. There have been many more localised estimates and observations for how hip fracture incidence will and has risen in different countries (Table 4.3).

Beyond the individual suffering caused by hip fracture there is also a large economic cost. In 1998 Dolan et al. [111] estimated the costs of osteoporotic fractures in the UK and found that hip fracture was not only the most common fracture in the over 50's, accounting for 87%, but also the most expensive to treat. The conservative estimate of the cost of a typical hip fracture was £4808 of acute care, directly received hospital treatment and surgery, along with £7152 of non-acute care, including social care, rehabilitation and long stay hospital costs. The total estimated cost to the UK was £942 million per year, which when inflation is corrected for is just over £1,100 million in 2008. This bill seems only set to rise in light of the global trends for an increasing elderly population.

Country	Comments	Source
U.S.A	512,000 fractures per year by 2040 (current estimates 250-300,000 [1])	Cummings et al. [138]
U.K.	60-117,000 fractures per year by 2015 (current estimates 50,000 [111])	Gullberg et al. [137]
China	Study in Beijing showed 33% increase in women and 33% increase in men between 1988-92.	Ling et al. [139]
Finland	3x increase between 1997-2030 based on trends observed between 1970 and 1997.	Kannus et al. [140]
Belgium	7x increase between 2003-2050 based on trends observed between 1984-1996.	Reginster et al. [141]
Greece	81% increase in recorded hip fracture between 1978 and 1992.	Paspati et al. [142]

Table 4.3: Table describing hip fracture observations in different countries.

4.1.7 In Vivo Studies of Hip Fracture Risk

As established, hip fracture has serious implications and therefore many studies have been undertaken to see if it is possible to predict whether a patient is at high risk of suffering the injury. In vivo or clinical studies, tend to use non-invasive imaging techniques to compare groups of patients who have suffered a fracture with control subjects. Age, gender, ethnicity and body size have an effect on bone geometry and quality, these factors are generally corrected for in final results and control groups are matched to the fracture group, as appropriate for the particular study. Almost all use Duel Energy X-Ray Absorptiometry (DXA) to examine femoral geometry and/or bone density as this is commonly used in medical practice, making large volumes of information available. In addition, if any useful metrics can be found within simple 2D scans then they will be easily implementable.

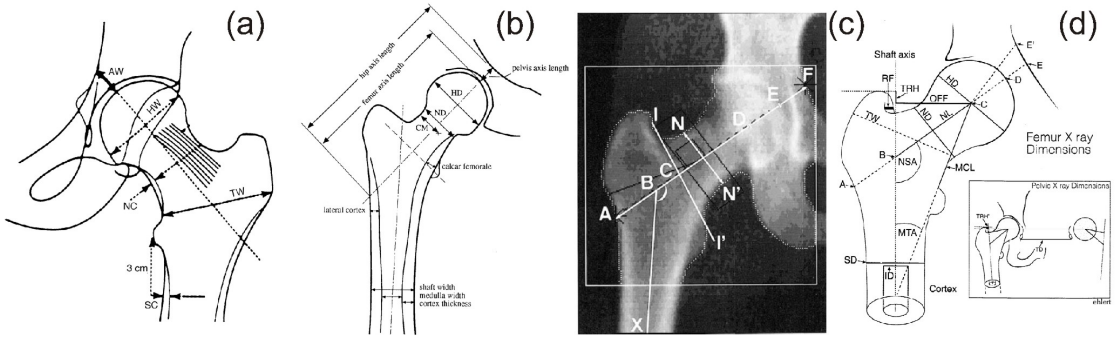


Figure 4.4: Illustrations of the geometric measurements taken by various clinical studies;(a) Theobald et al. [103],(b) Peacock et al. [143], (c) Bergot et al. [144] and (d) Michelotti et al. [106].

Several studies have investigated the influence of femoral geometry on fracture risk [145, 103, 106], whereas others have tried to gauge the influence of BMD in conjunction with geometry [143, 104, 146, 144]. There is some agreement in results but also a fair amount of contradiction. The studies which investigated bone density all agree that lower bone density is a significant differentiator between fracture and control groups in elevating fracture risk, and suggest that the incorporation of geometric indicators can increase predictive accuracy. However the choice of geometric parameter is not clear. Cortical thickness in the proximal femoral shaft was deemed the only significant geometric parameter by two studies [106, 143]. A long hip axis length (HAL), the length between the greater trochanter and pelvic brim, has been highlighted in several cases [145, 103, 104, 146], yet totally rejected by the others as insignificant. Bergot et al. [144] found HAL to be a reasonable indicator between healthy and fractured groups but unable to differentiate between non-fractured patients with low BMD, and so devised an alternative measurement between the femoral head center and intertrochanteric line. Other geometric parameters suggested by some papers, and often rejected by others, are; larger neck-shaft angle, greater femoral neck width and larger intertrochanteric width. Table 4.5 describes the aims, methods and findings of several related studies.

Inconsistencies in geometric results may be because these studies are attempting to measure a three dimensional object from a two dimensional image. The femur is 'twisted' in several planes, as explained in Section 2.2.1, therefore the perspective of an image taken of the femur can affect its appearance. Cheng et al. [147] compared DXA images of cadaveric femurs in both neutral and anteverted (mean 19.3°) positions and showed significant differences in both BMD and geometric measurements taken (Table 4.4). This shows the influence of patient morphology and particularly the large affect of intersubject variability. It is logical to also see how these findings relate to patient positioning during scanning, which currently is not standardised and has been shown to influence measurements [106, 104].

Measurement	Neutral (mean)	Anteverted (mean)	Difference (%)	Range
Neck BMD (g/cm^3)	0.705	0.723	+2.8	-5.3 - +9.8
Trochantic BMD (g/cm^3)	0.673	0.675	+0.2	-5.0 - +5.9
Neck Axis Length (cm)	9.97	9.73	-2.4	-7.6 - +4.3
Neck Width (cm)	3.36	3.36	0.0	-5.3 - +6.7

Table 4.4: Table showing the effect of anteversion on BMD and geometrical measurements of the proximal femur assessed by DXA [147].

Author	Aim	Methods	Main Findings
Faulkner et al. 1993 [145]	Evaluate whether simple **geometric** parameters can accurately predict hip fracture risk.	Took DXA scans of +8000 women over 67, of which 64 suffered hip fracture within 1.6 years. Compared geometric measures taken from 134 randomly sampled unfractured women to fractured women.	**Longer hip axis length** increased risk of both neck and trochantic fracture. No influence of neck width or neck shaft angle.
Theobald et al. 1998 [103]	Observed lower rate of hip fracture in African than Caucasian women. Investigated possible **racially related femoral differences** which could explain this.	Compared geometry and cortical thickness of 132 African and 175 Caucasian women using DXA.	Indicated African women have **thicker cortical bone** in the hip, **shorter hip axis length** and **smaller intertrochanteric width**.
Michelotti et al. 1999 [106]	Investigate effect of **femoral geometry** on fracture risk; when size, positioning and age are controlled.	Compared radiographs of contralateral hip of 43 fractured and 119 non fractured female hips.	Fracture group showed **thinner femoral cortex** (measured just below lesser trochanter). Showed that geometric measurements, specifically neck length, are highly sensitive to imaging position.
Center et al. 2008 [148]	Asses **volumetric bone density** as a measure of hip fracture risk.	Compared volumetric BMD calculated from DXA scans between patients with (73f, 23m) and without (852f, 635m) hip fracture.	Found **volumetric BMD** for fractured patients to be similar between the sexes, and a good indicator of fracture risk in men, unlike areal BMD. The sensitivity for women was comparable to areal BMD values.
Peacock et al. 1998 [143]	Examine **bone mass** and hip **geometry** in elderly (+60yrs) men and women, investigating measurement reproducibility and differences relating to gender, age and body size.	Took BMD and geometric measurements from DXA scans of healthy Caucasian subjects aged over 60 (336f, 141m).	Showed significant **diferences between men and women**, even with age and body size corrected for. Women show increase in medullary cavity width, at expense of shaft cortex thickness with age. However, for both sexes, BMD negatively related to age and positively to body size.
Gnudi et al. 1999 [104]	Investigate using proximal **femoral geometry** as an indicator of fracture risk in osteoporotic women.	Compared DXA scans of 111 post menopausal women with hip fracture to 329 healthy controls.	Fractured group showed **lower BMD**,**longer hip axis length** and **more valgus neck shaft angle**. **BMD** was best indicator of fracture risk, but inclusion of HAL and shaft angle could further increase prediction accuracy.
Gómez Alonso et al. 2000 [146]	Investigate the effect of **BMD** and **geometric measures** on hip fracture risk.	Assessed BMD and geometric parameters from DXA scans of 411 Spanish fracture patients (116m, 295f) and 545 control subjects (235m, 310f).	Fractured group showed **lower BMD**, **higher neck shaft angle** and **higher mean femoral neck width**.
Bergot et al. 2002 [144]	To find a **geometric** parameter able to better indicate hip fracture risk.	Used DXA scans of 49 non-trauma related neck fracture and 49 healthy female patients to compare BMD and various geometric parameters.	Showed **hip axis length** (greater trochanter to inner pelvic brim) to be significantly different between control and fracture groups, but not between fracture group and control group with low BMD. Suggests measure between femoral head center and intertrochanteric line (along neck axis) to be highly sensitive discriminator, the longer this length the higher the risk.

Table 4.5: Table describing several clinical studies investigating femoral fracture.

4.1.8 Computational Studies of Hip Fracture Risk

Hip fractures have been known to be a serious problem for a long time, therefore many studies have attempted to apply computational and experimental techniques to try to model and understand its mechanisms. The aims of this work have generally been to try to identify metrics by which fracture risk could be gauged by non-invasive techniques, several pieces of work are summarised in Table 4.6.

Early work was done by Lotz et al. [46] who understood that femoral neck fracture was most commonly the result of a fall and so investigated the effects of typical loading conditions seen during a fall as well as the more conventionally modelled stance loading, which is representative of spontaneous fracture during walking or possibly climbing steps. As with many computational studies which followed, finite element models were built from CT scans of a pair of cadaveric femurs which were subsequently subjected to mechanical testing to validate the FEA results. One femur was stance loaded and the other subjected to a fall impact, the load conditions created by Lotz are shown in Figure 4.5a-b. Remembering that this study only looked at one femur per load type, it suggested that stance or spontaneous fracture would result in a subcapital fracture location and a fall would lead to intertrochantic fracture. It was also seen that the load required to cause fracture was much lower for the fall configuration and therefore Lotz suggested that examining intertrochanteric bone quality may be the most useful fracture risk indicator. These findings have since been collaborated by several other studies who have tested a larger number of femurs [149, 102, 51, 150, 121, 151].

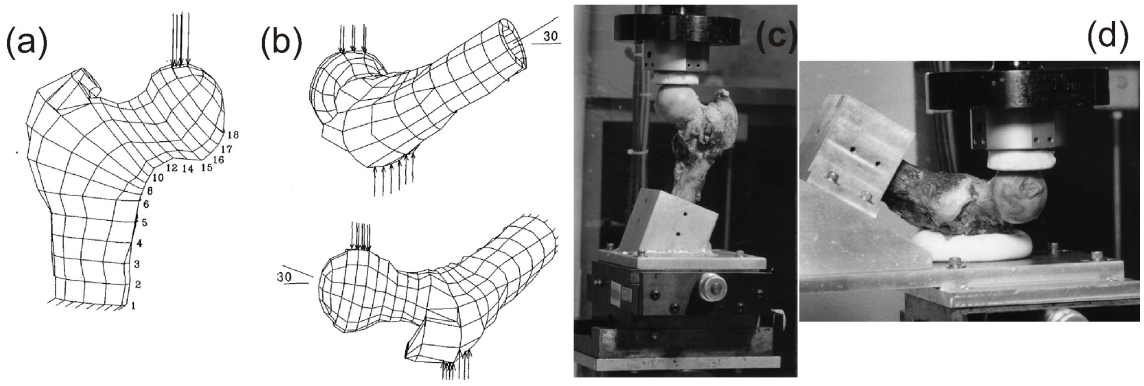


Figure 4.5: FE model loading conditions defined by Lotz et al. [46] for (a) stance and (b) fall, alongside images of corresponding experimental setup used by Keyak et al [51] for (c) stance and (d) fall.

Keyak et al. followed on from the work of Lotz, using their own automated finite element modelling techniques [38] to investigate femoral fracture and the stance and fall loading conditions previously developed. Using 18 pairs of cadaveric femurs the ability of the FE method to accurately predict fracture load was proved for both scenarios, also supporting the higher load requirement for stance fracture (Figure 4.5c-d) [51]. The same set up and methodology was later used to investigate fracture type and location, with strong agreement between the computational prediction and experimental result. The trend of stance loading resulting in cervical and subcapital fractures and falls leading to intertrochanteric fractures was again shown (Figure 4.6) [150].

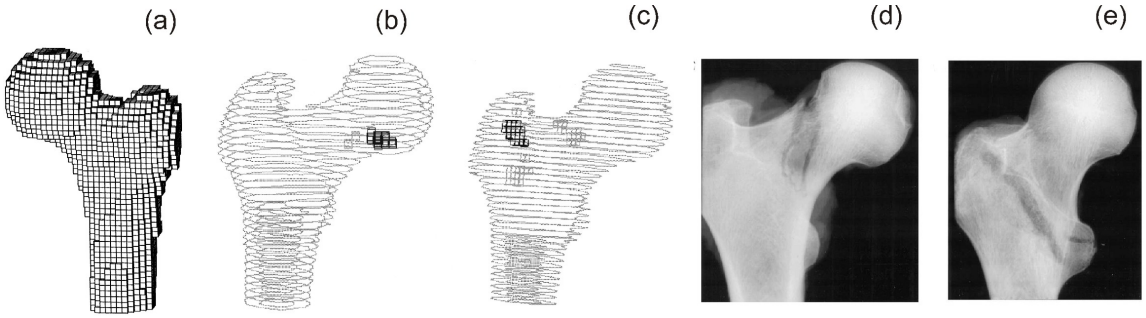


Figure 4.6: Keyak et al. [150]:(a) Example of FE models used, 8447 cubic elements and 10652 nodes. Fracture locations predicted by FE model for (b) stance and (c) fall loading, and radiographs of fractures produced in testing by (d) stance and (e) fall loading.

A further study by Keyak et al. [121] looked at the effect of force direction on fracture load, proving that the most dangerous configuration is an impact on the posterolateral aspect of the greater trochanter. While the most dangerous stance configuration exists where postero and lateral components act in a similar way to standing on one leg or walking up stairs. Bessho et al. [151] ran simulations at the most severe fall condition and in stance loading, modelling the healthy hip of 10 female patients who suffered cervical fractures. This study created FE models of the whole femur which were able to show the lower fracture load of the fall scenario and replicate similar fracture locations to those seen in reality.

A significant weakness of all of the studies discussed so far and the vast majority of all the fracture risk work is the fact that only the femur is considered. It is clear that in vivo there are many more factors involved, most significantly the interaction with the pelvis, damping affects of surrounding soft tissues (and impact surface for a fall) and inertia effects due to the subjects body weight. A single subject model of the pelvis and surrounding soft tissue has been created by Mujumder et al. [50] which attempts to incorporate all these factors and simulate an oblique fall to the

side. The model is complex due to its representation of the whole pelvis-femur-soft tissue structure, but a reasonably coarse mesh size of 4-7mm was used (Figure 4.7). Reassuringly, the result of the model supports the femur only models in suggesting fracture will occur in the trochanteric region.

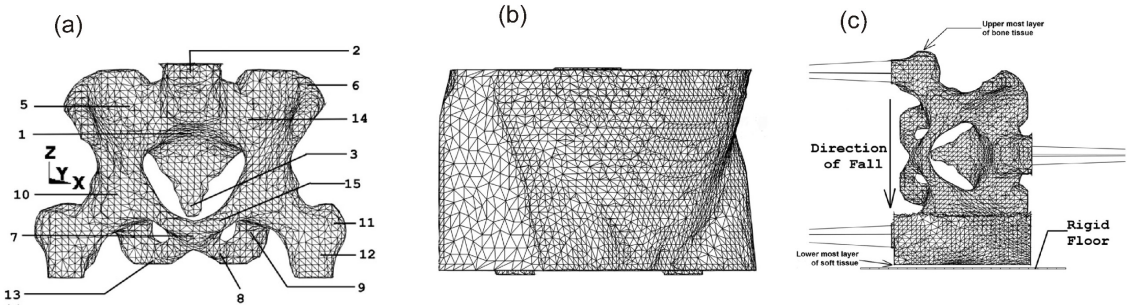


Figure 4.7: Illustrations of the models and simulations run by Majumder et al. [50].(a) FE pelvis-femurs complex, (b) pelvis-femur complex with surrounding soft tissue and (c) fall configuration showing body impact with the floor.

Author	Aim	Methods	Main Findings
Lotz et al. 1991 [46]	Use FE proximal femur models to compare stress distribution resulting from stance loading and from a fall impact. Also compare in vitro recorded and FE predicted strain and failure loads in the two loading conditions.	Used 2 cadaveric femurs to generate FE models and then for experimental simulation (1 for each loading condition). Used simple linear isotropic FE model, and looked at strains produced in different regions of the proximal femur.	Showed accurate prediction of yield onset and fracture load. Indicated highest stresses in stance are in subcapital region but in intertrochantic region for fall loading, which also failed at lower load. Suggests measurement of intertrochantic bone quality may be sensitive for hip fracture.
Cheng et al. 1997 [149]	Compare femoral BMD and geometry with femoral strength to try to identify potential fracture risk indicators.	64 cadaveric femurs (28f,36m) measured and their BMD assessed using Quantitative CT (QCT). Then mechanically tested to find their strength under a simulated fall load.	Resulted in 38 trochantic and 26 cervical fractures. Strong correlation between trochanteric cortical area and femoral strength, weak correlation with neck axis length and neck width. Neck-shaft angle not significant.
Cody et al. 1999 [102]	Test ability of FE models, QCT and DXA to predict femoral strength, with potential for in vivo patient monitoring.	51 cadaveric femurs (23f, 28m) imaged by QCT and DXA, QCT used to develop 3D FE femur models. Three methods of assessing femurs were trialled based on DXA, QCT and FE modelling. The predictions of each were compared to experimental results for fracture load of the femurs under a stance load.	Observed 69% subcapital, 18% subcaptial to mid-cervical and 13% oblique fracture. Showed FE model was best at predicting final fracture load.
Keyak et al. 1997 [51]	Investigate whether FE models automatically generated from CT scans can estimate in vitro fracture load.	Created FE models from 18 cadaveric femur pairs, applied stance loading to one of each pair and fall loading to the other. FE simulated the loading conditions on the matched FE models. Failure was defined when 15 contiguous elements exceeded a safety factor relating to their CT derived strength and the von Mises stress experienced.	Saw that the predicted and experimentally produced fracture loads were well correlated, with stance fracture load approximately twice that of the fall fracture load.
Keyak et al. 2001 [150]	Investigate whether FE models automatically generated from CT scans can estimate in vitro fracture location and type.	Used 18 cadaveric femora pairs, created FE models of them all, and applied stance loading to one of each pair fall loading to the other. FE simulation simulated the loading conditions on the matched FE models. Failure was defined as above.	Stance load: 72% agreement in predicted and actual location with 100% accuracy in predicting only cervical fractures. Fall load: 67% in predicted and actual location with 79% agreement in fracture types, predicted and saw majority trochanteric fractures.
Keyak et al. 2001 [121]	Attempt to quantify the effect of force direction on fracture load in terms of fracture risk.	Generated 4 FE proximal femoral models and tested two loading conditions on each; stance and fall. For each loading condition the force direction was varied and the fracture load recorded. Failure defined as above.	Showed that the fall configuration had a lower fracture load than any stance loading configuration tested. The most severe being a fall onto the posterolateral aspect of the greater trochanter. Comments that future testing should use this worst case scenario.
Bessho et al. 2004 [151]	Use FE models of the whole femur to predict fracture load and location, developed from CT scans taken of the contralateral hip of patients who recently suffered hip fracture.	10 subjects with hip fracture, classified as either subcapital or transcervical resulting from a fall, and 3 controls were scanned and models made of their femurs. FEA was used to apply a fall and stance load to the models until failure was defined when a 'cortical' element failed, crack propagation was predicted as perpendicular to the max. principle stress at this point.	Showed significantly lower load in fracture patients than controls. All configurations showed similar fracture patterns to those seen in reality, although fall loading also showed intertrochanteric fractures.
Majumder et al. 2007 [50]	Aim to improve conventional simulation of sideways falls by incorporating the entire bone structure of the hip and the surrounding soft tissue into FEA.	Generated an FE model of the pelvis and both proximal femurs, with surround soft tissue from a single subject CT scan. Modelled the effect of the whole body so total inertia, stiffness and damping, which control and affect the fall, could be included.	Found trochantic fracture resulting from a sideways backwards fall, and was able to incorporate the effect of trochantic soft tissue thickness variation on impact force and resulting strain.

Table 4.6: Table describing several experimental and computational studies investigating femoral fracture.

4.2 Methodology

The generation of a large number of femur models had several challenges which required solutions before it was possible to run a large scale, multi bone simulation. The study performed in this section aimed to generate 1000 unique femurs from the previously created statistical model of the femur and then run FE simulation on each of them to investigate femoral neck fracture. The stages required to do this are explained below. It is also important to comment at this stage that a major consideration of this work was the ability to automate the entire process from start to finish. If this was not achieved, and manual intervention was required then the usefulness of this tool would be lost.

1. Sample the statistical shape and intensity model to generate 1000 femur instances. This required the boundaries within which the statistical model was to be sampled to be set, and a method of ensuring that the probability space it described was appropriately sampled.

2. Check generated mesh quality to prevent later simulation failure or anomalous results.

3. Define loading conditions to each femur with individualised forces relating to predicted weight.

4. Produce an input file suitable for a finite element solver. This had to contain all nodal positions, elemental connectivity, material property information, applied loads and boundary conditions.

5. Automating the simulation of each file in a finite element solver and recording the results.

6. Automated post processing of the FEA results.

4.2.1 Generation of Large Set of Femur Models From Statistical Model

The first step of this study was to formulate a methodology for using the statistical model created in Chapter 3 to automatically generate large numbers of realistic femur models. Section 3.6 and Appendix B explain how the boundaries within which the model could be best sampled were found, concluding that the optimal choice for this model was the inclusion of 8 modes and stretching the data to $+/-1.5\sigma_{b_i}$ where σ_{b_i} is the standard deviation of the individual shape parameters, b_i, across the training set.

A Sobol sequence [152] was used to generate a set of vectors, S_j, with a value between 0 and 1. These vectors were then used to perturb the statistical model shape parameter, b_i, thus generating completely new and unique instances, as follows. The length of each Sobol generated vector was equal to the number of modes being used, 8, and the number of different vectors needed was equal to the number of femurs required, 1000 for this study. By setting the lower parameter limit, b_i^l, to $-1.5\sigma_{b^i}$ and the upper parameter limit to, b_i^u, to $+1.5\sigma_{b_i}$ the specific shape and modulus parameter vector for each generated femur, B_j, could be set by:

$$B_j = b_i^l + (b_i^u - b_i^l)S_j \quad \text{where } j = 1000, i = 8. \tag{4.1}$$

By using each of the generated shape vectors B_j it was possible to generate 1000 unique output vectors from the statistical model, each of which contained corresponding nodal coordinates and nodal modulus values that described a femur as a solid tetrahedral mesh with associated material properties (see Section 3.4.5).

Before the meshes could be fed into a solver, mesh quality had to be interrogated to ensure that there were no inside-out elements or elements so distorted as a result of the mesh morphing process that they could cause the later simulation to fail. Section B.0.5 discusses the mesh quality checks adopted for this model during the developments stage. It also explains that inside-out elements are a side effect of the mesh morphing process adopted, however node order is easy and quick to correct and so this simple check and repair step was written into the generation code. To prevent excessive element distortion a normalised shape ratio, Equation 3.3, and a shape distortion ratio, Equation 3.4, were calculated for each generated element, and compared to the warning and error limits defined by ANSYS©. If more than 1% recorded a warning, or any element showed an error, then the geometry failed and was regenerated. The process of generating and checking 1000 femurs took less than 5 minutes.

4.2.2 Subject Specific Fall Loading Conditions

The study aimed to investigate the strain resulting in the proximal femur following a fall. The FE loading conditions for this were created to emulate the experimental work of Keyak et al. [150], see figure 4.5d. To be able to set up this loading condition automatically for each femur model the correspondence between model elements and nodes was exploited. This allowed the conditions to be manually defined on the baseline femur mesh and the relevant nodes and elements fixed or loaded, as appropriate, in all subsequent generated models.

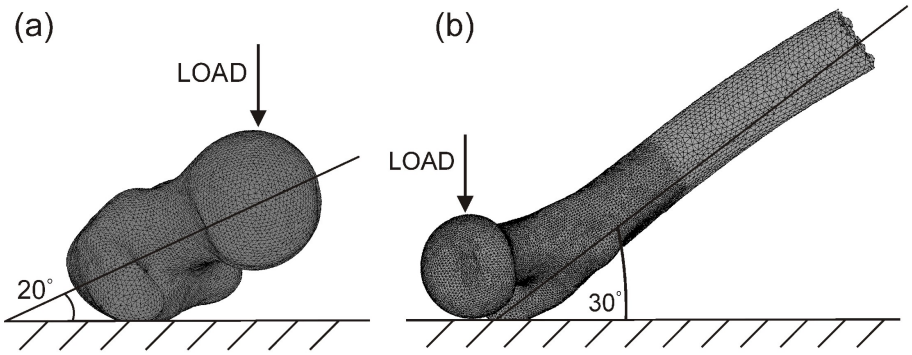
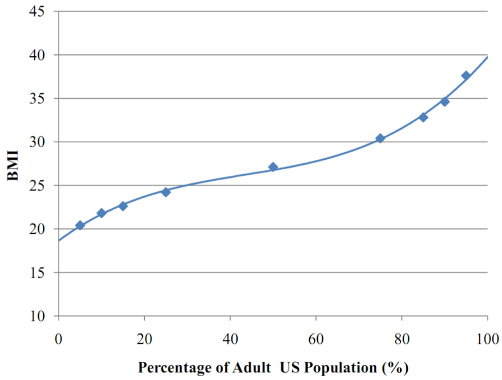


Figure 4.8: Illustration of the loading conditions applied to each femur to simulate a fall

The baseline femur was therefore manually rotated such that its shaft axis lay at 30° to the horizontal plane (Fig. 4.8a) and the neck axis in the transverse plane at 20° to the horizontal (Fig. 4.8b). The femur was fully restrained in two places; a short depth of the lowest part of the greater trochanter, replicating the polymethylmethacrylate(PMMA) cup holding the femur in the experimental test, and from the mid shaft of the femur down. A force was equally distributed over a ϕ 3cm area of the proximal, anterior femoral head. This area would vary slightly between models as it was translated into a loaded selection of surface nodes, however this area would be proportional in each generated model thanks to the surface registration technique applied (see section 3.5.2).

4.2.3 Subject Specific Loading



Percentile	BMI
5	20.4
10	21.8
15	22.6
25	24.2
50	27.1
75	30.4
85	32.8
90	34.6
95	37.6

Figure 4.9: BMI distribution curve for the adult US population, estimated from a national survey in 2002 [153].

The value of the applied force was set at one times body weight, due to the linearity of the model any strain results produced could be scaled so the choice of load magnitude was arbitrary. As all 1000 femurs were created statistically, no subject weight was known so this data was generated as follows. Femur length, taken as the distance from the most distal point of the lateral condyle to the most proximal point of the greater trochanter, was assumed to be 26.75% of subject height [154]. This was a generic relationship, ignoring gender and race with a subsequent possible error in predicted height reported at <0.6cm. A Body Mass Index (BMI) distribution curve was generated from data available from the National Health and Nutritional Examination Survey 1999-2002 (Fig. 4.9), conducted on all age groups within the U.S. population [153]. By randomly sampling a BMI value from the distribution, it was possible to calculate a subject weight in kilograms as BMI multiplied by the square of the predicted height in meters. The process is illustrated in Figure 4.10 for 21 femurs.

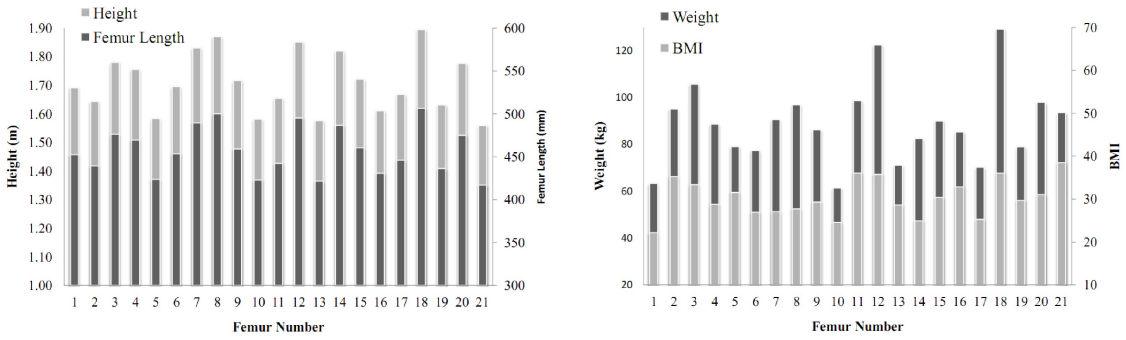


Figure 4.10: Illustration of assignment of BMI and weight to individual femurs, based on their length and related predicted subject height

4.2.4 Finite Element Simulations

At this point all the information needed by the finite element solver had been calculated. A simple script was created to write an input file for the FEA code MARC©(MSC, Santa Ana USA) , for each generated femur. After a few generic header lines defining element type and the size of the model the first data required was the geometry structure, consisting of the element connectivity which was consistent for each model, and the unique nodal coordinates. As each model contained over 600 000 elements it was not efficient to define material properties for each element individually. Therefore elements were grouped into 10MPa bands and given a modulus at the top end of this band, resulting in over 30 000 property bands per model. The lists of fixed and loaded node numbers were then defined, along with the instruction to run one loading cycle and then write out the result. Each input text file contained over 64Mb of data.

It was possible to batch the simulation runs such that they were controlled through MatLab© but operated through an MS-Dos shell. This enabled a simple fortran script to start the MARC simulations and then save out selected parameters. The extracted results were post loading element volumes and element strains. This data was then read back into the controlling program, Matlab, where it could be postprocessed. Due to the size of the models the time taken to load and write files was quite large, on average it took $\sim$3 minutes to write each input file and $\sim$9 minutes to load, run and save the results of each FE simulation.

4.2.5 Metrics Extracted from Statistical Model

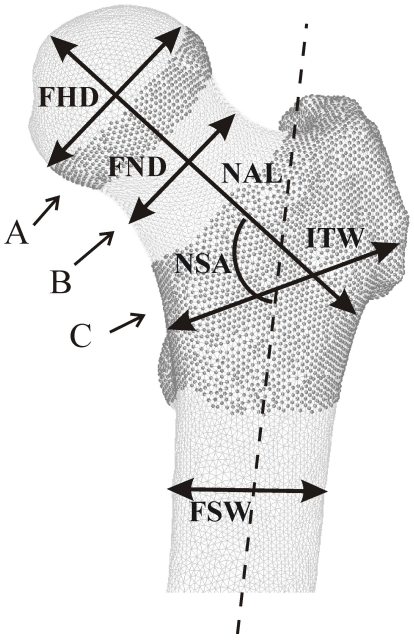


Figure 4.11: Illustration of metrics taken from femur models. Main areas of interest: A - lower femoral head, B - femoral neck, C - intertrochantic. Measures include: head and neck diameters, neck axis length, neck shaft angle, intertrochantic width, shaft width and anteversion angle.

Various metrics were devised to aid interrogation of the FE results, Fig. 4.11. A range of geometric parameters were automatically taken from each generated femur in an automated postprocessing stage, again it was possible to achieve this because of the elemental and nodal correspondence between models. The metrics were defined on the baseline femur model, with either measurements being found by lines between specific nodes (or the angles between these lines) and volumes being defined by groups of elements. The measurements were identified by a mixture of methods, but most are based around the femoral head center, found by fitting a sphere to the head, and the neck axis vector, found by fitting a cylinder to the femoral neck. From these measurements head diameter, neck diameter (both in anterior-posterior and proximal-distal directions), and neck axis length could be found. The neck shaft angle requires the prior calculation of the shaft axis, by finding the center of mass for several bands of the femoral shaft and fitting a line between them.

All the metrics used were based on parameters which have previously been used to analyse femoral shape, see Figure 4.4 [103, 106]. These were; Neck Axis Length (NAL), Neck-Shaft Angle (NSA), Femoral Head and Neck Diameters (FHD and FND), Intertrochantic Width (ITW), Femoral Shaft Width (FSW, measured ∼3cm below the lesser trochanter) and Anteversion Angle (AA). In addition, three key volumes were identified within the proximal femur to gauge bone quality and judge failure risk, these were; lower femoral head (A), femoral neck (B) and the intertrochantic region (C). They roughly relate to subcaptial, cervical and trochantic fracture regions. To highlight those femurs which were at highest risk of failure a conservative criterion was created identifying models where any of the three proximal sections experienced >10% volume exceeding yield strain, 0.7%. This criterion was developed only as a

way of isolating the worst performing femurs to examine their differences, as it was not felt that a precise failure limit could be verified.

4.3 Results

4.3.1 Overview of Results

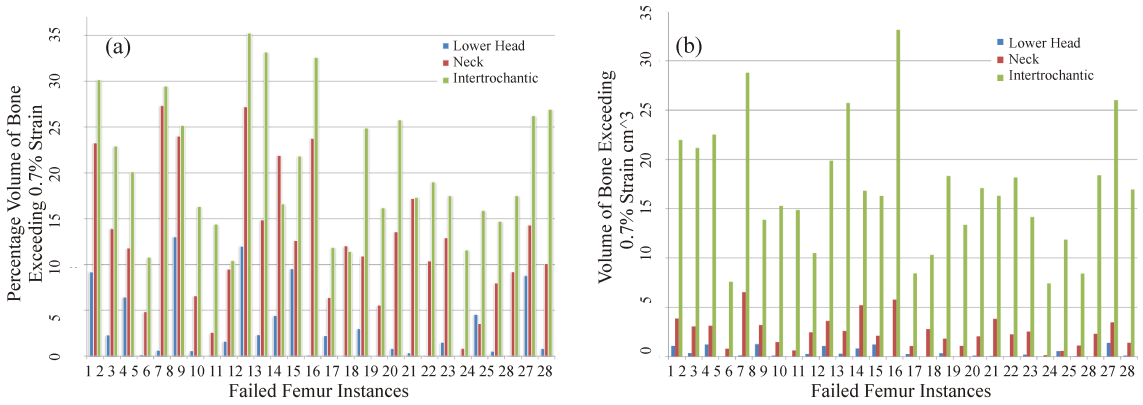


Figure 4.12: Plot of the volume of bone exceeding yield strain in each of the three defined regions for all 28 failed femurs. Shown as (a) the percentage volume of bone and (b) the actual volume of bone in cm^3.

The failure criteria established in this study, based on the previously defined criteria, found 28 of the 1000 generated femurs tested to be at risk of failure. This means that 28 femurs showed >10% bone volume in one or more of the three examined regions to exceed the yield strain of bone. Under closer examination, the trochantic and neck regions appeared to breach this criteria far more frequently than the lower head region. Figure 4.12 shows the percentage (a) and actual (b) bone volume exceeding this failure criteria in each of the three regions for all 28 femurs. As the intertrochantic area is much larger than the other two, it unsurprisingly dominates in terms of actual volume, but when the regional percentages are calculated it is still significantly higher in most cases. However there are clear differences between the strain distributions in each femur, some showing noticeably higher neck strains, for example femurs 5 and 10, and some showing higher strains in the lower head, femurs 6 and 10.

Having established which femurs performed worst in the fall load test and seen indications that the failure patterns produced in these femurs differ from each other, a further analysis was performed to visualise the volumetric strain distributions. The failed and non failed groups were separated and their volumetric elemental strain

results for regions A, B and C (head, neck and intertrochantic), were grouped into 0.02% strain sized bins (Figs. 4.13,4.14 and 4.15), clear differences were seen in the results. For the low risk group (center row) the mean volume strain above 0.3% is almost zero in all regions, with the majority of the bone volume below 0.1%. However the at risk group (top row) distribution is different. The >1% strain bin contains the largest volume of bone, although this is unsurprising due to the criteria used to separate this group from the larger data set. The key difference is how the strain distribution has flattened and been spread out, so there are notable bone volumes in all the strain bins up to 0.5% in the head, 0.6% in the neck and 1% in the intertrochantic region. The results for the training set femurs are shown alongside (lowest row), they match most closely to the non-risk group, with typically low strain results. These graphs show that the failed group have a shifted volumetric strain distribution throughout the strain range, indicating that the failure criteria is not being reached due to a pocket of high strain which could be caused by a boundary or loading condition.

Two main routes of investigation were performed on the data produced from the FE simulations. Firstly the fracture prone group of femurs was compared to the unfractured group to try to find any statistically significant difference between them using the geometric and material metrics previously defined. The second line of analysis was focused on the failed group, attempting to identify the likely origin site of any fractures. Both of these areas yielded results which could be compared to the existing literature from experimental and computational based studies.

4.3.2 Range of Femurs Generated

To analyse the results of the study a range of metrics were extracted, these enabled the characteristics of each femur to be identified. The metrics, which can be placed into three classes, are described in Table 4.7. Cortical bone was defined as bone with a modulus exceeding 3000 MPa, and so it follows that cancellous bone was classified as bone below 3000 MPa. Where appropriate, material property characteristics were calculated both in terms of regional percentage volumes and absolute volumes.

The immediate use of these results was to view the range of femurs generated. By compiling the statistics of all 1000 generated models it was quick and easy to ensure that the data being tested was realistic, i.e. no exceptionally over/under sized femurs. This was an important analysis, allowing the model set to be interrogated without manually visualising and checking each instance and showing that the training set

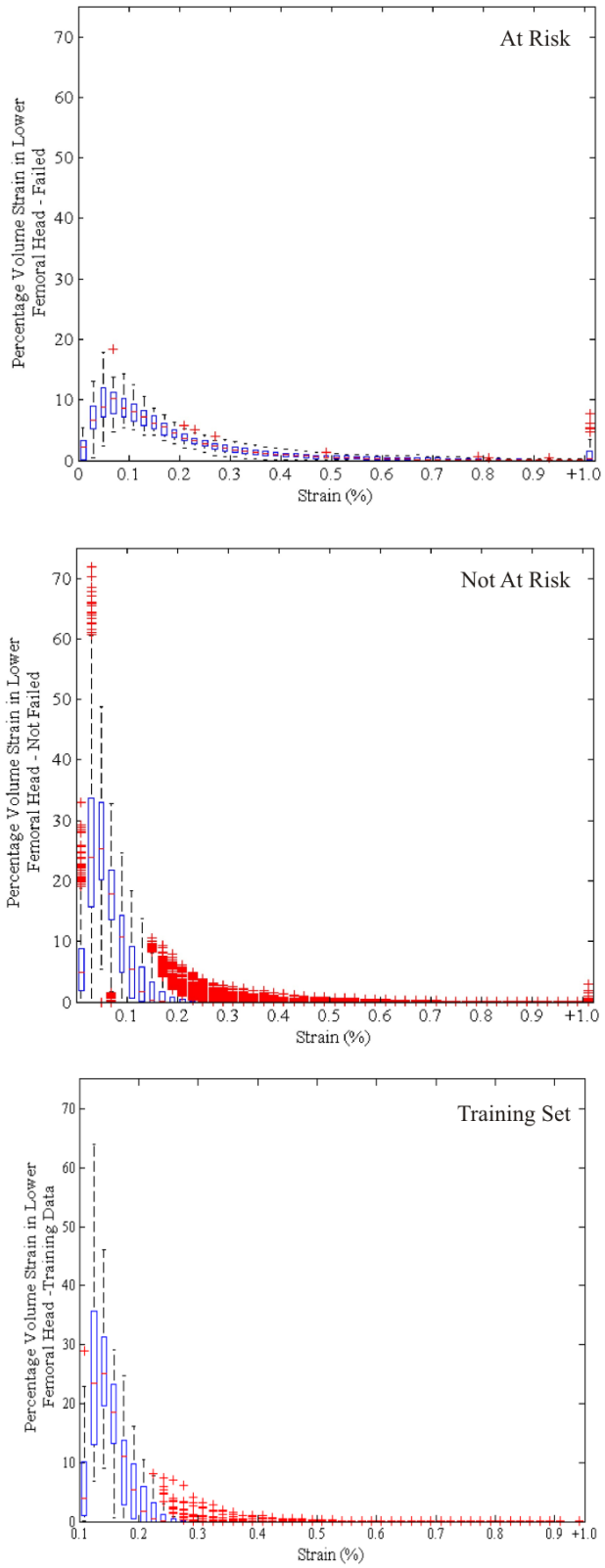


Figure 4.13: Box plots illustrating the strain in the head region by percentage volume. The top plot shows the 28 fracture risk group, the middle plot the 972 not at risk group and the lower plot the training set results. The box shows the median (red), upper and lower quartile (blue) and the whiskers extend to 1.5× the interquartile range, with values beyond this shown by crosses (red).

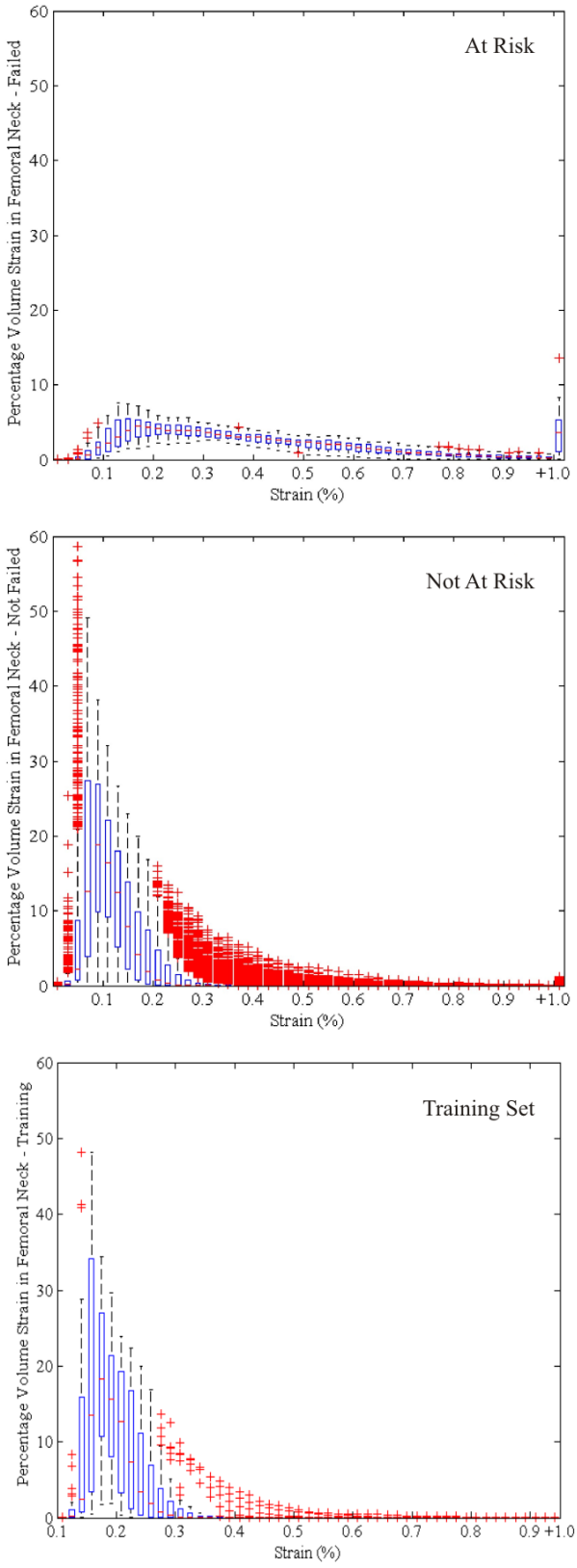


Figure 4.14: Box plots illustrating the strain in the neck region by percentage volume. The top plot shows the 28 fracture risk group, the middle plot the 972 not at risk group and the lower plot the training set results. The box shows the median (red), upper and lower quartile (blue) and the whiskers extend to 1.5× the interquartile range, with values beyond this shown by crosses (red).

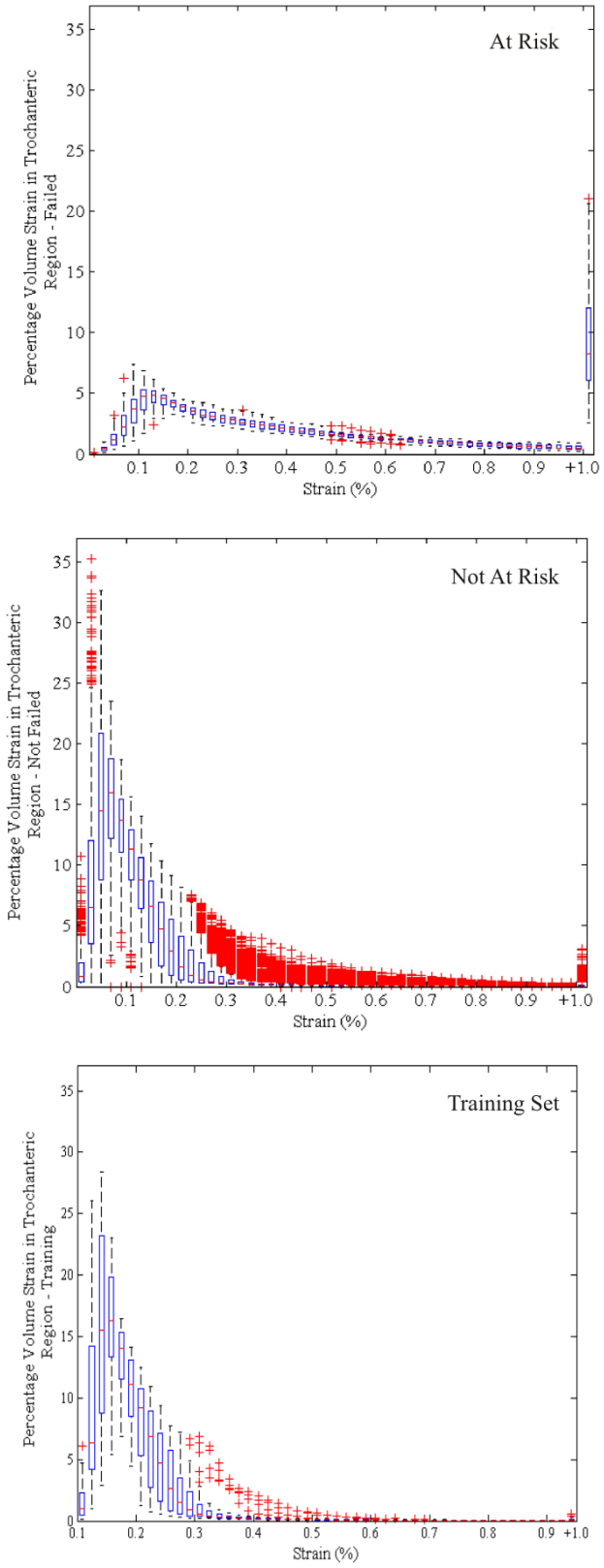


Figure 4.15: Box plots illustrating the strain in the head region by percentage volume. The top plot shows the 28 fracture risk group, the middle plot the 972 not at risk group and the lower plot the training set results. The box shows the median (red), upper and lower quartile (blue) and the whiskers extend to 1.5× the interquartile range, with values beyond this shown by crosses (red).

Subject Characteristics	Femoral Geometry Characteristics	Material Property Characteristics
Subject Height (mm)	Neck Axis Length (mm)	Lower Head Volume
Subject BMI	Neck Shaft Angle (mm)	Lower Head Cortical Volume
Subject Weight (kg)	Femoral Head Diameter (mm)	Lower Head Mean Cortical Modulus (MPa)
Load Applied (N)	Femoral Neck Diameter P-D (mm)	Lower Head Mean Cancellous Modulus (MPa)
	Femoral Neck Diameter A-P (mm)	Head Bone Volume + 0.4% Strain
	Neck-Head Diameter ratio	Head Bone Volume + 0.7% Strain
	Coronal-Sagital Neck dia. ratio	Lower Head Cortical Bone Volume %
	Anteversion (°)	Neck Volume
	Femoral Shaft Radius (mm)	Neck Cortical Volume
	Intertrochantic Width (mm)	Neck Mean Cortical Modulus
	Femoral Length (mm)	Neck Mean Cancellous Modulus
		Neck Bone Volume + 0.4% Strain
		Neck Bone Volume + 0.7% Strain
		Neck Percentage Cortical Bone Volume
		Intertrochanteric Volume
		Intertrochanteric Cortical Volume
		Intertrochanteric Mean Cortical Modulus
		Intertrochanteric Mean Cancellous Modulus
		Inter. Bone Volume + 0.4% Strain
		Inter. Bone Volume + 0.7% Strain

Table 4.7: Table listing the metrics extracted from all femurs.

was being fairly represented. The minimum, mean, maximum and standard deviation of this data is shown in Table 4.8. The geometric measurements can be seen to be very similar to the data previously reported for the training set in Tables B.2 and B.3.

Metric	Min	Mean	Max	Std
Neck Axis Length (mm)	83.212	100.214	118.881	6.508
Neck Shaft Angle (mm)	120.803	126.461	133.609	4.237
Femoral Head Diameter (mm)	36.887	47.446	58.682	3.914
Femoral Neck Diameter P-D (mm)	27.668	33.843	40.658	2.465
Femoral Neck Diameter A-P (mm)	24.911	33.189	44.052	3.432
Neck-Head Diameter Ratio	1.256	1.402	1.557	0.053
Neck Diameter Ratio	0.863	1.024	1.261	0.062
Anteversion (°)	15.052	22.010	28.752	2.273
Femoral Shaft Radius (mm)	12.987	16.136	19.455	1.220
Intertrochantic Width (mm)	44.585	55.551	66.958	4.040
Femoral Length (mm)	415.267	457.789	505.034	23.116
Subject Height (mm)	1449.804	1611.581	1791.211	79.481
Subject BMI	18.686	27.522	39.912	4.680
Subject Weight (Kg)	43.035	71.654	122.364	14.134
Load Applied (N)	422.027	702.688	1199.984	138.610
Lower Head Vol. (mm^3)	7680.431	15394.054	26563.900	3340.005
Lower Head Cortical Vol. (mm^3)	0.000	363.013	3431.368	547.201
Lower Head Mean Cort. Mod (MPa)	3002.252	3035.903	3864.665	865.972
Lower Head Mean Canc. Mod (MPa)	438.830	1320.628	1965.974	293.297
Vol. over 0.4 % Strain (mm^3)	0.000	57.538	3129.182	293.672
Vol. over 0.7 % Strain (mm^3)	0.000	14.656	1385.147	105.784
Lower Head Cortical Vol %	0.000	2.415	21.969	3.510
Per Vol. over0.4 % Strain	0.000	0.434	26.726	2.272
Per Vol. over 0.7 % Strain	0.000	0.113	13.024	0.850
Neck Vol. (mm^3)	11508.614	20310.235	32797.790	4199.371
Neck Cortical Vol. (mm^3)	1275.813	4505.211	12281.619	2025.212
Neck Mean Cort. Mod (MPa)	4197.257	5300.621	6640.665	367.375
Neck Mean Canc. Mod (MPa)	411.856	1346.733	2205.860	409.059
Vol. over 0.4 % Strain (mm^3)	0.000	415.461	15327.315	1599.733
Vol. over 0.7 % Strain (mm^3)	0.000	89.312	6531.974	497.330
Neck Per Cortical Vol. %	6.168	22.494	50.358	9.391
Per Vol. over 0.4 % Strain	0.000	2.075	64.675	7.918
Per Vol. over 0.7 % Strain	0.000	0.448	27.345	2.473
Introchantic Vol. (mm^3)	48139.601	83338.961	134340.522	16813.927
Intertroc. Cortical Vol. (mm^3)	6023.532	20565.793	46379.455	7030.814
Inter Mean Cort. Mod (MPa)	4915.590	5877.860	6973.287	355.784
Inter Mean Canc. Mod (MPa)	359.975	1098.351	1764.837	307.916
Vol. over 0.4 % Strain (mm^3)	0.000	3110.698	55034.264	6494.316
Vol. over 0.7 % Strain (mm^3)	0.000	946.778	33145.241	3037.308
Intertroc. Cortical Vol. %	8.494	24.938	46.012	7.596
Per Vol. over 0.4 % Strain	0.000	3.725	56.350	7.709
Per Vol. over 0.7 % Strain	0.000	1.139	35.257	3.653

Table 4.8: Minimum, mean, maximum and standard deviation of the geometric metrics, calculated across all 1000 generated femurs.

4.3.3 Predicted Risk Factors

As previously explained the results of the FE analysis were separated into failed and non failed groups on the basis of the strain produced in three areas of the proximal femur. This allowed the results for each metric extracted from the femur models to be compared between these two groups. This revealed seven metrics to be highly significant, as indicated by an F-Test analysis (Table 4.9). The most important of the parameters was the percentage of cortical bone in each section, especially significant in the lower femoral head where the mean cortical modulus was also highlighted. Three geometric parameters appeared to be important, neck shaft angle and to a lesser extent anteversion angle and femoral neck diameter ratio (the neck diameter ratio indicated the ovality of the neck, calculated as a ratio between neck diameters measured in the superior-inferior and anterior-posterior directions).

		Not At Risk			At Risk		
	F-Test	Min	Mean	Max	Min	Mean	Max
(A) Cortical Volume (%)	0.01	0.00	2.48	21.97	0.00	0.05	0.51
(B) Cortical Volume (%)	0.01	6.61	22.86	50.36	6.17	9.79	15.78
(C) Cortical Volume (%)	0.01	9.38	25.34	46.01	8.49	11.10	23.98
Neck-Shaft Angle(°)	0.025	120.8	128.7	133.6	121.3	124.4	127.8
(A) Mean Cort. Modulus (MPa)	0.025	3002	3274	3865	3016	3300	3675
Anteversion(°)	0.1	15.05	22.06	28.75	16.91	20.45	23.98
Neck Diameter Ratio	0.1	0.86	1.02	1.26	0.94	1.02	1.13

Table 4.9: Results of the most significant material property and geometric metrics found when comparing the failed and not failed model groups. The minimum, maximum and mean of each group is shown. A, B and C indicate the section of the femur.

4.3.4 Predicted Fracture Locations

The likely origin of any fracture was identified by interrogating the areas of highest strain in the 28 femurs which failed the fall simulation. The failed femurs were visualised and the areas experiencing high strain were highlighted. By gradually increasing the strain threshold, below which elements were not selected, it was possible to identify the probable origins of any fracture. The majority, 15 of 28, indicated failure in the trochantic region with 8 of these showing highest strain along the intertrochantic ridge (Fig. 4.16a). 4 femurs highlighted the anterior subcapital region and the remaining 9 had multiple regions of high strain making a specific location hard to identify (Fig. 4.16b+c). Most femurs showed some localised high strain

around the greater trochanter restraint, but no model showed this to be the only high strain location or any potential fracture lines stemming from this area.

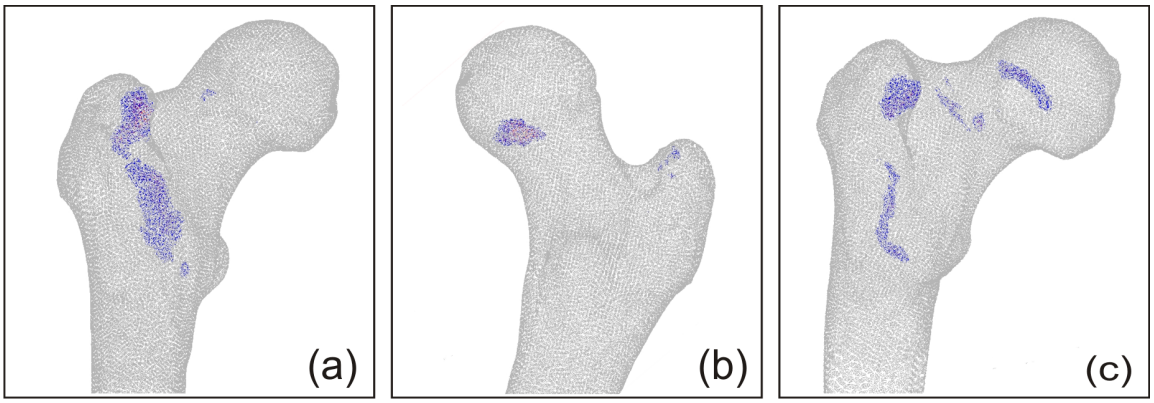


Figure 4.16: Illustration of the areas suffering highest strain following fall loading. (a) Intertrochantic, (b) anterior subcapital and (c) multiple regions. Areas highlighted exceed 1.5% strain

4.4 Discussion

This study was able to run a large scale, multi-bone, finite element analysis for the first time. In addition the process was made completely automated, needing only the number of femurs required as an input and subsequently using the existing statistical model to generate models with material properties, check their mesh quality, define a patient specific load, set up individual loading and boundary conditions, submit this to an FE solver and finally extract and analyse the elemental strains produced. In total this entire process took approximately 12 minutes per femur.

The FEA results were investigated to see if any geometric or material property metrics could be found to be significantly different between the group of femurs which were classed as failing under a fall load and those which were not. The model identified the overall percentage volume of cortical bone through the proximal femur, and the mean modulus of cortical bone in the lower femoral head as significant bone quality metrics. In terms of geometry; neck shaft angle, anteversion and the ovality of the femoral neck were seen to be important.

As discussed throughout Section 4.1, many previous studies have suggested femoral geometric and material features which may result in a predisposition towards femoral fracture with the exact features frequently contradicted between studies. The main feature which is agreed on is that a low BMD is a high indicating factor of risk [149, 104, 144, 146, 46], and also low cortical thickness (although usually defined in the

proximal femoral shaft) [149, 103, 106]. This was very clearly supported by the results of the model, with cortical bone percentage by far the most significant difference between the failed and non failed groups. There was no evidence that neck axis length was an indicator of risk, agreeing with some work [106, 146] but contradicting others [144, 103, 104, 145].

An interesting geometric parameter which was shown as significant was neck-shaft angle. Again this parameter had been shown to have little or no influence on fracture risk by some [144, 145] and yet important by others [104, 106, 146]. The studies which did indicate this measurement suggest that a larger angle increases the risk, however this study's results show a smaller angle in the failed group. Michelotti et al. [106] observed that this trend was seen in studies which took dimensions from three dimensional images as opposed to the common two dimensional X-Ray. Suggesting that subject positioning during imaging, particularly external femoral rotation, can result in apparent changes to neck axis length and neck shaft angle, a finding supported by work on the affect of anteversion by Cheng et al. [147]. This parameter may well be affected by the limited training set as it is known to be generally larger in women than men [146], however with only 21 femurs it was not possible to separate male and female subjects to generate gender specific models.

The present study corroborates previous findings that the majority of failures under fall loading occur in the intertrochantic region [149, 102, 51, 150, 121, 151]. Keyak et al. [150] published some details of the failure locations of femurs under a fall load which were tested experimentally as well as modelled computationally. The experimental tests done in Keyak's work were those replicated in the current study. The fracture site was identifiable for 15 tested femurs. Although the descriptions of fracture initiation sites are a little vague, it can be seen that a similar distribution of results has been found in both Keyak's work and this study (Table 4.10).

	Keyak - FE (%)	Keyak - Exp. (%)	Statistical Model (%)
Trochantic Intertrochantic	60	47	29 25
Cervical	13	40	14
Multiple	-	-	32
Subtrochantic	0	13	0

Table 4.10: Table showing the percentage of femurs identified with various fracture location origins. Comparing the results seen by Keyak et al [150] for the 15 femurs where experimentally identifiable failure locations were compared to FE predictions, with the failure locations predicted by this study using femur models generated from a statistical model.

There are limitations to this work. The model may suffer from the relatively small size of the training data set, as discussed in the previous chapter. The data set is taken from quite a general population group and so does not incorporate factors such as ostoeporosis, tumours or other pathologies which would weaken bone. Investigations performed on the model and on the generated femur geometries and material properties used in the current study do, however, show that realistic data was produced which was a fair representation of the training set. Ideally separate models would be generated for different genders, ages, ethnicities and pathologies, as these are known to affect femoral geometry and bone density [103, 143]. Therefore biases in the training set could, in theory, influence the statistical result of some geometric parameters.

The finite element analysis performed on the data was relatively simple, modelling bone - a material with anisotropic, non linear mechanical behaviour - as an isotropic, linear material and calculating the effect of static load. This simplification follows that of the study being replicated and greatly reduces the complexity and computational cost of the simulations. A later study by Keyak [155] repeated the computational simulations incorporating non linear behaviour, although still using isotropic material properties, stance loading the model over a number of displacement steps until a failure criteria was reached. A significant but small increase between predicted and actual fracture load correlation was seen. Since the aims of this study were to look for failure trends and not accurate fracture load prediction, the gains of this technique were outweighed by the additional computational complexity. In addition, the linear method and the fact that an impact rather than progressive load was being modelled, meant that the precise value of load applied was not crucial to the result. The load chosen, $1\times$ bodyweight, was a realistic value for a fall and proved sufficient to highlight an 'at risk' group from the data set. A further simplification to this FE analysis was the lack of inclusion of muscle forces, surrounding tissues and impact surface. Again this was possible as the work was aiming to replicate Keyak's previous work in order to validate the results produced, so showing that the model would replicate the same trends and not to accurately simulate real world scenarios.

This case study has shown the potential of this methodology to generate large numbers of models which describe the variations present in the data used to create it. The ability to characterise the population wide variability potentially has useful applications in both computational-experimental analysis and clinical settings. Keyak et al. [150] is a good example of the type of experimental-computational work which could be enhanced by incorporating this statistical modelling technique, where relatively small number of cadaveric femurs were tested, 18, and compared to computational models. If the statistical model was used to replicate the experimental

test results accurately, the model could then be extended to a wider population of femur models with some confidence. Another possible use of being able to run such large scale simulations is the ability to gain an understanding of how factors affect a population, such that parameters taken from any patient can be compared to these to see how they fit. This could give a more sophisticated indicator of risk factors than current methods such as the World Health Organisation's arbitrary cut off, set at 2.5 standard deviation from the mean, to quantify predictions for osteoporotic hip fracture.

4.5 Conclusions

This study of femoral fracture has shown that the previously generated statistical femoral model is capable of being used to drive a large scale, multi femur analysis. The process was fully automated and completed within ∼12 minutes per femur. The results extracted from this 1000 femur analysis are meaningful, in that they agree with trends highlighted in existing clinical and computational studies. However, due to the limitations of the model's training set, it would be unwise to try to draw population wide conclusions at this stage. There is clear potential for this to be possible in the future, suggesting statistical models could be used as the means of incorporating interpatient variability into computational analysis.

Chapter 5

Automated Implantation of Femoral Resurfacing Implant [1]

5.1 Computational Analysis of Femoral Resurfacing Implants

Computational analysis of orthopaedic implants has become commonplace in biomedical research and preclinical implant testing. A severe limitation of the vast majority of studies is the use of a single, often generic, bone model - failing to acknowledge the differences in bone geometry and quality seen between people. This interpatient variability has been recognised as having a significant influence on results in all areas of assessment from clinical data to experimental and computational analysis, and in a range of studies whether investigating implant performance [31, 52, 6, 3, 7, 37] or natural occurrences such as falls [51, 150, 156].

The previous two chapters have shown how the statistical model of the femur developed in this work has been able to produce a wide variety of bone models, representing interpatient differences in morphology and bone quality. In addition it has been shown how this statistical model can be used to drive a large scale analysis of fracture risk in the intact femur. This was possible due to the computational analysis being fully automated. Femoral implant assessment would benefit from the ability to incorporate intersubject variation to better establish performance in the patient

[1] Bryan, R., Nair, P.B., and Taylor, M. An Automated, Large Scale Finite Element Study into the Influence of Femoral Head Resurfacing on Load Transfer in the Proximal Femur, submitted to Journal of Biomechanics: October 2009 - based on the work in this chapter

population. In particular femoral head resurfacing success is known to be sensitive to patient selection. The procedure is regarded as an attractive option over total hip replacement for younger, active patients as it offers greater bone conservation and more natural joint kinematics [157]. It has proved markedly less successful in older, female patients.

In vivo, femoral neck fracture is known to be a common reason for revision following Resurfacing Arthroplasty (RA), often occurring in the first few months after surgery. Computational studies have highlighted two possible issues following resurfacing which could act as failure modes; namely bone resorption due to stress shielding and increased strain in the neck around the proximal implant rim. However there is debate as to the extent to which these mechanisms influence failure, with some believing factors relating to surgical errors, such as notching, disrupted blood supply or poor positioning, are the major cause of early failure. With evidence showing that load transfer is altered by the insertion of a resurfacing component the ability to model a wide variety of femur representations could help to indicate whether there are patient traits which make failure more likely or whether it is more probable that surgical influences are the key culprit. The statistical modelling approach allows sufficient numbers of analyses to be run for meaningful statistical interrogation to be performed, which is not feasible in traditional computational studies.

Aim of Automated Implantation Study

The aim of the work described in this chapter is to establish a robust methodology for allowing fully automated computational testing of orthopaedic implants over a large number of models. In addition, it is hoped that this will demonstrate a feasible approach to using statistical modelling techniques in implant analysis. The adopted study will carry out a multi femur FE study into the influence of femoral head resurfacing on load transfer through the proximal femur. To achieve this, a system for correctly sizing and positioning the implants into any given femur geometry will need to be established which does not require any manual intervention, as well as a method of correctly performing the Boolean operations on the femurs to simulate surgical 'cuts' and generate a cement layer between the bone and implant.

In order to evaluate the success of the developed automated implantation methodology the study will emulate the analysis performed on 16 femurs by Radcliffe and Taylor [7]. In this previous work each example femur was analysed in both a resurfaced and intact configuration, relying heavily on manual model generation and implantation. The trends in strain change seen following femoral resurfacing were then

investigated using FEA. If similar trends are observed in the current work using a generated, automatically implanted set of femurs then this would help to suggest that the methodology is sound. In addition, it is hoped that testing a large number of models will allow statistically meaningful interrogation of femurs with outlying strain results. Therefore it will be possible to identify any geometric or material characteristics which are related to unusual strain patterns, particularly those in areas which have been related to failure such as the femoral neck and head.

5.2 Resurfacing Arthroplasty

5.2.1 Brief History of Joint Replacements

Joint replacements are primarily regarded as a final treatment option for degenerative diseases such as arthritis and osteoporosis. These diseases are often age related, therefore it is unsurprising that with the increasing number of people over the age of 60 in the population these types of procedures are becoming more and more common. Joint replacements may also be considered for the treatment of other conditions, as discussed in Section 5.2.2.

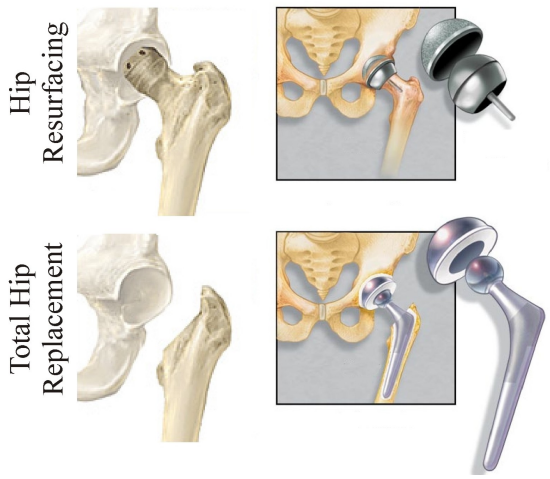


Figure 5.1: Illustration of resurfacing hip replacement (top) and total hip replacement (bottom), showing required surgical cuts (left) and representation of successful implantation (right). Adapted from images available at www.mayoclinic.com and www.portlandhipresurfacing.com.

The most prevalent, long standing and successful type of joint replacement is hip arthroplasty (although in recent years knee replacement rates have increased to match those for hips and even exceed them in some countries). Sir John Charnley is regarded to have pioneered what is recognisable as a modern Total Hip Replacement (THR) in the 1960's, developing a prosthesis consisting of a metal stem with a small ball acting as a femoral head which was fixed into the femur and an Ultra High Molecular Weight Polyethylene (UHMWPE) cup, which was fitted into the acetabulum (Figure 5.1). The components were secured in place using medical bone cement, made from a polymer called polymethyl metharylate (PMMA), which acts as a grout or putty in preventing movement

between the implant and bone. Cementing has the advantage of making the joint weight bearing almost immediately, avoiding a period of immobility while the bone 'heals' to stabilise the components. This is particularly useful in elderly patients who may never rebuild the muscle strength and mobility lost if subjected to prolonged inactivity. However problems can arise due the use of cement. In the short term thermal damage can occur in the surrounding bone during the exothermic cement curing process and in the longer term the cement mantle can suffer brittle failure. This is more common in active patients and can result in implant loosening.

In the 1980's cases of osteolysis were seen to be rising, this is where pockets of bone are resorbed by the body leading to a loss of bone stock around components. The condition was blamed on bone cement and thus developments were made towards uncemented implant designs, relying on bone in-growth into the prosthesis surface to provide stability [158]. This was encouraged by the inclusion of textured surfaces and the use of hydroxyapitite coatings. Bone cement was not abandoned as its omission did not reduce osteolysis rates and cementless designs had several drawbacks. Firstly they were bulkier, requiring greater initial loss of bone stock, and secondly in order to stabilise the joint it must be unloaded for a period of time so bone in-growth can occur. These issues make them unsuitable for some patients, specifically those who are older with poor bone quality where the likelihood of cement failure is low. A further alternative aiming to optimise the benefits of cement is the hybrid configuration, where the acetabular side is cemented but the femoral side is not (a reverse hybrid being the opposite way around).

It is now understood that osteolysis is due to the body's reaction to wear debris, namely particles of UHMWPE. In light of this it follows that bearing surfaces have been a focus of development in order to reduce the wear debris produced by the conventional metal ball on polyethylene socket. Improvements in wear properties were made by using either a ceramic ball, a metal on metal bearing surface or, with the lowest wear rate, ceramic on ceramic. The advances in metal on metal bearing surfaces allowed resurfacing designs, originally trialled unsuccessfully in the 1960's, to be implemented (fig. 5.1). These showed reliably good results in the patient group least satisfied by conventional THR, the young and active. The minimal bone resection and large bearing diameter allow more normal movement and load transfer through the femur and relatively straight forward revision options for conversion to a THR when required [157].

5.2.2 Reasons for Joint Replacement and Procedure Rates

Several national joint registries exist around the world, each collecting valuable patient data relating to a range of aspects surrounding joint replacement. The data stated in the following section have been taken from the 2008 Australian Orthopaedic Association National Joint Replacement Registry [61], the 2008 National Joint Registry for England and Wales [159] and the 2007 Swedish Hip Arthroplasty Register [160].

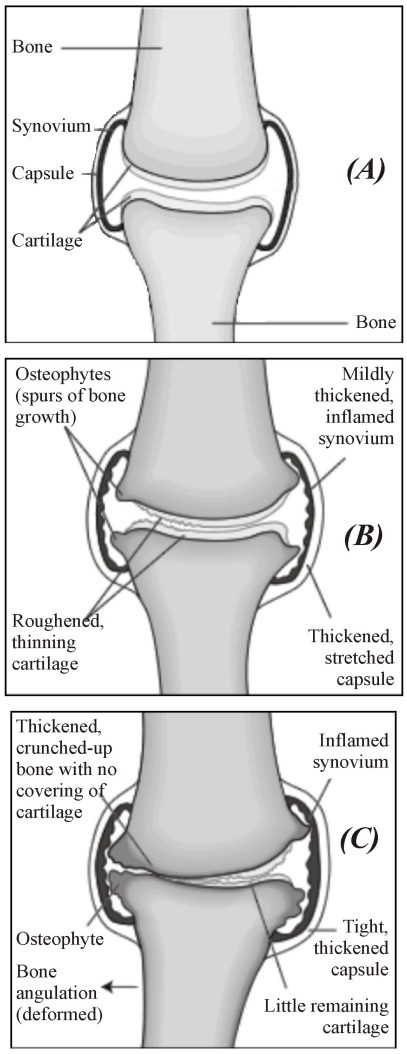


Figure 5.2: Illustration of how the progression of osteoarthritis affects a joint, from healthy (A) through mildly affected (B) to severely osteoarthritic (C), images adapted from Campaign [161].

Joint replacement is an irreversible procedure involving traumatic and highly invasive surgery. As a result it is only performed as a last resort when a patient's pain and/or mobility reaches such a point as to be seriously detrimental to quality of life. If it is possible alternative treatments are adopted in preference to surgery, these include; physiotherapy, to improve strength and mobility, and pharmacological treatments to control inflammation and pain. However it is usually when such treatments are no longer able to manage pain that joint replacement is considered.

By far the most common reason for hip replacement is the treatment of osteoarthritis, reportedly lying behind 78-95% of primary THR and 94-95% of hip resurfacings. Additional reasons include avascular necrosis of the femoral head (2-4%), femoral neck fracture (2-11%-typically unsuitable for treatment with RA), dysplasia (1-3%) and rheumatoid arthritis (1-2%). Osteoarthritis is a degenerative disease which leads to the damage of joint surfaces and changes in the underlying bone, currently causing approximately 2 million people a year to seek medical attention in the UK alone [161]. The disease is characterised by roughening and thinning of the articular cartilage which facilitates smooth joint motion and osteophyte formation, boney growths at the joint edges. Externally the effects of this noticeably alter the joints appearance as the synovium

swells while the capsule and surrounding ligaments thicken and change shape. The effect of osteoarthritis is depicted pictorially in a generic joint in figure 5.2. In severe cases of the disease the cartilage will often be worn away completely, leaving bone rubbing against bone which causes extreme pain, while the soft tissue changes and swelling drastically limit mobility. The disease is commonly associated with age, as joints become worn, healing processes slow and muscle strength decreases. However, there are other factors which can increase the risk of its development. After age the next most notable indicator is gender, women are more likely to suffer than men, as well as underlying hereditary factors, obesity and previous joint injury in decreasing order of importance [161].

5.2.3 Failure of Femoral Resurfacings

Despite extensive testing and development of prosthesis design, material and implantation methods, there are still a notable percentage of hip replacement failures. Failure does not only relate to catastrophic breaking of the implant or bone surrounding it, but also covers those cases where function is so poor or pain so severe that reoperation is required. Revision rates vary depending on a combination of factors including implant type, age and gender, with surgical error also having an influence. Revision rates for resurfacings are a little higher than those seen in THR (reported rates for resurfacing are stated to be 2.8% in UK [159]). This is indicative of the younger more active age group involved. However, with suitable patient selection alongside appropriate and skilful surgical technique, survivorship is comparable to conventional THR [162].

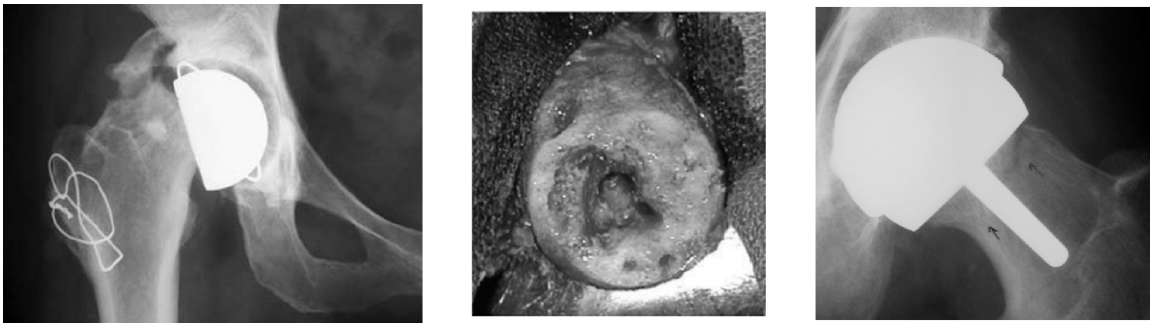


Figure 5.3: Radiographs of femoral resurfacing components showing: left - neck fracture with underlying aseptic loosening [163], center - degeneration of bone stock in femoral head beneath resurfacing component [164] and right - notched femoral neck (superiorly) with fracture propagating from the superior head-neck junction [162].

The most common reason for RA, and THR, revision is aseptic loosening, 45-68%. This is an umbrella term which describes loss of fixation leading to implant

instability when no signs of infection are evident (fig. 5.3-left and center). There are several main failure mechanisms which are recorded in this bracket;

- Osteolysis is a biological response to wear debris. Wear particles originating from sources such as a UHMWPE cup, worn cement mantle or metallic debris, infiltrate the bone-implant interface they can be attacked by the bodies natural immune system response. This engulfs the particles at the expense of the bone at the boundary, which is transformed into a far less stiff and less supportive soft tissue, reducing implant fixation in these areas.
- Failed bonding can occur if sufficient integration at the bone-implant (if uncemented) or bone-cement interface is not achieved. The result is excessive micro-motion of the implant which can, if exceeding 150μm, lead to bone resorption and damage accumulation at the interface. In addition cracks can then provide a path for wear debris to reach the bone interface and accelerate the process through osteolysis.
- Stress shielding occurs as a result of load paths through bone being altered by the implantation of a prosthesis. Due to the relative stiffnesses of the metal implant and bone a greater proportion of joint load is transmitted through the implant, thus unloading the bone which can lead to resorption. Stress bypass is similar although is often associated with poor proximal and good distal fixation, thus unloading the proximal bone which is resorbed. Both scenarios can lead to weakening of the bone, reduced implant fixation and a risk of fracture.
- Remaining notable failure mechanisms are dislocation, deep infection, fracture (of the bone or implant), pain and component wear.

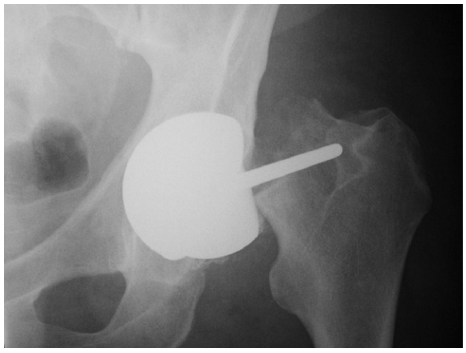


Figure 5.4: Radiograph of femoral neck fracture following resurfacing procedure [165]

According to the 2008 National Joint Registry for England and Wales [159] revision rates for resurfacing show there was a notable rate of fracture (25%), pain (23%) and a relatively low occurrence of dislocation (5%). Femoral neck fracture is a concerning complication which, due to the nature of THR, is a unique risk to resurfacing arthroplasty (fig. 5.4). The prevalence of neck fracture is indicated to be even greater by the Australian Orthopaedic Association National Joint Replacement Registry [61] which states that neck fracture accounts

for over 40% of resurfacing revisions, while Shimmin and Back [166] quote the incidence of fracture at ∼2%. Clinically, neck fractures are often seen within the first few months of surgery [165, 167].

There is debate as to the dominant reason behind fracture risk, as ever it is likely to be a multi-factorial problem where combinations of observed influences contribute [45, 168, 4, 169, 170, 166, 162, 165, 171]. Often damage done during surgery or poor surgical technique are blamed for weakening the femoral neck. While preparing the femoral head there is potential to cut/disrupt blood supply which would lead to necrosis (death of the blood starved bone), this would then be resorbed and effectively disappear as a load bearing entity or surface for fixation. Notching of the neck, occurring proximally through inaccurate cutting of the femoral head, is also blamed for establishing a stress concentration liable to be the source of fracture (Figure 5.3). Amstutz et al. [172] describes the link between poor implant seating and neck fracture, where any exposed reamed bone would act as a stress riser which could lead to fracture. Damage to the underlying bone has also been observed to occur through trauma during impaction and thermal damage from cement curing. Surgical variability is recognised as influencing fracture and potential range of motion. The related surgical factors include; implant malalignment (varus) which is seen to lead to less natural load transmission through the proximal femur, incorrect implant sizing and poor cementing - where uneven, insufficient or too generous cement layers can accelerate failure [172, 4].

Strain shielding following resurfacing has been linked to bone resorption in the femoral head, and potential failure through head collapse and loosening. Evidence of resorption has been observed in clinical retrievals [45, 168, 173]. However, this failure mechanism has been widely accepted as not being the main reason for loosening or fracture. Yet it is logical that pockets of resorbed bone will weaken the underlying bone structure which could lead to elevated stresses and so fracture (Figure 5.3). This process would take time to evolve, so it is unlikely that it would lead to the short term failures observed in many resurfacing neck fractures. If a patient already has poor bone quality or cysts then the chances of successful resurfacing are reduced - thus appropriate patient selection is very important.

5.2.4 Computational Studies of Femoral Resurfacing Arthroplasty

Resurfacing hip replacements have not escaped the computational interrogations which have become widely used through bioengineering research and development.

There have been a range of finite element studies since the mid 1980's to analyse the strengths and weaknesses of the design concept [174]. In 1985 Huiskes et al. [45] found that their FE model of a resurfaced proximal femur did not predict the physiological strain patterns which were expected to result from this implant design. Their model showed a significant proportion of joint load being transmitted through the implant to the bone at its superior rim, as opposed to directly through the head to the medial cortex. It suggested that this change in load path would inevitably lead to stress shielding in the head and resorption, particularly at the bone-implant interface, thus providing a possible mechanism explaining the observed early RA failure through loosening. Watanabe et al. [175] generated a three dimensional finite element femur model in an intact and an implanted form, allowing comparison of the stress and strain patterns of both when subjected to identical loading conditions relating to ambulation with crutches. This study again showed high stress at the implant rim and stress shielding in the antero-superior region beneath the implant, which could lead to long term loosening.

These early studies were simplified, using generic or single bone representations and describing bone material as either cancellous or cortical with a single modulus value assigned to each. Later work has increased in complexity, using patient specific models with geometry and material properties infered from CT, although on the whole still only using a single model. An example is Taylor [84], where the changes in load transfer through the proximal femur after RA were examined - in particular the influence of the metaphyseal stem (diameter and degree of contact) and cement mantle thickness. Rather than investigating the femur in isolation, Ong et al. [49] developed a patient specific model incorporating the pelvis and sacrum as well as the femur. This model was then used to investigate the influence of fixation and interface conditions on stress and strain alteration through the femur. Radcliffe and Taylor [37, 7] applied a multi model analysis to two aspects of resurfacing surgical variability: cementing technique and varus-valgus alignment. Their analyses each used 16 manually generated and implanted patient specific models, incorporating the influence of different geometries and material properties.

The results of the majority of recent computational studies agree on the main features of load alteration following RA, namely strain shielding in the superior head and elevated strain in the femoral neck with peak strain at the implant rim. There is also agreement that excessive, insufficient or uneven cement mantles are detrimental to survivorship as is a varus implant alignment [174]. However, as with the clinical data, there is no clear consensus as to whether these changes are sufficient on their own to initiate a failure mechanism or whether an additional influence is required. These could include poor patient selection, where low bone quality leads to

poor results, or an accidentally introduced notch in the neck. Although Long et al. [167] argue that short term neck fracture is not always accompanied by evidence of notching, suggesting that high strain concentrations at the implant rim may cause damage accumulation leading to fatigue failure. The influence of interpatient and surgical variation has rarely, if at all, been incorporated into these analyses yet it is widely felt that differences in bone quality, geometry and implant positioning will affect implant performance.

5.3 Methodology

Establishing a methodology for the automated generation, implantation and analysis of a large number of different femur geometries required the solution of a range of problems which will be discussed as the stages of the process are explained. The study aimed to use the statistical model developed in Chapter 3 to generate a population of 400 unique, realistic FE ready femur models to examine the change in strain in the proximal femur following resurfacing (this number was chosen to satisfy the time restrictions the study was subject to). In order to achieve this the entire process had to be robust, reliable and fully automated from initial model generation through to post processing. The stages necessary to achieve this are as follows, all of which were controlled using Matlab:

1. Sample the statistical shape and intensity model to generate 400 FE femur models using the sampling boundaries described in Chapter 3.
2. Analyse each femur to calculate geometric measures and landmark points, indicating model specific implant size and position.
3. Perform Boolean operations to simulate the cuts made to the femur during surgery and to create a cement layer between the bone and implant, and finally position the implant.
4. Remesh the newly cut femur and reallocate elemental material properties.
5. Apply model specific loads to simulate level gait.
6. Run each model through an FE solver and extract resulting elemental volume and strain results in both their implanted and natural state.
7. Post process the strain results to calculate the strain changes between the intact and implanted instances, as well as the geometric and material characteristics of the femurs.

5.3.1 Generation of a Set of Femurs for Analysis from a Statistical Model

A statistical model of the whole femur was generated using principal component analysis from a training set of 46 available CT scans. The set consisted of 15 females and 31 males aged between 43 and 91, describing a wide range of femoral morphologies (as described in detail in Section 3.5.1). The statistical model was able to generate synthetic femur models as high quality FE-ready, tetrahedral meshes with element specific material properties. As detailed in Section 3.6, the statistical model was used to produce a population of femurs by creating new PCA weightings based on the training data. This was done by varying the weightings between $\pm 3\sigma_{b_m}$, where σ_{b_m} was the standard deviation of the individual shape parameters across the training set, using statistically independent random numbers drawn from truncated Gaussian distributions. The created models reproduced the variability in the set whilst maintaining realistic bone geometries and material property distributions. Automatic mesh quality checking was incorporated into the model generation process to ensure high mesh quality was maintained.

5.3.2 Technique for Virtual Implantation

Model Specific Alignment and Sizing

In order to create a totally automated process for modelling femoral resurfacing it was key to establish a system for identifying anatomical landmarks by which a set of reference points could be constructed to position an implant. It was also essential to be able to measure the femur to choose the correct prosthesis size to fit. The starting point for characterising each femur was therefore made by specifying groups of nodes and elements marking areas of the proximal femur, i.e. the femoral neck and head, on the baseline geometry. This was possible despite each generated femur having a unique geometry because they were all described by an identical mesh with nodal and elemental correspondence. Using the sets of elements defining the femoral head and neck, model specific positioning points were established by fitting three spheres using a least squares method.

The first sphere was fitted to the femoral head, the center of this sphere indicated the natural head center and its diameter was used to make an initial guess at the correct implant size (Fig. 5.5a). The second and third spheres were fitted within

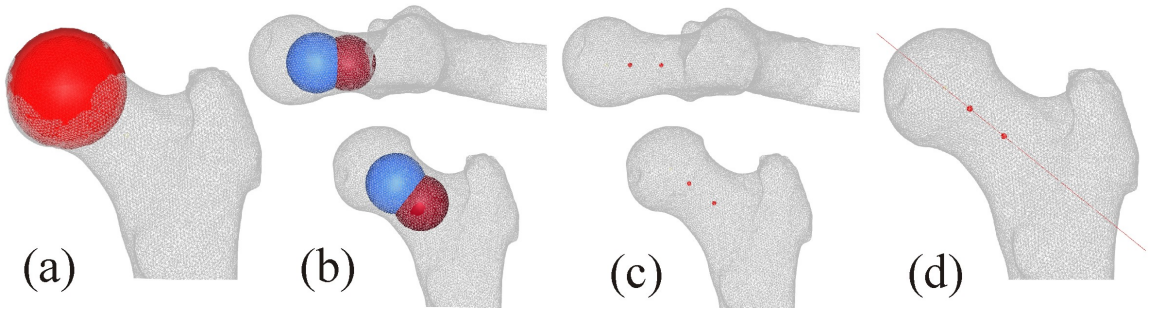


Figure 5.5: Illustration of methodology used to identify implant alignment by fitting spheres to the proximal femur. (a) sphere fitted to the femoral head - indicating femoral head diameter and natural head center, (b) two spheres fitted to the upper and lower parts of the femoral neck, (c) the centres of the two spheres are found and (d) indicate the neck axis line.

the femoral neck, one to the proximal side and one to the distal. The centres of the spheres were joined to indicate the femoral neck axis (Fig. 5.5b-d), this axis was used as the neutral axis for the femoral implant thus fixing its alignment. By translating the natural head center onto the neck axis to give the implant head center the final position of the implant was set.

A concern with the technique was that it would result in notching of the femoral neck, as a result of poor alignment or incorrect implant sizing. Therefore once the implant size was selected and its position fixed the configuration was checked to ensure that notching did not occur. This was done by calculating the neck diameter at the proposed location of the implant rim and comparing this to the implant rim inner diameter. As the study did not intend to incorporate implant positioning as a variable, instead maintaining alignment along the neutral axis for all implants, this problem was solved by altering the size of the component. If the difference between the neck and implant rim diameters was too great then a step up or step down in size was made and the measurement repeated.

Boolean Operations- Simulating Surgery

Hypermesh^{TM}(Altair Engineering Inc, USA) was used to perform the Boolean operations required to carry out virtual implantation. This process was automated using a command script written for each femur individually, containing the implant alignment coordinates found in the previous step and the details of which implant size to import. In addition to a geometric model of the implant, a cutter geometry was also imported which was matched to the relevant implant size. Each cutter consisted of two parts: an outer profile and a stem guide, which produced the external head

profile needed for resurfacing. The final geometry imported was a cement layer blank which would be cut to match the femur at a later stage, resulting in a 3mm thick layer between the implant and bone with no cement present around the implant stem which remained out of contact with the surrounding bone.

The statistical model generated femur representations as solid, tetrahedral meshes which could be imported directly into Hypermesh. However, Boolean operations on meshes were not possible so the models were converted into solid geometries. In order to ensure this process was robust it was done in three steps; shrink fitting a surface mesh to the femur, then fitting a series of geometric surface patches to the surface mesh before finally converting the surface patches into a solid geometry. To reduce computational complexity the femur was partitioned one third of the way down its length. This allowed a higher mesh density to be specified in the proximal femur (the region of most interest) and a lower mesh density in the distal femur, thus reducing the overall size of the model for later FEA.

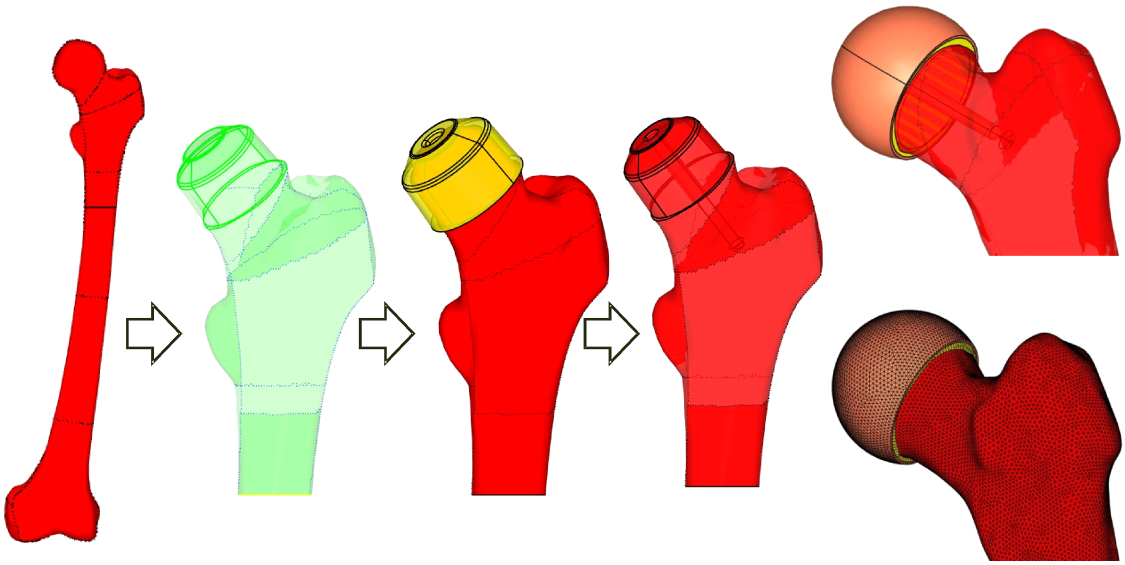


Figure 5.6: Images of implantation process. From left to right: Converting femur to a solid geometry and splitting the femur into proximal and distal sections, cutting femoral head external profile, generating cement layer, cutting the guide stem hole, position femoral implant and finally meshing components.

Figure 5.6 shows the stages of Boolean operations carried out after aligning the implant and cutters over the femur. From left to right: the external profile of the femoral head was cut using the first part of the implant specific cutter, the cement blank was then aligned and its internal surface cut using the cut femoral head (the external profile matches the implant) and finally the stem guide was cut into the femur. By performing the steps in this order it made certain that the three components fitted together and cement filled any defects in the femoral head even if they were

deeper than the 3mm target thickness. At this stage the implant was positioned, leaving the implant, cement mantle and bone as solid geometries. Each of the four parts of the model were then converted into surface meshes so that they could be exported to dedicated meshing software and converted into the final solid, four-noded tetrahedral meshes needed for analysis.

The Boolean stage was the most vulnerable to failure due to the complexity of the operations being performed on the intersecting shapes, especially as the femoral head was far from a regular geometry. It was found that these errors preventing Boolean cuts could often be solved by applying minor transformations to the cutter/implant geometries' positioning. A fail safe loop was therefore coded into the process to shift the imported geometries 0.001mm in the x,y and z directions if an intersection error caused the code to crash. This could be performed up to five times before the geometry was abandoned and the next model was attempted. The stage of converting an intact model to a set of surface meshes representing an implanted configuration was processed in between 6 to 10 minutes, depending on the model in question and if any implant shifts were required.

Remeshing and Reassignment of Material Properties

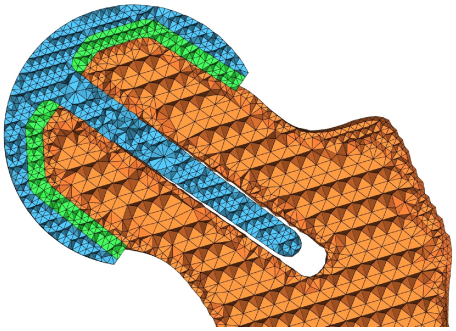


Figure 5.7: Cross section through a generated, automatically implanted femur after remeshing, illustrating the changes in mesh resolution from implant to cement to bone.

The four surface meshes representing the implanted model as proximal femur, distal femur, cement mantle and implant were imported into Ansys ICEM CFD^{TM}(Ansys Inc., Cannonsburg, PA). This software was scriptable and allowed the components to be meshed with different element densities while retaining rigidly bonded interfaces. This reduced computational cost in later FEA by limiting the ultimate number of elements in the model and by avoiding the need for contact analysis, thus reducing later computational complexity which would be introduced by having to model contact in order to join the components together. The distal femur was meshed with 2-4mm elements, the proximal femur with 0.5-1.5mm and the implant and cement layer with 0.75-1.5mm elements (Figure 5.7).

As the femur had been cut and completely remeshed from the original instance generated by the statistical model it was necessary to reassign material properties

to the new elements. This was done by interpolating the original model's modulus at its nodes onto the new bone nodal positions, from which elemental modulus could be determined by averaging over the four nodes making up each element. The interpolation used an in-built Matlab three dimensional data gridding function - griddata3. The implant was given a modulus of 200,000 MPa and the cement mantle 2,800 MPa [37]. The Ansys remeshing process and material reassignment was completed in ∼1-2 minutes per model.

5.3.3 Subject Specific Finite Element Analysis

Figure 5.8: Illustration of the loading condition defines by Heller et al. [33]

A static, linear elastic FEA was performed on each synthetic femur in both implanted and intact forms. Identical, subject specific loading conditions were applied to simulate level gate following the patterns described by Heller et al. [33] (figure 5.8). These loads are prescribed in terms of body weight, which was unknown for the synthetic femurs and therefore had to be estimated. This was done by predicting stature, using the 0.2675:1 ratio of femur length to height defined by Feldesman and Fountain [154], then randomly assigning a BMI from a distribution curve generated from data from the National Health and Nutritional Examination Survey (NHANES) 1999-2002 [153]. Subject weight, in kilograms, could then be calculated by multiplying this assigned BMI by the square of the predicted height in meters (a similar approach to applying subject specific loads was used in the femoral neck fracture study, Section 4.2.3).

As shown in Figure 5.8 each femur was constrained distally and subjected to two loads proximally, which applied a hip contact force to the femoral head and an abductor force to the greater trochanter [33]. The distal boundary condition was set by fixing the lowest 30mm of nodes on the femur surface. The contact force was centred on the highest point of the femoral head and evenly distributed over a 20mm diameter area, the abductor force was also evenly distributed over an equivalent area centred on the most lateral node of the greater trochanter.

The models were now fully defined and ready for analysis. The model geometry, described by nodal coordinates and element connectivity, with associated elemental modulus and boundary conditions were written out as an input file for Marc$^{\circ c}$. The processes of writing this file, performing the FE and saving the results were completed in ∼10 minutes.

5.3.4 Automated Post Processing

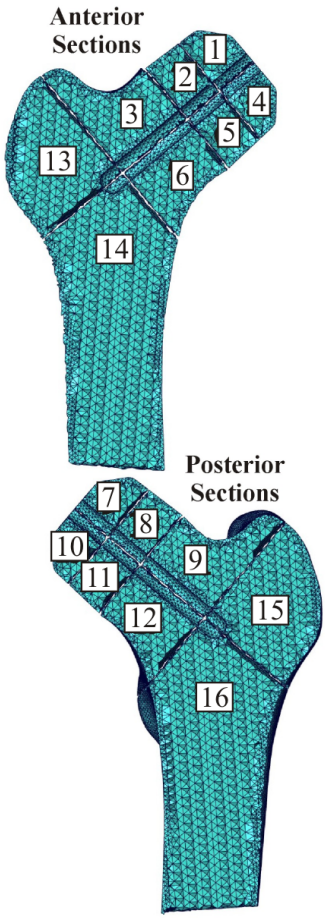


Figure 5.9: Illustration of the sections defined in each femur to allow the changes in strain through the proximal femur to be analysed.

Post processing extracted the elemental strains and elemental volumes recorded for each analysed model from both the intact and implanted configurations. The change in load transfer through the femur which occurred as a result of resurfacing was assessed by comparing these results, however this could not be done directly as the implanted configuration had been cut and remeshed. The approach devised was to interpolate the strains from the intact femur onto the implanted femur mesh by nodal position, allowing the same volumes and areas of bone to be directly and fairly compared.

To obtain a full picture of how strain varied through the proximal femur, 16 sections were automatically defined in the implanted model using the neck axis identified for implant alignment and planes along and perpendicular to it. This gave an anterior and a posterior section, each of these having 4 superior and 4 inferior portions (Figure 5.9). Planes were defined to cut through at the following points; the femoral head center, the base of the implant, base of the femoral neck and a horizontal cut one third of the way down the length of the femur. The sections partitioned by these planes were used to identify groups of elements and so the strain and volume information for these could be found, allowing the mean change in strain in each section to be calculated. Examining the data in this way allowed the potential for

further analysis to be done such as looking at the strain distributions through sections and volumes of bone exceeding strain thresholds, as well as potential for femur by femur comparisons for those models with extreme results.

In addition to the strain results a range of geometric and material property metrics were extracted to allow the performance of resurfacing to be compared to the characteristics of the femurs being examined. These were measured using element volumes and nodal markers identified on the baseline, as in previous work to study the range of variation in the femurs generated by the statistical model (Section 3.6.4) and in the femoral neck fracture investigation (Section 4.2.5).

5.4 Results

5.4.1 Assessment of Implantation Methodology

In total 400 models were successfully generated and analysed as both intact femurs and following a virtual femoral resurfacing procedure. The stages involved in this were; model specific sizing and alignment of resurfacing implant, Boolean operations to perform virtual implantation, remeshing and reassignment of material properties, calculation of model specific loading conditions, finite element analysis of the system and finally complete post processing of strain results and model characterisation. Both intact and implanted analyses, as well as post processing comparisons between the two, were performed at a rate of more than four models per hour (using an Intel Xeon X5355 2.66GHz processor). The methodology previously described was able to achieve this as a fully automated process with a success rate of $\sim$85% of generated femurs completing all stages through to fully processed strain results without error. Nearly all of the analyses which failed to complete came to a halt during the Boolean operations, these failures occurred randomly through the generated set. The small number of further failures were related to model name confusions on importing the components to Ansys ICEM for volume meshing.

Assessing the success of the implantation methodology was approached in three ways; firstly by visually inspecting each model to manually check that the sizing and orientation of the implanted component was acceptable, secondly comparing the distribution of implant sizes to those used clinically and thirdly by comparing the change in strain trends with previous work where implantation was performed manually. The later of these is described in the next section.

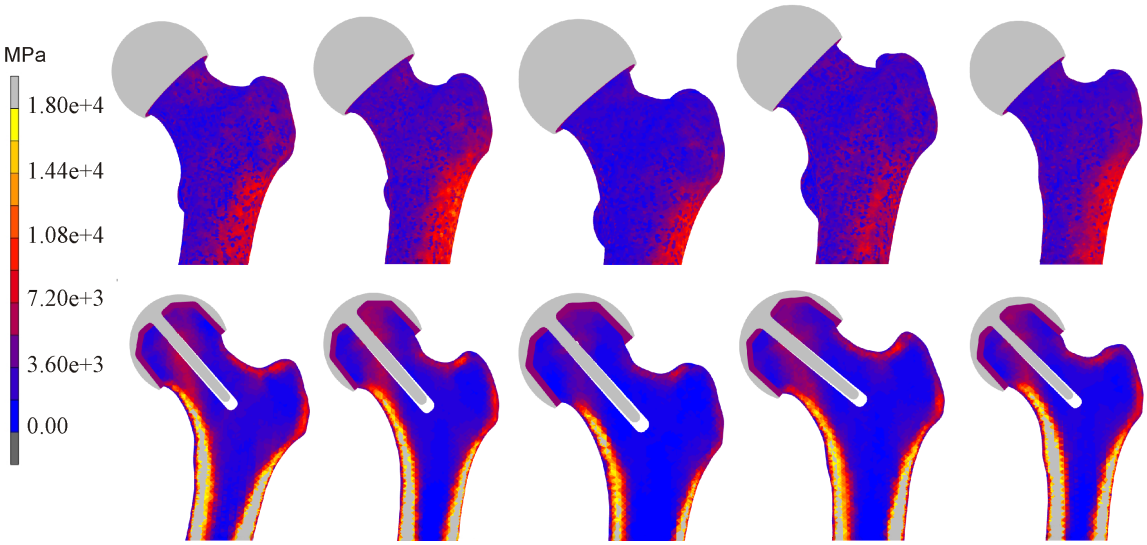


Figure 5.10: Illustration of a sample of generated femur instances of varying morphology with automatically sized, aligned and implanted femoral resurfacing components. Top: external view of each model. Bottom: section views through femoral head showing modulus distribution, implant and cement layer.

Manual inspection was potentially a tedious task, therefore to reduce this an automated script was run to capture screen shots of each processed model - an external view and sections showing modulus and strain distribution. Once generated it was straight forward to assess each model using these images. Figure 5.10 shows a set of randomly chosen resurfaced models with varying size and geometry as both an external view and as a section view through the proximal femur. The images indicated that the resurfacing procedure had been carried out appropriately in all cases.

The distribution of assigned implant sizes for this study and the proportion of global sales accounted for by implant size (excluding the US) are illustrated in Figure 5.11. The trends displayed in the two graphs relating to implant size popularity are very similar, although the extremes are different. As discussed previously (Section 3.6.4) the generated data set is slightly bigger than the population indicated by the NHANES survey, which is a reflection of the training data. The sales data, however, covers a global market and therefore a range of ethnicities. It is likely that this has resulted in a range of smaller patients, for example from Asia, who are not currently present in the statistical model.

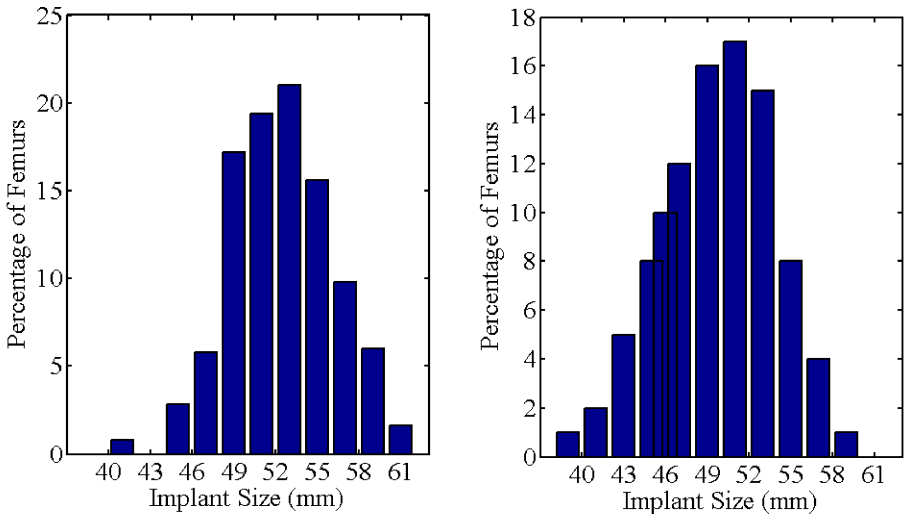


Figure 5.11: Graph illustrating the percentage distribution of implant sizes assigned by the automated methodology described to the femur population developed in this study (left), and the percentage of implant sizes used clinically (right). Clinical data is taken from global sales figures (excluding data from the US).

5.4.2 Predicted Trends in Strain Alteration Through Proximal Femur

Strain changes were calculated between the intact and implanted models in the 16 predefined sections making up the proximal portion of each generated femur, thus enabling the effect of femoral resurfacing on load transfer to be investigated. Figure 5.12 shows the average percentage strain changes found by Radcliffe et al. [176] in 16 manually implanted femurs (top) along with the strain changes observed in this study of 400 automatically implanted, synthetic femurs (bottom). Although Radcliffe sectioned the proximal femur into 20 parts it is still evident that both show a similar trend in strain change and a similar magnitude of change, in the majority of sections.

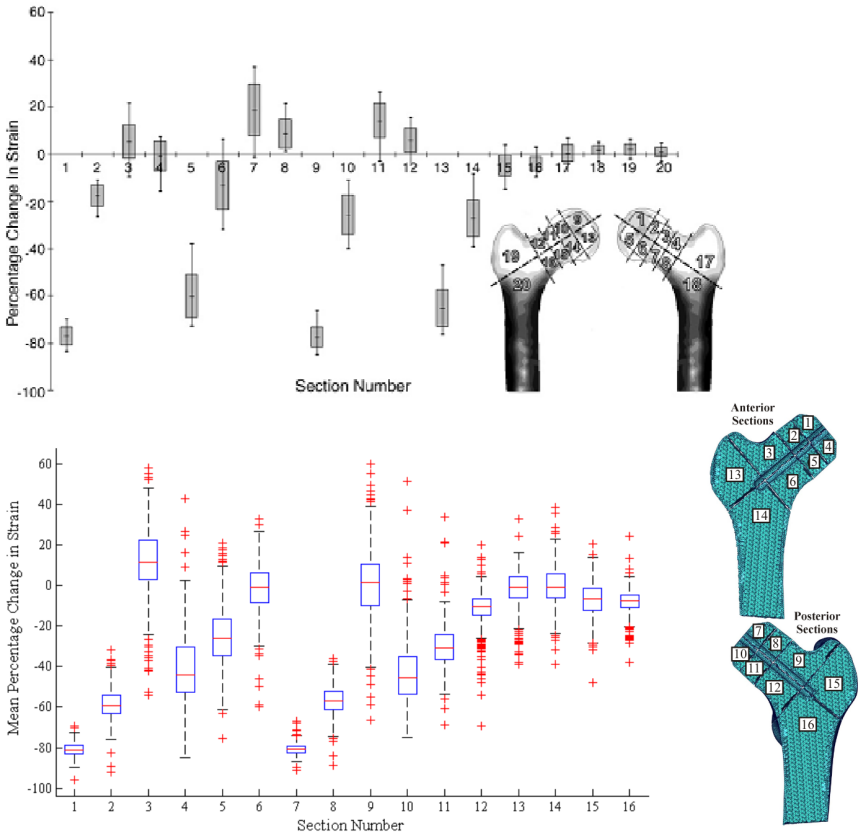


Figure 5.12: Boxplot of the average change in strain seen by Radcliffe et al. [176] across 18 proximal sections for 16 manually implanted femurs (top) and across 16 sections of the proximal femur illustrated for 400 automatically implanted, generated femur models (lower). Positive values indicate an increase in strain, negative values a decrease.

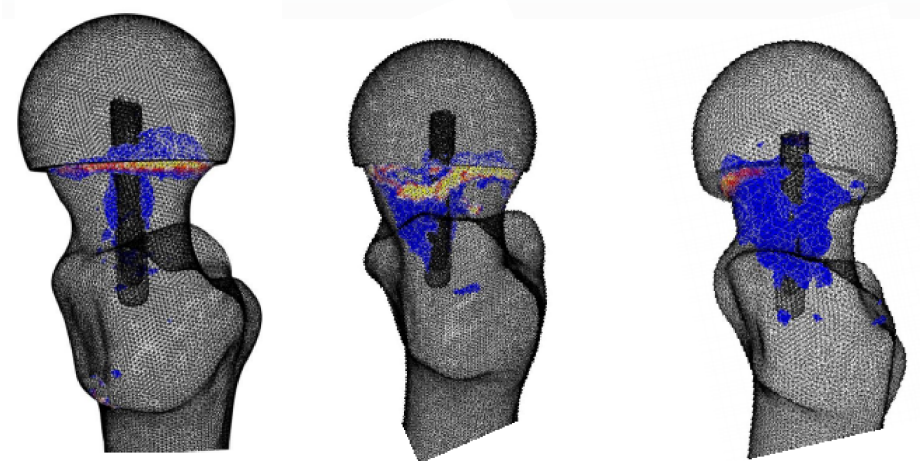


Figure 5.13: Images of highest strain regions following resurfacing, showing elevated strains through the femoral neck at the proximal implant rim and on either side of the implant stem.

The automated study showed that following resurfacing there was clear evidence of strain shielding within the femoral head in all models, most dramatic in the su-

perior sections 1 and 7 where average strain reductions of 81% and 59% were seen. The other main change in strain indicated was an increase in the superior femoral neck, sections 3 and 9, most notably in the anterior portion. On closer inspection this trend tended to relate to a high strain concentration at the proximal implant rim and elevated strains on the anterior and posterior sides of the implant stem within the femoral neck (Figure 5.13). Although the mean increases seen in the neck sections of this group were not large (11% in section 3, 1.5% in section 9) the largest strain increases observed were, at almost 60%.

5.4.3 Analysis of Outlying Strain Results

The greatest strength of the technique described in this chapter was that the large number of models run allows the possibility of analysing those poorly performing, vulnerable instances which would appear as outlying or extreme strain results. The femur areas of most clinical interest in relation to failure mechanisms following resurfacing are the femoral head and femoral neck, therefore models suggesting excessive bone remodelling in these sections were interrogated to attempt to find any significantly different material or geometric characteristics between these and the main group of femurs. Four outlier groups were identified, containing those femurs exceeding a two standard deviation increase/decrease from the mean change in strain in any of the femoral head or neck sections respectively, and comparisons were performed using one-way ANOVA tests. The outlier groups contained between 18 and 59 models, each group was analysed in respect to the extensive list of metrics listed in Section 4.2.5.

Examination of instances with outlying strain results in the head showed some notable differences between those femurs exhibiting greater or less strain shielding than the majority of the group, the metrics found to be statistically significant are detailed in Table 5.1. The femurs with the least strain shielding exhibited a higher percentage of cortical bone throughout (head $p<0.0005$, neck $p<0.025$ and intertrochanter $p<0.005$), the findings also indicate that they were bigger than the others (head volume $p<0.025$ and neck volume $p<0.025$). Another key set of results related to the proportions of the femur, where this outlier group showed a higher head-neck ratio (bigger head to neck, $p<0.00005$) and on average a round femoral neck, greater neck diameter in anterior-posterior dimension as compared to normal ($p<0.025$). The femurs experiencing the most strain shielding appeared to have lower mean cancellous modulus than the main group (head and intertrochantic $p<0.001$, neck $p<0.025$) and a lower percentage of cortical bone in the intertrochantic region ($p<0.025$). The analysis also suggested that this group were smaller than the rest (shaft radius $p<0.025$,

intertrochantic volume p<0.025), with more oval necks (larger in posterior-distal dimension p<0.01).

Minimum $\downarrow \varepsilon$ in head (n=59)	p-value	Outliers mean	Main Group mean
Neck-Head Diameter Ratio	<0.00005	1.24	1.18
Lower Head % Volume Cortical	<0.0005	5.95	3.27
Intertrochantic % Volume Cortical	<0.005	28.36	24.86
Femoral Neck Diameter A-P (mm)	<0.025	36.09	33.92
Femoral Neck Diameter Ratio	<0.025	1.00	1.03
Lower Head Volume (mm^3)	<0.025	17329.84	15762.39
Neck % Volume Cortical	<0.025	27.07	22.98
Neck Volume (mm^3)	<0.025	23940.93	21490.5
Maximum $\downarrow \varepsilon$ in head (n=18)	p-value	Outliers mean	Main Group mean
Intertrochantic Mean Cancelous Modulus (MPa)	<0.001	724.22	1158.93
Lower Head Mean Cancelous Modulus (MPa)	<0.01	1117.71	1415.10
Femoral Neck Diameter Ratio (mm)	<0.01	1.11	1.03
Femoral Shaft Radius (mm)	<0.025	14.71	16.39
Intertrochantic % Volume Cortical	<0.025	18.26	25.13
Intertrochantic Volume (mm^3)	<0.025	67952.23	86588.44
Neck Mean Cancelous Modulus (MPa)	<0.025	1073.13	1391.49

Table 5.1: Table of the mean results for the most significant metrics compared between the group of femurs with outlying results in the femoral head compared to the rest of of the femurs analysed, showing the results for two outlier groups, those with the most strain shielding in the head (maximum) and those with the least (minimum).

Examination of the outliers in the neck found a clear differences between those with increased and decreased strain. The main observation was that those showing an unusually large increase in strain in the neck were smaller in size over a comprehensive range of shape metrics; neck axis length (p<0.00025), femoral shaft radius (p<0.0005), neck diameter anterior-posterior (p<0.001), femoral length (p<0.0025), intertrochantic width (p<0.0025), head diameter (p<0.005), intertrochantic volume (p<0.0025), head volume (p<0.005) and neck volume (p<0.025). The most significant metric was neck diameter ratio (p<0.00001) indicating a more oval neck with a smaller diameter in the anterior-posterior direction, as with the group exhibiting the most strain shielding. No material metrics were found to have statistical significance.

In contrast, for those outliers where strain decreased in the femoral neck material properties did appear as significant features, showing lower mean cancellous modulus throughout (head p<0.005, neck p<0.0025, intertrochantic p<0.005) and a lower percentage of cortical bone in the head (p<0.05) and neck (p<0.025). In addition

Maximum ↑ ε in neck (n=37)	p-value	Outliers mean	Main Group mean
Femoral Neck Diameter Ratio (mm)	<0.00001	1.12	1.03
Neck Axis Length (mm)	<0.00025	92.04	101.30
Femoral Shaft Radius (mm)	<0.0005	14.73	16.41
Femoral Neck Diameter A-P (mm)	<0.001	30.15	34.16
Femoral Length (mm)	<0.0025	404.39	437.31
Intertrochantic Width (mm)	<0.0025	51.26	56.25
Intertrochantic Volume (mm^3)	<0.0025	66096.72	86719.38
Subject Height (mm)	<0.0025	1511.73	1634.79
Anteversion	<0.005	10.31	18.60
Femoral Head Diameter (mm)	<0.005	37.70	41.10
Lower Head Volume (mm^3)	<0.005	12080.67	15930.91
Subject BMI	<0.01	24.71	28.98
Subject Weight (kg)	<0.01	61.56	77.50
Neck Shaft Angle (mm)	<0.025	132.27	128.14
Neck Volume (mm^3)	<0.025	16949.45	21729.62
Neck-Head Diameter ratio	<0.025	1.231	1.18
Maximum↓ ε in neck (n=20)	p-value	Outliers mean	Main Group mean
Neck Shaft Angle (mm)	<0.00001	135.38	128.08
Neck Mean Cancellous Modulus (MPa)	<0.0025	1019.63	1393.378
Inter Mean Cancellous Modulus (MPa)	<0.005	873.22	1160.30
Lower Head Mean Cancellous Mod (MPa)	<0.005	1147.41	1415.25
Femoral Neck Diameter Ratio	<0.01	0.96	1.03
Neck % Volume Cortical	<0.025	16.06	23.30

Table 5.2: Table of the mean results for the most significant metrics compared between the group of femurs with outlying results in the femoral neck compared to the rest of the femurs analysed, showing the results for two outlier groups, those with the most reduced neck strain (maximum ↓) and those with the most increased (maximum ↑).

they showed a reduced neck diameter ratio ($p<0.01$) with a greater anterior-posterior neck diameter, again opposing the trend of the previously described neck outlier group. Neck shaft angle appeared as significant for neck outliers at each end of the spectrum, being higher than the rest of the group.

5.4.4 Influence of Implant Size on Load Transfer Through the Femur

Analysis of the change in strain in 400 femurs following femoral resurfacing indicated that the size of the femur influenced the strain change in the femoral neck, suggesting

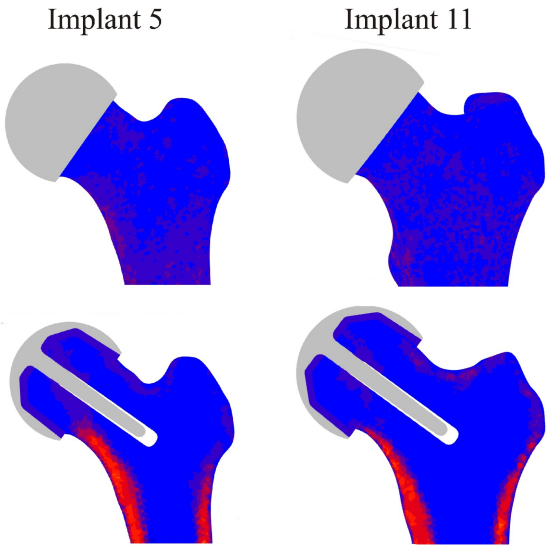


Figure 5.14: Illustration of both implant size 5 and 11 implanted into correctly sized femurs, showing external view and a cross section through the proximal femur.

that larger femurs showed reduced strain in the neck and smaller femurs suffered an increase. As previously discussed, femoral neck fracture is a significant failure mechanism following resurfacing so any factor which may increase strain in this area is of interest. Further to this, there have been clinical observations of poorer performance with smaller implant sizes as compared to larger [61, 172, 177]. In light of this two subgroups were extracted from the processed data set. These related to femurs with implant size 5 and femurs with implant size 11. i.e. two groups representing a small implant and a very large implant (see Figure 5.14). These two implant sizes were chosen as they corresponded to a reasonable number of femurs and accounted for approximately the same number of models, 25 and 20 respectively (Figure 5.11). Analysis was then carried out to investigate whether there where any differences in performance between the large and small implant groups, the method used was again a one way ANOVA test.

Figure 5.15 shows the results of the areas with statistically significant differences in the change in strain distribution following resurfacing. These plots clearly show two trends; firstly that the smaller implant size suffers greater strain shielding in the superior femoral head (anterior: section 1 $p<0.00001$, section 2 $p<0.005$ and section 4 $p<0.05$, posterior: section 7 $p<0.0005$ and section 8 $p<0.01$) and secondly that the smaller implant sizes show a greater increase in strain in the femoral neck (anterior-superior: section 3 $p<0.05$, posterior-superior: section 9 $p<0.0005$ and posterior-inferior: section 12 $p<0.01$). The actual value of strain change these relate to (mean and maximum) are detailed in Table 5.3. As with the outlier analysis a range of geometric and material metrics were taken from the groups, although due to the

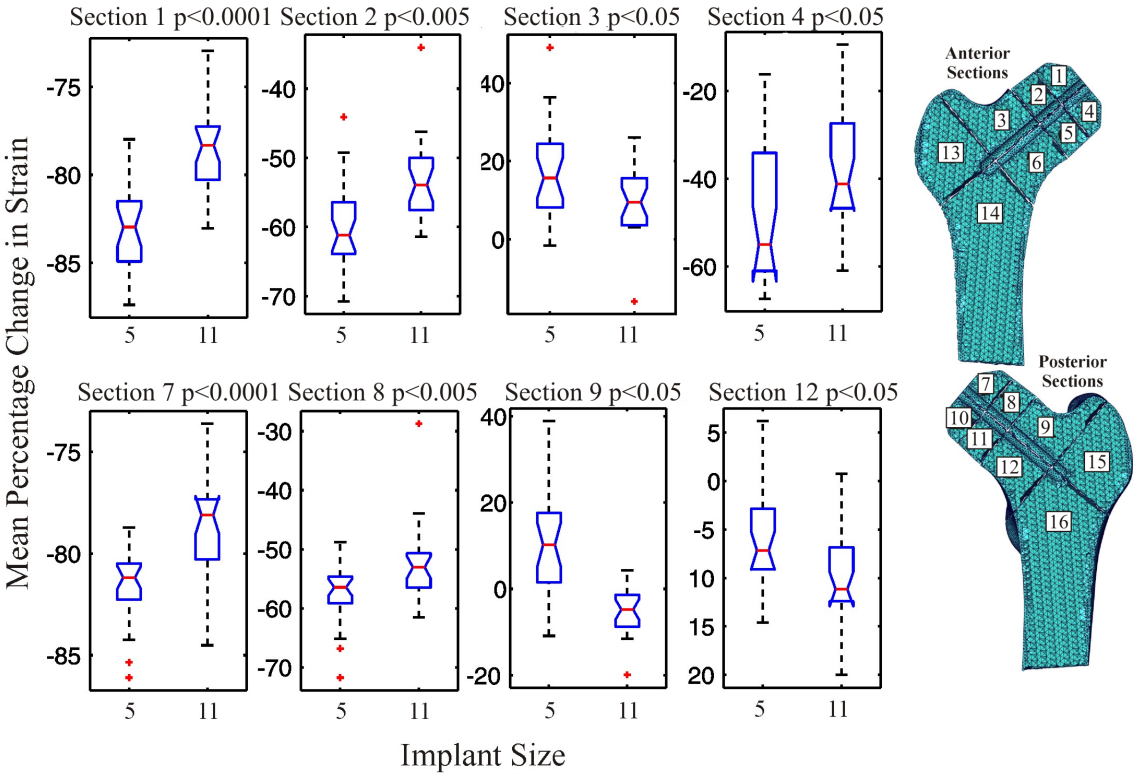


Figure 5.15: Boxplots of mean percentage change in strain results by section for the implant 5 and implant 11 groups of femurs, for those sections where statistical significant differences were seen.

criteria by which the groups were defined the majority of geometric metrics were ignored as they related to femur size which was inevitably different. Of the remaining metrics no statistically significant difference was seen between the groups, including material quality, loading/BMI and femur proportions such as neck diameter ratio or head-neck ratio, which were highlighted in the previous outlier analysis.

5.5 Discussion

In this study 400 femurs were generated and analysed to see the effect of femoral head resurfacing on load transfer through the femur. This involved the fully automated implantation of appropriately sized and orientated femoral components. The success of this was evident when each case was visually checked. The methodology ran without error for 85% of the femur geometries attempted. The instances which were not completed failed at the Boolean operation stage, a complication which was not unexpected due to the inherent complexity of the shapes being cut and the trouble this can cause even if being performed manually. The entire process, including FEA of the related intact instance, was able to run at a rate of 4 models per hour (on

Femoral Head	Mean % change		Maximum % change	
	Implant 5	Implant 11	Implant 5	Implant 11
Section 1	-82.8	-78.3	-87.4	-83.0
Section 2	-59.7	-53.2	-70.8	-61.4
Section 7	-81.5	-78.6	-86.1	-84.5
Section 8	-57.2	-51.8	-71.8	-61.5
Femoral Neck	Mean % change		Maximum % change	
	Implant 5	Implant 11	Implant 5	Implant 11
Section 3	16.7	9.2	49.1	26.1
Section 9	10.3	-5.5	38.9	4.3
Section 12	-5.8	-10.4	6.2	0.7

Table 5.3: Table containing the mean and extreme results for the changes in strain in the femoral head and neck sections which were found to show significantly different results between the large and small implant sizes.

a duel quad core processor). This mainly reflects the time taken in transferring data between different pieces of software and the intricacy of the models generated, hindering speed despite the relative simplicity of the final FE analysis. The time constraints were considered entirely manageable for this scenario of 400 models. In future it is hoped that it will be possible to perform the majority of these steps within a single program to improve this, however currently there is no one piece of software available to the study which is capable of completing all model preparation tasks.

Comparison of the change in strain predicted through the proximal femur to those published by Radcliffe et al. [176] for identical loading conditions provided a means of checking the results found in this work. Each of the 16 femurs modelled by Radcliffe were manually generated and implanted, a painstaking process meaning that implant positioning was entirely reliable. Therefore the comparable trends in strain change seen through the proximal femur are a further indication of the success of the automated implantation technique (Figure 5.12). As hoped the increase in sample size between the studies, made possible thanks to the use of a statistical model to generate large numbers of unique femurs, has increased the range of predicted outcomes in each section of the femur. This shows the potential of the technique in identifying instances where specific ranges of characteristics lead to unusual outcomes. The number of models being tested allow trends for these less common cases to be observed which other approaches may never find.

The main strain alterations observed were severe strain shielding in the femoral head, particularly in the superior sections, and increased strain in the femoral neck, again most notable superiorly. These patterns of load transfer change have been

reported in previous FE and experimental work and have been associated with failure mechanism theories. Predicted strain shielding in the femoral head was linked to bone resorption, and potential failure through head collapse which has been observed in clinical retrievals [45, 168, 178, 84]. However many studies have not indicated loosening through altered load transfer to be a major cause of revision and have proposed reasons such as disrupted blood supply, thermal damage during cement curing, trauma during implantation and underlying poor bone quality to explain the changes seen [166, 169, 4, 170, 174].

Femoral neck fracture is now regarded as a serious risk following femoral resurfacing. By exploiting the large number of analyses run in this study it was possible to interrogate the features of those femurs with increased strain in the femoral neck. Despite the magnitude of the increases in most cases not seeming large enough to directly result in fracture there have been a range of reasons proposed for how the neck may become weakened or subjected to higher loads. Thus the combination of contributing factors may be sufficient. In addition outlying results were found where average strain increases approaching 60% were seen (with strain pattern trends shown in figure 5.13). Damage done during surgery or poor surgical accuracy are often blamed for weakening the femoral neck; cutting blood supply leading to necrosis, notching the neck producing a stress concentration, poorly seating an implant leaving uncovered cancellous bone or malalignment of the implant (varus) resulting in unfavourable loading [4, 168, 166, 165]. Clinically, neck fractures are commonly seen within the first few months of surgery [165, 167]. Previous FE and experimental work have shown increased strain in the superior femoral neck although not over a volume sufficient to cause fracture, but the addition of a notch could be enough to change this [175, 49, 84, 179]. Although Long et al. [167] argued that short term neck fracture is not always accompanied by evidence of notching. Long suggesting that in bones with lower modulus where higher loads were applied, the areas around the implant rim suffer high strain concentrations which may lead to damage accumulation and fatigue failure before the bone adaption cycle can complete in response.

The average strain changes predicted in the neck by the 400 femur analysis are of a similar magnitude to earlier studies. Examination of the outliers where strain increased by more than two standard deviations from the mean indicated that these femurs were significantly smaller over a range of measurements than the main set of models, with slightly higher modulus in the femoral neck. The measurements highlighted included intertrochantic width, femoral shaft diameter, neck axis length, femoral head diameter and overall proximal volume. Also highlighted were proportional features, suggesting these outliers had more oval necks (larger in superior-inferior dimension) and a larger difference between head and neck diameters. There

were several outlying femurs showing a notable decrease in strain in the neck, when compared to those models with the largest strain increase a statistically significant difference in size was seen. The femurs with the largest reduction in strain had lower mean cancellous modulus values than the overall generated population. These statistically significant geometric and material differences were only identifiable thanks to the number of models this study was able to run.

An intriguing set of observations relating to RA were made in the Australian National Joint Replacement Registry [61]. These included an inverse relationship between implant size and revision rate, with smaller implant sizes (<44mm) suffering 9% revision while the larger sizes (>55mm) only 1.7%, with no significant gender difference in failure rate once implant size was accounted for. This was alongside a clear picture of the typical RA patient demographic, being male (76.4% and rising) and young (91.9% <65 years), with the revision rate at 5 years for female patients more than twice that for males (6.5% to 2.6%). Many studies have analysed risk factors for femoral resurfacing to try to identify reasons why the procedure may be more or less likely to require revision. Smaller implant sizes and/or female gender (normally associated with smaller implant sizes) have been flagged as having a higher revision risk by previous clinical studies [163, 180, 172, 157], with Mai et al. [177]'s 12 year follow up study showing an 82% survivorship rate with the large implant group compared to 56% with the small. Although some studies have not found implant size to be significant [166, 165].

Due to the evidence found from examining outliers in the previous analysis linking femur size to adverse traits such as increased strain in the femoral neck and greater strain shielding, and the clinical evidence of poorer performance in smaller implant sizes, an investigation was performed to compare the altered strain patterns produced by a small and a large implant size. By identifying instances from the 400 generated model set which had been assigned size 5 or size 11 implants it was possible to gather two subgroups of 25 and 20 femurs respectively. These groups were analysed using the same methods as the outliers, allowing direct comparisons to be drawn between them and any statistically significant differences found. The results of this analysis showed two potential trends, linked to suggested failure mechanisms, which were exaggerated in the smaller implant group. These were strain shielding in the femoral head, higher in the proximal head (sections 1, 2, 7 and 8), and elevated strains in the femoral neck, again largest proximally (sections 3, 9 and 12). The superior-posterior neck (section 9) showed a mean increase in strain for the small implants while a mean decrease for the large. Unlike the trends observed in the femoral head where the differences between the means were a few percent and the extremes of the data were similar, in the neck these mean changes in strain and extremes were very different (Table 5.3).

As with the overall results the typical mean changes in strain in the smaller implant size are not large. However the highest results seen, ∼40-50%, may be enough to lead to damage if a weakness was present such as a notch or cyst and so these femurs could be at higher risk of neck fracture than the others assessed. The only metrics, other than those relating to scale, which proved to segregate the implant size results concerned mean cortical bone modulus in the neck and intertrochantic regions and the percentage of cortical bone in the neck. All of these were higher in the smaller femurs.

A potential limitation of the study was the initial training data set. This was male dominated with a mean age of 70 years, no preselection of the data was done due to the difficulty in sourcing whole femur CT scans. Ideally the statistical model would be tailored to model a specific population group by segregation on the grounds of gender, age, ethnicity and so on, which all affect femur morphology. However, the range of femurs which can be produced by this model has been demonstrated previously (Section 3.6.4), and in that study are comparable to the population data found in the NHANES survey (Figure 3.27). This study's focus was on generating a methodology to allow the automated modelling of orthopaedic implants, therefore the model was run to produce a wide range of femoral morphologies. A further implication of this was that a simplified implantation strategy was employed which aimed to prevent neck notching and restore approximately natural biomechanics. A more rigorous simulation of surgical procedure would be possible in future work, for example this could incorporate the more complex positioning recommendations detailed by Shimmin et al. [162]. An extension to the study could also be to introduce a notch to the neck, similar to one which may be made in error while cutting the femoral head, to investigate whether this caused sufficiently elevated strains to lead to fracture.

5.6 Conclusions

The feasibility of using a statistical model to enable large scale multi model testing of an intact femur was demonstrated in Chapter 4 in a case study of femoral neck fracture risk. The current work has shown how this concept has been extended to allow the benefits of analysing a wide range of models to be applied to orthopaedic implant testing. A fully automated technique performs the complex tasks required to implant and analyse given geometries with a high success rate. A wide range of investigations are possible with simple alterations to the method. Modelling a wide population of bones will allow meaningful statistical analyses to be carried out, infeasible in conventional studies, and enable more comprehensive testing of orthopaedic implants in the future.

Chapter 6

Conclusions and Future Work

6.1 Background to the problems motivating this work

Computational analysis has become an increasingly intrinsic tool in biomechanics, used as a fundamental part of research and orthopaedic design. In-silico techniques have evolved as computer technology has developed and allowed the potential advantages of the methods to be exploited, advantages which most experimental alternatives can not feasibly match. For example the ability to test the same model multiple times, under different conditions or to compare the performance of a range of models under identical conditions. The detailed results which computational analyses are capable of producing, allowing interrogation throughout a three dimensional structure, are also not able to be replicated by in-vitro or in-vivo tests.

However, there are still some significant weaknesses in conventional in-silico experimental design which limit their ability to model biomechanical scenarios completely. One important weakness is the lack of incorporation of variability - namely differences in material properties and morphology between subjects. In addition, when orthopaedic implants are modelled it is rare for surgical variability to be accounted for in assessment, usually perfect positioning and implantation is assumed despite this being hard to achieve in reality. When the effect of subject or surgical variations have been investigated they have been found to dramatically alter the outcomes of the analysis [3, 6, 52, 37, 7].

In the case of inter-subject variability there are good reasons why computational testing has not been able to incorporate the differences known to exist in real life.

These boil down to two key issues; 1) the time consuming, labour intensive and tedious task of manual model generation and 2) the limited availability of high quality imaging data (CT or MRI). It is these issues which this work has aimed to address such that intersubject variability can be incorporated into in-silico analysis without massive computational or labour costs.

Of previous studies where attempts have been made to better model the true clinical circumstances around biomedical scenarios, such as joint replacement, by incorporating variability only a few have developed patient specific models. Some studies have taken an approach of simplifying the problem and only varying one parameter, for example artificially varying material properties on an identical geometry [3] or scaling a single model and related parameters to generate a range of models [58].

Publications which have generated subject specific models include Radcliffe et al [176, 7, 37] and Keyak et al [51, 150] where sets of models (16 and 18 respectively) were manually generated from CT data, extracting geometry and material properties, and manually put through FE analysis. These produced a wealth of information but were limited by what was still a statistically small data set, where it was likely that they related to CT scans which were available rather than representative of patients suitable for the study. The process of performing the analyses was time consuming and even if more image data had been available this limitation would have made it infeasible for a substantially larger group of models to be tested this way.

6.2 Aims of this work and assessment of whether they have been achieved

The fundamental aim of this project was to investigate the possibility of using statistical modelling techniques to generate a population of femur models and to show whether these could be efficiently integrated into a large scale computational stress analysis.

The initial stage of generating the statistical model, detailed in Chapter 3, demanded a range of challenges to be met. The first of these was to acquire a large enough database of CT data sets with which the statistical model could be trained, such that the variability in femoral morphology and material distribution could be captured. In total 46 individual CT scans were collected, segmented and used to

construct the statistical model. The ultimate number of available scans was to a large extent beyond the control of the authors, however every whole femur CT set which could be used in this work was used. Reconstruction errors and eigenvalue decay were used to indicate how well the femoral variability in this training set was represented by the model, where rapid decay of either would indicate high correlation between them (figs. 3.21). These results suggested that the number of femurs used was effective at capturing geometric variability but insufficient to completely capture the complex changes in material properties. This would reduce the ability of the model to generate specific instances accurately, an ambitious ask of just 46 training examples. However the proposed use in this work was not to do that but to generate new realistic instances and it was shown that the patterns of material distribution in generated femurs were realistic and the statistical model (SM) was able to produce legal instances - provided suitable sampling limits were set.

A necessary aim of the model generation process was to find a method of establishing accurate, three dimensional correspondence between each training set member such that they were each described in an identical form suitable for PCA. With the ultimate aim that instances generated from the SM would be used in FEA it was also key for mesh quality to be high throughout the construction process. The conventional approach for achieving registration in complex 3D data sets like the femur is the computationally demanding free form deformation method developed by Rueckert et al. [13]. The technique adopted in the current work was a combination of surface matching and mesh morphing to allow a baseline tetrahedral mesh to be morphed onto each other femur instance. The elastic matching strategy adopted for surface matching required adaption to be able to handle the large number of points used in this model to represent each femur. The limitations of the conventional algorithms [91] would have made this process extremely slow.

The developed strategy initially led to reasonable meshes over the majority of the matched instances but some areas, particularly those with rapidly changing curvature such as around the greater trochanter, suffered from severe mesh degradation/folding. The problem was resolved, without significant computational cost, by the introduction of Laplacian smoothing at each surface registration step. The accuracy of the matched meshes was high, with a mean error within the voxel resolution of the original image (figs. 3.14, 3.9).

The sampling method used to generate new instances from the SM was ultimately found to be the key factor in deciding whether or not the goal of FE ready, realistic femur models were produced. The parameters concerned were; the number of eigenmodes to include, the range over which the models should be sampled and the

sampling technique (see Section 3.6.3). These parameters can not truly be treated individually as they can work with and against each other, therefore the choice of these values was challenging and they were found to be fundamental to the quality and variety of models generated. The potential limitations and problems of this sampling challenge will be discussed later in this chapter. By careful consideration and analysis of the models produced it was possible to achieve the aim of creating a set of unique, realistic femurs with a mesh quality of a high enough value to allow direct use in FE. The values of a range of meaningful morphometric measurements showed that the synthetic population generated by the sampling process was comparable to a true human population, as detailed by Centers for Disease Control & Prevention (CDC) [107].

The femoral neck fracture study, documented in Chapter 4, was performed to investigate whether the second part of the overall objective could be achieved, namely using the SM as a source of femur models to drive a large scale FE analysis. The objectives of this analysis were to show how this could be integrated and run, and to test that the results generated from these synthetic femurs were realistic. The methodology created to perform the study was robust and scripted throughout, requiring only the number of femurs to be tested to start and finishing with fully post-processed strain results alongside a full range of metrics characterising each model. This lack of manual interaction with the analysis, together with automated checks to ensure the process was running correctly, allowed a large number of models to be tested. It would not have been feasible to perform a similar analysis on 1000 femurs manually. Thus the correlation of strain results and geometric/material properties over a statistically significant number of models achieved by this method would have been impossible.

A useful outcome of the statistical modelling technique is that all generated femurs were described by the same number of nodes/elements with each of them being in the same relative position on every femur. This provided a predetermined landmarking system upon which it was straight forward to base boundary conditions, loading conditions and measurement positions. The metrics extracted in post processing were able to be compared to previous clinical, experimental and computational studies where similar loading conditions had been observed - i.e. a fall. The agreement of the distribution of fracture initiation sites over the data set and the agreement of several known risk characteristics (mainly related to material metrics) help to suggest that the results of the analysis are realistic. This aids the case for the possibility of using a SM to generate data sets for use in larger, population based studies as opposed to the single or small set of bone models which currently is the conventional basis for analysis.

A vast number of orthopaedic computational analyses are carried out to investigate the interaction of an implant on the stress/strain distribution through their host bone. In order for analyses using large numbers of models to be useful in biomechanics a methodology to allow automated implantation was needed to be proved possible (i.e. virtual implantation without manual intervention). Chapter 5 discusses the development of such a methodology for automatically implanting a femoral resurfacing component into any femur model generated from the statistical model. This study required the accurate alignment, sizing and virtual implantation of resurfacing components, as well as subsequent FE analysis. The implanted model generation procedure was complex and meant that the entire process was not as robust as the previous neck fracture study. There was a notable failure rate, ∼15%, mainly at the Boolean stage which was disappointing but not unpredictable due to the complexity of the shape interactions. The failure of an individual model did not crash the entire simulation, i.e. the script would simply move on to the next instance, thus the study was able to run without manual intervention in the same way as the previously described study.

The results of the study were able to replicate the trends of a previous analysis by Radcliffe and Taylor [37], where resurfacing was manually simulated on 16 femur models, as well as highlighting models with poorer outcomes and relating these back to the physical characteristics of those models. This analysis achieved its aim of allowing accurate, automated implantation over a large number of models and illustrated the potential of performing computational analyses of orthopaedic implants in multiple models.

6.3 Limitations

The quality of any model is always heavily dependent on the quality of data which is used to create it. The statistical model which forms the basis of this work is no exception and therefore the information which is used to train it will always be a fundamental limitation. As previously mentioned the ultimate number of individual CT scans available to the project was out of the author's control, however every usable set was included. The population described by these 46 femurs turned out to be male biased, with a mean age of ∼70 and from a Caucasian population. It was clear from simply viewing the scans that there were variations in bone quality (with an age range of 43 to 91 this was inevitable), but none of the subjects were known to have a pre-existing degenerative joint disease.

Due to this relatively small training set size, and its potential bias, the range of femur representations which can be created from it will undeniably reflect this and thus careful sampling and observation of these femurs needs to be made. It is key to remember that this study was intended to be a proof of concept rather than aiming to produce a definitive model which either covers the entire possible population of human femurs or a particular subgroup. Extending the project along these lines is discussed further in the context of future work.

The quality of the data fed into the statistical model was vulnerable to the introduction of errors during the process of extracting each femur representation from their respective image sets. The geometries were generated by semi-automated segmentation, based on greylevel thresholding, and manual slice-by-slice correction. During segmentation a certain amount of judgement is required to follow the bone contours, particularly where bones are close together (e.g. femoral head and acetabulum), where image resolution is poor or where bone quality is low - making boundaries less defined. Certain areas of the femur are harder to segment neatly than others due to this issue, including areas of the femoral head where the cortical shell is thin and the greater trochanter where very dense soft tissues connect to the femur. It is entirely possible for low level geometric errors to be introduced to the final statistical model due to these sorts of issues.

It is evident from the various analyses done to assess reconstruction error in chapter 3 that the statistical model is successful in capturing geometric variations but has some difficulty with material properties. The culprit behind this could again be the quality of the initial training data, although the distribution and complexity of material distribution undoubtedly varies significantly and is most likely not sufficiently captured by the 46 base examples. The main challenge with assigning material properties in this work was that the CT data sets used were clinical and so uncalibrated, meaning there were no calibration phantoms in the images which could be used to correct greylevel for the particular machine set up. Instead an alternative calibration method was devised which relied on picking two materials of known density in the image and correcting the entire greylevel range based on these (Section 3.3.3). The assumption required to do this was that there was some area of cortical bone in every data set which had the textbook maximum cortical bone density of $1.73 g/cm^3$. It is likely that some of the subjects, particularly the most elderly, did not quite achieve this value even in the densest areas and therefore this assumption may be introducing errors.

The most important factor found to influence the success of the statistical model in generating a widely varying, unique and yet realistic set of femur models with high

mesh quality was sampling. The correct choice of sampling parameters and sampling method must be established for each statistical model and are dependent on the characteristics of the training data. A simplification made in this work has been that the probability distributions of each individual coefficient of the PCA model have not been used as the basis for sampling, instead superimposing a Gaussian sampling curve (Section 3.6.3). The potential effect of this is the possibility of combinations of eigenvector weightings producing unrealistic femurs, i.e. over/under sized or with extreme material properties. The risk was understood in this work and so care was taken to ensure that models were suitable through a combination of automated pass/fail filters (e.g. based on femur length), manual interrogation and limiting the sampling range with appropriate parameters. For future use of the model more investigation may be needed into the choice and control of sampling methods and parameters so that it can be relied upon completely without the potential need for further levels of filtering. An increase in training data could also help to alleviate this problem.

The FE analyses performed in this work have been simplistic, this was again because the aim of the work has been to prove that the concept of running a large number of FE studies on a widely varying set of bone models, generated from a SM, was feasible. The trends of the results follow those of previous work with similar investigations which was sufficient for this purpose, however in future it may be possible to increase the complexity of the model.

Currently the main limiting factor in the automated implantation methodology is the requirement for a succession of different pieces of software, each of which is suited to a particular process which the other can not perform as reliably. The result is a time cost in writing out/reading in at each stage of the process and a potential for errors as information is translated. The failure rate of Boolean operations is hard to completely avoid. A series of steps were taken to prevent this occurring due to any irregular elements (which would prevent surface fitting) or unfavourable intersections (by shifting the cutter by fractions). Alternative software or further investigation into the reason for failure at this stage may reduce the failure rate. At this proof of concept stage these limitations were accepted and are a clear direction for future work in refining the techniques.

A final notable limitation to the large scale analyses described in this work, but by no means restricted solely to it, is the issue of establishing a comprehensive interpretation of results. As each of the many models run produces reams of data, relating to elemental strains and femur characteristics, the true results of the analysis can quickly become swamped under the weight of numbers. This volume of data is the

strength of the method, allowing a wide range of post processing to be performed over numbers sufficient to show statistical significance, yet it has proved all too easy for important trends to be missed. Amongst the large matrix of data, results will not be found unless they are looked for.

6.4 Future Work

As this work has not been a continuation of a previous project it has been approached as a proof of concept study, investigating the potential of applying statistical modelling techniques to generating synthetic bone models for large scale, multi model FE analyses. This aim has been achieved and thus opens up the possibilities for a range of future work directions, including both refining the techniques which have been suggested and extending them to new problems.

There are several clear ways in which the statistical model could be improved which have been previously mentioned, including; an increase in the size and diversity of the training data, the use of calibrated CTs and the development of more sophisticated sampling techniques. The creation of more focused statistical models could prove extremely useful for orthopaedic research where risk factors have already been identified and a better understanding of their influence required. For example the development of gender and/or age specific models to match a patient demographic of interest or a disease specific model such as one built from osteoporotic bones.

There is potential for improvement in the registration scheme. Currently the surface and volume registration stages are performed separately but in future it may be possible to integrate these into one piece of code. As discussed, the scheme is currently less able to capture modulus variability than geometric which could be directly linked to the approach taken. Investigations into the possible improvements of registration schemes which use intensity information as well as shape to establish correspondence could improve this problem [13, 82, 81, 70].

The complexity of the FE analyses in this work was minimised due to the aims of the study, however in future it may be of interest to look at increasing this to try to model situations more realistically. This may include investigating more complex interface conditions between implant-cement-bone, introducing suboptimal implant alignments, modelling poor fixation or uneven cement coverage. The incorporation of surgical variations in combination with intersubject variability is an intriguing possibility which may better emulate clinical conditions and lead to greater insight

into the affect of joint replacement than current studies. The methodology described is easily adjustable to incorporate such variability, for example the alteration of a single line of code can result in an oversized or undersized component being implanted (details in Appendix E).

The process trialled for resurfacing arthroplasty could be applied to other femoral implants, at both the proximal and distal ends. The automated implantation of a total joint replacement component would require a new alignment strategy and undoubtedly raise new problems, however if achieved it would open up the potential for comparative assessment of implant designs over a large number of femurs.

This work has focused solely on the femur, but there is no reason why the statistical modelling techniques developed can not be applied to other bones which host implants, such as; the tibia, acetabulum or vertebra. It is extremely common for femoral and acetabular components to be analysed separately, yet this is removing the implant and bone from the surrounding system. Some studies have begun to include the acetabulum, femur and sometimes surrounding musculature into analyses of the hip [49, 50]. It would be a logical development to attempt to apply these statistical modelling techniques to multiple bodies, such as the acetabulum and proximal femur, distal femur and proximal tibia or a set of lumbar vertebra. The challenges of this progression may exceed the capabilities of the current methodology but this direction is an ambitious and potentially useful aim for the future.

Appendix A

Further Details of the Registration Strategy

A.1 Registration Scheme

The first stage of establishing correspondence between each member of the training set was to register a common baseline mesh to each femur model surface. The median length femur of the set was chosen to be the baseline femur, with the rationale that this would lead to the minimum element distortion when stretching or compressing the mesh to fit another instance. It has been shown that the choice of example from which the template mesh is created has only a minimal effect on the final model accuracy [73].

The chosen femur surface mesh was imported into meshing software, Ansys ICEM CFD (Ansys Inc., Canonsburg, PA), and converted into a high quality, solid tetrahedral mesh with a global element size of 3mm. In order to balance model definition and computation cost the model was then split into three regions and the upper and lower thirds' mesh size refined to 1–1.5mm at the surface (Fig. A.1). This was justified for two reasons: 1) these areas are of more clinical interest so a fine mesh will be important for future use of the model and 2) these areas contain the most rapidly changing geometry and hence require a higher density of surface points to achieve accurate registration. The baseline tetrahedral mesh, and therefore any subsequent mesh produced by the model, consisted of 615,523 elements and 117,225 nodes, of which 27,171 were on the surface (Fig. A.1).

Every femur in the training set was initially aligned to the same orientation as the baseline femur using ICP [10].

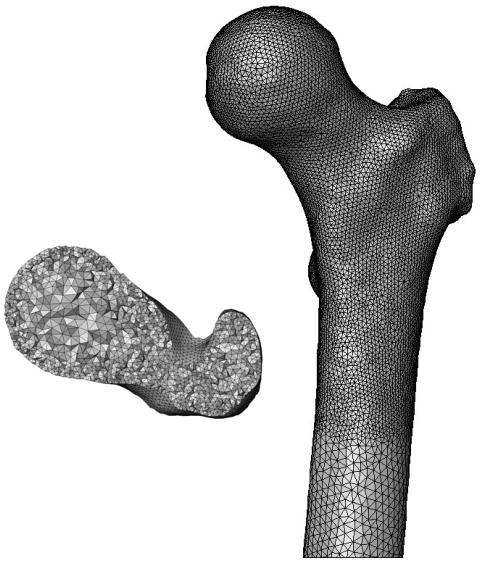


Figure A.1: Figure A1:The baseline meshed geometry. Shows a section of mesh through the femoral head and the transition in mesh size between the proximal femur and femoral shaft

The registration scheme developed has two steps; 1) surface registration and 2) volumetric registration via mesh morphing, to achieve three dimensional correspondence between a baseline and a target femur model. Surface matching iteratively deformed the baseline surface vertices to match a target surface and was able to achieve a smooth, accurate final mesh through user defined inputs controlling the magnitude and speed of the deformation. These parameters controlled the area affected by moving a single vertex and the magnitude of each iterative step, both decreasing over the specified number of iterations. This initially allowed almost global deformations to act which roughly aligned the surfaces before narrowing the regions of influence to produce fine, local manipulations of the mesh until a stopping criteria was reached. The accuracy of the surface registration achieved using this algorithm is illustrated in Figure A.1, showing the morphed and target surfaces superimposed. The volumetric registration step used the registered surface points as a scaffold and positioned the internal mesh points based on the surface node displacement vectors.

A.1.1 Surface Registration

The surface registration technique employed in the present study was based on the three dimensional generalization of Burr's elastic registration algorithm proposed by Moshfeghi et al. [91]. The inputs required by the scheme were the ordered vertex

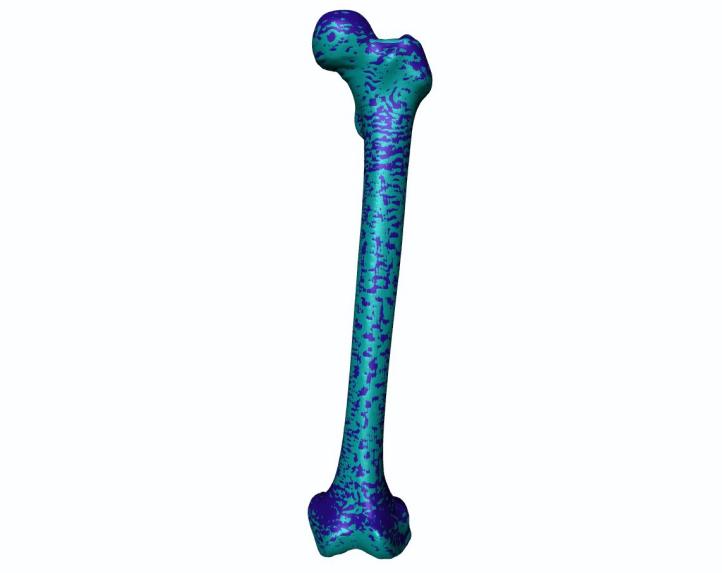


Figure A.2: Target mesh (dark) with morphed baseline mesh (light) superimposed, showing the accuracy achievable by the registration scheme

and triangle patch lists of a baseline femur ($\vec{S}_1$) and a target femur ($\vec{S}_2$) to which the baseline mesh would be mapped. The baseline femur was represented as follows

$$\vec{S}_1 = \{\{x_{1i}, y_{1i}, z_{1i}\}, \{\Delta_c\}\}, \quad \text{where } 1 \leqslant i \leqslant N_1, 1 \leqslant c \leqslant T_1. \tag{A.1}$$

here (x_{1i}, y_{1i}, z_{1i}) were the coordinates of the ith vertex and Δ_c referred to the triangle patch c. Here, N_1 indicated the total number of vertices and T_1 the number of triangle patches on the baseline surface $\vec{S}_1$. The target femur was similarly represented, where N_2 and T_2 were the total number of vertices and triangle patches on $\vec{S}_2$ respectively

$$\vec{S}_2 = \{\{x_{2j}, y_{2j}, z_{2j}\}, \{\Delta_d\}\}, \quad \text{where } 1 \leqslant j \leqslant N_2, 1 \leqslant d \leqslant T_2. \tag{A.2}$$

The key steps of the registration algorithm to elastically deform the baseline surface ($\vec{S}_1$) to match the target surface ($\vec{S}_2$) are outlined below.

1. Inputs specified: baseline surface $\vec{S}_1 = \{\{x_{1i}, y_{1i}, z_{1i}\}, \{\Delta_c\}\}$, target surface $\vec{S}_2 = \{\{x_{2j}, y_{2j}, z_{2j}\}, \{\Delta_d\}\}$, nearest neighbour parameter m, normalization factor α, smoothing parameters γ, σ^0 and f, maximum number of iterations k_{max} and stopping criteria ϵ. The required parameters used in this study were; m=50, γ=2, σ^0=10, f=1.0715 and k_{max}=100.

2. Coarse registration performed to align the surfaces $\vec{S}_1$ and $\vec{S}_2$.

3. Set $k = 1$ and defined $\vec{W}^{k-1} = \vec{S}_1$ before iteratively performing the following steps **WHILE** $k \leq k_{max}$:

4. Constructed k-d tree representations, $\mathcal{KDT}_1$ and $\mathcal{KDT}_2$, of the centroids of surface triangulations of the deforming baseline mesh $\vec{W}^{k-1}$ and the target mesh $\vec{S}_2$.

5. **DO FOR** $i = 1, 2, \ldots, N_1$

   (a) For each vertex $\vec{P}_1(i)$ in the deforming baseline mesh $\vec{W}^{k-1}$, the m nearest surface triangles on the target $\vec{S}_2$ were identified using the k-d tree $\mathcal{KDT}_2$.

   (b) The closest target surface triangle patch to the considered vertex depends on the location of the intersection point, $\vec{G}$, produced between a perpendicular drawn from the vertex $\vec{P}_1(i)$ to the target triangle patch Δ_d. If $\vec{G}$ lay inside Δ_d the distance measure, γ, was defined by $\gamma(i,d) = |\vec{G} - P_1(i)|$, otherwise by $\gamma(i,d) = |\vec{G} - \vec{P}_1(i)| + |\vec{P} - \vec{G}|$. Here $\vec{P}$ was the closest triangle vertex on the target surface $\vec{S}_2$ to the intersection point $\vec{G}$. It should be noted that only a distance measure is used to determine the nearest triangle patch. This is different to Moshfeghi et al. [91] who incorporate a weighted directional incompatibility measure, ω, in the similarity measure $\delta = \gamma + \alpha\omega$. Using this similarity measure often resulted in mesh folding after registration, particularly in the case of the fine mesh used for the femur model.

   (c) Having found the location of the nearest surface triangle on the surface $\vec{S}_2$ to $\vec{P}_1(i)$, the displacement vectors required to align the surfaces could be calculated as the deformation field vector $\vec{D}_1(i,d) = \vec{G} - \vec{P}_1(i)$ if the point $\vec{G}$ was inside or on the Δ_d, or $\vec{D}_1(i,d) = \vec{P} - \vec{P}_1(i)$ if it was not.

   END DO

6. Processes (a-c) within the previous step were then repeated with $\vec{S}_2$ in place of $\vec{S}_1$ and $\vec{W}^{k-1}$ replacing $\vec{S}_2$ to find the deformation field $\vec{D}_2(j,c)$ describing the displacement vectors required to align $\vec{S}_2$ to $\vec{W}^{k-1}$.

7. The smoothed displacement field was computed as

$$\vec{DS}^{k-1}(x,y,z) = \frac{1}{\gamma}\left[\frac{\sum_{i=1}^{N_1} G_{1i}(x,y,z)\vec{D}_1(i,d_i)}{\sum_{i=1}^{N_1} G_{1i}(x,y,z)} - \frac{\sum_{j=1}^{N_2} G_{2j}(x,y,z)\vec{D}_2(j,c_j)}{\sum_{j=1}^{N_2} G_{2j}(x,y,z)}\right], \tag{A.3}$$

   where G_1 and G_2 could be chosen to be radial basis functions of the form $G_{1i}(x,y,z) = \phi(r_{1i})$ and $G_{2j}(x,y,z) = \phi(r_{2j})$. r_{1i} was the Euclidean distance between (x,y,z) and the ith vertex of the baseline surface $\vec{S}_1$ and r_{2j} was the

Euclidean distance between (x, y, z) and $(x_{2j} + D_{2x}(j, c_j), y_{2j} + D_{2y}(j, c_j), z_{2j} + D_{2z}(j, c_j))$. Moshfeghi et al. [91] suggested the use of the following Gaussian radial basis functions

$$G_{1i} = \exp\left(-\frac{1}{(\sigma^k)^2}\left[(x - x_{1i})^2 + (y - y_{1i})^2 + (z - z_{1i})^2\right]\right) \quad (A.4)$$

$$G_{2j} = \exp\left(-\frac{1}{(\sigma^k)^2}\left[(x - x_{2j} - D_{2x}(j, c_j))^2 + (y - y_{2j} - D_{2y}(j, c_j))^2 + (z - z_{2j} - D_{2z}(j, c_j))^2\right]\right), \quad (A.5)$$

where the smoothing parameter σ^k was decreased at each iteration using the update rule $\sigma^k = \sigma^0 f^{-k}$ with $1 \leq f \leq 2$. The factor γ was set to 2 in our numerical studies and can be interpreted as a damping factor. By increasing f and σ^0 a greater deformation was achieved but mesh quality could be reduced. A good trade off between achieved deformation and mesh quality was achieved by using a relatively high value of σ with a low value of f, over a large number of iterations ($k \approx 100$).

8. Iteratively deformed $\vec{S}_1$ to $\vec{S}_2$ using the following update rule

$$\vec{W}^k(i, c) = \vec{W}^{k-1}(i, c) + \vec{DS}^{k-1}(x, y, z). \quad (A.6)$$

9. Apply improved Laplace smoothing [93] to help maintain the mesh integrity. A set number of iterations of smoothing are applied to $W^k(i, c)$ to stop mesh folding occurring. Observations evidence showed that only using Gaussian smoothing on the displacement field is not enough to stop distortion of the mesh (Fig. A3).

10. The iterations were terminated if the following stopping criterion was reached, if not then set $k = k + 1$ and returned to step 4.

$$\frac{1}{N_1}\sum_{i=1}^{N_1} |\vec{DS}(x_{1i}, y_{1i}, z_{1i})| \leq \epsilon. \quad (A.7)$$

When registration was completed and convergence achieved, say after k iterations, the elastic matching algorithm provided the deformed surface mesh (A.8), with the baseline mesh $\vec{S}_1$ mapped onto the target surface $\vec{S}_2$.

$$\vec{W}^k(i, c) = \{(x'_{1i}, y'_{1i}, z'_{1i}), \Delta_c\}, \quad \text{where } 1 \leqslant i \leqslant N_1, 1 \leqslant c \leqslant T_1. \quad (A.8)$$

The displacement vectors required to register the vertices of $\vec{S}_1(i,c)$ with the surface $\vec{S}_2(j,d)$ were then given by equation A.9;

$$\vec{u}_i = \{(x'_{1i} - x_{1i}), (y'_{1i} - y_{1i}), (z'_{1i} - z_{1i})\}, \quad \text{where } 1 \leqslant i \leqslant N_1. \tag{A.9}$$

A.1.2 Volume Mesh Morphing

The problem of deforming a volume mesh to track moving/perturbed surfaces has been extensively studied in several areas including computational geometry, moving surface fluid flow simulation and engineering design optimisation [94, 95, 96]. The central idea underpinning many existing approaches is to solve partial differential equations (for example the elasticity or diffusion equations), for the coordinates of the deformed mesh with the known surface displacements imposed as essential boundary conditions.

In the present work the diffusion based mesh deformation strategy was used. It is computationally efficient and known to work well in a variety of application domains [97, 94]. This approach solved the following decoupled three-dimensional Laplace equations to deform the baseline volume mesh

$$\nabla^2 \delta x = 0, \ \nabla^2 \delta y = 0, \ \nabla^2 \delta z = 0, \tag{A.10}$$

where δx, δy and δz were the displacement fields applied to the baseline mesh coordinates in the $x-$, $y-$ and $z-$ directions, respectively, and ∇^2 was the standard Laplacian operator defined on the baseline volume mesh.

The x, y, z components of the surface registration vectors $\vec{u}_i$, specified by equation (A.9), were used to specify appropriate Dirichlet boundary conditions for equation A.10. The result was a morphing of the internal points of the baseline volume mesh onto the target volume. Once the whole registration process was applied to all members of the training set, each was described by a solid tetrahedral mesh, with direct correspondence established between each point and element.

It is important to note that when the magnitudes of $\vec{u}_i$ become large some elements of the morphed volume may experience reversal, leading to negative signed volumes. One straightforward way to circumvent this problem is to tightly couple the mesh morphing strategy with the surface registration algorithm. For example, by carrying out volume meshing after surface registration step 8 where the Laplace operator is defined on the volume corresponding to the surface W^{k-1}, with Dirichlet boundary

conditions imposed accordingly. The surface registration vector at iteration k then becomes

$$\vec{u}_i^k = \left\{(x_{1i}^k - x_{1i}), (y_{1i}^k - y_{1i}), (z_{1i}^k - z_{1i})\right\}, \quad \text{where } 1 \leqslant i \leqslant N_1. \tag{A.11}$$

where $(x_{1i}^k, y_{1i}^k, z_{1i}^k)$ is the ordered vertex list of $\vec{W}^k$. It was found that this incremental strategy did not suffer from element reversal issues and also provided high-quality meshes. Similar observations were previously made by Shontz and Vavasis [96]. However, this improvement was achieved at the expense of increased computational cost. In the present study, due to computational constraints, we did not employ the incremental morphing strategy and instead chose to repair the registered volume mesh if any elements were produced with a negative signed volume. Mesh check and repair was fully automated and incorporated into the volume meshing procedure.

Appendix B

Characteristics of 21 Femur Statistical Model

B.0.3 Interpretation of Eigenmodes

	Covariance Method		Correlation Method	
Mode	Variance Captured	Total Variance	Variance Captured	Total Variance
1	34.85	34.85	31.97	31.97
2	10.13	44.98	25.37	57.33
3	6.16	51.14	17.58	74.92
4	5.76	56.90	4.26	79.18
5	4.64	61.54	2.86	82.04
6	4.28	65.81	2.34	84.42
7	3.77	69.59	1.69	86.53
8	3.26	72.85	1.54	88.21
9	3.17	76.01	1.26	89.76
10	2.88	78.90	1.20	91.02

Table B.1: Percentage variation captured by first 10 eigenmodes computed using the covariance and correlation methods.

Comparing the eigenvalues produced from PCA of the data set from the covariance and correlation approaches shows significant differences (Table B.1). The first mode contains a comparable proportion of the variation, $\sim 35\%$ with the covariance based and $\sim 32\%$ with the correlation based approach. However, when considering the first three modes the correlation approach captures close to 75% of the total variation compared to just 51% described by the more traditional method. This significant initial improvement together with a more rapid decay of eigenvalue suggests

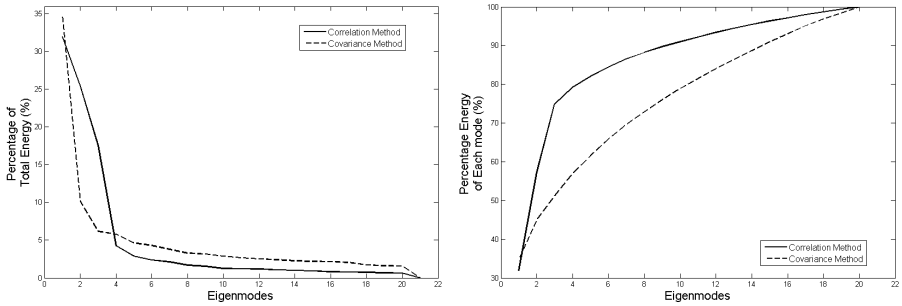


Figure B.1: Plots of eigenvalue percentage. Shows percentage for each mode (left) and cumulative percentage variation explained (right) for the covariance and correlation methods.

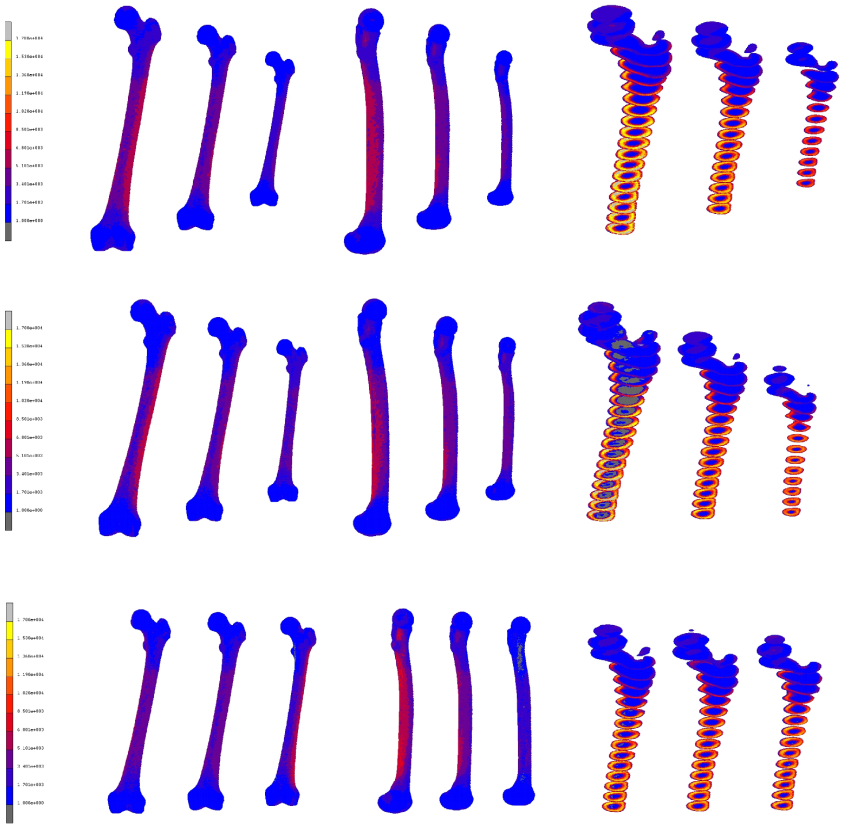


Figure B.2: Geometry and material property changes with first 3 modes of variation acting in isolation, produced using the correlation method.

that the correlation approach is more suited for the analysis of data with mixed units (Fig. B.1).

The physical effect on the model shape and material modulus of each eigenmode was investigated by varying each mode in isolation and visualising the femur produced (Fig. B.2). The first mode shows a straightening of the femoral axis towards the mid line in the frontal plane and a reduction in the bowing of the femoral shaft in the sagittal plane. Another significant geometric feature is a rotation of the femoral head in relation to the body of the femur, known as anteversion angle. The average modulus of the bone appears to rise, particularly in the shaft where there is a clear thickening of the cortex. The second mode was seen to be dominated by scaling affects and an increase in femoral shaft and neck diameter. Similarly the third mode shows an increase in shaft and neck thickness along with a large increase in anteversion angle as opposed to the reduction seen in the first mode. The influence of the modes is an insight into the significant ways in which this set of femurs vary, however it must be noted that these modes will never occur in isolation. In reality any femur will be the product of the combined effect of a number of modes, which may result in the features observed being cancelled out or exaggerated.

Reconstruction Error

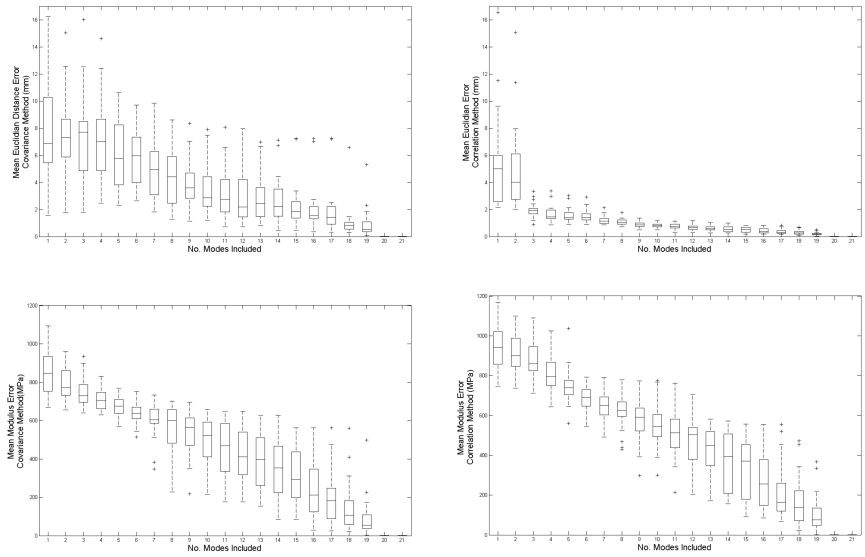


Figure B.3: Boxplot of mean Euclidean distance and mean modulus error between corresponding points when reconstructing a known instance with an increasing number of eigenmodes for both the covariance (left) and correlation methods (right).

Reconstruction error was calculated as a more appropriate test of the statistical model created in this study due to the small size of the training set. This was done by taking the known shape parameter of each training example and reconstructing it with an increasing number of included eigenvectors from the statistical model.

The mean Euclidian distance error and mean absolute modulus error at each point in each model were found as the number of included eigenvectors was increased using both the correlation and covariance methods (Fig. B.3). The tests show a significant improvement in reconstruction error with the correlation approach, where a mean error of <2mm is achieved by just 3 modes, corresponding to ∼8mm with the covariance which has a much higher error spread. The modulus error reflects the noisiness of modulus data being read in from the CT scans, showing little sensitivity to either method. The test results indicate the optimum number of eigenmodes required for the generation of new femur models, suggesting 3 modes were sufficient to construct geometry alone but at least 8 modes would be needed to fairly reproduce modulus values as well.

B.0.4 Ability to Describe Training Data

Reconstruction error is an excellent mathematical test of a statistical model, providing details of how accurate the model is in reproducing a specific instance. The prospective application of the model in this project is not, however, to regenerate specific geometries but to generate new instances with realistic characteristics. In view of this, it seems logical to investigate the generated geometries to ensure that this is what is happening by seeing how well they describe the variation present within the training data.

		1 S.D.				1.5 S.D.			
	Training	7	8	9	10	7	8	9	10
Neck Axis Length	100.19	100.05	100.09	100.07	100.08	100.12	100.18	100.17	100.20
Neck Shaft Angle	126.50	126.73	126.70	126.70	126.70	126.66	126.64	126.56	126.11
Femoral Head Dia.	47.43	47.33	47.36	47.35	47.35	47.37	47.41	47.42	47.41
Femoral Neck Dia. P-D	33.88	33.72	33.73	33.71	33.71	33.77	33.80	33.79	33.80
Femoral Neck Dia. A-P	33.29	33.03	33.04	33.03	33.04	33.15	33.18	33.18	33.20
Anteversion	22.03	22.02	22.01	22.00	22.00	21.98	21.98	21.96	21.98
Femoral Shaft Radius	16.26	16.05	16.06	16.06	16.07	16.11	16.12	16.13	16.15
Femoral Length	457.41	457.01	457.05	457.06	457.08	457.50	457.50	457.64	457.69

Table B.2: Table showing the mean value of various geometric measurements taken from 100 femurs generated with bounds of 1 and 1.5 standard deviations, at 7, 8, 9 and 10 included modes, alongside the mean values of the statistical model training data.

The reconstruction tests indicated that 8 modes would be needed to reproduce modulus and geometry, and that the correlation method was more suited to this data set. Therefore the correlation based model was used from this point forward and the optimum model limits investigated by using it to randomly generate 100 femurs using 7, 8, 9 and 10 modes and setting the bounds to +/- 1, 1.5 and 2 σ_{b_i}. Anatomically meaningful measurements were then extracted from each generated femur and from

		1 S.D.				1.5 S.D.			
	Training	7	8	9	10	7	8	9	10
Neck Axis Length	29.51	19.20	19.91	19.95	20.24	28.85	29.90	29.96	30.38
Neck Shaft Angle	10.59	6.58	6.50	6.53	8.21	10.00	10.62	10.41	56.46
Femoral Head Dia.	16.45	11.64	11.91	12.04	12.17	17.44	17.86	18.03	18.26
Femoral Neck Dia. P-D	12.31	8.37	8.52	9.57	9.13	12.55	12.75	14.21	13.55
Femoral Neck Dia. A-P	15.96	10.68	10.82	12.39	11.93	15.93	16.16	18.49	17.79
Anteversion	14.30	6.90	8.25	7.95	7.92	9.93	12.22	11.89	11.69
Femoral Shaft Radius	6.01	4.01	4.05	4.53	4.42	5.94	6.06	6.74	6.58
Femoral Length	89.53	58.13	58.48	62.14	61.67	85.34	86.21	88.40	89.59

Table B.3: Table showing the spread in values of various geometric measurements taken from 100 femurs generated with bounds of 1 and 1.5 standard deviations, at 7, 8, 9 and 10 included modes,alongside the mean values of the statistical model training data.

every training example, these were measurements commonly used to describe femoral geometry, including femoral neck axis length, femoral head diameter, neck diameter and neck-shaft angle [103, 143]. As each example was based on the same baseline mesh there was accurate point to point correspondence between models, making this process relatively simple, involving only the identification of the key node numbers and some simple trigonometry.

The mean values and spread of results produced with 7-10 included modes at 1 and 1.5 σ_{b_i} alongside the statistics of the training data the model is attempting to represent, are shown in Tables B.2 and B.3. Even with a low number of eigenmodes the mean measurements were consistent with the training set, this remained true as the number of modes increased. However, by 10 modes some parameters began to exceed the spread of values seen in the training data and the quality of the meshes produced began to deteriorate slightly. The spread of geometric values generated shows a clearer difference between the different boundary limits and to a limited extent the number of modes. At +/-1 standard deviations the range is significantly smaller than the training set, whereas at +/-1.5 the range is very similar. Overall 8 modes and a boundary of +/-1.5 σ_{b_i} appears to match the mean and spread of training data geometry measurements most closely. The results for 2 s.d. are not shown, although they also showed comparable mean values to the original data, the spread of the geometries exceeded the training set, at the detriment of the realistic appearance of some models and mesh quality was seen to suffer.

All femur geometries produced at +/-1.5 s.d. gave realistic anatomical measurements, comparable to the training set. Although generally on the smaller side, the measurements extracted were within the bounds of data reported from clinical papers investigating femur geometry [103, 143]. The differences most likely reflect the characteristics of the training data and differences in measuring styles, as most clinical

studies report measurements from two dimensional imaging techniques such as X-ray and the current study calculated distances between three dimensional positions.

B.0.5 Mesh Quality Checks

Mesh quality is a key concern when automating the generation of finite element models. As the statistical model is created by stretching and morphing a baseline mesh some level of mesh distortion is inevitable. It is crucial that despite this, mesh quality is maintained such that meshes can be used in finite element simulation without causing errors or producing inaccurate results.

To investigate the general mesh quality resulting from the statistical model generation scheme, the distribution of the quality scores produced by both the NSR and SDR metrics were recorded for 100 generated femurs (see section B.0.5). Figure B.4 shows the mean quality scores (solid lines) with a bounding envelope of the minimum and maximum scores seen (dashed). In addition, the quality of the baseline mesh is also shown for comparison (dotted line). Mesh quality is seen to drop from the baseline, which is expected, but the quality remains high for both metrics.

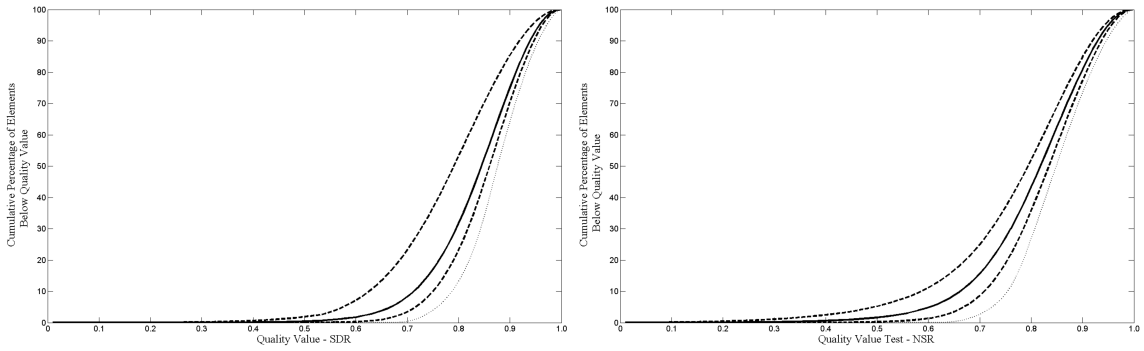
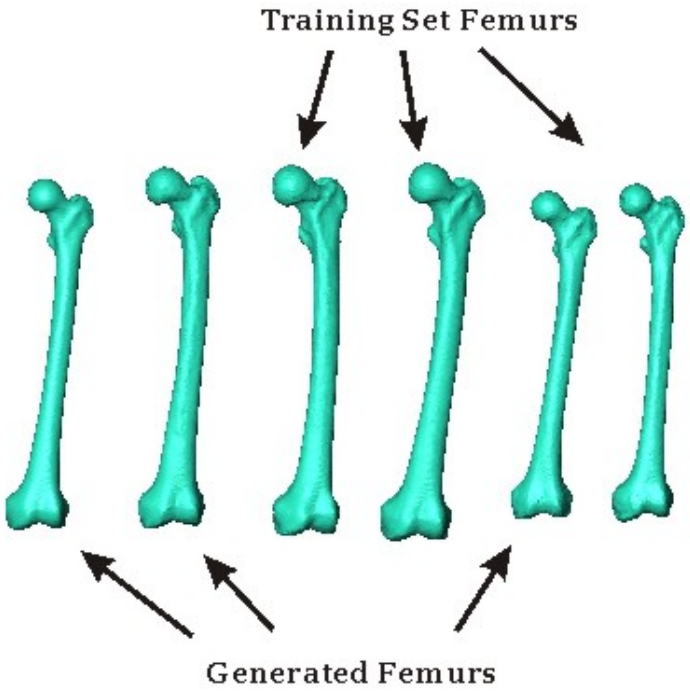


Figure B.4: Graphs of the percentage cumulative element mesh quality for (a) distortion metric and (b) normalised shape ration. Solid lines shows mean value, dashed lines indicate min and max bounds and dotted line shows original baseline mesh quality.

Appendix C

Images of Training and Generated Femurs



Appendix D

Table of Assumptions

Stage	Assumptions
Data	No Pathology in the data set. i.e. all healthy subjects 46 subjects are sufficient to represent the variation in geometry and material distribution in the femur Over representation of males in training set due to available data Lack of ethnic diversity in training set due to available data
Segmentation	Possibility of introduction of manual errors due to data interpretation Possible inaccuracy due to smoothing process required to generate smooth surfaces from voxelised segmentation
Material Property Allocation	Possible smoothing effect due to sampling and averaging method of elemental greylevel interpretation Possible artificial partial volume effects on element modulus around material boundaries Assumed maximum bone density present in all bones to allow calibration as no phantoms present in images Reliant on single density to modulus relationship from the literature
Registration	Accurate surface and volume representation when morphing baseline mesh onto each example High mesh quality after morphing Nodal and elemental correspondence between all training exampled following registration and morphing
PCA Sampling	Appropriate number of modes included to represent data set without including noise leading to mesh degradation Sampling limits appropriate to represent data set while not extending too far and generating infeasible/unrealistic models Normalised sampling technique suitable to represent the training set distribution

Appendix E

An Automated, Large Scale Finite Element Study into the Influence of Under and Over Sizing Femoral Resurfacing on Load Transfer Through the Proximal Femur

Bryan R. [1*], Nair P. B.[2*], Taylor M. [3*]

[1*] Bioengineering Sciences Research Group, University of Southampton, UK. RB102@soton.ac.uk
[2*] Computational Engineering and Design Centre, University of Southampton, UK. P.B.Nair@soton.ac.uk
[3*] Bioengineering Sciences Research Group, University of Southampton, UK. M.Taylor@soton.ac.uk

Introduction: Resurfacing arthroplasty (RA) is regarded as an attractive alternative to full hip replacement for younger active patients, offering more natural replication of joint kinematics and greater bone conservation. However, it is known that the success of RA is highly influenced by patient selection and surgical skill, with poorer outcomes seen particularly in older females with smaller implant sizes[1].

The focus of this work was the influence of incorrect implant size selection on load transfer through the proximal femur. A novel technique involving statistical modelling of the femur was used to generate large numbers of synthetic femurs, each with unique geometries and material properties[3]. All models were automatically implanted to analyse the effects of under and over sizing resurfacing components on a wide variety of femurs.

Materials and methods: 290 FE ready femur models, with varying and unique geometry and material properties, were generated by sampling a statistical model built from 46 Computer Tomography scans[3]. A fully automated methodology virtually resurfaced each femur firstly with a correctly sized implant, then with implants 1 size too big and 1 size too small. The intact femur was also analysed as a baseline.

The automated model generation process ran as follows; Implant sizing and position was calculated on a model by model basis using Matlab®. HyperMesh™ (Altair Engineering Inc, USA) was used to perform Boolean operations, cutting the femur and creating a 3mm thick cement mantle. Ansys ICEM™ (Ansys Inc., Cannonsburg, PA) then generated tetrahedral meshes of the bone (0.5-1.5mm proximal, 2-4mm distal), cement (0.75-1.5mm) and implant (0.75-1.5mm) with nodal correspondence across boundaries. Elemental material properties of the cut bone were reassigned from the original femur mesh. Identical subject specific load were applied simulating level gait[4]. The load was defined by relating femur length to stature and assigning a BMI to give a unique subject weight. A static, linear elastic FE analysis was then run and the resulting elemental strain extracted.

Results and Discussion: Comparison of strain change from the intact to implanted configurations, for all 3 scenarios, were calculated over 16 proximal sections. These were defined identically on each femur by dividing the bones along and perpendicular to the neck axis (Fig.1). One-way ANOVA tests revealed the main difference was observed in strain shielding in the head sections. A statistically significant difference between the groups was seen in all but one of the head sections (Fig.1).

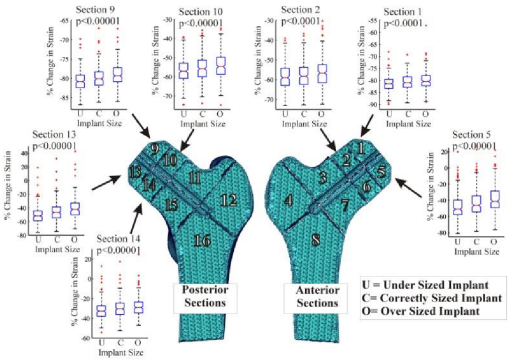


Fig. 1:Illustrating proximal femur sections with statistically significantly different mean strain changes following resurfacing with under- and over- sized implanted as compared to the nominal size. Boxplots contain: median, 25th and 75th percentiles, notches showing comparison intervals and whiskers extending to 3s.d.

The undersized group consistently showed the greatest strain shielding in the head while oversizing showed the least. Interrogation of the femurs indicated that in all scenarios smaller implants performed worse than larger implants, with greater strain shielding in the anterior head (sec.1&5 $p<0.0025$) and greater strain increase in the posterior neck (sec.11&15 $p<0.00001$).

Conclusion: This study has run a fully automated comparative analysis of nearly 300 different femurs to assess one possible source of surgical error, combining surgical and intersubject variability. The findings illustrated, with statistical significance in 7/8 head sections, that using an undersized component was likely to increase strain shielding in the femoral head.

The technique described offers a method for population based analysis in computational biomechanics in the future, as opposed to the subject by subject approach which is currently used. The method also has the flexibility to be applied to a wide range of future investigations, such assessing the impact of incorrect implant positioning or comparing implant designs.

Acknowledgements: Thank you to Andrew Hopkins, Surya Mohan, Francis Galloway and Technology Strategy Board UK.

References: [1] Australian Orth. Assoc. NJRAR. 2008. [2] Amstutz H.C. et al., JBJS.86(1).2004 [3] Bryan R. et al., ME&P. 2009 [4] Heller et al. J. Biomech;34 883-93. 2001

References

[1] R. Marks, J.P. Allegrante, C. Ronald MacKenzie, and J.M. Lane. Hip fractures among the elderly: causes, consequences and control. *Ageing Research Reviews*, 2(1):57–93, 2003.

[2] S. Kobayashi, N. Saito, H. Horiuchi, R. Iorio, and K. Takaoka. Poor bone quality or hip structure as risk factors affecting survival of total-hip arthroplasty. *The Lancet*, 355(9214):1499–1504, 2000.

[3] A. S. Wong, A. M. R. New, G. Isaacs, and M. Taylor. Effect of bone material properties on the initial stability of a cementless hip stem: a finite element study. *Proceedings of the Institution of Mechanical Engineers Part H-Journal of Engineering in Medicine*, 219:265–275, 2005.

[4] M. Morlock, N. Bishop, W. Rüther, G. Delling, and M. Hahn. Biomechanical, morphological, and histological analysis of early failures in hip resurfacing arthroplasty. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 220(2):333–344, 2006.

[5] P. J. Prendergast. Finite element models in tissue mechanics and orthopaedic implant design. *Clinical Biomechanics*, 12(6):343–366, 1997.

[6] A. Pancanti, M. Bernakiewicz, and M. Viceconti. The primary stability of a cementless stem varies between subjects as much as between activities. *Journal of Biomechanics*, 36(6):777–785, 2003.

[7] I.A.J. Radcliffe and M. Taylor. Investigation into the effect of varus-valgus orientation on load transfer in the resurfaced femoral head: A multi-femur finite element analysis. *Clinical Biomechanics*, 22(7):780–786, 2007.

[8] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[9] T.F. Cootes and C.J. Taylor. Statistical models of appearance for medical image analysis and computer vision. *Procedings SPIE Medical Imaging*, 42, 2001.

[10] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[11] F.M. Vos, P.W. Bruin, J.G.M. Aubel, G.J. Streekstra, M. Maas, L.J. van Vliet, and A.M Vossepoel. A statistical shape model without using landmarks. *Proceedings of the 17th International Conference on Pattern Recognition. ICRP.*, 3, 2004.

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

[13] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: Application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999. 45.

[14] B. Couteau, Y. Payan, and S. Lavellée. The mesh matching algorithm: an automatic 3d mesh generator for finite element structures. *Journal of Biomechanics*, 33(8):1005–1009, 2000.

[15] L.B. Querol, P. Buchler, D. Rueckert, L.P. Nolte, and M.G. Ballester. *Statistical finite element model for bone shape and biomechanical properties.* Medical Image Computing and Computer-Assisted Intervention. 2006.

[16] R.L. Drake, W. Vogl, and A. Mitchell. *Gray's Anatomy for Students.* Elsevier/ Churchill Livingstone, Philipdelphia; London, 2005. 3.

[17] C. Norkin and D. White. *Measurement of Joint Motion: a Guide to Goniometry.* Philadelphia : F.A. Davis, 2nd ed edition, 1995.

[18] NASA RP. Anthropometric source book: Volume 1: Anthropometry for designers anthropology staff/webb associates. *NASA*, pages 7–78.

[19] J.A. DeLisa. *Gait Analysis In The Science Of Rehabilitation.* DIANE Publishing, 1998.

[20] P. Levangie and C. Norking. *Joint Structure and Function.* F.A. Davis, Philidelphia, 3rd edition, 2001.

[21] O. Husmann, P. J. Rubin, P. F. Leyvraz, B. deRoguin, and J. N. Argenson. Three-dimensional morphology of the proximal femur. *Journal of Arthroplasty*, 12(4):444–450, 1997. 6.

[22] J.C. Koch. The laws of bone architecture. *American Journal of Anatomy*, 21 (2):177–298, 1917.

[23] M. Nordin and V. H. Frankel. *Basic biomechanics of the musculoskeletal system.* Lippincott, Williams and Wikins, Philadelphia, 2001. 8.

[24] S.C. Cowin. *Bone mechanics.* CRC Press Boca Raton, FL, 1989.

[25] E.F. Morgan and T.M. Keaveny. Dependence of yield strain of human trabecular bone on anatomic site. *Journal of Biomechanics*, 34(5):569–577, 2001.

[26] R.L. Braddom, R.M. Buschbacher, L. Chan, K.J. Kowalske, E. Laskowski, D.J. Matthews, and K.T. Ragnarsson. *Physical Medicine and Rehabilitation.* Saunders, 2000.

[27] J.P. Paul. Forces transmitted by joints in the human body. *Proceedings of Institution of Mechanical Engineers*, 181(3J):8–15, 1967.

[28] G.N. Duda, E. Schneider, and E.Y.S. Chao. Internal forces and moments in the femur during walking. *Journal of Biomechanics*, 30(9):933–941, 1997.

[29] H. Röhrle, R. Scholten, C. Sigolotto, W. Sollbach, and H. Kellener. Joint forces in the human pelvis-leg skeleton during walking. *Journal of Biomechanics*, 17 (6):409–424, 1984.

[30] G. Bergmann, F. Graichen, and A. Rohlmann. Hip joint loading during walking and running, measured in 2 patients. *Journal of Biomechanics*, 26(8):969–990, 1993.

[31] G. Bergmann, G. Deuretzbacher, M. Heller, F. Graichen, A. Rohlmann, J. Strauss, and G. N. Duda. Hip contact forces and gait patterns from routine activities. *Journal of Biomechanics*, 34(7):859–871, 2001.

[32] R.A. Brand, D.R. Pedersen, D.T. Davy, G.M. Kotzar, K.G. Heiple, and V.M. Goldberg. Comparison of hip force calculations and measurements in the same patient. *The Journal of Arthroplasty*, 9(1):45–51, 1994.

[33] MO Heller, G. Bergmann, G. Deuretzbacher, L. Dürselen, M. Pohl, L. Claes, NP Haas, and GN Duda. Musculo-skeletal loading conditions at the hip during walking and stair climbing. *Journal of biomechanics*, 34(7):883–893, 2001.

[34] N.W. Rydell. Forces acting on the femoral head-prosthesis. A study on strain gauge supplied prostheses in living persons. *Acta Orthopeadica Scandinavica*, 37(88):1–132, 1966.

[35] TA English and M. Kilvington. In vivo records of hip loads using a femoral implant with telemetric output (a preliminary report). *Journal of Biomedical Engineering*, 1(2):111–5, 1979.

[36] WA Brekelmans, HW Poort, and TJ Slooff. A new method to analyse the mechanical behaviour of skeletal parts. *Acta Orthopaedica Scandinavica*, 43(5): 301, 1972.

[37] IAJ Radcliffe and M. Taylor. Investigation into the affect of cementing techniques on load transfer in the resurfaced femoral head: A multi-femur finite element analysis. *Clinical Biomechanics*, 22(4):422–430, 2007.

[38] J. H. Keyak, J. M. Meagher, H. B. Skinner, and C. D. Mote. Automated 3-dimensional finite element modeling of bone - a new method. *Journal of Biomedical Engineering*, 12(5):389–397, 1990.

[39] M. Viceconti, L. Bellingeri, L. Cristofolini, and A. Toni. A comparative study on different methods of automatic mesh generation of human femurs. *Medical Engineering and Physics*, 20(1):1–10, 1998.

[40] M. Viceconti, M. Davinelli, F. Taddei, and A. Cappello. Automatic generation of accurate subject-specific bone finite element models to be used in clinical studies. *Journal of Biomechanics*, 37(10):1597–1605, 2004.

[41] E. Schileo, F. Taddei, A. Malandrino, L. Cristofolini, and M. Viceconti. Subject specific finite elements can accurately predict strain levels in long bones. *Journal of Biomechanics*, 40(13):2983–2989, 2007.

[42] H. Lamecker, M. Seebaß, H.C. Hege, and P. Deuflhard. A 3d statistical shape model of the pelvic bone for segmentation. In *Proc. SPIE Medical Imaging: Image Processing*, volume 5370, pages 1341–1351. Citeseer, 2004.

[43] D. Kainmueller, T. Lange, and H. Lamecker. Shape constrained automatic segmentation of the liver based on a heuristic intensity model. In *Proc. MICCAI Workshop 3D Segmentation in the Clinic: A Grand Challenge*, pages 109–116. Citeseer, 2007.

[44] J. Schmid and N. Magnenat-Thalmann. MRI bone segmentation using deformable models and shape priors. In *Proceedings of the 11th international*

conference on Medical Image Computing and Computer-Assisted Intervention-Part I, pages 119–126. Springer, 2008.

[45] R. Huiskes, P.H.G.E. Strens, J. van Heck, and T.J.J.H. Slooff. Interface stresses in the resurfaced hip: finite element analysis of load transmission in the femoral head. *Acta Orthopaedica*, 56(6):474–478, 1985.

[46] J.C. Lotz, E.J. Cheal, and W.C. Hayes. Fracture prediction for the proximal femur using finite element models: Part I - linear analysis. *Journal of Biomechanical Engineering*, 113:353, 1991.

[47] R. Huiskes and B. Van Rietbergen. Preclinical testing of total hip stems. *Clin Orthop Relat Res*, pages 64–76, 1995.

[48] J.R. Britton and P.J. Prendergast. Preclinical testing of femoral hip components: an experimental investigation with four prostheses. *Journal of biomechanical engineering*, 127:872, 2005.

[49] KL Ong, SM Kurtz, MT Manley, N. Rushton, NA Mohammed, and RE Field. Biomechanics of the Birmingham hip resurfacing arthroplasty. *Journal of Bone & Joint Surgery, British Volume*, 88(8):1110, 2006.

[50] S. Majumder, A. Roychowdhury, and S. Pal. Simulation of hip fracture in sideways fall using a 3D finite element model of pelvis–femur–soft tissue complex with simplified representation of whole body. *Medical Engineering and Physics*, 29(10):1167–1178, 2007.

[51] J.H. Keyak, S.A. Rossi, K.A. Jones, and H.B. Skinner. Prediction of femoral fracture load using automated finite element modeling. *Journal of Biomechanics*, 31(2):125–133, 1997.

[52] M. Lengsfeld, R. Burchard, D. Gunther, T. Pressel, J. Schmitt, R. Leppek, and P. Griss. Femoral strain changes after total hip arthroplasty patient-specific finite element analyses 12 years after operation. *Medical Engineering and Physics*, 27(8):649–654, 2005.

[53] S.K. Easley, S. Pal, P.R. Tomaszewski, A.J. Petrella, P.J. Rullkoetter, and P.J. Laz. Finite element-based probabilistic analysis tool for orthopaedic applications. *Computer methods and programs in biomedicine*, 85(1):32–40, 2007.

[54] P.J. Laz, J.Q. Stowe, M.A. Baldwin, A.J. Petrella, and P.J. Rullkoetter. Incorporating uncertainty in mechanical properties for finite element-based evaluation of bone mechanics. *Journal of Biomechanics*, 40(13):2831–2836, 2007.

[55] M.T. Bah and M. Browne. Effect of Geometrical Uncertainty on Cemented Hip Implant Structural Integrity. *Journal of Biomechanical Engineering*, 131: 054501, 2009.

[56] D.P. Nicolella, B.H. Thacker, H. Katoozian, and D.T. Davy. Probabilistic risk analysis of a cemented hip implant. *ASME-PUBLICATIONS-BED*, 50:427–428, 2001.

[57] C. Dopico-González, A.M. New, and M. Browne. Probabilistic analysis of an uncemented total hip replacement. *Medical Engineering and Physics*, 31(4): 470–476, 2009.

[58] M. Viceconti, G. Brusi, A. Pancanti, and L. Cristofolini. Primary stability of an anatomical cementless hip stem: A statistical analysis. *Journal of biomechanics*, 39(7):1169–1179, 2006.

[59] National joint registry: Annual report 2005 - 2006. *National Joint Registry Steering Committee and National Joint Registry*, 2006.

[60] Canadian joint replacement registry 2006 report - hip and knee replacements in canada. *Toronto: Canadian Institute for Health Information*, 2006. 12a.

[61] Australian Orthopaedic Association National Joint Replacement Registry. Annual report. 2008.

[62] National joint registry for england and wales 5th annual report. *National Joint Registry Steering Committee and National Joint Registry*, 2008.

[63] D. Terzopoulos, J. Platt, A. Barr, and K. Fleischer. Elastically deformable models. In *B Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 205–214. ACM Press, 1987.

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

[65] G. Zsemlye. *Shape prediction from partial infotmation.* PhD thesis, Swiss Federal Institute of Technology, 2005.

[66] H. Bornfleth, G. Vincent, M. Slomczykowski, M. Bowes, and A. Ashby. Three-dimensional statisitcal proximal femur modelling from two x-ray images for a new generation of prostheses, 19-21 April 2007. 37.

[67] M. Sermesant, C. Forest, X. Pennec, H. Delingette, and N. Ayache. Deformable biomechanical models: Application to 4d cardiac image analysis. *Medical Image Analysis*, 7(4):475–488, 2003.

[68] J. Hladuvka. Establishing point correspondence on training set boundaries. Technical report, Technical Report TR-VRVis-2003-040, 2003.

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

[70] K.O. Babalola, T.F. Cootes, C.J. Twining, V. Petrovic, and C. Taylor. 3D brain segmentation using active appearance models and local regressors. In *Proceedings of the 11th international conference on Medical Image Computing and Computer-Assisted Intervention-Part I*, page 408. Springer, 2008.

[71] M. A. G. Ballester, X. Pennec, M. G. Linguraru, and N. Ayache. Generalized image models and their application as statistical models of images. *Medical Image Analysis*, 8(3):361–369, 2004.

[72] Y. H. Yang, A. Bull, D. Rueckert, and A. Hill. 3d statistical shape modeling of long bones. In *Biomedical Image Registration, Proceedings*, volume 4057 of *Lecture Notes In Computer Science*, pages 306–314, 2006.

[73] M.R. Kaus, V. Pekar, C. Lorenz, R. Truyen, S. Lobregt, and J. Weese. Automated 3-D PDM construction from segmented images using deformable models. *IEEE Transactions on Medical Imaging*, 22(8):1005–1013, 2003.

[74] K.T. Rajamani, S.C. Joshi, and M.A. Styner. Bone model morphing for enhanced surgical visualisation. *International Symposium on Biomedical Imaging : Macro to Nano*, pages 1255–1258, 2004.

[75] K.T. Rajamani, M.A. Styner, H. Talib, G. Zheng, L.P. Nolte, and M.A.G. Ballester. Statistical deformable bone models for robust 3d surface extrapolation from sparse data. *Medical Image Analysis*, 11:99–109, 2007.

[76] J. Hohe, S. Faber, R. Muehlbauer, M. Reiser, K. H. Englmeier, and F. Eckstein. Three-dimensional analysis and visualization of regional mr signal intensity distribution of articular cartilage. *Medical Engineering and Physics*, 24(3): 219–227, 2002. 49.

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

[78] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46(3):R1–R45, 2001.

[79] H. L. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, 2003.

[80] V. E. Garzon and D. L. Darmofal. Impact of geometric variability on axial compressor performance. *Journal of Turbomachinery-Transactions of the Asme*, 125(4):692–703, 2003. 57.

[81] M. Staring, S. Klein, and J.P.W. Pluim. Nonrigid registration with tissue-dependent filtering of the deformation field. *Physics in Medicine and Biology*, 52(23):6879–6892, 2007.

[82] J. Yao and R.H. Taylor. Tetrahedral mesh modeling of density data for anatomical atlases and intensity-based registration. *Lecture notes in computer science*, pages 531–540, 2000.

[83] RH Davies, CJ Twining, TF Cootes, JC Waterton, and CJ Taylor. A minimum description length approach to statistical shape modeling. *IEEE Transactions on Medical Imaging*, 21(5):525–537, 2002.

[84] M. Taylor. Finite element analysis of the resurfaced femoral head. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 220(2):289–297, 2006.

[85] A.F. Frangi, D. Rueckert, J.A. Schnabel, and W.J. Niessen. Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling. *IEEE Transactions On Medical Imaging*, 21(9): 1151–1166, 2002.

[86] C.S.K. Chan. Cadaver validation of the use of ultrasound for 3d model instantiation of bony anatomy in image guided orthopaedic surgery. *Medical Image Computing And Computer-Assisted Intervention - Miccai 2004, Pt 2, Proceedings.*, pages 397–404, 2004. 35.

[87] M. Fleute, S. Lavallee, and L. Desbat. Integrated approach for matching statistical shape models with intra-operative 2d and 3d data. pages 364–72, Tokyo, Japan, 2002. Springer-Verlag.

[88] C. Zannoni, R. Mantovani, and M. Viceconti. Material properties assignment to finite element models of bone structures: a new method. *Medical Engineering and Physics*, 20(10):735–740, 1998. 38.

[89] D.R. Carter. Mechanical loading histories and cortical bone remodeling. *Calcified Tissue International*, 36:19–24, 1984.

[90] E. Morgan, H. Bayraktar, and T. Keaveny. Trabecular bone modulus-density relationships depend on anatomic site. *Journal of Biomechanics*, 36(7):897–904, 2003.

[91] M. Moshfeghi, S. Ranganath, and K. Nawyn. Three-dimensional elastic matching of volumes. *IEEE Transactions on Image Processing*, 3(2):128–138, 1994.

[92] H. Samet. *Foundations of Multidimensional and Metric Data Structures*. Morgan-Kaufmann, San Francisco, 2006.

[93] J. Vollmer, R. Mencl, and H. Muller. Improved laplacian smoothing of noisy surface meshes. *Computer Graphics Forum*, 18:131–138, September 1999.

[94] I. Robertson and S. Sherwin. Free-surface flow simulation using hpSpectral elements. *Journal of Computational Physics*, 155(1):26–53, 1999.

[95] O. Hassan, E.J. Probert, and K. Morgan. Unstructured mesh procedures for the simulation of three-dimensional transient compressible inviscid flows with moving boundary components. *International Journal for Numerical Methods in Fluids*, 27(14):41–55, 1998.

[96] S. Shontz and S. Vavasis. An algorithm based on finite element weights for warping tetrahedral meshes. *Submitted to SIAM Journal on Scientific Computing*, 2008.

[97] B.T. Helenbrook. Mesh deformation using the biharmonic operator. *International Journal for Numerical Methods in Engineering*, 56:1007–1021, 2003.

[98] D. Field. Qualitative measures for initial meshes. *Inernaitonal Journal for Numerical Methods in Engineering*, 47:887–906, 2000.

[99] A. Liu and B. Joe. On the shape of tetrahedra from bisection. *Mathematics of Computation*, 63(207):141–154, 1994.

[100] I. Joliffe. *Principal component analysis*. Springer - Verlag, New York, 1986.

[101] C.B. Ruff and W.C. Hayes. Cross-sectional geometry of pecos pueblo femora and tibiae-a biomechanical investigation: I. method and general patterns of variation. *American Journal of Physical Anthropology*, 60(3), 1983.

[102] D.D. Cody, G.J. Gross, F. J. Hou, H.J. Spencer, S.A. Goldstein, and D. P. Fyhrie. Femoral strength is better predicted by finite element models than QCT and DXA. *Journal of Biomechanics*, 32(10):1013–1020, 1999.

[103] T. Theobald, J. Cauley, C. Gluer, C. Bunker, F. Ukoli, and H. Genant. Black-white differences in hip geometry. *Osteoporosis International*, 8(1):61–67, 1998.

[104] S. Gnudi, C. Ripamonti, G. Gualtieri, and N. Malavolta. Geometry of proximal femur in the prediction of hip fracture in osteoporotic women. *British Journal of Radiology*, 72(860):729, 1999.

[105] R.G. Ghanem and A. Doostan. On the construction and analysis of stochastic models: characterization and propagation of the errors associated with limited data. *Journal of Computational Physics*, 217(1):63–81, 2006.

[106] J. Michelotti and J. Clark. Femoral neck length and hip fracture risk. *Journal of Bone and Mineral Research*, 14(10):1714–1720, 1999.

[107] Centers for Disease Control & Prevention (CDC). National health and nutrition examination survey data. *National Center for Health Statistics (NCHS)*, 2003.

[108] D.R. Carter and W.C. Hayes. Bone compressive strength: the influence of density and strain rate. *Science*, 194(4270):1174–1176, 1976.

[109] Julia A Schnabel, Daniel Rueckert, Marcel Quist, Jane M Blackall, Andy D Castellano-Smith, Thomas Hartkens, Graeme P Penney, Walter A Hall, Haiying Liu, Charles L Truwit, Frans A Gerritsen, Derek L G Hill, and David J Hawkes. *A Generic Framework for Non-rigid Registration Based on Non-uniform Multi-level Free-Form Deformations*, volume 22, chapter Medical Im, pages 573–581. Springer-Verlag, 2001. doi: 10.1007/3-540-45468-369.

[110] M. Fluete and S. Lavellee. *Medial Image Computing and Computer-Assisted Intervention - MICCAI'98. First International Conference*, pages 879–87, 1998.

[111] P. Dolan and D.J. Torgerson. The cost of treating osteoporotic fractures in the United Kingdom female population. *Osteoporosis International*, 8(6):611–617, 1998.

[112] J.G. Evans, V. Seagroatt, and M.J. Goldacre. Secular trends in proximal femoral fracture, Oxford record linkage study area and England 1968-86. *Journal of Epidemiology and Community Health*, 51(4):424–429, 1997.

[113] United Nations Population Division. *World population aging 1950-2050*. New York: United Nations, 2003.

[114] R.Y. Hinton. The association of age, race, and sex with the location of proximal femoral fractures in the elderly. *The Journal of Bone and Joint Surgery*, 75 (5):752–759, 1993.

[115] J.D. Michelson, A. Myers, R. Jinnah, Q. Cox, and M. Van Natta. Epidemiology of hip fractures among the elderly. Risk factors for fracture type. *Clinical Orthopeadic and Relatated Research*, 311:129–35, 1995.

[116] K.M. Fox, J. Magaziner, J.R. Hebel, J.E. Kenzora, and T.M. Kashner. Intertrochanteric versus femoral neck hip fractures: differential characteristics,

treatment, and sequelae. *Journals of Gerontology Series A: Biological and Medical Sciences*, 54(12):635–640, 1999.

[117] A.J. van den Kroonenberg, W.C. Hayes, and T.A. McMahon. Hip impact velocities and body configurations for voluntary falls from standing height. *Journal of Biomechanics*, 29(6):807–811, 1996.

[118] D.P. Rice, E.J. MacKenzie, Centers for Disease Control (US, and National Highway Traffic Safety Administration. *Cost of injury in the United States: a report to Congress*. Produced by Institute for Health & Aging, University of California, and Injury Prevention Center, School of Hygiene and Public Health, the Johns Hopkins University, 1989.

[119] W.C. Hayes, E.R. Myers, J.N. Morris, T.N. Gerhart, H.S. Yett, and L.A. Lipsitz. Impact near the hip dominates fracture risk in elderly nursing home residents who fall. *Calcified Tissue International*, 52(3):192–198, 1993.

[120] J. Parkkari. Majority of hip fractures occur as a result of a fall and impact on the greater trochanter of the femur: A prospective controlled hip fracture study with 206 consecutive patients. *Calcified Tissue International*, 65(3):183–187, 1999.

[121] J.H. Keyak, H.B. Skinner, and J.A. Fleming. Effect of force direction on femoral fracture load for two types of loading conditions. *Journal of Orthopaedic Research*, 19:539–544, 2001.

[122] P. Kannus, J. Parkkari, S. Niemi, M. Pasanen, M. Palvanen, M. Jarvinen, and I. Vuori. Prevention of Hip Fracture in Elderly People with Use of a Hip Protector, 2000.

[123] L.D. Smith. Hip Fractures: The role of muscle contraction or intrinsic forces in the causation of fractures of the femoral neck. *The Journal of Bone and Joint Surgery*, 35(2):367, 1953.

[124] AL Armstrong and WA Wallace. The epidemiology of hip fractures and methods of prevention. *Acta Orthopaedica Belgica*, 60(1):85–101, 1994.

[125] T.N. Gardner, A. Simpson, C. Booth, P. Sprukkelhorst, M. Evans, J. Kenwright, and J. Grimley Evans. Measurement of impact force, simulation of fall and hip fracture. *Medical Engineering and Physics*, 20(1):57–65, 1998.

[126] S.R. Cummings, J.L. Kelsey, M.C. Nevitt, and K.J. O'Dowd. Epidemiology of osteoporosis and osteoporotic fracatues. *Epidemiologic Reviews*, 7(1):178–208, 1985.

[127] C. Cooper, G. Campion, and L.J. Melton. Hip fractures in the elderly: A world-wide projection. *Osteoporosis International*, 2(6):285–289, 1992.

[128] K.J. Koval and J.D. Zuckerman. Hip fractures: I. overview and evaluation and treatment of femoral-neck fractures. *Journal of American Acadademy of Orthopaedic Surgery*, 2(3):141–149, 1994.

[129] P. Lips. Epidemiology and predictors of fractures associated with osteoporosis. *American Journal of Medicine*, 103(2A):3S–8S, 1997.

[130] R. Berntsen, V. Fonnebo, A. Tollan, A. Johanne Sogaard, and J.H. Magnus. Forearm Bone Mineral Density by Age in 7,620 Men and Women The Tromso Study, a Population-based Study. *American Journal of Epidemiology*, 153(5): 465, 2001.

[131] A.R. Lyons. Clinical outcomes and treatment of hip fractures. *American Journal of Medicine*, 103(2A):51S–63S, 1997.

[132] G.S. Keene, M.J. Parker, and G.A. Pryor. Mortality and morbidity after hip fractures. *British Medical Journal*, 307(6914):1248, 1993.

[133] C.W. Davidson, M.J. Merrilees, T.J. Wilkinson, J.S. McKie, and N.L. Gilchrist. Hip fracture mortality and morbidity - can we do better. *New Zealand Medical Journal*, 114(1136):329, 2001.

[134] A.H. Myers, E.G. Robinson, M.L.V. Natta, J.D. Michelson, K. Collins, and S.P. Baker. Hip fractures among the elderly: factors associated with in-hospital mortality. *American Journal of Epidemiology*, 134(10):1128–1137, 1991.

[135] C. Cooper. The crippling consequences of fractures and their impact on quality of life. *American Journal of Medicine*, 103(2A):12S–17S, 1997.

[136] M. Karlsson, JÅ Nilsson, I. Sernbo, I. Redlund-Johnell, O. Johnell, and K.J. Obrant. Changes of bone mineral mass and soft tissue composition after hip fracture. *Bone*, 18(1):19–22, 1996.

[137] B. Gullberg, O. Johnell, and JA Kanis. World-wide projections for hip fracture. *Osteoporosis International*, 7(5):407–413, 1997.

[138] S.R. Cummings, S.M. Rubin, and D. Black. The future of hip fractures in the United States: numbers, costs, and potential effects of postmenopausal estrogen. *Clinical Orthopaedics and Related Research*, (252):163–166, 1990.

[139] X. Ling, L. Aimin, Z. Xihe, C. Xiaoshu, and S.R. Cummings. Very Low Rates of Hip Fracture in Beijing, People's Republic of China: The Beijing Osteoporosis Project, 1996.

[140] P. Kannus, S. Niemi, J. Parkkari, M. Palvanen, I. Vuori, and M. Jarvinen. Hip fractures in Finland between 1970 and 1997 and predictions for the future. *Lancet*, 353(9155):769–70, 1999.

[141] J.Y. Reginster, P. Gillet, and C. Gosset. Secular increase in the incidence of hip fractures in Belgium between 1984 and 1996: need for a concerted public health strategy. *Bulletin of the World Health Organization*, 79:942–946, 2001.

[142] I. Paspati. Hip fracture epidemiology in Greece during 1977–1992. *Calcified Tissue International*, 62(6):542–547, 1998.

[143] M. Peacock, G. Liu, M. Carey, W. Ambrosius, C.H. Turner, S. Hui, and C.C. Johnston Jr. Bone mass and structure at the hip in men and women over the age of 60 years. *Osteoporos Internaitonal*, 8(3):231–9, 1998.

[144] C. Bergot, V. Bousson, A. Meunier, M. Laval-Jeantet, and JD Laredo. Hip Fracture Risk and Proximal Femur Geometry from DXA Scans. *Osteoporosis International*, 13(7):542–550, 2002.

[145] KG Faulkner, SR Cummings, D. Black, L. Palermo, CC Gluer, and HK Genant. Simple measurement of femoral geometry predicts hip fracture: the study of osteoporotic fractures. *Journal of Bone Mineral Research*, 8(10):1211–7, 1993.

[146] C.G. Alonso, M.D. Curiel, F.H. Carranza, R.P. Cano, and A.D. Perez. Femoral bone mineral density, neck-shaft angle and mean femoral neck width as predictors of hip fracture in men and women. Multicenter Project for Research in Osteoporosis. *Osteoporos International*, 11(8):714–20, 2000.

[147] XG Cheng, PHF Nicholson, S. Boonen, P. Brys, G. Lowet, J. Nijs, and J. Dequeker. Effects of anteversion on femoral bone mineral density and geometry measured by dual energy X-ray absorptiometry: A cadaver study. *Bone*, 21 (1):113–117, 1997.

[148] J.R. Center, T.V. Nguyen, N.A. Pocock, and J.A. Eisman. Volumetric bone density at the femoral neck as a common measure of hip fracture risk for men and women, 2004.

[149] XG Cheng, G. Lowet, S. Boonen, PHF Nicholson, P. Brys, J. Nijs, and J. Dequeker. Assessment of the strength of proximal femur in vitro: Relationship

to femoral bone mineral density and femoral geometry. *Bone*, 20(3):213–218, 1997.

[150] J.H. Keyak, S.A. Rossi, K.A. Jones, C.M. Les, and H.B. Skinner. Prediction of fracture location in the proximal femur using finite element models. *Medical Engineering and Physics*, 23(9):657–664, 2001.

[151] M. Bessho, I. Ohnishi, H. Okazaki, W. Sato, H. Kominami, S. Matsunaga, and K. Nakamura. Prediction of the strength and fracture location of the femoral neck by CT-based finite-element method: a preliminary study on patients with hip fracture. *Journal of Orthopaedic Science*, 9(6):545–550, 2004.

[152] I.M. Sobol. *A Primer for the Monte Carlo Method*. CRC Press, 1994.

[153] M.A. McDowell, C.D. Frayer, R. Hirsch, and C.L. Ogden. Anthropometric reference data for children and adults: U.s. population, 1999-2002. *US Department of Health and Human Services*, 361, 2005.

[154] M.R. Feldesman and R.L. Fountain. Race specificity and the femur/stature ratio. *American Journal of Physical Anthropology*, 100(2):207–224, 1996.

[155] J.H. Keyak. Improved prediction of proximal femoral fracture load using nonlinear finite element models. *Medical Engineering and Physics*, 23(3):165–173, 2001.

[156] R. Bryan, P.B. Nair, and M. Taylor. Use of a statistical model of the whole femur in a large scale, multi-model study of femoral neck fracture risk. *Journal of Biomechanics*, 42(13):2171–2176, 2009.

[157] D. McMinn and J. Daniel. History and modern concepts in surface replacement. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 220(2):239–251, 2006.

[158] G. X. Ni, W. W. Lu, K. Y. Chiu, and D. Y. Fong. Cemented or uncemented femoral component in primary total hip replacement? a review from a clinical and radiological perspective. *Journal of Orthopeadic Surgery (Hong Kong)*, 13 (1):96–105., 2005. 11.

[159] National Joint Registry for England and Wales. 5th annual report. 2008.

[160] Swedish Hip Arthroplasty Register. Annual report. 2007.

[161] Arthritis Reseach Campaign. Osteoarthritis: An information booklet. 2005. www.arc.org.

[162] A. Shimmin, P.E. Beaule, and P. Campbell. Metal-on-metal hip resurfacing arthroplasty. *The Journal of Bone and Joint Surgery*, 90(3):637, 2008.

[163] E.J. Yue, M.E. Cabanela, G.P. Duffy, M.G. Heckman, and M.I. O'Connor. Hip Resurfacing Arthroplasty: Risk Factors for Failure Over 25 Years. *Clinical Orthopaedics and Related Research*, 467(4):992–999, 2009.

[164] S.T. Ball, M.J. Le Duff, and H.C. Amstutz. Early results of conversion of a failed femoral component in hip resurfacing arthroplasty. *The Journal of Bone and Joint Surgery*, 89(4):735, 2007.

[165] R.T. Steffen, P.R. Foguet, S.J. Krikler, R. Gundle, D.J. Beard, and D.W. Murray. Femoral Neck Fractures After Hip Resurfacing. *The Journal of Arthroplasty*, 24(4):614–619, 2009.

[166] AJ Shimmin and D. Back. Femoral neck fractures following Birmingham hip resurfacing. A national review of 50 cases. *Journal of Bone & Joint Surgery, British Volume*, 87(4):463–464, 2005.

[167] J.P. Long, T.J. Santner, and D.L. Bartel. Hip resurfacing increases bone strains associated with short-term femoral neck fracture. *Journal of Orthopaedic Research*, 2009.

[168] D. McMinn, R. Treacy, K. Lin, and P. Pynsent. Metal on Metal Surface Replacement of the Hip Experience of the McMinn Prosthesis . *Clinical Orthopaedics*, pages S89–S98, 1996.

[169] CP Little, AL Ruiz, IJ Harding, P. McLardy-Smith, R. Gundle, DW Murray, and NA Athanasou. Osteonecrosis in retrieved femoral heads after failed resurfacing arthroplasty of the hip. *Journal of Bone & Joint Surgery, British Volume*, 87(3):320–323, 2005.

[170] RBC Treacy, CW McBryde, and PB Pynsent. Birmingham hip resurfacing arthroplasty a minimum follow-up of five years. *Journal of Bone & Joint Surgery, British Volume*, 87(2):167–170, 2005.

[171] A.S. Dickinson, A.C. Taylor, and M. Browne. Performance of the resurfaced hip Part 1: the influence of the prosthesis size and positioning on the remodelling and fracture of the femoral neck. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, DOI:10.1243/09544119JEIM679, 2009.

[172] H.C. Amstutz, P.A. Campbell, and M.J.L. Duff. Fracture of the neck of the femur after surface arthroplasty of the hip. *The Journal of Bone and Joint Surgery*, 86(9):1874, 2004.

[173] CB Hing, DL Back, M. Bailey, DA Young, RE Dalziel, and AJ Shimmin. The results of primary Birmingham hip resurfacings at a mean of five years: an independent prospective review of the first 230 hips. *Journal of Bone & Joint Surgery, British Volume*, 89(11):1431, 2007.

[174] K.L. Ong, M.T. Manley, and S.M. Kurtz. Have contemporary hip resurfacing designs reached maturity? A review. *The Journal of Bone and Joint Surgery*, 90(Supplement 3):81, 2008.

[175] Y. Watanabe, N. Shiba, S. Matsuo, F. Higuchi, Y. Tagawa, and A. Inoue. Biomechanical study of the resurfacing hip arthroplasty finite element analysis of the femoral component. *The Journal of Arthroplasty*, 15(4):505–511, 2000.

[176] I.A.J. Radcliffe, P. Prescott, H.S. Man, and M. Taylor. Determination of suitable sample sizes for multi-patient based finite element studies. *Medical Engineering and Physics*, 29(10):1065–1072, 2007.

[177] M.T. Mai, T.P. Schalzried, F.J. Dorey, P.A.T.A. Campbell, and H.C. Amstutsz. The Contribution of Frictional Torque to Loosening at the Cement-Bone Interface in Tharies Hip Replacements. *The Journal of Bone and Joint Surgery*, 78 (4), 1996.

[178] CB Hing, DA Young, RE Dalziel, M. Bailey, DL Back, and AJ Shimmin. Narrowing of the neck in resurfacing arthroplasty of the hip: a radiological study. *Journal of Bone & Joint Surgery, British Volume*, 89(8):1019, 2007.

[179] L. Cristofolini, M. Juszczyk, F. Taddei, RE Field, N. Rushton, and M. Viceconti. Stress shielding and stress concentration of contemporary epiphyseal hip prostheses. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 223(1):27–44, 2009.

[180] H.C. Amstutz, P. Grigoris, and F.J. Dorey. Evolution and future of surface replacement of the hip. *Journal of orthopaedic science*, 3(3):169–186, 1998.