A novel framework for assessing variations in 3D geometry of mouse tibiae in longitudinal preclinical studies

Stamatina Moraiti

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

INSIGNEO institute for in silico Medicine
The University of Sheffield
Faculty of Engineering
Department/School of Mechanical Engineering

August 2024

ABSTRACT

The mouse tibia is commonly used for testing osteoporosis treatments preclinically. *In vivo* micro-Computed Tomography (microCT) is the standard for monitoring longitudinal bone changes. Standard morphometry applied on images provides temporal changes in scalar geometric properties. However, these properties lack precision due to their limited dimensionality and inability to describe non-uniform 3D geometric changes. This PhD project aimed to develop a novel statistical framework enabling comprehensive assessment of spatio-temporal bone changes. The proposed framework combined longitudinal microCT imaging, image processing, Principal Component Analysis (PCA) and post-processing statistical analysis. First, the accurate, robust PCA-based model was developed, with less than one voxel error in describing bone shapes. Initially, the framework was applied to longitudinal microCT images of the tibial midshaft in an osteoporotic mouse model, elucidating 3D changes induced by *in vivo* mechanical loading (ML) including bone turnover at the anterior crest (0.103 mm) and posterior-lateral compartment. Next, it was applied to the entire cortex of osteoporotic mice, which received ML, Parathyroid hormone (PTH), and a combination of both (PTHML). The model revealed high bone turnover proximally, with posterior thickening, medial expansion and anterior thickening, highlighting increased anabolic effects of the PTHML. The first modes detected axial shape variations, coupled with positional misalignments. The latter was up to 0.142 mm in the axial position between the two ages in the diseased group. Motivated by this, the next study demonstrated non-negligible instance alignment effects in the shape analysis, suggesting a registration protocol that allowed meaningful biological inferences. Finally, a novel Partial Least Squares regression model revealed that bone length increase, posterior thickening and area expansion of the diaphysis described increased bone strength due to combined PTH and ML (R^2=81%). This novel statistical framework offers a precise strategy for assessing treatment efficacy by measuring 3D bone geometry and strength changes in preclinical studies.

ACKNOWLEDGMENTS

It is my honour to dedicate these acknowledgements to all the people who supported me through my PhD journey. Each of them has contributed, in their way, to who I was, who I am and who I will be.

Firstly, I would like to express my gratitude and thank my supervisory team: Dr. Pinaki Bhattacharya, Prof. Enrico Dall'Ara, and Prof. Visakan Kadirkamanathan. I thank Pinaki for his continued professional and personal support. With his positivity, smile, encouragement and guidance, he has fundamentally helped me become a confident professional scientist. I also thank Enrico for his scientific direction, which has contributed to my development. He is one of the most brilliant and passionate researchers I have worked with, and his *Meraki* has inspired me a lot. Finally, I thank Visakan for his support, and our fruitful and insightful conversations.

I also kindly acknowledge my funding sources: the Engineering and Physical Sciences Research Council (EPSRC: EP/K03877X/1, EP/S032940/1 and EP/V050346/1) and the National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs: NC/R001073/1).

I have been very fortunate to be a member of the Integrative MusculoSkeletal Biomechanics (IMSB) group at the INSIGNEO Institute and interact with excellent and wonderful researchers. Many thanks to our amazing C+14 office for the coffee breaks in the kitchen, the chats and the laughter. Special thanks to Saira, Vee, Saleh, Frederik, Lin, Fiona and Jack.

I also thank my best friends Eirini, Nasia, Ioanna and Denie−four amazing women who have inspired and supported me with their 14-year-long friendship. Additionally, I thank my friends Germayne, Mahnoor, Julian, Stelios and Persa who opened their arms and made Sheffield a new home for me.

I want to express my deep appreciation to my sister, Effie, my mother, Ioanna, and my father, George. Without you, I would not be who I am today. Thank you from the bottom of my heart! You have been my starting point and my home. Being away from you has been difficult sometimes.

Finally, I would like to express my gratitude to my partner, Robert, who has been tirelessly and constantly beside me. With your unconventional, artistic eyes, love for nature and incredible sense of humour, you have shown me new ways of feeling, thinking, and living.

PUBLICATIONS & CONTRIBUTIONS

Journal entries:

Under review

1) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kadirkamanathan & P. Bhattacharya, 2024, A novel framework for elucidating the effect of mechanical loading on ovariectomized mouse tibia geometry using Principal Component Analysis.

In preparation

1) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kadirkamanathan & P. Bhattacharya. Tentative title: "Image rigid registration influences the assessment of the 3D shape variation in mouse tibia bones".

2) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kadirkamanathan & P. Bhattacharya. Tentative title: "Assessment of treatment effects on osteoporotic mouse tibia reveals novel geometric features with important influence in bone strength".

Conference abstracts:

Oral presentations

1) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kadirkamanathan & P. Bhattacharya. "Principal Component Analysis for elucidating important changes in mouse tibia geometry", ESB 2022, Porto, Portugal.

2) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kadirkamanathan & P. Bhattacharya. "Principal Component Analysis for elucidating important treatment effects on ovariectomised mice", BioMedEng 2022, London, United Kingdom.

3) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kadirkamanathan & P. Bhattacharya. "3D statistical shape modeling for classification of treatment effects on osteoporotic mouse bone geometry", ESB 2023, Maastricht, the Netherlands.

4) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kardikamanathan & P. Bhattacharya. "A new stochastic approach for assessment of treatment effects on osteoporotic mouse bone geometry", BioMedEng 2023, Swansea, Whales, United Kingdom.

5) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kardikamanathan & P. Bhattacharya. "Rigid registration protocol influences the assessment of 3d geometric variations in the mouse tibia", ESB 2024,

Edinburgh, Scotland, United Kingdom.

Poster presentations

1) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kardikamanathan & P. Bhattacharya. "Principal Component Analysis for elucidating important treatment effects on ovariectomised mice", INSIGNEO showcase 2022, University of Sheffield, United Kingdom.

2) **S. Moraiti**, E. Dall'Ara, V. Kardikamanathan & P. Bhattacharya. "Task-parallelised elastic registration of 3-dimensional mouse tibia images", ISC High Performance 2023, Hamburg, Germany.

3) **S. Moraiti**, V.S. Cheong, E. Dall'Ara, V. Kardikamanathan & P. Bhattacharya. "Three-dimensional changes in tibia cortical geometry that explain variance in strength in a murine osteoporotic bone model", BRS 2024, Sheffield, United Kingdom.

Awards:

1) Women in HPC bursary up to £3000 to attend ISC High Performance2023, Hamburg

2) ESB travel award 2023, Maastricht

3) GTA award 2024 for "Proactivity in teaching", Multidisciplinary Engineering Education, Faculty of Engineering, University of Sheffield.

Other career activities:

1) Supervision of two INSIGNEO summer research placements- Posters published on INSIGNEO website: i) "The effect of spatial resolution on accuracy and precision errors: A comparative study using elastic image registration", 2022.

ii) "Investigating three-dimensional localised effects of age, disease and treatment on mouse bone geometry using Principal Component Analysis", 2023.

2) Co-supervision of master's dissertation: "Finite-element analysis of the impact of material heterogeneity on the mechanical competence of the mandible subject to masticatory forces".

3) Graduate Teaching Assistant in the Structures and Mechanics Lab.

4) Marking Assistant in the "Introduction to Medical Device Regulation" BIE6432 module.

5) Applied for Fellow Higher Education Academy certificate, recognising teaching experience (under review).

DECLARATION

I, Stamatina Moraiti, confirm that the Thesis is my own work. I am aware of the University's Guidance on the Use of Unfair Means (www.sheffield.ac.uk/ssid/unfair-means). This work has not been previously been presented for an award at this, or any other, university.

TABLE OF CONTENTS

NOMENCLATURE

MicroCT	Micro Computed Tomography
ShIRT	Sheffield Image Registration Toolkit
PCA	Principal Component Analysis
PLS	Partial Least Squares
NS	Nodal Spacing
OP	Osteoporosis
NHS	National Health Service
BMD	Bone Mineral Density
OVX	Ovariectomy
PTH	Parathyroid Hormone
ML	*In vivo* mechanical loading
PTHML	Combined treatment of Parathyroid Hormone and *in vivo* mechanical loading

Chapter 1

Introduction and thesis overview

1.1 Motivation

In our ageing societies, a high percentage of the population suffers from osteoporosis (OP). This disease affects over 200 million people worldwide, resulting in patient suffering, and deterioration of life quality, and leading to 37 million fragility fractures annually across the globe (Wu et al., 2021). Statistically, 1 to 3 women and 1 to 5 men over the age of 50 will sustain osteoporotic fractures (Kanis et al., 2000; Melton III et al., 1998; Melton III, 2020). In the UK, the patients number over 3 million and the National Health Service (NHS) spends approximately £4 billion to treat osteoporotic fractures (NICE, 2018). The recovery after such fractures has proven to be highly challenging. Even when following the rehabilitation protocols, patients struggle to regain their independence and quality of life.

Given the substantial health implications arising from osteoporosis and the subsequent socioeconomic burden, professionals in the domain of healthcare and medicine aim to develop and apply efficient strategies for prevention, diagnosis and treatment (Khosla & Hofbauer, 2017). Various prevention strategies include general suggestions about balanced diet, healthy lifestyle and exercise, as well as regular checkups and bone density measures to monitor bone quality in advance ages.

As osteoporosis is associated to aging, more targeted strategies are needed for efficient disease management. Anabolic drugs can provide therapeutic effects on osteoporotic bones, while exercise can regularly decline the rate of bone loss. Such strategies aim at stimulating molecular mechanisms for bone formation, adapting the bone structure, enhancing the bone quality and subsequently increasing the bone strength.

The drug development process, which is expensive and time-consuming, involves various stages from preclinical research to sequential clinical trials and approvals (Nuffield Council on Bioethics, 2005). During this extensive process, assessing the effect size of a bone-disease drug on bone geometry is crucial for determining the treatment's efficacy. Therefore, the primary objective when evaluating the drug’s efficacy is to precisely assess the changes in bone geometry over time, while considering the natural variability within the studied population.

In preclinical research, murine bone models are often used to test novel osteoporosis treatments. The choice of these species is justifiable as humans and mice have been shown to share the same bone loss pathways (Jilka, 2013). Several bone sites, including long bones, such as femur or tibia, and vertebra, are tested. Tibia is highly preferable due to its accessibility and clinical relevance as a weight-bearing bone. Research has developed extensive experimental models to quantify the disease and treatment effects on bone structure and mechanics. Such models are created by subjecting the mice to several interventions to induce osteoporosis phenotypes and simulate treatments.

In the last decades, with the evolution of imaging protocols and image processing techniques, *in vivo* micro Computed Tomography (microCT) has enabled the conduction of longitudinal image snapshots of bones of different ages. This approach monitors the bone geometry alterations over time in a low-invasive experimental manner while providing a 3D digitisation of the bone samples. The images are used to extract the geometry information, characterise the bone turnover and measure the subsequent changes in bone strength (Dall'Ara et al., 2016). However, the characterization of the bone geometry in 4D (space and time) is still challenging due to its high dimensionality. Current state-of-the-art geometric assessment is limited to using scalar quantifiers that have several shortcomings: they are averaged and scalar and fail to capture localised important bone structural changes; they are correlated with each other describing interdependency in changes; they do not comprehensively describe the shape variations in the entire bone volume. Hence, more sophisticated statistical shape models that boost the precision and the efficiency on the assessment of the bone remodelling in OP murine bone models are yet to be developed.

With the emerging development of modelling in Medicine (*in silico*), research endeavours aim to reduce the use of animals in bone research. However, this aim comes with challenges such as developing a model that mimics the behaviour of nature reliably enough to describe the variations of natural structures. Given the variability in bone structure and mechanics, the assessment or the prediction of bone changes usually necessitates the development of numerous individual-specific models to quantify accurate effect sizes. This, in turn, leads to increased use of resources, and utilisation of more

experiments, since the current individual-specific bone models serve as point-based estimators or predictors of bone adaptation. The present study works in the direction of the 3Rs (Replacement, Refinement and Reduction of Animals in Research) (Viceconti & Dall'Ara, 2019) by proposing a statistical computational model that provides a precise and compact assessment of 3D bone adaptation patterns to interventions in an examined population of mouse tibiae. The ultimate motivation is to contribute to the ongoing data sharing and collaboration in the scientific community for developing and improving *in silico* Medicine.

1.2 Aim and objectives

The current PhD project aims to propose a novel data-based statistical shape model for treatment investigation in preclinical studies. This model:

provides a complete and localized assessment of the systematic temporal variations of the mouse tibia shape induced by an intervention

The aim has been accomplished by reaching the following objectives:

1. Develop an accurate and robust statistical framework for elucidating treatment and osteoporosis effects in a section of murine tibia bones
2. Extend the framework to investigate bone changes of the entire cortex induced by various treatment strategies
3. Evaluate how bone positioning and alignment techniques affect the assessment of shape variations of the entire tibia cortex
4. Identify the geometric variations due to osteoporosis/treatment that explain changes in strength

1.3 Outline

To approach the objectives given above, this thesis has been divided into six chapters. The description of the content of each chapter is described below and the distribution of the objectives within the thesis is shown in the schematic illustration in Figure 1.1.

Chapter 1 provides the research topic, the motivation, the aim and objectives and finally gives the content of this thesis.

Chapter 2 details the scientific background about bone tissue, and the models used to evaluate bone turnover and adaptation to external intervention in preclinical studies. More details are given about mouse models, experimental protocols, imaging, image processing tools, image registration techniques and morphometric analysis. It also provides a dedicated description of previous related longitudinal mouse bone studies. It finally introduces the concept of natural variability in bone structures and random variables and describes useful methodologies that this project applied. A specific literature review related to the objectives of each chapter while highlighting the research gaps and the scope of each study is given separately in the introduction sections of each following chapter.

Chapter 3 presents the proposed statistical framework to elucidate geometric changes in tibia midshaft in osteoporotic mice during a course of treatment. It also details the robustness analysis to identify and measure errors and uncertainties in the framework. This chapter demonstrates a feasibility study associated with Objectives 1 and it stands as a proof of concept.

Chapter 4 presents the application of the proposed framework in the entire tibia cortex considering more treatment interventions and more time points. It highlights the method refinements applied to the framework in order to enable the comparison between different treatment strategies and identify the different effects in the bone geometry of the entire tibia cortex. This chapter is related to Objectives 2.

Chapter 5 is dedicated to one major challenge of statistical shape modelling related to rigid registration errors and bias to reference structures. Particularly, it demonstrates the comparison between different shape instance alignment protocols to assess their effects on the reliability of the shape model. This chapter is associated with Objective 1 and Objective 3 suggesting better performance of the statistical model when the data are processed using one specific registration protocol.

Chapter 6 provides further statistical model development. It suggests a statistical linear regression model to quantify treatment effects on the tibia cortex of osteoporotic mice that explain changes in bone strength (Objective 4).

Chapter 7 concludes the current work, highlighting the contributions to related research and proposing future work. Future work suggestions specific to the individual studies are also detailed in each study separately.

The Chapters 3, 4, 5 and 6 follow a typical paper format. Chapters 1 (Introduction), Chapter 2 (Scientific background) and Chapter 7 (Conclusions of the thesis) are written in a narrative structure and serve as connecting pieces that integrate the separate studies presented in the other chapters.

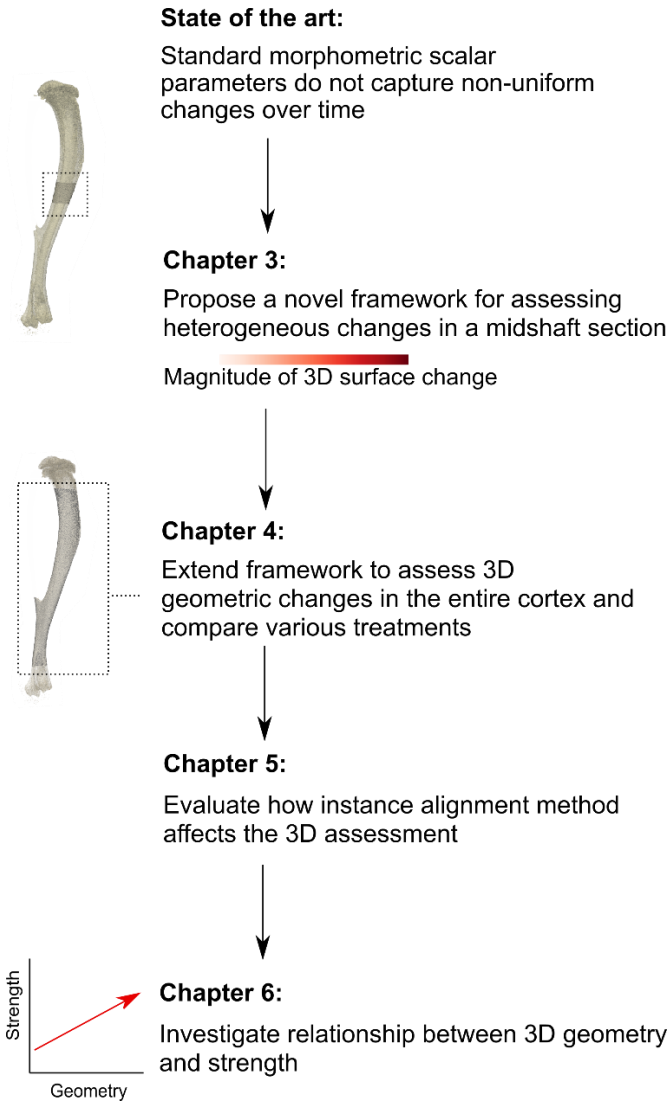


Figure 1.1: The story of this thesis.

1.4 References

Dall'Ara, E, Boudiffa, M., Taylor, C., Schug, D., Fiegle, E., Kennerley, A. J., Damianou, C., Tozer, G. M., Kiessling, F., & Müller, R. (2016). Longitudinal imaging of the ageing mouse. *Mechanisms of Ageing and Development*, *160*, 93–116. https://doi.org/10.1016/j.mad.2016.08.001

Jilka, R. L. (2013). The Relevance of Mouse Models for Investigating Age-Related Bone Loss in Humans. *The Journals of Gerontology: Series A*, 68(10), 1209–1217. https://doi.org/10.1093/gerona/glt046

Kanis, J. A., Johnell, O., Oden, A., Sernbo, I., Redlund-Johnell, I., Dawson, A., De Laet, C., & Jonsson, B. (2000). Long-term risk of osteoporotic fracture in Malmö. *Osteoporosis International*, 11, 669–674. https://doi.org/10.1007/s001980070064

Khosla, S., & Hofbauer, L. C. (2017). Osteoporosis treatment: recent developments and ongoing challenges. *The Lancet Diabetes & Endocrinology*, *5*(11), 898–907. https://doi.org/10.1016/S2213-8587(17)30188-2

Melton III, J. L. (2020). Perspectives: how many women have osteoporosis now? *Journal of Bone and Mineral Research*, *10*(2), 175–177. https://doi.org/10.1002/jbmr.5650100202

Melton III, L. J., Atkinson, E. J., O'Connor, M. K., O'Fallon, W. M., & Riggs, B. L. (1998). Bone density and fracture risk in men. *Journal of Bone and Mineral Research*, *13*(12), 1915–1923.

National Institute for Health and Care Excellence, NICE Impact Reports: Falls and fragility fractures UK; 2018

on Bioethics, N. C. (2005). The ethics of research involving animals. Nuffield Council of Bioethics.

Wu, A.-M., Bisignano, C., James, S. L., Abady, G. G., Abedi, A., Abu-Gharbieh, E., Alhassan, R. K., Alipour, V., Arabloo, J., Asaad, M., & others. (2021). Global, regional, and national burden of bone fractures in 204 countries and territories, 1990--2019: a systematic analysis from the Global Burden of Disease Study 2019. *The Lancet Healthy Longevity*, 2(9), e580--e592. https://doi.org/10.1016/S2666-7568(21)00172-0

Viceconti, M., & Dall'Ara, E. (2019). From bed to bench: How in silico medicine can help ageing research. *Mechanisms of Ageing and Development*, *177*, 103–108. https://doi.org/10.1016/j.mad.2018.07.001

Chapter 2

Scientific Background

2.1 Bone and osteoporosis

Bone is the building stone of the skeleton in vertebrate animals, which serves an assortment of duties; it protects soft tissues and organs, supports the body, provides movement, and is also mineral storage, vital for a wide variety of metabolic functions (Florencio-Silva et al., 2015).

Regarding the shape, it can be classified as a long or flat bone. Long bones are the femur, tibia and fibula, humerus, and forearm bones. In the context of its structure, it is categorised at the macroscopic level as cortical (compact) bone and trabecular (cancellous or spongy) bone. Long bones exhibit several macroscopic features. They consist of three regions: diaphysis, epiphysis, and metaphysis (Figure 2.1). The diaphysis is the central section (shaft) of the bone, and it is essentially composed of compact bone. The epiphysis represents the top and bottom regions and exhibits an inner trabecular structure. The metaphysis is the transitional region between the diaphysis and the epiphysis that transfers the external loads from the joints to the dense and hard tissue of the diaphysis.

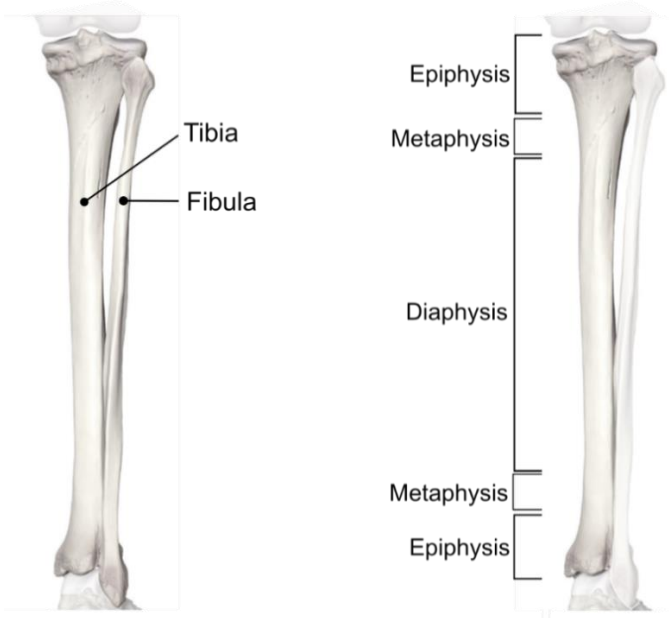


Figure 2.1: Tibia and fibula in humans. Epiphysis, metaphysis and diaphysis of the tibia structure are indicated as regions. Image reproduced from Tibia, or shinbone (anatomystandard.com) and further edited by the author of the current thesis.

In humans, the tibia and fibula are two separate long bones of the lower leg (Figure 1). It is the largest bone in the lower leg, and it is considered a weight-bearing bone. Its proximal end articulates with the femur and its distal one with the talus. It is hollow in the centre, forming two free-bone surfaces. The

interior surface is called the endosteum, and the exterior is called the periosteum. The trabecular region fills the ends. The cortical bone forms a protective shell externally of the entire structure. More details about the anatomy of the bone and the relevance of mouse bones to humans are given in Section 2.2 and 2.3, respectively.

Osteoporosis (OP) is a progressive bone disease that deteriorates bone quality by declining bone mass and impairing bone architecture (Rodan & Martin, 2000) (Figure 2.2). It is associated with the imbalance of bone formation and resorption, affecting the ageing population and increasing the susceptibility to fragility fractures. Particularly in women, bone mineral density (BMD) in bone reduces gradually with the postmenopausal drop in oestrogen levels (Russo et al., 2003). Fractures most frequently occur in the hip, wrist or spine but other sites are also affected by the disease (Larsson, 2002) (Figure 2).

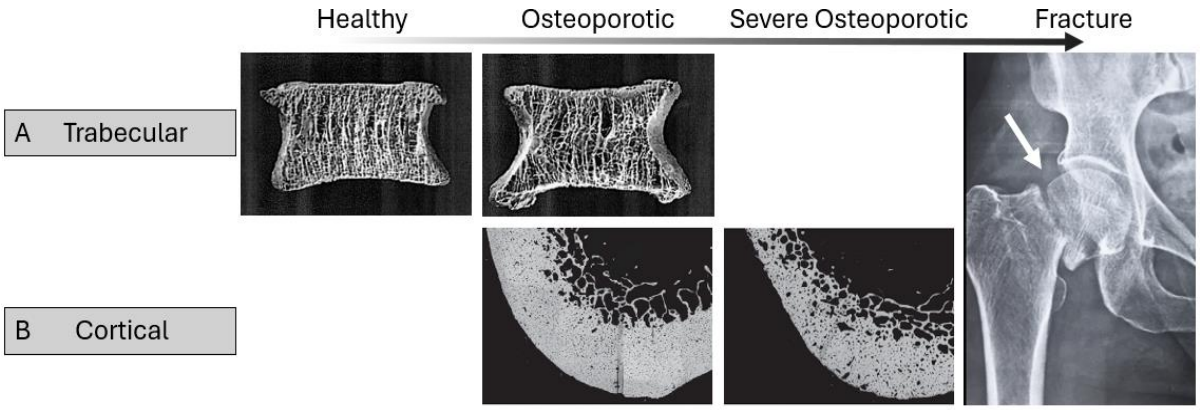


Figure 2.2: Osteoporosis effects at different levels of severity in flat and long bones with trabecular (A) and cortical (B) bone regions. One common type of osteoporotic fracture occurs at the femur neck. Figures extracted from Ferguson & Steffen, (2005), Neck of femur fracture x ray - wikidoc, and Zebaze et al., (2010) (License No. 5840721490320).

The clinical significance of osteoporosis arises from bone fractures. The bone fracture occurs when the bone cannot withstand the applied external loads. Because osteoporosis is asymptomatic, bone fractures may happen without a diagnosis of osteoporosis. Thus, most clinical cases cannot be addressed on time. As well as health issues, it leads to a financial burden on the NHS. Specifically, the estimated annual cost exceeds £4.4 billion, while it is expected that the osteoporotic fractures will increase significantly in the next decades (Compston et al., 2017).

In the context of OP treatments, there is a range of pharmacological strategies that aim at the treatment of osteoporosis. Recombinant parathyroid hormone [1–34] (PTH (1–34)) as a basis of some anabolic drugs, such as Teriparatide and Abaloparatide, establishes widened therapeutic effects in diseased bones. It reduces the risk of fracture by improving the quantity of the BMD and increasing bone formation (Hodsman et al., 2005). However, its long-term efficacy is questionable, with PTH provoking resorptive effects after 2 years of use (Aslan et al., 2012). Other side effects of such drugs include atypical bone fractures related to hypercalcemia and bone cancer such as osteosarcoma (Reeve, 1996).

2.2 Bone structure and bone remodelling

Bone is a hierarchically organised structure (Figure 2.3). This study works on the organ level of the bone; however, a brief overview of all scales is given below.

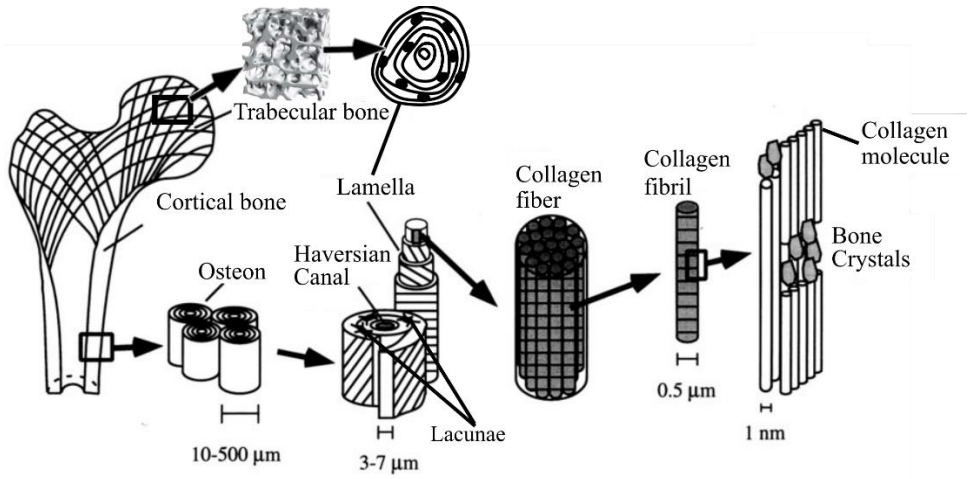


Figure 2.3: Hierarchical bone structure in large mammals. Figure of the structure extracted from Rho et al., (1998)(License Number: 5913650749681) and figure of the trabecular bone extracted from Wirth et al., (2010). Further edited by the author.

In the nanoscale, bone is a multi-phase composite material comprising collagen fibrils (organic), bone mineral (inorganic), and water (Olszta et al., 2007). At the sub-microscale, mineralised collagen fibrils are organised in bundles or arrays, forming structures such as parallel lamellae (Weiner & Wagner, 1998), which are mechanically anisotropic, or woven fibres, which are more flexible but weaker. Shapiro and Wu (2019) illustrated these structures in normal mouse neonatal femoral cortical bone.

Moving to the next hierarchic level (microscale), compact bone is primarily made up of osteons, or Haversian systems, which are prominent in large mammals but less so in smaller animals. A deficiency

of an osteon-composed Haversian system in geckos, Nile monitors, sparrows, ducks and geese femora has been shown by Ahmed et al., (2017), while the rodent bone is also not organised to osteons. The absence of further spatial arrangements in mouse and rat bones is probably due to the small-scale separation of the bone cells and small bone macroscale structure (Bianco et al., 2000). Thus, the main characteristic of the mouse's long bone is the cortex which is structured as one large osteon.

In the mesoscale, two different bone types can be detected: the cortical and the trabecular bone. The cortical bone is a compact and dense structure with low porosity. It mainly forms the external shell of all bones, providing support. In long bones such as tibia, it is deposited primarily in the diaphysis, occupying 80% of the total body bone mass. The cortical bone is the main structure which mainly supports body weight and sustains external forces. In contrast, the trabecular bone exhibits high porosity of 50%-90%, high heterogeneity and an increased turnover rate. It is allocated inside the cortical shell at the ends of the long bones. It accounts for approximately 20% of long bones. It is composed of irregular thin rods known as trabeculae.

At the compositional level, the building stone of cortical and trabecular bone is primarily the same, but some fundamental differences exist in micro and macrostructural levels. Goodyear et al., (2009) demonstrated that the mouse cortical bone is preloaded and more mineralised than the trabecular bone ('older bone tissue'). Specifically, although there are locations in the cortex of the mouse long bone that are constantly remodelling during a lifetime, others never turnover after bone maturity (Bianco et al., 2000).

Bone is a dynamic tissue, which constantly turns over. Let us then introduce the existence of the "builders" and "breakers" cells, that are responsible for bone remodelling. Although this project does not focus on the cell level, the demonstration of the function of the bone cells contributes to describing the bone remodelling at higher bone levels. More information could also be found in Currey, (2002).

Osteoblasts are the "builders" bone cells which create new bone tissue. The osteoblastic function is fundamental for bone formation in postnatal growth and bone remodelling to repair damage. On the

other hand, the osteoclasts are “breakers” cells and are the largest in size. They are extremely aggressive, and they “demolish” the damaged bone tissue. Conversely to osteoblasts, the osteoclastic function is responsible for bone absorption.

From childhood to adulthood and ageing, the interaction of these cells determines both apposition and resorption (Currey, 2002) and affects the bone mass as shown in Figure 2.4. Its ability to sustain the applied loads during an individual’s everyday life is determined by bone remodelling, which is a mechanically stimulated process (Lanyon et al., 1982). Capable of adapting its tissue, bone can sense external mechanical stimuli and repair itself where bone microdamage is detected. Osteoblasts and osteoclasts work synergically in the same bone remodelling unit to replace the damaged bone with new and healthy bone tissue. In adulthood, bone remodelling is balanced, preserving the high peak bone mass stable for some years around the age of 25-35 years. During the lifespan, this equilibrium constantly changes. Following adulthood, the number of osteoblasts decreases and the coupling between apposition and resorption does not occur, tipping the scales in favour of the osteoclasts. Thus, progressive age-related bone loss deteriorates the bone quality and composition and leads to the osteoporotic cohort (Demontiero et al., 2012).

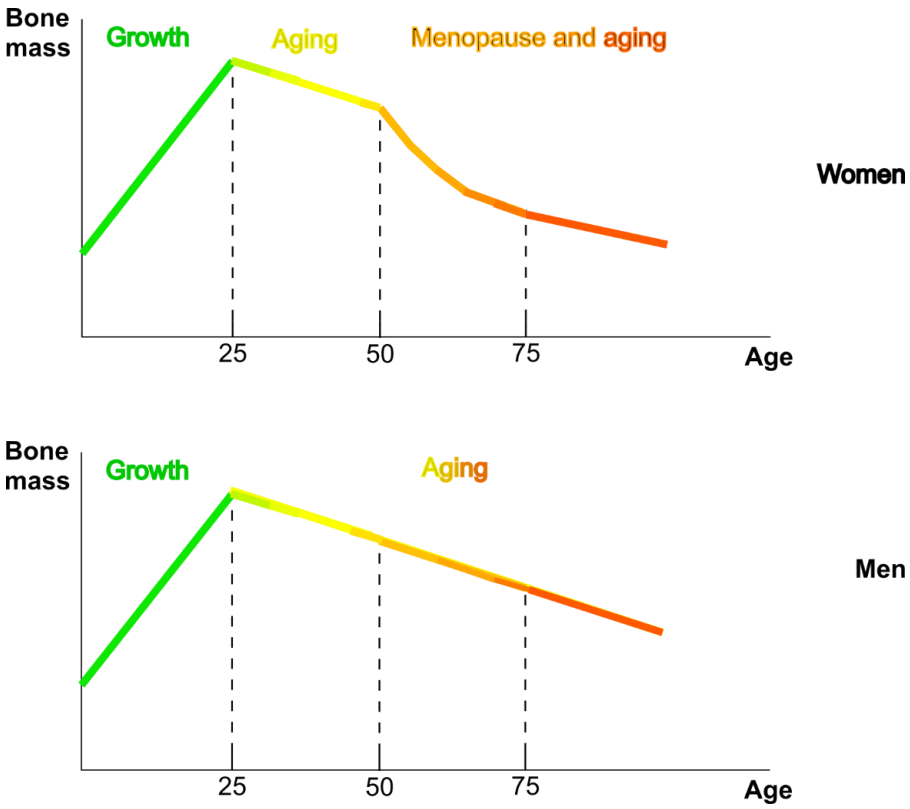


Figure 2.4: Changes in bone mass over age for both women and men.

2.3 Murine bone models

In preclinical studies, mice are primarily used as bone models to investigate osteoporosis progression and treatment effects on bone geometry (Sommer et al., 2019). The researchers' preference for these models is mostly based on their relevance to bone loss in human bones (Jilka, 2013). Like humans, mice experience bone loss in advanced age, targeted on cancellous and cortical bone, leading to thinning of bone regions and increasing bone porosity. In both mammals, loss of bone strength is associated with loss of BMD. Additionally, the loss of BMD in mice starts around 16-25 months of their age, which is a similar mechanism as observed in humans at the age of 35 years old.

The utilisation of mouse bone models also comes with some great advantages associated with the study design. First and foremost, murine life expectancy is limited. As small mammals, the mouse metabolism is very active, and the mouse is highly adaptable to external interventions. This attribute enables the researchers to observe the effects on the bones in the short term. Particularly, it has been reported that the remodelling cycle typically lasts 2 weeks (Weinstein et al., 1998). Therefore, the first evidence of an intervention can be depicted on murine bones after 2 weeks. Additionally, due to their small size, the

control of the experimental environment is easier and the need for space and resources is less. Moreover, their cost is low. Hence, mice are justifiably considered one of the most favourably used mammalian species for investigating bone loss or temporal changes due to pathology and treatment effects.

Across all bone sites, flat bones e.g. vertebrae or long bones such as tibiae and femurs are the most widely used bones for this class of murine studies due to their clinical relevance. The primary focus of examining vertebra is the microarchitecture of trabeculae and their local apposition and resorption (Brandi, 2009). Research on vertebrae provides insights into how interventions may impact overall skeletal function.

On the other hand, other research studies the mechanical importance of weight-bearing bones such as leg bones and understands the impact of the disease and treatment on bone structure, material and subsequently mechanics. Therein, the choice of tibia or femur is justified, but the specific choice of the one over the other depends on several factors. For instance, accessibility to tibia for *in vivo* mechanical loading and *in vivo* scanning enables an ease of generating longitudinal data. Since tibia holds these advantages, it has become a standardised model, making it easier to keep consistency in the experimental protocols and compare findings across labs and research groups.

Unlike the separate human tibia and fibula bones, the mouse tibia is fused with the fibula toward the distal metaphysis. This is called the tibia-fibula complex. The tibia length is approximately 15-20 mm and 80% of it is cortical bone. The anatomical coordinate system for a mouse tibia is illustrated in Figure 2.5. The anatomical features are demonstrated in Figure 2.6. The distal and proximal ends are filled with trabeculae inside and with a very thin protecting cortical shell from outside (proximal image slice in Figure 2.5). The growth plate is located in the proximal epiphysis (Figure 2.6, proximal zoom-in). Proximally to the midshaft, the tibia exhibits a curvature and a crest at the anterior aspect, called the tibial ridge. The midshaft is primarily characterised by a rounded triangular hollow structure of a diameter around 800-1200 μm that gradually approaches a cylinder shape as moving distally (midshaft image slice in Figure 2.5). The interior surface is called the endosteum, and the exterior is called the periosteum. Distal to the midshaft, the tibia and fibula are fused for approximately 20% of the tibia

length. At the distal end, the two bones are separated again, as highlighted in the distal image slice in Figure 2.5 and the 3D shape in Figure 2.6.

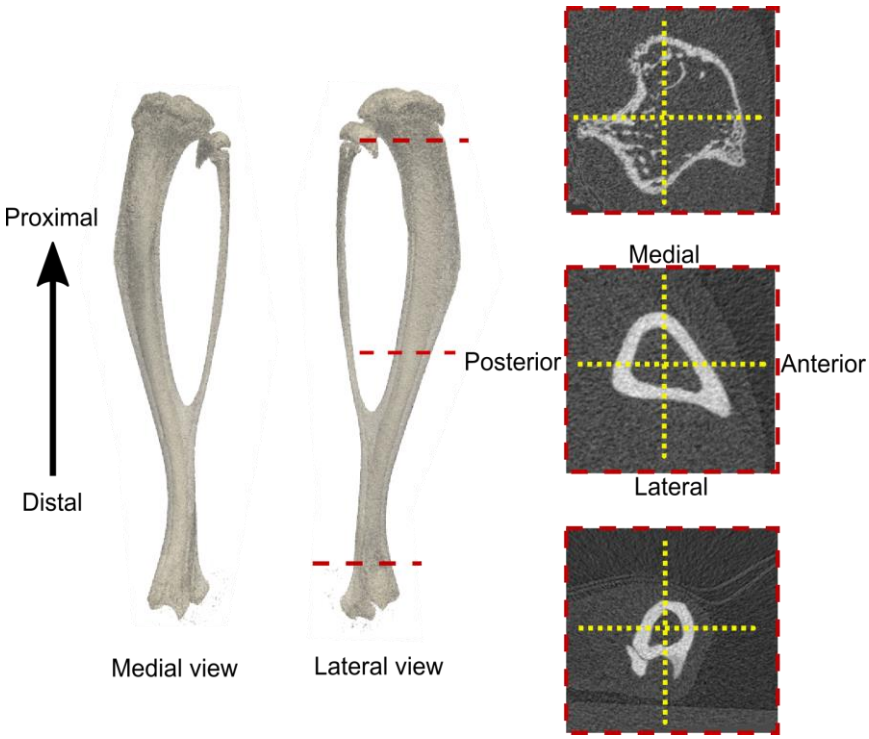


Figure 2.5: Anatomical orientation of the mouse tibia.

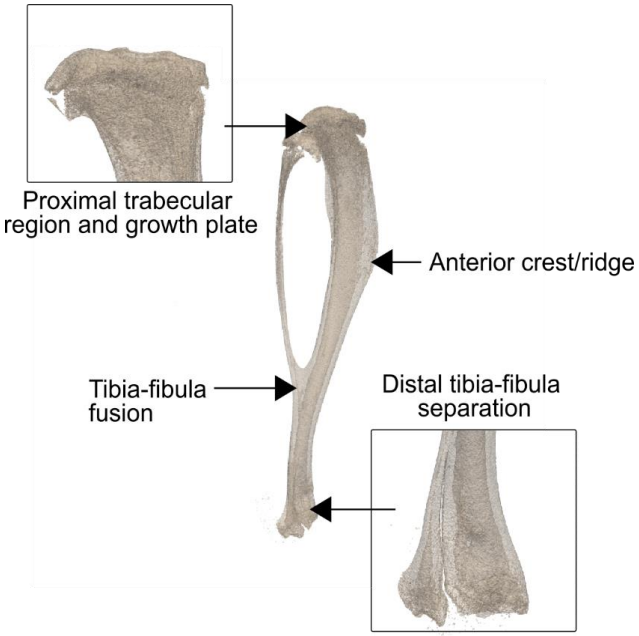


Figure 2.6: Anatomical features of mouse tibia (lateral to medial view)

2.4 Treatment administrations in preclinical studies

Simulating osteoporosis treatment strategies in mice involves several steps combining the induction of the osteoporosis phenotype and administration of the different treatment strategies. The next paragraphs detail the design and conduction of the appropriate experimental procedures.

Ovariectomy:

To simulate post-menopausal osteoporosis, female mice typically undergo ovariectomy (OVX) to surgically remove the ovaries. This intervention initiates the osteoporosis phenotype, by provoking oestrogen deficiency. As the remodelling cycle lasts around 2 weeks, the first osteoporotic evidence on bones can be observed in the short term. As a result, 2 weeks after the OVX surgery, osteoporosis seems to have a catabolic effect on growth progression leading to lower changes in tibia length and other cross-sectional changes (Roberts et al., 2019).

Injection of Parathyroid Hormone PTH (1-34):

Injecting Parathyroid Hormone (PTH (1-34)) in mice is a common method used in preclinical studies to investigate the bone response to osteoporosis treatments. PTH (1-34) simulates drugs based on synthetic forms of PTH which are prescribed to postmenopausal women and men with high fracture risk due to osteoporosis.

***In vivo* passive mechanical loading:**

Mechanical loading of long bones simulates the exercise which constitutes another treatment strategy with anabolic effects on bones. In preclinical studies, murine bone models are uniaxially compressed to simulate the loading conditions in human long bones. Passive compression is currently the most common configuration, since *in vivo* low-invasive protocols have been developed and it is also clinically relevant to physical exercise. It is performed by applying loading between the proximal and the distal end. The choice of the right loading protocol is crucial for simulating physiological loading and provoking anabolic effects to both cortical and trabecular bone without micro-damages (De Souza et al., 2005). Figure 2.7 is a schematic representation of *in vivo* mechanical loading in mouse tibia.

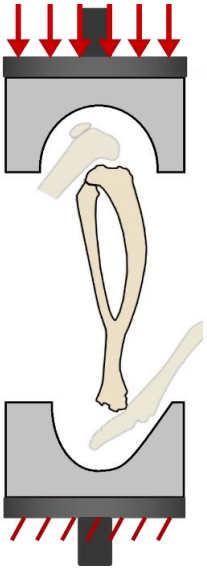


Figure 2.7: Schematic of in vivo mechanical loading of mouse tibia. Figure extracted from Cheong et al., (2020) and further edited by the author.

2.5 Bone imaging and processing techniques in mouse models

With the advancement of imaging technology in previous decades, more detailed and informative data have been used to feed bone research in understanding ageing, disease progression and treatment effects. Micro-Computed Tomography was first introduced in the late 1980s. Now it is considered the gold standard to assess bone structures. Imaging could be *ex vivo* or *in vivo*. *In vivo* imaging is commonly preferred because of two major advantages in preclinical studies about treatment investigation: dynamic assessment of disease progression or/and treatment response in individuals, and limited invasive approach. On the other hand, it may have a lower resolution (9-20 micrometres) than the *ex vivo* image acquisition.

This technology generates cross-sectional images of the 3D structure. With the specimen placed in the imaging equipment as shown in Figure 2.8, the 3D image set is a set of 2D images aligned with the length axis of the sample, with each image slice capturing the cross-section of the sample at a specific location of the length axis. The left top corner of the 3D is called the origin and the number of the pixels in each dimension defines the image dimensionality. Although the resolution of the images varies among studies, high-resolution *in vivo* imaging of 9-19μm voxel ("3D voxels") size has been used for assessing trabecular and cortical architecture in rodent bones (Akhter & Recker, 2021). Each voxel has a grey level which is related to the local mineral density of the bone. Hard and soft tissue are characterised by

different grey level ranges. The main components of the microCT are the X-ray source, the filter, the sample holder, and the 2D charge-coupled device (CCD) array (Figure 2.9).

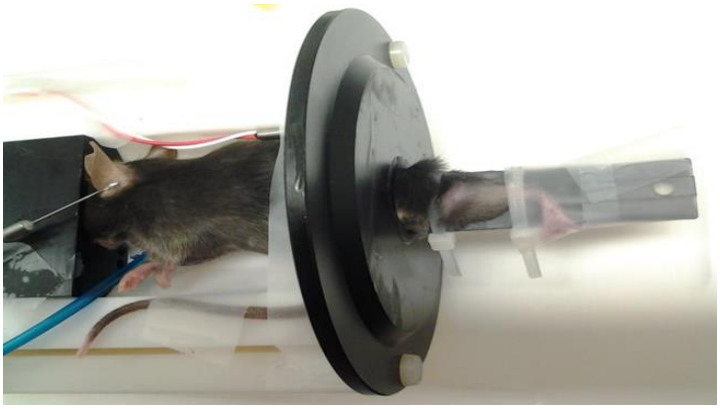


Figure 2.8: In vivo micro–Computed Tomography of mouse tibia. Kindly provided by Prof. Enrico Dall'Ara's team.

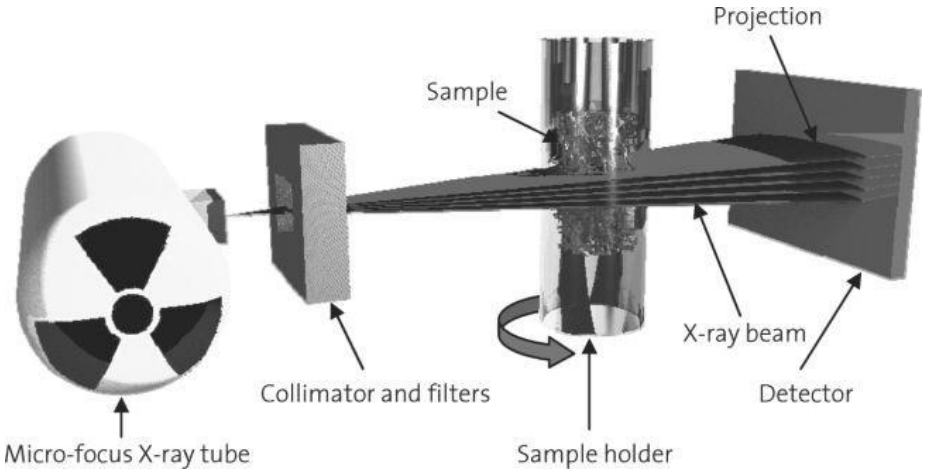


Figure 2.9: Micro-Computed Tomography Reproduced with permission from (Stauber & Müller, 2008) with Licence No. 5840740713714

To extract geometric and densitometric longitudinal information of a dataset of *in vivo* bone images, there are a number of processing approaches one could apply to the images including registration, segmentation, morphometric and densitometric analysis, etc. There are various types of registration, depending on the required deformations. For the scope of this study, the thesis focuses on the rigid, affine and deformable registration.

2.5.1 Rigid and affine registration

Aiming to align image samples, rigid registration allows the translation and rotation of the image to all or some directions. One could also enable scaling and shearing of image objects. In the latter case, the

registration is called affine. For aligning the image samples, a similarity measure between the images is defined. Different similarity metrics include mutual information and normalised mutual information, mean squared difference of intensities and normalised correlation. Among those, the normalised mutual information was demonstrated to better perform in rigid registrations for microCT images of C57BL/6 mouse tibia (Oliviero, 2019). In this procedure, one image is defined to be the reference. The other image is then translated and rotated through an optimisation process with iterations till the similarity measure is maximised. After the optimisation algorithm of the registration finishes, the registered 3D image is resampled to the reference coordinate system by applying interpolation to compute the new voxel intensities of the registered image.

The choice between these approaches depends on the objective of the study. In bone research investigating spatiotemporal changes, rigid registration is widely used to align all bone samples to a baseline coordinate system. Affine registration is usually used for observations of different sizes.

2.5.2 Deformable registration

Deformable registration encompasses a wider range of deformations enabling non-rigid and local deformations. It is widely used in medical imaging to enable direct localised comparison between images acquired at deformation states. As in any other registration type, there is the source and the target image: the source (alternatively “moved”) is deformed to overlap with the target image (alternatively “fixed”). One specific type of algorithm, available for 2D or 3D images, is grid-based.

An example of a grid-based algorithm is a toolkit called Sheffield Image Registration Toolkit (Barber & Hose, 2005). In the case of 3D registration, the input images are 3D. In order to map the images, a cubic grid is superimposed on both. The structural characteristics of the grid are that its elements are hexahedral, and the distance between the nodes (Nodal Spacing) is bigger than one voxel size. The algorithm deforms the grid of the “oved image in iterations until the similarity function between the image pair is optimised. Specifically, defines the displacement in the three directions x,y and z and the intensity change of each grid node so the difference between the intensity functions of the moved and fixed image are minimized. This is accomplished in iterations where a smoothing parameter and Nodal

Spacing are also altered. More details on the algorithm could be found in (Barber & Hose, 2005). The displacements on the nodes are usually applied to the bone surfaces to provide the local displacements on the bone. In this project, the smoothing parameter is let to be automatically optimized by the algorithm. However, Nodal Spacing is a parameter of significant interest as the spatial resolution of the grid has a direct impact on the accuracy and precision of the output mapping (Dall'Ara et al., 2017).

As deformable registration is a localised approach, it is commonly used to detect local structural and material differences. For example, deformable registration is primarily part of the Digital Volume Correlation (DVC) which measures strain fields by comparing deformed with the original undeformed bone structures (Bay et al., 1999). This is calculated by a displacement field calculated by tracking the changes of the bone features between the microCT images acquired from a non-deformed and deformed configuration after loading.

2.5.3 Segmentation/Binarization

Interested in the bone tissue solely, one could perform segmentation to separate the region of interest from the surrounding area. Segmentation is based on the grey levels of the images. This approach consists of three major steps. It first creates the histogram of the grey values. The histogram represents the number of pixels associated with a specific grey value in the image. A typical example of the histogram created by a microCT bone image exhibits two peaks as shown in Figure 2.10. The low grey level peak corresponds to the background and the high one to the bone. The algorithm continues with measuring these grey values and defining a threshold. Typically, a single-level threshold is used by computing the average of the peak grey values. Following this, a condition is conducted in which the pixels with higher pixel intensity than the threshold take the value of 1, and the pixels with lower grey values take the value of 0. The processed output image is the binary (black and white) image with 0 for the background and 1 for the bone, as shown in Figure 2.10. For this reason, the described process is also called binarization.

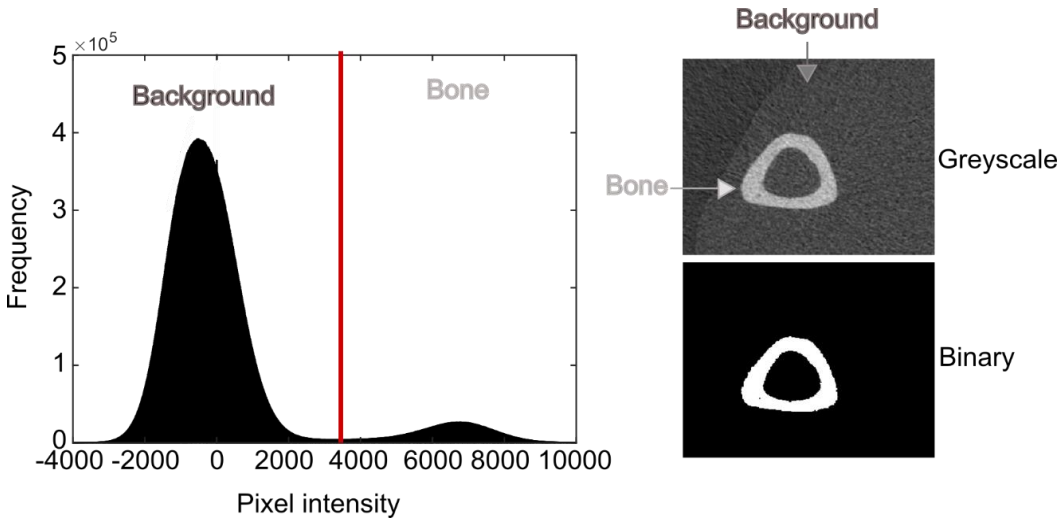


Figure 2.10: Segmentation (binarization) of greyscale microCT images using single-level thresholding.

2.5.4 Geometric Assessment

The yielded binarized images are then typically used to apply standard morphometric analysis which characterises the bone geometry, i.e., trabecular and cortical regions. This analysis uses a predefined set of scalar morphometric parameters which were introduced from the histomorphometry of *ex vivo* images and later from morphometric analysis of the *in vivo* images to ensure consistency between studies and enable comparisons.

In histomorphometry, *ex vivo* sections are extracted from bone samples, and they undergo microscopy and imaging when thickness, area and other geometric features are measured. Morphometric analysis of *in vivo* images was introduced as an advancement for geometry assessment enabling similar measures but in a non-invasive longitudinal manner. Both methods focus on limited sections. As in histomorphometry, where cross-sectional thin slices of the bone samples are analysed, standard 3D morphometric analysis of *in vivo* bone images is also typically performed on limited regions. The extracted parameters are then averaged across the Region of Interest. More details regarding the implementation of the morphometric analysis could be found in Bouxsein et al., (2010).

The standard morphometric properties that characterise the cortical bone are typically extracted from midshaft segments (Figure 2.11), and they include cortical thickness, total cross-sectional area, cortical area fraction and cortical bone area. Other common parameters include eccentricity, marrow area, maximum moment of inertia, minimum moment of inertia and polar moment of inertia. Temporal

changes within the examined population of bones are typically reported as differences in the morphometric properties. Despite the advancement of longitudinal imaging in capturing bone geometries at different ages, the simplified description of the geometry at limited regions is still a significant shortcoming of the current works, not exploiting the entire information that images could possibly give.

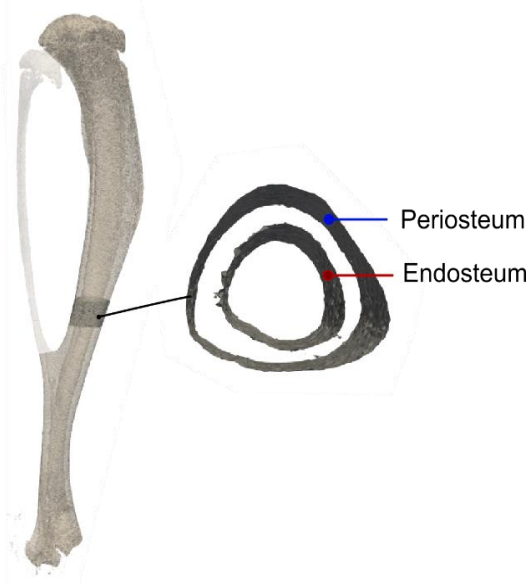


Figure 2.11: Cortical Region of Interest for the morphometric analysis (lateral to medial view)

2.6 Longitudinal mechanical assessment using Finite Element (FE) Modelling

Understanding the mechanics of materials is of great interest to engineers. Analytical solutions to describe how the structure deforms under a load or displacement are feasible for simple geometries and materials. However, in real life complex structures such as buildings, bridges, air wings, and complex materials such as wood, bones, muscles etc. are characterised by complex structural or /and compositional features. In such cases, analytical solutions are impossible to acquire. The Finite Element (FE) method was proposed to obtain approximate solutions for such structures (Burnett, 1987). It was originally performed in traditional engineering, such as aircrafts, but it has emergently expanded to other fields such as biomechanics. With the advancement of imaging technologies, enabling the accurate acquisition of complex bone geometries, FEM is now the gold standard for obtaining bone mechanics through simulations and out of labs.

In FE models, the structures are discretised into elements. Each element is composed of faces and a number of nodes, with three degrees of freedom (x, y, z coordinates). Some examples of 3D elements include the 8-noded hexahedral and the 4-noded tetrahedral ones. In bone research, both element types have been widely used, with the latter being a good choice for discretising irregular geometries and avoiding stair-edged surfaces.

In bone research, mechanical properties can also be measured with experiments. For instance, in preclinical studies, bone strength and stiffness are often evaluated by testing mouse bone specimens using three-point bending or compression tests. When using three-point bending, there are concerns such as the aspect ratio not covering the entire length of the bone and the lack of accounting shear (Wallace et al., 2014). In contrast, compression is more clinically relevant, with the loading direction aligned to the length. Regardless of the loading mode, this experimental process is destructive, and it only offers mechanical assessment cross-sectionally at the end of the testing phase. Thus, a longitudinal assessment of the intervention effects on bone strength is impossible in an experimental manner. For this reason, compressive micro-FEMs are created by longitudinal microCT 3D bone images and serve as predictive models. These mechanical models provide a non-invasive measurement of the mechanical properties while reducing the use of animals in the drug development process (Oliviero, 2019; Viceconti & Dall'Ara, 2019). Extensive validation against experimental data has proven the substantial contribution of these models in longitudinal studies to investigate intervention effects on bone geometry and mechanics.

In murine bone models, micro–Finite Element models are derived from microCT images considering the grey pixel intensity information. Specifically, this approach converts each bony voxel to a hexahedral element (stair-like surfaces). Although local density is sometimes considered to infer the calculation of the heterogenous young modulus, homogenous material properties are commonly used. That is because with rodent bones being less compositionally complex than human bones, the BMD variations across the bone structure are not prominent. Due to their smaller size, the spatial arrangement of the bone tissue is simpler, and heterogeneity in bone mineral density across the entire bone is

negligible. For instance, a previous study by Oliviero et al., (2021) demonstrated that the variations of tissue mineral density in mouse tibia, which result in heterogeneous Young Modulus, produce similar bone mechanics as models that assume homogenous material properties. Thus, geometry primarily has a direct impact on mechanical properties. The current study investigates two courses of treatments, the PTH (1-34) and/or the mechanical loading and their effects on mouse bone geometry and bone strength, with the assumption of homogenous material properties.

To investigate intervention impact on both geometry and mechanical properties, an association of the standard morphometric parameters and mechanical properties is typically assessed via correlation analysis.

2.7 Related *in vivo* longitudinal mouse studies

Previous *in vivo* murine studies in the research group generated osteoporotic mouse bone models which received various treatments (Roberts et al., 2019, 2020). In those studies, all experiments were ethically approved by the local Research Ethics Committee of the University of Sheffield, performed with the British Home Office licenses (PPL 40/3499 and PF61050A3) and complied with the UK Animals (Scientific Procedures) Act 1986. In the first study by Roberts et al. (2019), five C57BL/6 female mice at age 14 weeks were subjected to surgical removal of their ovaries to induce the osteoporotic phenotype. The mice remained untreated until the age of 24 weeks when they were euthanised. This group of mice is labelled as "OVX". In the next study by Roberts et al. (2020), 18 mice from the same in-bred genetic strain underwent ovariectomy, following the same procedure reported in the previous study. However, from week 18 onwards, the mice were separated into three groups (each containing six mice), namely "OVX+ML", "OVX+PTH" and "OVX+PTHML". These groups, respectively, received *in vivo* mechanical loading, injection of Parathyroid Hormone PTH (1-34) and the concurrent combination treatment with both.

In vivo mechanical loading was performed by uniaxially compressing the mouse tibia 3 days/week (Monday, Wednesday, and Friday). Waveform load cycles with a peak physiological load of 12 N over 10-second intervals were performed. Mechanical loading was applied at weeks 19 and 21 to all mice in

the “OVX+ML” and “OVX+PTHML” groups. Parathyroid hormone (PTH (1-34)) was administered by injection at 100μg/kg/day and 5 days/week (Monday−Friday). Injections were carried out every week (weeks 18−22) in all mice in the “OVX+PTH” and “OVX+PTHML” groups. In the combined treatment group, i.e., “OVX+PTHML”, the two treatments overlapped at weeks 19 and 21.

All time points, beginning from the start of the treatment intervention window were examined every other week, i.e., weeks 18, 20, 22 and 24. Thus, it is worth noting that all murine bones at week 18 presented osteoporotic evidence for 4 weeks, but after this age, each treatment group represented different treatment strategies.

In these studies, the right tibiae were scanned every two weeks between weeks 14 and 24 using *in vivo* micro–CT (Roberts et al., 2019, 2020). The longitudinal scanning protocol was previously proposed by Oliviero et al. (2019) which aimed to reduce the radiation dose (Oliviero et al., 2017). Specifically, the scanning parameters were: 10.4-μm isotropic voxel size, voltage of 55 keV, intensity of 145 μA, field of view of 32 mm, 1500/750 samples/projections, and integration time of 100 ms.

In later studies, the *in vivo* bone images were extensively used for geometric and mechanical assessments advancing the understating in how different OP treatments affect bone geometry and mechanics over time. To facilitate such assessments, image rigid registration was the fundamental step enabling the inter and intra comparisons within the mouse population. The rigid registration was typically accomplished by following the suggested protocol detailed in Lu et al. (2016). An arbitrary reference image was chosen and was anatomically oriented. Specifically, the *z*-axis was aligned to the reference tibia anatomical axial direction and the *x*-axis was aligned to the anterior–posterior direction, such that the *x*–*z* plane bisected the midpoint of the line joining the centres of the articular surfaces of the medial and lateral condyles. Next, the growth plate, fibula and condyles were removed from the images to improve the rigid registration (Cheong et al. 2020b). This led to a cropping that resulted in 80% of the total tibia length, starting from under the growth plate. All remaining 3D image samples were finally rigidly registered to the reference.

The standard geometric assessment protocol to elucidate the osteoporosis and treatment effect in cortical and trabecular regions included the standard morphometric analysis applied on the 3D *in vivo* microCT scans, as detailed in section 2.5. Roberts et al. (2019, 2020) demonstrated that treatment increases cortical thickness and area, where as osteoporosis reveals no statistically significant changes in these parameters over time. Another geometrical model on the same dataset, combined a mechano-regulation algorithm with micro-finite element models, which highlighted the localised geometrical changes on the active surfaces of the mouse tibia (Cheong et al., 2020a, 2020b, 2021).

The mechanical assessment included the utilisation of the micro-finite element models to predict the strength of the bone samples. This model was validated against experimental data (Oliviero et al., 2018). Additionally, the failure criterion was optimised (Oliviero et al., 2021a) and hexahedral/ tetrahedral and heterogeneous/homogeneous material properties were compared (Oliviero et al., 2021b). This model was then used to investigate the osteoporosis and treatment effects on tibia strength, showing that combined pharmacological and mechanical treatment has increased anabolic effects, whereas the osteoporotic mice had stable strength over time (Roberts et al., 2024). This study also developed regression models for investigating the relationship between geometric parameters and the strength and showed high correlations of the strength with the cortical thickness.

2.8 Statistical methods in bone research

2.8.1 Variability of bone structures

The complexity of bone models arises from the heterogeneity of the geometric properties of such structures. Among the complexity of the bone geometry, the bone adaptation is also heterogeneous, and it can be very localised. While the state-of-the-art geometric assessment is primarily associated with scalar quantifiers, the assessment of treatment effects on the entire geometry remains a challenge. Hence, complete discretisation of bone shape is needed to describe the 3D complexity of the shape variations.

Additionally, sources of variability: genetic, biological factors, age or external environment determine the formation and the remodelling of the bone, causing compositional variation among the organisms.

Since natural variability is inseparably bonded with biological structures (Currey et al., 2007), there is an inherent variability in the intervention's impact on bones.

Finally, in experiments, the specimens' preparation, testing methods, equipment, and protocols add another source of variance in the data. Standardisation in the models and protocols try to deal with this uncertainty across different laboratories. However, even in the same laboratory, some of these uncertainties cannot disappear and they are carried out within the entire analysis. Some examples include day of birth, genetic phenotypes, weight, success of external interventions, the placement of the specimens in the imaging equipment, the placement in the *in vivo* mechanical loading, etc.

Considering the above factors, a sample size of specimens is typically defined using power analysis. This is determined based on a desired effect size that one study expects to observe between the control cohort and the examined intervention group. Since the effect size of the interventions is a key aspect of the treatment investigation and drug development, more and more animals are used in preclinical studies for testing novel treatments. Therefore, there is an unmet need to develop sophisticated statistical computational models that can measure variations, provide a complete assessment of systematic spatiotemporal geometric changes induced by an intervention and determine representative effect sizes that describe these changes. This advancement would be in line with the ultimate purpose of *in silico* Medicine and the 3Rs, i.e., replacement, refinement, and reduction of the number of animals used in bone research. To accomplish this, this project identified the appropriate statistical approaches for a comprehensive statistical representation of bone geometry.

Bone research usually involves some statistics to measure the variability of bone properties like material characteristics, mechanical properties, and morphometric features. In statistical terms, these properties are treated as one-dimensional random variables, where observations are used to quantify variability. For more complex cases, such as bone geometry, the random variables are in more than one dimension. Herein, statistical approaches examine shape variability by discretising the bone surface into points, treating each point's coordinates as random variables. Since these points are linked with each other and form biological smooth surfaces that represent natural shapes, they are correlated with each other.

Mathematically speaking, this means that the variations of them are interdependent, with non-zero non-diagonal elements in the covariance matrix. In other words, these points cannot vary in an uncorrelated way. The key aspects of the statistical methodology are elaborated in Chapter 2.

2.8.2 Variability of Shapes - Principal Component Analysis

To understand the variations in shapes, this paragraph introduces the fundamental terms in Statistical Shape Modelling. An insightful book and a great introduction to Shape analysis is given by Davies et al., (2008). More technical details and a literature review specific to this research topic are given in Chapter 2.

In statistical learning of Shapes, the aim is to collect a database of shape observations which represent the same physical shape and learn from it by extracting 3-dimensional patterns of variability. The physical shape that these examples represent is usually called a “system” and in this study, the system is the mouse tibia.

The physical shape of tibia is represented by a cloud of points in the 3D space. Each point describes a unique anatomical feature of the mouse tibia and therefore this discretisation is called representation of the system. It is usually referred to as a “template” or “reference”. For now, and onwards, this thesis uses the word “reference”.

To compare the different shape observation with each other, any physical shape within the database should correspond to the points of the reference, with each point describing the same anatomical features between the reference and the shape sample. This characteristic is called point correspondence, being a fundamental requirement for developing statistical learning of shapes. The process used to achieve point correspondence is called mapping. Although there are many algorithms to perform mapping and meet the point correspondence requirement, they all finally yield mapped reference points to the shape observations.

To elucidate the modelling of shape variations within the database, the definition of distance between shapes is fundamental. Distances describing the difference between corresponding points are directly

linked with the point correspondence through the mathematical formula of point-to-point distance. For instance, in the case of two distinct shapes, each point of one shape is linked with the corresponding point of the other shape by a vector which is given by the distance. Subsequently measuring the distance for all points results in the construction of a 3D vector field of distances, that describes the 3D differences between the two shapes.

Differences in shapes measured by the shape distances can be classed into three different categories. The first category could be the pose or orientation differences. For instance, in the case of two same shapes, these differences could occur when the one is moved and rotated differently than the other. However, these differences in pose or orientation are not considered to form differences and, in many cases, such transformations should be eliminated. The second category is the size which is associated with the scaling of the shapes. For instance, in the case of two same shapes, these differences could be introduced when one shape is scaled up or down compared to the other. This transformation is typically of interest when statistical learning tools are used to understand how shapes change longitudinally. The third category is the local form differences which occur in specific areas of the shapes. In this case, the distance profile between two shapes would be non-uniform with higher magnitudes in the most different areas. The second and the third categories are considered as form differences between shapes and constitute the primary interest of the current project.

When considering a population of shape observations, modelling of the distribution of differences between shapes is typically accomplished using statistical learning tools. One widely used procedure to model the distribution of shape instances in the 3D space is the Principal Component Analysis (PCA).

PCA describes the shapes as a linear combination of the orthogonal axes of variations, called principal directions, principal components or modes (Davies et al., 2008). Specifically, each shape observation is modelled as a function of the mean shape and some vector fields of deformation describing the differences of this shape from the mean shape. Let's denote the data matrix **X** with each row representing a shape observation and each column representing a coordinate of a point on the shape.

Mathematically, PCA is the Singular Value Decomposition of the centred data **X**, i.e., subtracting the mean shape (Jolliffe & Cadima, 2016), as follows:

$$\mathbf{X} = \mathbf{U\Sigma V}^{T} \tag{2.1}$$

Where **V** contains the modes. Each column corresponds to one mode. **UΣ** represent the projection of the cantered data **X** onto the principal directions, and they are called scores. The matrix S encompasses information about the variance captured by each mode. Specifically, the square of the k-th singular value σ_k^2 is proportional to the corresponding mode as described by:

$$\text{Explained variance (mode k)} = \frac{\sigma_k^2}{\sum_{j=1}^{p} \sigma_j^2} \text{ , j=1, …, p: number of modes} \tag{2.2}$$

2.8.3 Variability of Shapes and Mechanics - Regression models

Simple regression models investigate the relationship between two scalar variables. They are usually applied to examine how one variable Y is affected by the variation of the other variable X. In this context, the Y variable is called response (predicted variable), and the X variable is called predictor. One example of a linear regression model could include the relationship between the thickness –scalar variable– (predictor X) and the bone strength (response Y) as conducted by Roberts et al., 2024.

In more complex cases, when the system has multiple predictors (X1, X2, …Xn, n= number of predictors) that are also correlated with each other, more sophisticated regression models should be utilised. One striking example is Partial Least Squares (PLS). An application in which this model could be useful is the investigation of how the 3D bone geometry affects bone strength. In this context, the geometry is modelled using discretised surfaces, with its nodes being the PLS predictors.

This paragraph details the PLS algorithm based on Singular Value Decomposition (De Jong, 1993). Let's denote the predictor matrix **X** and the response matrix **Y**. PLS identifies modes that maximise the covariance between X and Y and computes the scores, i.e., projections of both variables, that align along the modes. Firstly, PLS computes the cross-covariance matrix **C**:

$$\mathbf{C} = \mathbf{X}^{T}\mathbf{Y} \tag{2.3}$$

and performs the Singular Value Decomposition on the **C**:

$$\mathbf{C} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{T} \tag{2.4}$$

Where **U** represent the predictor modes, **V** the response modes, and finally, the diagonal elements of the **Σ** represent the strength of the relationship between **X** and **Y**. The predictor scores are the projections onto the predictor modes:

$$\mathbf{T}=\mathbf{X}^{T}\mathbf{U} \tag{2.5}$$

Similarly for the response scores. To calculate the portion of variance in the predictor variable **X** explained by a predictor mode (a column in **U**), PLS calculates the underlying total variance in **X,** $||\mathbf{X}||^2$. It then computes the reconstructed X_k using the k-th predictor mode:

$$\mathbf{X}_k=\mathbf{t}_k\mathbf{u}_k \tag{2.6}$$

Where $\mathbf{u}_k$, $\mathbf{t}_k$ represent the k-th mode and the corresponding scores, respectively. Following this, it quantifies the variance explained by each reconstructed $\mathbf{X}_k$, i.e., $||\mathbf{X}_k||^2$, and finally calculates the fraction, i.e., dividing it with the total variance:

$$\text{Explained variance (predictor mode k)} = \frac{\left|\left|\mathbf{X}_k\right|\right|^2}{\left|\left|\mathbf{X}\right|\right|^2} \tag{2.7}$$

Similarly for the response Y. There are alternative algorithms to compute the modes and scores for both response and predictor, but the objective of the PLS remains the same. This thesis details the algorithm that is used in the following chapters.

2.9 References

Ahmed, Y., Abdelsabour-Khalaf, M., & Khalil, F. (2017). Absence of typical haversian system from the compact bone of some reptile and bird species. *ArXiv Preprint ArXiv:1707.07444*. https://doi.org/10.48550/arXiv.1707.07444

Akhter, M. P., & Recker, R. R. (2021). High resolution imaging in bone tissue research-review. *Bone, 143*, 115620. https://doi.org/10.1016/j.bone.2020.115620

Aslan, D., Andersen, M. D., Gede, L. B., de Franca, T. K., Jørgensen, S. R., Schwarz, P., & Jørgensen, N. R. (2012). Mechanisms for the bone anabolic effect of parathyroid hormone treatment in humans. *Scandinavian Journal of Clinical and Laboratory Investigation, 72*(1),

14–22. https://doi.org/10.3109/00365513.2011.624631

Barber, D. C., & Hose, D. R. (2005). Automatic segmentation of medical images using image registration: diagnostic and simulation applications. *Journal of Medical Engineering & Technology*, *29*(2), 53–63. https://doi.org/10.1080/03091900412331289889

Bay, B. K., Smith, T. S., Fyhrie, D. P., & Saad, M. (1999). Digital volume correlation: three-dimensional strain mapping using X-ray tomography. *Experimental Mechanics*, *39*, 217–226. https://doi.org/10.1007/BF02323555

Bianco, P., Robey, P. G., & others. (2000). Marrow stromal stem cells. *The Journal of Clinical Investigation*, *105*(12), 1663–1668. https://doi.org/10.1172/JCI10413.

Boruah, S., Subit, D. L., Paskoff, G. R., Shender, B. S., Crandall, J. R., & Salzar, R. S. (2017). Influence of bone microstructure on the mechanical properties of skull cortical bone--A combined experimental and computational approach. *Journal of the Mechanical Behavior of Biomedical Materials*, *65*, 688–704. https://doi.org/10.1016/j.jmbbm.2016.09.041

Bouxsein, M. L., Boyd, S. K., Christiansen, B. A., Guldberg, R. E., Jepsen, K. J., & Müller, R. (2010). Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. *Journal of Bone and Mineral Research*, *25*(7), 1468–1486. https://doi.org/10.1002/jbmr.141

Brandi, M. L. (2009). Microarchitecture, the key to bone quality. *Rheumatology*, *48*(suppl_4), iv3--iv8. https://doi.org/10.1093/rheumatology/kep273

Burnett, D. S. (1987). Finite element analysis: from concepts to applications.

Cheong, V.S., Roberts, B.C., Kadirkamanathan, V. and Dall'Ara, E. (2021). Positive interactions of mechanical loading and PTH treatments on spatio-temporal bone remodelling. *Acta Biomaterialia*, *136*, 291-305. https://doi.org/10.1016/j.actbio.2021.09.035

Cheong, V.S., Campos Marin, A., Lacroix, D. and Dall'Ara, E. (2020a). A novel algorithm to predict bone changes in the mouse tibia properties under physiological conditions. *Biomechanics and modeling in mechanobiology*, *19*, 985-1001. https://doi.org/10.1007/s10237-019-01266-7

Cheong, V.S., Roberts, B.C., Kadirkamanathan, V. and Dall'Ara, E. (2020b). Bone remodelling in the mouse tibia is spatio-temporally modulated by oestrogen deficiency and external mechanical

loading: A combined in vivo/in silico study. *Acta Biomaterialia*, *116*, 302-317. https://doi.org/10.1016/j.actbio.2020.09.011

Compston, J., Cooper, A., Cooper, C., Gittoes, N., Gregson, C., Harvey, N., Hope, S., Kanis, J. A., McCloskey, E. V, Poole, K. E. S., & others. (2017). UK clinical guideline for the prevention and treatment of osteoporosis. *Archives of Osteoporosis*, *12*, 1–24. https://doi.org/10.1007/s11657-017-0324-5

Currey, J. D. (2002). *Bones: structure and mechanics*. Princeton university press.

Davies, R., Twining, C., & Taylor, C. (2008). *Statistical models of shape: Optimisation and evaluation*. Springer Science \& Business Media.

Dall'Ara, E., Peña-Fernández, M., Palanca, M., Giorgi, M., Cristofolini, L., & Tozzi, G. (2017). Precision of digital volume correlation approaches for strain analysis in bone imaged with micro-computed tomography at different dimensional levels. *Frontiers in Materials*, *4*, 31. https://doi.org/10.3389/fmats.2017.00031

De Jong, S. (1993). SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and intelligent laboratory systems*, *18*(3), 251-263. https://doi.org/10.1016/0169-7439(93)85002-X

De Souza, R. L., Matsuura, M., Eckstein, F., Rawlinson, S. C. F., Lanyon, L. E., & Pitsillides, A. A. (2005). Non-invasive axial loading of mouse tibiae increases cortical bone formation and modifies trabecular organization: A new model to study cortical and cancellous compartments in a single loaded element. *Bone*, *37*(6), 810–818. https://doi.org/10.1016/j.bone.2005.07.022

Demontiero, O., Vidal, C., & Duque, G. (2012). Aging and bone loss: new insights for the clinician. *Therapeutic Advances in Musculoskeletal Disease*, *4*(2), 61–76. https://doi.org/10.1177/1759720X11430858

Ferguson, S. J., & Steffen, T. (2003). Biomechanics of the aging spine. *European Spine Journal*, *12*(SUPPL. 2), S97–S103. https://doi.org/10.1007/S00586-003-0621-0/FIGURES/3

Florencio-Silva, R., Sasso, G. R. da S., Sasso-Cerri, E., Simões, M. J., & Cerri, P. S. (2015). Biology of bone tissue: structure, function, and factors that influence bone cells. *BioMed Research International*, *2015*(1), 421746. https://doi.org/10.1155/2015/421746

Goodyear, S. R., Gibson, I. R., Skakle, J. M. S., Wells, R. P. K., & Aspden, R. M. (2009). A comparison of cortical and trabecular bone from C57 Black 6 mice using Raman spectroscopy. *Bone*, *44*(5), 899–907. https://doi.org/10.1016/j.bone.2009.01.008

Hodsman, A. B., Bauer, D. C., Dempster, D. W., Dian, L., Hanley, D. A., Harris, S. T., Kendler, D. L., McClung, M. R., Miller, P. D., Olszynski, W. P., & others. (2005). Parathyroid hormone and teriparatide for the treatment of osteoporosis: a review of the evidence and suggested guidelines for its use. *Endocrine Reviews*, *26*(5), 688–703. https://doi.org/10.1210/er.2004-0006

Jilka, R. L. (2013). The Relevance of Mouse Models for Investigating Age-Related Bone Loss in Humans. *The Journals of Gerontology: Series A*, *68*(10), 1209–1217. https://doi.org/10.1093/gerona/glt046

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Lanyon, L. E., Goodship, A. E., Pye, C. J., & MacFie, J. H. (1982). Mechanically adaptive bone remodelling. *Journal of Biomechanics*, *15*(3), 141–154. https://doi.org/10.1016/0021-9290(82)90246-9

Larsson, S. (2002). Treatment of osteoporotic fractures. *Scandinavian Journal of Surgery*, *91*(2), 140–146. https://doi.org/10.1177/145749690209100202

Lu, Y., Boudiffa, M., Dall'Ara, E., Bellantuono, I., & Viceconti, M. (2016). Development of a protocol to quantify local bone adaptation over space and time: Quantification of reproducibility. *Journal of Biomechanics*, *49*(10), 2095–2099. https://doi.org/10.1016/j.jbiomech.2016.05.022

Oliviero, S., Owen, R., Reilly, G.C., Bellantuono, I. and Dall'Ara, E. (2021a). Optimization of the failure criterion in micro-Finite Element models of the mouse tibia for the non-invasive prediction of its failure load in preclinical applications. *Journal of the Mechanical Behavior of*

Biomedical Materials, *113*, 104190. https://doi.org/10.1016/j.jmbbm.2020.104190

Oliviero, S, Roberts, M., Owen, R., Reilly, G. C., Bellantuono, I., & Dall'Ara, E. (2021b). Non-invasive prediction of the mouse tibia mechanical properties from microCT images: comparison between different finite element models. *Biomechanics and Modeling in Mechanobiology*, *20*, 941–955. https://doi.org/10.1007/s10237-021-01422-y

Oliviero, S. (2019). *Non-invasive prediction of bone mechanical properties of the mouse tibia in longitudinal preclinical studies*. University of Sheffield. https://etheses.whiterose.ac.uk/22718/

Oliviero, S., Giorgi, M. and Dall'Ara, E. (2018). Validation of finite element models of the mouse tibia using digital volume correlation. *Journal of the Mechanical Behavior of Biomedical Materials*, *86*, 172-184. https://doi.org/10.1016/j.jmbbm.2018.06.022

Oliviero, S., Lu, Y., Viceconti, M. and Dall'Ara, E. (2017). Effect of integration time on the morphometric, densitometric and mechanical properties of the mouse tibia. *Journal of Biomechanics*, *65*, 203-211. https://doi.org/10.1016/j.jbiomech.2017.10.026

Olszta, M. J., Cheng, X., Jee, S. S., Kumar, R., Kim, Y.-Y., Kaufman, M. J., Douglas, E. P., & Gower, L. B. (2007). Bone structure and formation: A new perspective. *Materials Science and Engineering: R: Reports*, *58*(3–5), 77–116. https://doi.org/10.1016/j.mser.2007.05.001

Reeve, J. (1996). PTH: a future role in the management of osteoporosis? *Journal of Bone and Mineral Research*, *11*(4), 440–445. https://doi.org/10.1002/jbmr.5650110404

Roberts, B.C., Cheong, V.S., Oliviero, S., Arredondo Carrera, H.M., Wang, N., Gartland, A. and Dall'Ara, E. (2024). Combining PTH (1-34) and mechanical loading has increased benefit to tibia bone mechanics in ovariectomised mice. *Journal of Orthopaedic Research*, *42*(6), 1254-1266. https://doi.org/10.1002/jor.25777

Roberts, B. C., Arredondo Carrera, H. M., Zanjani-pour, S., Boudiffa, M., Wang, N., Gartland, A., & Dall'Ara, E. (2020). PTH (1–34) treatment and/or mechanical loading have different osteogenic effects on the trabecular and cortical bone in the ovariectomized C57BL/6 mouse. *Scientific Reports*, *10*(1), 8889. https://doi.org/10.1038/s41598-020-65921-1

Roberts, B. C., Giorgi, M., Oliviero, S., Wang, N., Boudiffa, M., & Dall'Ara, E. (2019). The longitudinal effects of ovariectomy on the morphometric, densitometric and mechanical properties in the murine tibia: A comparison between two mouse strains. *Bone*, *127*, 260–270. https://doi.org/10.1016/j.bone.2019.06.024

Rodan, G. A., & Martin, T. J. (2000). Therapeutic Approaches to Bone Diseases. *Science*, *289*(5484), 1508–1514. https://doi.org/10.1126/science.289.5484.1508

Russo, C. R., Lauretani, F., Bandinelli, S., Bartali, B., Di Iorio, A., Volpato, S., Guralnik, J. M., Harris, T., & Ferrucci, L. (2003). Aging bone in men and women: beyond changes in bone mineral density. *Osteoporosis International*, *14*, 531–538. https://doi.org/10.1007/s00198-002-1322-y

Shapiro, F., & Wu, J. Y. (2019). Woven bone overview: structural classification based on its integral role in developmental, repair and pathological bone formation throughout vertebrate groups. *Eur Cell Mater*, *38*, 137–167.

Sommer, N. G., Hahn, D., Okutan, B., Marek, R., & Weinberg, A.-M. (2019). Animal models in orthopedic research: the proper animal model to answer fundamental questions on bone healing depending on pathology and implant material. In *Animal Models in Medicine and Biology*. IntechOpen. DOI: 10.5772/intechopen.89137

Stauber, M., & Müller, R. (2008). Micro-computed tomography: a method for the non-destructive evaluation of the three-dimensional structure of biological specimens. *Osteoporosis: Methods and Protocols*, 273–292.

Unnanuntana, A., Rebolledo, B. J., Michael Khair, M., DiCarlo, E. F., & Lane, J. M. (2011). Diseases Affecting Bone Quality: Beyond Osteoporosis. *Clinical Orthopaedics and Related Research®*, *469*(8), 2194–2206. https://doi.org/10.1007/s11999-010-1694-9

Viceconti, M., & Dall'Ara, E. (2019). From bed to bench: How in silico medicine can help ageing research. *Mechanisms of Ageing and Development*, *177*, 103–108. https://doi.org/10.1016/j.mad.2018.07.001

Wallace, R.J., Pankaj, P. and Simpson, A.H.R.W. (2014). Major source of error when calculating bone mechanical properties. *Journal of Bone and Mineral Research*, *29*(12), 2697-2697. https://doi.org/10.1002/jbmr.2304

Weiner, S., & Wagner, H. D. (1998). The material bone: structure-mechanical function relations. *Annual Review of Materials Science*, *28*(1), 271–298. https://doi.org/10.1146/annurev.matsci.28.1.271

Weinstein, R. S., Jilka, R. L., Parfitt, A. M., Manolagas, S. C., & others. (1998). Inhibition of osteoblastogenesis and promotion of apoptosis of osteoblasts and osteocytes by glucocorticoids. Potential mechanisms of their deleterious effects on bone. *The Journal of Clinical Investigation*, *102*(2), 274–282. https://doi.org/10.1172/JCI2799.

Wirth, A. J., Müller, R., & van Lenthe, G. H. (2010). Computational analyses of small endosseous implants in osteoporotic bone. *Eur Cells Mater*, *20*, 58–71.

Wu, A.-M., Bisignano, C., James, S. L., Abady, G. G., Abedi, A., Abu-Gharbieh, E., Alhassan, R. K., Alipour, V., Arabloo, J., Asaad, M., & others. (2021). Global, regional, and national burden of bone fractures in 204 countries and territories, 1990--2019: a systematic analysis from the Global Burden of Disease Study 2019. *The Lancet Healthy Longevity*, *2*(9), e580--e592. https://doi.org/10.1016/S2666-7568(21)00172-0

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

Chapter 3

A novel framework for elucidating the effect of mechanical loading on the geometry of ovariectomized mouse tibiae using Principal Component Analysis

The study in this chapter has been originally submitted for publication and it is under review.

ABSTRACT

Murine models are used to test the effect of anti-osteoporosis treatments as they replicate some of the bone phenotypes observed in osteoporotic (OP) patients. The effect of disease and treatment is typically described as changes in bone geometry and microstructure over time. Conventional assessment of geometric changes relies on morphometric scalar parameters. However, being correlated with each other, these parameters do not describe separate fractions of variations and offer only a moderate insight into temporal changes. The current study proposes a novel image-based framework that employs deformable image registration on *in vivo* longitudinal images of bones and Principal Component Analysis (PCA) for improved quantification of geometric effects of OP treatments. This PCA-based model and a novel post-processing of score changes provide orthogonal modes of shape variations temporally induced by a course of treatment (specifically *in vivo* mechanical loading). Errors associated with the proposed framework are rigorously quantified and it is shown that the accuracy of deformable image registration in capturing the bone shapes (~1 voxel) is of the same order of magnitude to the relevant state-of-the-art evaluation studies. Applying the framework to longitudinal image data from the midshaft section of ovariectomized mouse tibia, two mutually orthogonal mode shapes are reliably identified to be an effect of treatment. The mode shapes captured changes of the tibia geometry due to the treatment at the anterior crest (maximum of 0.103 mm) and across the tibia midshaft section and the posterior (0.030 mm) and medial (0.024 mm) aspects. These changes agree with those reported previously but are now described in a compact fashion, as a vector field of displacements on the bone surface. The proposed framework enables a more detailed investigation of the effect of disease and treatment on bones in preclinical studies and boosts the precision of such assessments.

3.1 Introduction

Osteoporosis is one of the most severe and common skeletal diseases that reduces bone mineral density (BMD), diminishes bone quality and structural integrity, leading to weaker bones and higher risk of fracture. Treatment strategies such as exercise aim to enhance bone structure by stimulating new bone tissue formation and reducing bone loss (Rodan & Martin, 2000). Research in this area focuses on elaborating how bones respond to external mechanical stimuli. Murine models play a crucial role in investigating treatment strategies for osteoporosis due to their rapid response to interventions (Jilka, 2013). In particular, the ovariectomized murine model is an accepted model of oestrogen deficiency that accelerates bone resorption as observed in post-menopausal osteoporotic patients (Bouxsein et al., 2005). *In vivo* micro–Computed Tomography (microCT) enables the longitudinal acquisition of high-resolution images of peripheral bones (e.g., the tibia, caudal vertebrae) in mice (Bouxsein et al., 2010). For example, this approach has been used to study the effect of aging, ovariectomy, mechanical loading and pharmacological interventions (Akhter & Recker, 2021; Dall'Ara et al., 2016; Levchuk et al., 2014; Viceconti & Dall'Ara, 2019).

Established imaging and image processing protocols for *in vivo* microCT images of murine bones enable the quantification of spatio-temporal changes in geometry, microstructure and bone adaptation over time (Birkhold et al., 2015; Javaheri et al., 2020; Jepsen et al., 2015; Zhang et al., 2019; Roberts et al., 2019; Oliviero et al., 2019; van't Hof & Dall'Ara, 2019). Mouse bone geometric variations are often reported as variations of scalar geometric properties obtained from standard morphometric analysis of 3D images (Bouxsein et al., 2010). Such geometric properties, for e.g. cortical thickness, area, volume, eccentricity, moments of inertia, allow comparisons with other similar histomorphometric measures (Iida-Klein et al., 2007; Zhou et al., 2003). As histomorphometry analysis involves 2D *ex vivo* bone samples, standard morphometric analysis is also focused on short, pseudo-2D regions, such as the tibia midshaft, even when applied to 3D images. As such, standard morphometric analysis is based on assumptions and simplifications of the geometry and produces measures that are averaged over the examined section. Despite the advancement of *in vivo* longitudinal imaging, the methodology to analyse such images for a complete assessment of bone shape changes remains underdeveloped.

A limitation of the state-of-the-art approach is that it provides averaged scalar quantifiers to characterize the structure which effectively ignores the 3D complexity of bone shape changes. However, other preclinical studies provide evidence of localized bone geometry changes for e.g. due to external mechanical loading. Miller et al. (2021) measured the local thickness on four cross sections (25%, 37%, 50% and 75% of length) of murine tibiae and showed that the local changes due to mechanical loading depended on the location and the loading magnitude. Zhang et al. (2019) introduced a novel wavelet transform framework and found that over an 8-week period, changes observed in mouse tibia geometry were heterogeneous and contained both low- and high-frequency components in space. Javaheri et al (2020) used superimposition and rigid registration measures and showed that the mouse tibia adapted to a short-term loading regime, but after the regime concluded, these adaptations were only partially retained. Other computational studies predict local bone adaptation that also cannot be described by standard morphometric parameters alone. For example, Cheong, et al. (2020) suggested that mechanical loading of mouse tibia predominantly impacts the periosteum. Although the aforementioned techniques have revealed new insights regarding the local bone adaptation, it remains unclear which geometrical features explain the most variability within a population and which consistently change due to disease or treatment.

Another drawback of the scalar morphometric properties is their mutual interdependence i.e., non-negligible covariance. One example is the correlation between cortical thickness and cortical area, with the latter being described as a function of the former. This explains why studies report concurrent changes over time in both parameters (Roberts et al., 2020). Therefore, not only do the variations in individual standard morphometric parameters overlap with each other, but also, when combined, these also only partially explain the total variation in bone geometry. This also points to a potential challenge in using the effect sizes of the standard morphometric parameters (changes in bone shapes), when this assessment aims to represent the efficacy of a candidate treatment on bone geometry.

A potential alternative to the currently incomplete assessment of geometric changes is to use Principal Component Analysis (PCA) to extract ‘patterns’ of bone geometry variations. This is promising because

each ‘pattern’ (or PC mode) encodes local variations throughout the bone surface and is guaranteed to explain an independent fraction of the total bone geometry variation. More details on methodology and application of PCA in shape modelling can be found in monographs such as Davies et al. (2008).

PCA has been used in various applications. One common use in bone research is to identify shape and/or intensity variations on osteoporotic bones to improve osteoporosis care. The literature reviews of Castro-Mateos et al. (2014) and Grassi et al. (2021) detail the major methodological characteristics of the PCA models. Their various applications include segmentation, preoperative planning, 2D-to-3D intensity and shape reconstruction using 2D clinical images (Dual-energy X-ray Absorptiometry, DXA), Finite Element Modelling and investigation of bone fracture. More specifically, in the context of the human femur, PCA has been used to predict hip fracture risk combining shape and intensity modes derived by DXA (Aldieri et al., 2020). Additionally, it has been used to reduce the dependence on expensive imaging modalities (e.g. CT/MRI) in pre-operative planning (Rajamani et al., 2007, Barratt et al., 2008). Moreover, it was developed to generate statistical models of geometry and material properties of human femur (Bryan et al., 2010). In mouse bone research, the use of PCA is relatively limited. Killian et al. (2019) used PCA to investigate the deformity of dysplastic murine hips and Chan et al. (2012) used PCA to identify the major distal femoral geometrical features that temporally vary during bone maturation. Both these studies focused on sections of bones. A larger scale model developed by Brown et al. (2017) investigated abnormalities in the murine hind paws with rheumatoid arthritis.

Assessing shape variations requires image processing to enable direct comparison between the shape observations (Castro-Mateos et al., 2014). Chan et al. (2012) developed an automatic landmarking approach for murine bones and obtained an atlas with 412 anatomical landmarks. This bone structure was a segment of distal femur extracted from *ex vivo* images of mice at different ages (resolution 9 μm). Since the images were acquired in a cross-sectional experimental design, captured variations due to bone maturation could not be classed into temporal changes. Therefore, they were reported as variations from the mean shape. Additionally, the performed rigid registration protocol involved isotropic scaling

and eliminated the size variations related to length and area. Killian et al. (2019) also focused on subsections of bones in the hip joint, scanned using *in vivo* micro–Computed Tomography (resolution 21 μm). In that study, the surfaces were scaled up and an automatic landmarking algorithm was used to discretize the shapes using 2048 points. Although a comparison between disease severity also appears feasible using this approach, Killian et al. (2019) examined only one disease stage (severe) in their study. Brown et al. (2017) applied an automated method for discretizing and registering the meshes (~200,000 vertices) from volumetric microCT images of the mouse hind paw (resolution ~14 μm). Finally, Hoshino et al. (2023) analysed murine skull variations to investigate dysmorphology (resolution ~62 μm), using 33 landmarks to define the skull shape. The summary of this application is given in Table 3.1.

The findings from the literature suggest that instance alignment, shape correspondence and reliability analysis emerge as three important aspects common to all methodologies. The protocol of instance alignment varies among different studies. This step starts with defining an atlas (alternatively called template or reference) that can be either the mean shape or a random choice of one shape observation. Then the shapes are aligned to the reference. This alignment is either applied to the images, using rigid body movements, or applied to meshes, using algorithms based on distance metrics. Isotropic scaling of shapes is also sometimes used in the literature to disregard size variations. The number of landmarks, used to achieve correspondence among shapes, varies depending on factors such as examined structure, image resolution, discretization and mapping method. Correspondence is primarily accomplished via affine registration or other mathematical toolkits. Past studies highlight that the choices made in each step depend on the goal of each study, and the reliability of each step can influence the reliability of any shape model derived from them. For example, a small number of landmarks may lead to unrepresentative shape descriptions, or missing morphologic information of the examined structure, resulting in a lack of robustness to the shape analysis.

Table 3.1: PCA applications in mouse bones

Study	Research question	System (size)	Imaging	Reference	Number of landmarks	Alignment
Chang et al. (2012)	Murine bone maturation	Distal femur (segment)	*Ex vivo* (9μm)	Atlas (mean shape)	412	Image alignment with scaling
Killian et al. (2019)	Deformity of dysplastic murine hips	Hip joint (segments)	*Ex vivo* (21μm)	No	2048	Iterative closest point with scaling
Hoshino et al. (2023)	Murine skull dysmorphology	Entire skull	*In vivo* (2 μm)	No	33	Generalized Procrustes with normalized size
Brown et al. (2017)	Rheumatoid arthritis in murine hind paws	Hind paw	*Ex vivo* (14μm)	Sample	200,000	Mesh alignment

The objective of the present study is to develop a robust and accurate framework that can be used to assess the effects of disease (here, ovariectomy-induced bone loss) and treatment (here, passive mechanical loading) on the 3D geometry of long bones (here, tibial midshaft section) in a murine population (here, female adult C57BL/6 mice). This novel framework utilizes principal component analysis, a method that has not been used to assess the effect of treatment on mouse bone geometry.

3.2 Materials and methods

The framework developed in this study is used to analyze image data obtained from past animal experiments. Sections 3.2.1 and 3.2.2 recall the specific past studies concerning the mice used, the interventions performed, and longitudinal microCT images acquired following a well-established imaging protocol (Lu et al., 2016a). The details are also provided in Chapter 2, section 2.7. The details of the new framework follow from section 3.2.3 onwards. Sections 3.2.3–3.2.6 describe how the existing images are processed to extract bone geometry information. A validated and robust deformable registration technique named Sheffield Image Registration Toolkit (ShIRT) (Barber & Hose, 2005) is used here for the first time to map bones within a population of mice over a course of treatment. The mapping leads to differences in coordinates of "anatomically similar" locations on the endosteal and periosteal surfaces (see also schematic in Figure 1). In section 3.2.7, these differences in bone surface are decomposed into mutually orthogonal mode shapes using PCA, and section 3.2.8 describes a novel post-processing statistical analysis of PCA scores that is used to identify effects on geometry due to disease and treatment.

3.2.1 Animals and interventions

The experimental data were taken from two previous murine studies published elsewhere (Roberts et al., 2019, 2020). Briefly, female virgin C57BL/6 mice (n=5, Roberts et al. (2019); n=6, Roberts et al. (2020)) were housed starting at 13 weeks old. All mice were skeletally mature at the time of purchase and were ovariectomized at 14 weeks old to generate oestrogen deficiency (Turner 2001). The five mice comprising the control group "OVX" (Roberts et al., 2019) remained untreated throughout the study. The six mice in the treatment group "OVX+ML" (Roberts et al., 2020) were subjected to mechanical loading using the tibia loading model (Nepal et al., 2023) three days per week at 19 and 21 weeks old as shown in Figure 3.1.

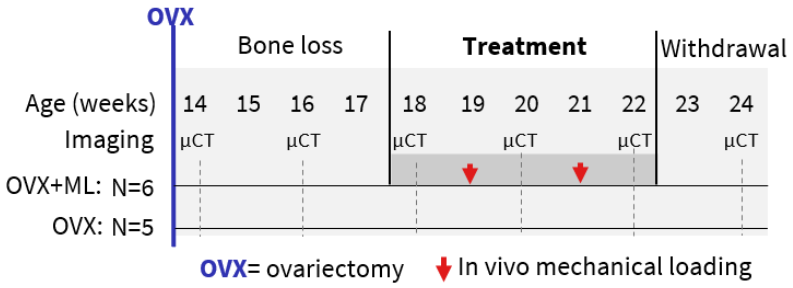


Figure 3.1: Examined groups of mice and the different interventions.

3.2.2 *In vivo* imaging and image alignment

The right tibia of each mouse was scanned using *in vivo* micro–Computed Tomography every 2 weeks between weeks 14 and 24 (Roberts et al., 2019, 2020). The microCT images of all mice and time-points were co-registered as follows (Cheong, et al., 2020; Cheong et al., 2021). First, the growth plate, the fibula and the condyles were removed from all microCT images. The proximal and distal cropping resulted in examining 80% of the total length. Next, the 3D image of one random tibia at 14 weeks was considered as reference and it was manually oriented to the anatomical coordinate system (Lu et al., 2016a). The remaining images were rigidly registered to the reference tibia image using Normalized Mutual Information as the similarity metric and Lanczos interpolation (Amira, v5.4.3, FEI Visualization Sciences Group, France).

In the present study, the registered images of each mouse from two time-points (18 and 24 weeks) constituted the input database comprising (5 "OVX" + 6 "OVX+ML") × 2 ages = 22 observations. Note that all mice at 18 weeks present evidence of ovariectomy untreated for a period of four weeks. The bones in the "OVX" group at 24 weeks present evidence of ovariectomy untreated for a period of 10 weeks. Lastly, the bones in the "OVX+ML" group at 24 weeks present evidence of ovariectomy untreated for a period of four weeks followed by mechanical loading treatment at weeks 19 and 21.

3.2.3 Midshaft section and further alignment

The 80% of the total length was divided in 10 equal longitudinal sections and slices corresponding to a midshaft section (Figure 3.2, Step 0) measuring 8% of the tibia length were considered. To suppress differences in relative position due to differences in whole tibia length, the cropped tibial midshaft images were rigidly registered to a randomly selected reference cropped image (Figure 3.2, Step 1).

Rigid registration was performed in Amira (v6.3, Thermo Fisher) using Mutual Normalized Information as similarity metric and Lanczos interpolation.

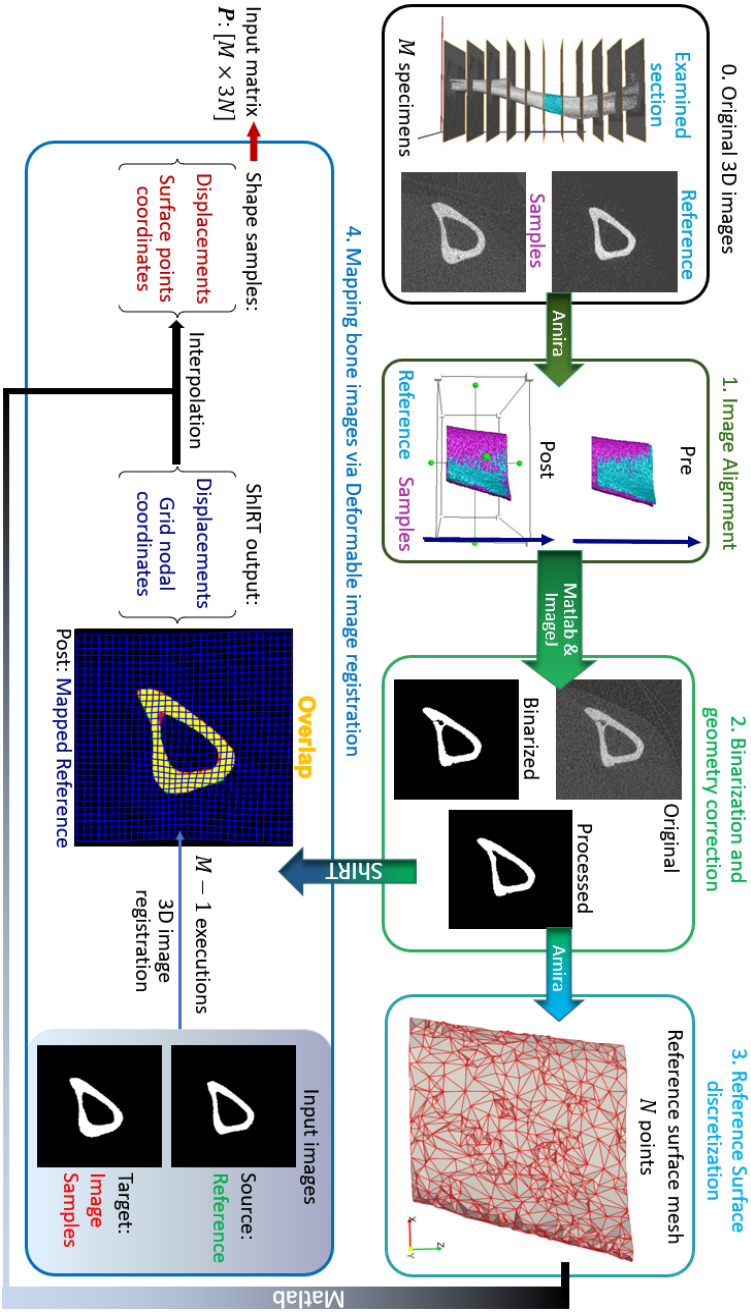


Figure 3.2: Framework flowchart. Step 0: The grayscale image slices corresponding to the midshaft section are extracted for each of the M images of mouse tibia. Step1: All images are aligned to a reference image sample previously registered to its anatomical coordinate system. Step 2: The greyscale images are binarized and corrected to enforce topological equivalence. Step 3: The reference surface mesh, consisting of N nodes with x, y and z coordinates, is extracted from the binarized reference image. Step 4: The reference bone is mapped to all other image samples applying a deformable image registration algorithm using the Sheffield Image Registration Toolkit (ShIRT). The deformation field is applied to the reference surface mesh leading to individual surface meshes for each bone shape observation. The periosteum and endosteum surface mesh for each mouse image constitutes the rows of the PCA input ***P****.*

3.2.4 Binarization and geometry correction

The midshaft sections typically contain cortical pores and trabeculae. These isolated features do not occur at anatomically similar locations across bones, and as such their presence poses challenges in mapping the periosteal and endosteal surfaces as desired. To suppress these features, the following semi-automatic process was used (Figure 3.2, Step 2). Firstly, every image was binarized using a single-level threshold. The arithmetic means of the peaks in the intensity histogram corresponding to the background and bone voxels was used as the threshold value. Then all 2D image slices were individually considered and holes with perimeter smaller than 50 pixels were automatically identified. The holes were then artificially 'filled' by applying simultaneous dilation and erosion algorithms (Matlab v2021, functions: 'imdilate', 'imerode'). The remaining big hole-like features such as the cavities formed between the endosteum and the trabecula as in Figure 3.3, were treated manually. The image data at both ages for that mouse were considered, and the number of adjacent slices containing the feature was counted in these images. Since bone remodelling over time may have an impact on trabeculae's thickness, if both counts were smaller than 7 slices (~70 micrometers), then the trabecular feature bounding the hole was identified and deleted from all slices in both images, if both counts were smaller than 7 slices (~70 micrometers), then the trabecular feature bounding the hole was identified and deleted from all slices in both images. The remaining hole-like features were classified as either being bound by trabeculae or being cortical pores, looking at their structural changes between the two time points. Features that were not holes, but appeared as notches and transverse gaps in the cortex were classified as early-stage formed trabeculae and transverse cortical pores respectively. After classification, trabeculae were deleted, and cortical pores were filled, leading finally to binarized images of the registered section with well-defined endosteal and periosteal surfaces. The manual processing of one set of image sample required approximately 10-15 minutes.

This geometry correction step was evaluated to demonstrate the bone geometry and mechanical alterations due to the simplification of the bone sections. The sensitivity study is elaborated in the Supplementary Material 1. Briefly, micro–Finite Element analyses were conducted on the original and processed bone sections. The number of common and different elements were assessed to indicate the

fraction of the geometric differences between the two types of models. The differences in the strain distribution across the entire bone section and locally around the cortical pores and trabeculae and in the highly deformed areas were also assessed. It was found that the removed features occupied an average 0.40% the total number of elements across all mouse samples and the strain distribution of the processed section is not different from the ones of the original bones.

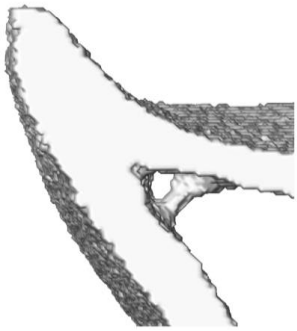


Figure 3.3: Example of the trabecula feature.

3.2.5 Reference surface discretization

The endosteal and periosteal surfaces bounding the bone volume in the binarized reference image were discretized (Figure 3.2, Step 3). The discretization was produced using the default settings in Amira v6.3 and is denoted as Mesh-0. It contained 166639 triangular faces with a mean edge length of 13.9 μm. A coarser discretization of the reference bone surface can potentially reduce the computational demand in the trilinear interpolation and PCA steps of the framework described later. Therefore, the above mesh was coarsened successively using Amira v6.3 such that the final chosen mesh (denoted as Mesh-15) included 2978 vertices, 5802 faces and a mean edge length equal to 68.2 μm. The sub-study reported in Supplementary Material 2 demonstrated that such coarsening did not result in a loss of important geometrical features. Specifically, the coarsening induced error in geometry was randomly distributed over the bone surface, the mean error was 0.22 μm and the maximum error of 1.5 μm was located in a relatively flat area of the bone surface.

3.2.6 Mapping bone images using deformable image registration

Three-dimensional deformable image registration was used to individually map the binarized reference bone image to each of the remaining 21 (also binarized) individual bone images (Figure 3.2, Step 4). The Sheffield Image Registration Toolkit (ShIRT) was used for mapping, employing Nodal Spacing

(NS) of 5 voxels. Details on this methodology can be found in section 2.5.1 and in Barber & Hose, (2005). Briefly, ShIRT superimposed a cubic registration grid with a given NS on both the fixed (an individual mouse tibia sample) and the moved (reference to be mapped) images. The mapping between the two images was determined as a displacement vector field on the grid. Hence, the displacement at any reference image voxel location or reference surface mesh vertex was then computed by tri-linearly interpolating the displacements at its eight closest registration grid nodes. Adding the displacement of a reference surface mesh vertex to its position on the moved image gave the position of an "anatomically similar" location on the individual bone surface of the fixed image.

Since the performance of the algorithm is sensitive to its spatial resolution (and hence NS) and the inherent bone surface differences in the input images, an evaluation study optimized the choice of the NS considering a combination of sources of complexities in the images. Similar to past studies (Dall'Ara et al., 2014), this was done by applying known "virtual" displacement fields to representative images, predicting this displacement using ShIRT, and quantifying the difference between known and predicted displacements for a range of NS values. Six studies, described below, define virtual displacement fields of increasing complexity by successively including new sources of uncertainty. The images of the six mice at 18 weeks in the "OVX+ML" group are representative of the full dataset. The known virtual displacements were used to synthetically generate six corresponding fixed bone images. ShIRT was used to register the fixed/moved image pairs with NS ranging from 5 to 50 voxels in steps of 5 voxels. The absolute difference between the imposed displacement and that estimated by ShIRT was calculated at bone boundary voxels, i.e. voxels whose voxel neighbours were not all bone or not all background. For a given NS, the average and standard deviation of the absolute differences, taken over all bone boundary voxels of all six moved images, are referred to as errors associated with accuracy and precision of registration, respectively. The median of the errors associated with accuracy and precision over mice was finally computed. As the execution time and memory demand to register a typical pair of images was ~15 minutes and 2.26GB RAM with NS=5 voxels, but ~35 minutes and 12.33GB RAM with the smallest allowable NS=2 voxels, respectively, NS lower than 5 voxels were not considered in the following. Similar evaluations, using the same algorithm for registering images

of same resolution and same bone structure, considered Nodal Spacings larger than 10 voxels (Dall'Ara et al., 2017). For all the six studies, the bone images obtained after binarization and geometry correction in section 3.2.4 above were considered as the initial moving images.

1. **Uniform translation of binarized images**: This study investigated the effects of rigid misalignment of the bone samples on the performance of the elastic registration. It tested the ideal scenario of uniform translations, equal to an integer number of voxel size. Displacements of either 2, 4 or 6 voxels each in the three Cartesian directions were individually applied to the moved images to obtain the fixed images.
2. **Non-uniform translation on binarized images**: Similar to Study #1, this study investigated the effects of the non-uniform rigid misalignments of the bone samples on the elastic registration. Thus, a displacement of 2 voxels along both *x* and *y* directions and 4 voxels in the *z* direction was applied to obtain the fixed images.
3. **Non-uniform, fractional voxel translation on binarized images**: This study measured the effect of rigid registration, including interpolation errors due to the translations being equal to non-integer number of voxels. To test this, a translation of 2.5 voxels along both *x* and *y* directions were first applied to the images. These were then resampled using bilinear interpolation and finally a translation of 2 voxels in the *z* direction was applied to obtain the fixed images.
4. **Translation on greyscale images followed by binarization**: This study evaluated the effect of the rigid translation, combined with binarization and geometry correction, as in the sequence of steps in the proposed framework. The same translation as in study #3 was applied, but to the greyscale images of the six mice at 18 weeks in the "OVX+ML" group, as obtained prior to binarization and geometry correction steps in section 3.2.4. Binarization and geometry correction was applied to the translated images to obtain the fixed images. Study #4 is similar to Study #3, with the only difference being the binarization coming after the translation and interpolation. Therefore, this study includes the effect of applying the rigid registration firstly on grayscale images, and also investigates how the binarization and geometry correction influences the deformable registration performance. This case is more representative of the order of the image processing steps as proposed in the current PCA framework.

5. **Local deformation**: This study evaluated the performance of the ShiRT to capture the non-uniform and local shape differences within the examined population of different mice and ages. To do so, simulations of several local deformation fields were tested. Three affine transformations along the *x* direction were separately applied on the binarized images as obtained at the end of study #4. The three individual affine transformations were: (i) compression of the posterior half of the bone image by 0.95x (referred as 'Study #5: Posterior, Half, Tc=0.95'), (ii) compression of a smaller posterior section of the image by 0.85x, 0.95x or stretching by 1.15x (referred as 'Study #5: Posterior, Smaller Part, Tc=...'), and (iii) stretching the anterior half by 1.2x (referred as 'Study #5: Anterior, Half, Tc=1.2'). The third case was motivated by the observation that when any individual mouse bone image was overlaid on the reference image, the horizontal distance between the bone surfaces at the anterior crests were at most 11 voxels apart. The stretching factor of 1.2 achieved an imposed displacement of similar magnitude along the *x*-direction at the anterior crest. Figure 3.4 illustrates examples of virtual translation and affine transformation successively added on the original images, as evidenced by incomplete overlap between the obtained binarized deformed images (fixed) and the binarized original images (moved).
6. **Image noise:** Finally, this study examined the effect of the image noise in the performance of the elastic registration. To do so, original image noise was measured from the original input image and subsequently added to the image to create a simulated noisy image. These steps were accomplished as follows. For each mouse in the "OVX+ML" group at 18 weeks, 3D image masks were created to separate the bone tissue and background regions in the grayscale images. The standard deviations of the pixel intensity in these regions were separately computed and then averaged over the six mice, leading to one standard deviation value each for bone and background. These were used to define two normal distributions (both with zero mean) from which samples were drawn and added to the bone and background regions of the greyscale mouse images to create simulated noisy images. The simulated images were then deformed as described in study #5.

All virtual translations described above were performed using ImageJ 1.53s. Images were manipulated with affine transformation and noise in Matlab R2022b.

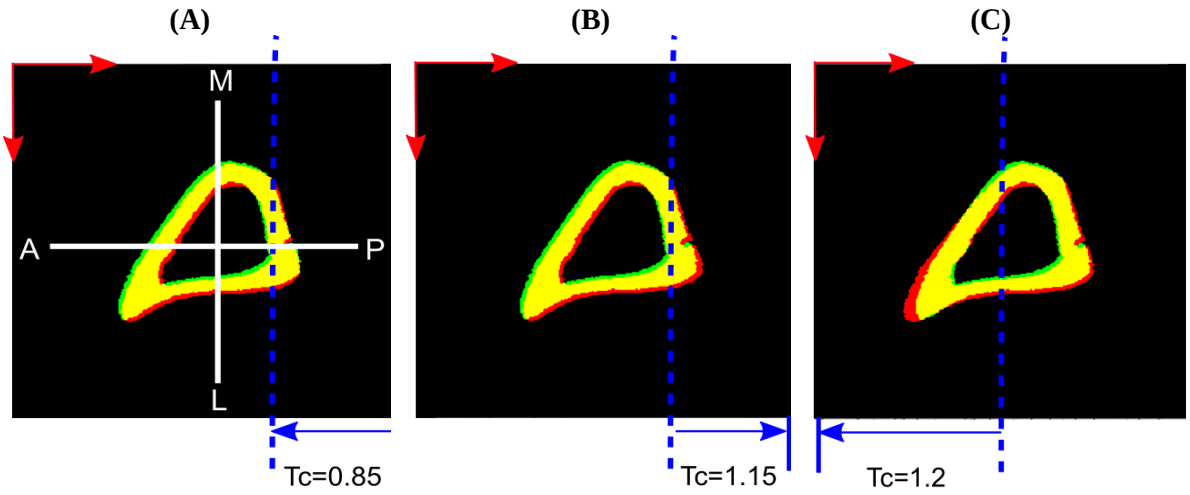


Figure 3.4: Examples of applied virtual deformation fields: (A) Study #5: Posterior, Smaller Part, Tc=0.85; (B) Study #5: Posterior, Smaller Part, Tc=1.15; (C) Study #5: Anterior, Half, Tc=1.2. The deformations comprise: a virtual translation with components 2.5 voxels in horizontal (anterior, A to posterior, P) and vertical (medial, M to lateral, L) directions (red arrows), and 2 voxels in the length direction (proximal to distal, not shown here); and an affine transformation that deforms only the part of the image spanned by the blue arrows, displaces the dashed blue line by zero and reaches its full magnitude (Tc) at the corresponding edge of the image. Regions of overlap between the fixed (red) and moved (green) bone images are shown in yellow.

Once the coordinates of the "anatomically similar" locations of the discretized endosteal and periosteal surfaces were obtained for each bone shape observation, these were concatenated to construct a single "shape vector"

$$\boldsymbol{x}_j = \{x_j^1, y_j^1, z_j^1, x_j^2, y_j^2, z_j^2, \ldots, x_j^N, y_j^N, z_j^N\}, \tag{3.1}$$

where (x_j^i, y_j^i, z_j^i) are the Cartesian coordinates of the i^{th} point (i=1, 2, …, N=2978) of the j^{th} bone shape observation (j=1, 2, …, M=22). Here, N and M are the number of vertices in Mesh-15 and the number of tibia observations respectively. A data matrix **X** is constructed where the j^{th} row constitutes the centred shape vector $\boldsymbol{x}_j - \boldsymbol{x}_0$ where $\boldsymbol{x}_0$ is a (3N)-element row vector denoting the mean shape. The shape variance was calculated from the product of the **X** with its self-transpose. The sum of all the diagonal elements divided by the N-1 gives the total variance, whereas the sum of the diagonal elements corresponding to a cohort group or age gives the proportional variance of a specific subgroup within the examined population.

3.2.7 Decomposition into mode shapes and validation

The matrix **X** is decomposed using PCA as

$$\mathbf{X} = \boldsymbol{a} \cdot \boldsymbol{Y} \tag{3.2}$$

where $\boldsymbol{a}$ is the matrix of mode scores (dimension M x (M–1)) and $\boldsymbol{Y}$ is the matrix of mode vectors (dimension (M–1) x (3N)). The M scalar values in the k^{th} column of matrix $\boldsymbol{a}$ (k=1, 2, …, M–1) represent the contributions to the M tibia observations due to the k^{th} Principal Component (or mode). The shape of this mode is given by the k^{th} row of matrix $\boldsymbol{Y}$.

Leave-one-out tests were conducted to assess the accuracy of reconstructing any tibia not belonging to the original database. In turn, the two shape observations (corresponding to weeks 18 and 24) of each of the six mice in the "OVX+ML" group, were randomly removed. PCA was performed on the remaining data, i.e., (5 "OVX" + 5 "OVX+ML") × 2 ages. For each PCA mode, the score values of the two left-out sample tibiae were found by projecting these on the mode shape. The left-out samples were reconstructed as the sum of the mean shape and the linear superposition of the mode shapes, weighted by the score values. The reconstruction error was defined as the distance between corresponding surface nodes of the two pairs of reconstructed and left-out samples. Note that reconstruction errors were assessed only for the six pairs of shape observations in the "OVX+ML" group to evaluate the sensitivity of the PCA model to describe treatment-related variations over time.

3.2.8 Temporal variations and treatment effects categorization

In the current application, longitudinal data of treated and untreated mice were analysed, i.e., 22 observations, to investigate the disease progression for the control group, "OVX", and the treatment progression for the "OVX+ML" group. The 21 PCA modes describe multiple sources of variations. These variations can be categorized (Figure 3.5) as either a combination of natural variability between individual mice and random errors in the image processing framework or a temporal change. For modes that are not associated with temporal changes, it is expected that their scores will remain similar over time for an individual mouse. Therefore, the change over time in the scores of each mode were computed for all individual mice. Only those modes are considered as associated with temporal change for which median score changes over time are statistically significantly different from zero in either or both the "OVX+ML" and "OVX" groups. Two-sided Wilcoxon signed rank test was used to test if the temporal score changes in any group are significantly different from zero. A nonparametric test was

used because of the small data sample size.

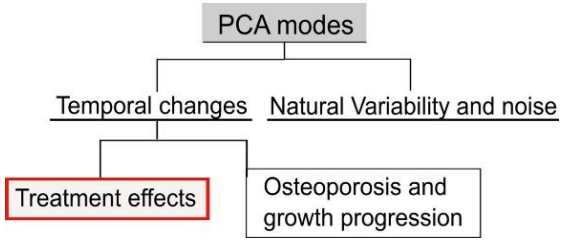


Figure 3.5: Types of variations in the examined population. The population includes longitudinal data of treated and untreated mice. The proposed PCA score processing uncovers all the sources of variations and differentiates them into classes. These classes are generic variations among groups and systematic variations with respect to the time. The latter can be further categorized into disease progression and growth, and treatment effects.

Temporal changes were further distinguished into either an effect of treatment or a combined effect of other temporal factors, e.g. ovariectomy and growth. Here, three possibilities arise: 1) effect only due to treatment: evidenced by a non-zero median score change in only the "OVX+ML" group; 2) effect not due to treatment: evidenced by a non-zero median score change in only the "OVX" group, 3) effect partially due to treatment: evidenced by non-zero median score changes in both "OVX" and "OVX+ML" groups. The Mann–Whitney U test was performed to compare the two cohorts with each other. Where a mode is established to be associated with a temporal change, this change was quantified in two ways. First, Cohen's d effect size was computed as the ratio of the average and standard deviation of the score changes in the group (Lakens, 2013). Second, the surface change that each mode describes was computed by considering the median score changes of the "OVX+ML" group between two ages and scaling the mode vectors. Specifically, a centred shape vector $\boldsymbol{P}_{\boldsymbol{j18}}^{\boldsymbol{k}}$, of the j-th shape observation at week 18, can be reconstructed using a treatment-related mode $\boldsymbol{Y}^{\boldsymbol{k}}$, k=1, …, M–1 using the PCA formula. Similarly for the j-th shape observation at week 24. Therefore, the two reconstructions of the j-th shape observation of a group at the two ages can be written as

$$\boldsymbol{P}_{j18}^{k} = a_{j18}^{k} \cdot \boldsymbol{Y}^{k}$$

$$\boldsymbol{P}_{j24}^{k} = a_{j24}^{k} \cdot \boldsymbol{Y}^{k} \tag{3.3}$$

The surface change $\boldsymbol{SC}$ of the j-th shape reconstruction using the k-th model is the difference between the coordinates of $\boldsymbol{P}_{j24}^{k}$ and $\boldsymbol{P}_{j18}^{k}$, and it can be written as:

$$\boldsymbol{SC}_j^k = \boldsymbol{P}_{j24}^k - \boldsymbol{P}_{j18}^k = \left(a_{j24}^k - a_{j18}^k\right) \cdot \boldsymbol{Y}^k \tag{3.4}$$

To represent a group, the median of the surface change is:

$$\boldsymbol{MSC}^k = median\left[\boldsymbol{SC}_j^k\right] = median\left[\left(a_{j24}^k - a_{j18}^k\right)\right] \cdot \boldsymbol{Y}^k \tag{3.5}$$

The median surface change of a group is henceforth called surface change for the sake of brevity. Note that the surface change has dimensions of length. The direction of the vectors indicates bone formation or bone deletion as geometric changes on the active surfaces.

3.3 Results

3.3.1 Mapping bone images using deformable image registration

In study #1, the registration algorithm ShIRT determined the known displacement of uniform rigid translation by an integer number of voxels, leading to a maximum error (across the range of NS investigated) of 2.5×10^{-5} voxels associated with accuracy and of 1.3×10^{-4} voxels associated with precision. In study #2, where a non-uniform rigid translation (2 voxels in *x* and *y* directions and 4 voxels in *z*) was applied, the registration errors were similarly negligible, as a maximum error associated with accuracy of 2.8×10^{-5} voxels and associated with precision of 1.7×10^{-4} voxels were obtained.

In study #3, rigid translation by a non-integer number of voxels led to constant errors (across the range of NS investigated) of ~0.15 voxels in magnitude associated with accuracy (Figure 3.6A) and ~0.05 voxels in precision (Figure 3.6B). Note that translation with a non-integer number of voxels involved blurry transition due to the linear interpolation of black and white pixels and then forced back to binary values (thresholding). This possibly explains the higher error magnitudes in this study compared with those in studies #1–2.

On the other hand, the same translation field on grayscale images revealed lower errors associated with accuracy (~0.08 voxels, Figure 3.6). Note that in this study #4 linear interpolation is performed on a spatially smooth field of grey values, which results in a smoother final bone surface than in study #3, which probably explains the lower errors associated with accuracy. The errors associated with precision

were similar to those in study #3. The slight decrease in errors associated with accuracy and precision with large NS in both studies was because the output displacement fields were smoother, possibly reducing a noise effect and getting closer to the uniform imposed virtual translation.

The displacements applied in 'Study #5: Posterior, Half, Tc=0.95' led to almost constant errors of ~0.87 voxels associated with accuracy and ~1.1 voxels associated with precision. The displacements applied in 'Study #5: Posterior, Smaller Part, Tc=…', led to errors associated with accuracy of ~0.20 voxels for affine transformation coefficient Tc=0.85, ~0.12 voxels for Tc=0.95 and ~0.18 voxels for Tc=1.15. Errors associated with precision were nearly constant around 0.34, 0.08 and 0.26 voxels when Tc=0.85, 0.95 and 1.15 respectively. Comparing 'Study #5: Posterior, Half, Tc=0.95' and 'Study #5: Posterior, Smaller Part, Tc=0.95' suggests that errors associated with both accuracy and precision increase as a larger part of the image is deformed. Comparing within 'Study #5: Posterior, Smaller Part, Tc=…' suggests that the errors associated with accuracy and precision are higher when the magnitude of the difference between the affine transformation coefficient and unity increases (Tc=0.85/1.15 vs 0.95, i.e., |Tc–1|=0.15 vs 0.05). It also revealed that the errors increase when the direction of the deformation and of rigid transformation oppose each other (shrinkage: 0.85, Figure 3.4A) instead of being aligned (expansion: 1.15, Figure 3.4B).

For the case 'Study #5: Anterior, Half, Tc=1.2', the errors associated with accuracy varied between 0.52 and 0.97 voxels, whilst those associated with precision varied between 0.68 and 1.45 voxels. Both types of errors increased with increase in NS, in a monotonic fashion. The higher errors in precision are due to the localization of large magnitude errors around the anterior crest, which is expected because the imposed displacement magnitude is highest in that region (Figure 3.7A). For NS=5, the maximum error at the anterior crest was 2.8 voxels, which is less than 25% of the maximum imposed simulated local deformation (12 voxels). The 75% (interquartile range, IQR=Q3–Q1) of the surface locations were successfully registered with a systematic error smaller than 0.47 voxels. The median registration error over the bone surface and for all mice was 0.18 voxels (Figure 3.7B).

Addition of image noise (Study #6) led to 0.56 voxels and 0.73 voxels in errors associated with accuracy

and precision respectively for NS=5, averaged across all mice, and to 0.59 voxels and 0.78 voxels in errors associated with accuracy and precision respectively for NS=10. For NS=5, the median registration error over the bone surface was 0.33 voxels, and its spatial distribution was similar to when image noise was absent for the same nodal spacing with slightly higher standard deviation ('Study #5: Anterior: Half, Tc=1.2').

After this step, the mesh samples were consistently discretized, with a fixed number of nodes. The shape variance as described by the centred mesh data in the groups "OVX" week 24 and "OVX+ML" week 24 amounted to 18% and 44% respectively of the total variation of geometry in the data set. The remaining variation in the dataset was due to all mice at week 18.

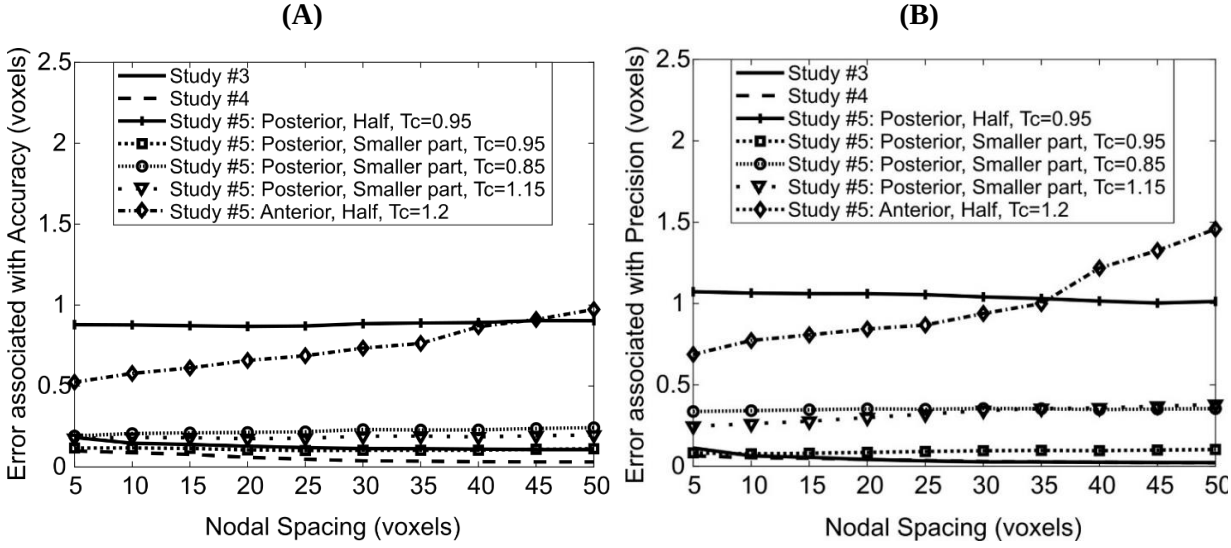


Figure 3.6: Dependence of errors associated with accuracy (A) and precision (B) of the deformable registration on Nodal Spacing. Each line corresponds to a different imposed displacement field; the details of these are found in the main text.

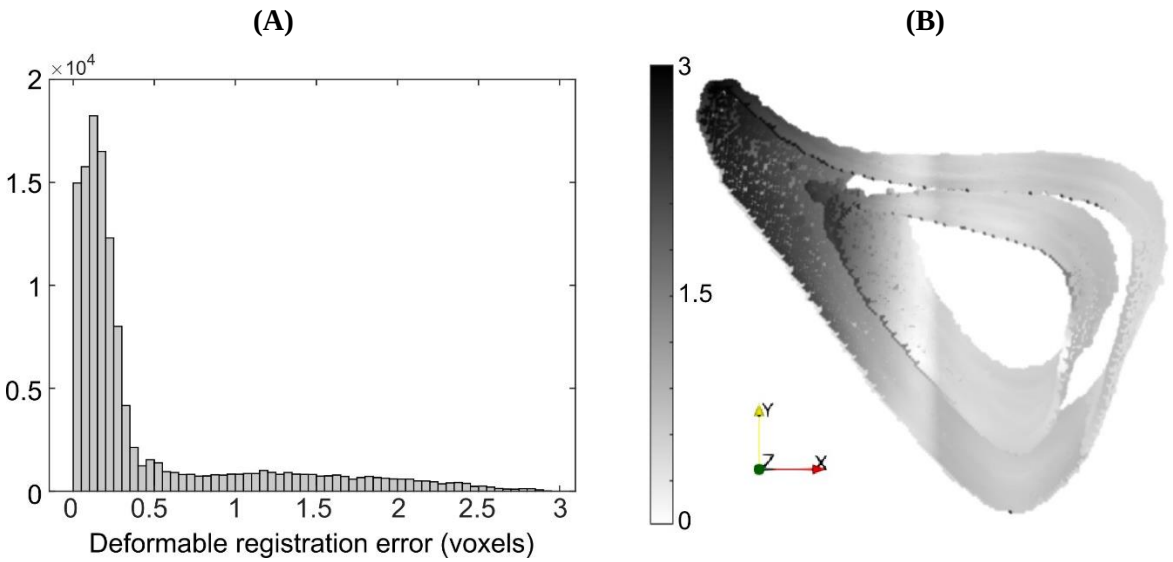


Figure 3.7: Histogram of registration errors and their spatial distribution on the bone surface for the simulated displacement field given by 'Study #5, Anterior, Half, Tc=1.2'. Errors are shown for a representative specimen taken from the "OVX+ML" group at 18 weeks of age. Contour darkness indicate error magnitude at the specific location on the bone surface.

3.3.2 Decomposition into mode shapes and validation

The first 6 PCA modes explained 91% of the total variation in tibia geometry (Figure 3.8, percentage explained is given by Equation 2.2 in subsection 2.8.2). All remaining modes (7–21 modes) individually explained 2% or less of the total variation in tibia geometry.

Figure 3.9 illustrates the 3D profiles of the modes. The modes are plotted as vector fields on the mean shape. The longer and darker vectors of a specific mode indicate the 3D features of the bone midshaft that exhibits the largest variations within the examined population. As shown, Mode 1 (explaining 49% of total shape variation) describes variations in the thickness and "sharpness" of the anterior crest as the darkest arrows are allocated at both endosteum and periosteum of the anterior crest. Mode 2 (explaining 20%) describes variations at the endosteum at the medial aspect. As shown in the Supplementary Material 3, mode 3 (explaining 8%) describes variations in the curvature of the medial and lateral aspects and mode 4 (explaining 6%) describes local variations at the lateral aspect of the distal end. Figure 3.9 also shows that Mode 5 (explaining 5%) describes local variations at the distal anterior crest and Mode 6 (explaining 3%) describes variations in local features scattered all over the bone midshaft, of which the variations at the posterior periosteal and posterior-lateral endosteal surfaces are prominent.

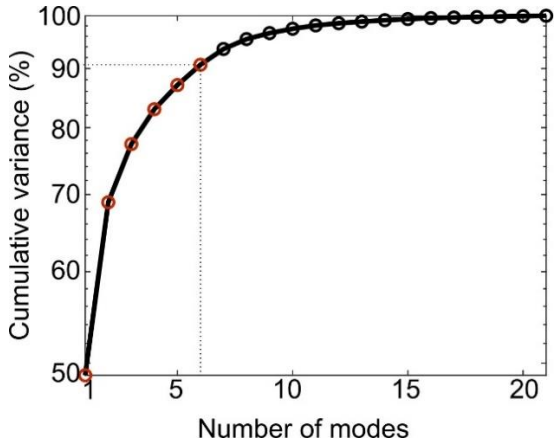


Figure 3.8: Cumulative variance (%) explained by the PCA modes. The first 6 modes (red circles) describe up to 91% of the total geometric variance within the examined population.

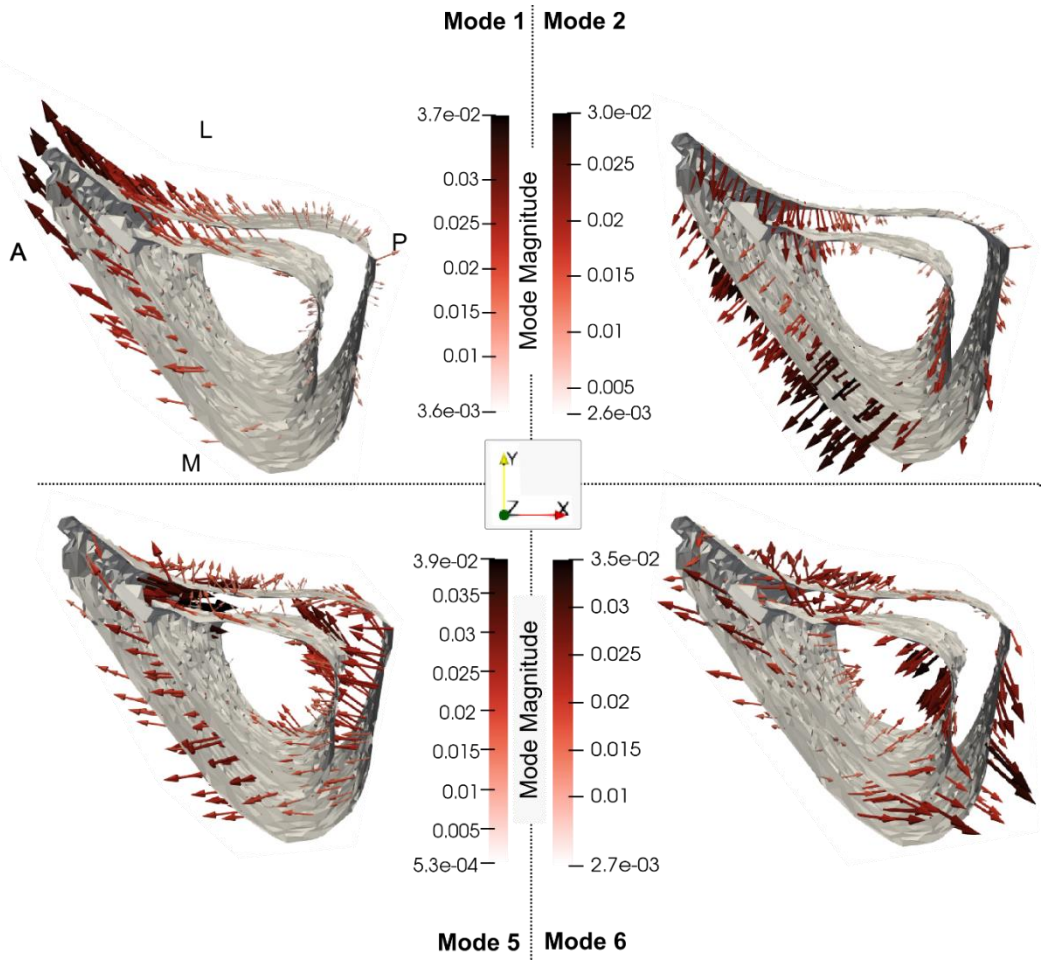


Figure 3.9: The 3D profiles of the treatment-related modes 1, 2, 5 and 6, depicted as vectors plotted on the mean shape. The vectors are scaled and coloured by the mode magnitude at each point of the mesh. Darker and longer arrows indicate higher variability in shape across different bone specimens at that surface location. All profiles share the same viewpoint from proximal to distal and the same anatomical orientation indicated by A (anterior), L (lateral), P (posterior), M (medial) as in Mode 1.

Reconstruction errors in the leave-one-out tests had an average, standard deviation and maximum of 1.5, 1.2 and 9.5 voxels respectively, when considered over all surface points and all mouse samples.

Figure 3.10 shows the boxplots of the error distribution over the tibia surface for specimens grouped by age. The large number of outliers in both distributions highlights their skewness. The median errors in week 18 (1.2 voxels) were lower than those in week 24 (1.4 voxels) (Figure 3.10A, $p<0.05$). Figure 3.10B shows the median error distribution on the 3D bone profile over the six mice at 24 weeks of age. Relatively higher error magnitudes are found around the proximal medio-posterior and anterior edges and at the distal latero-posterior aspect in the endosteum only.

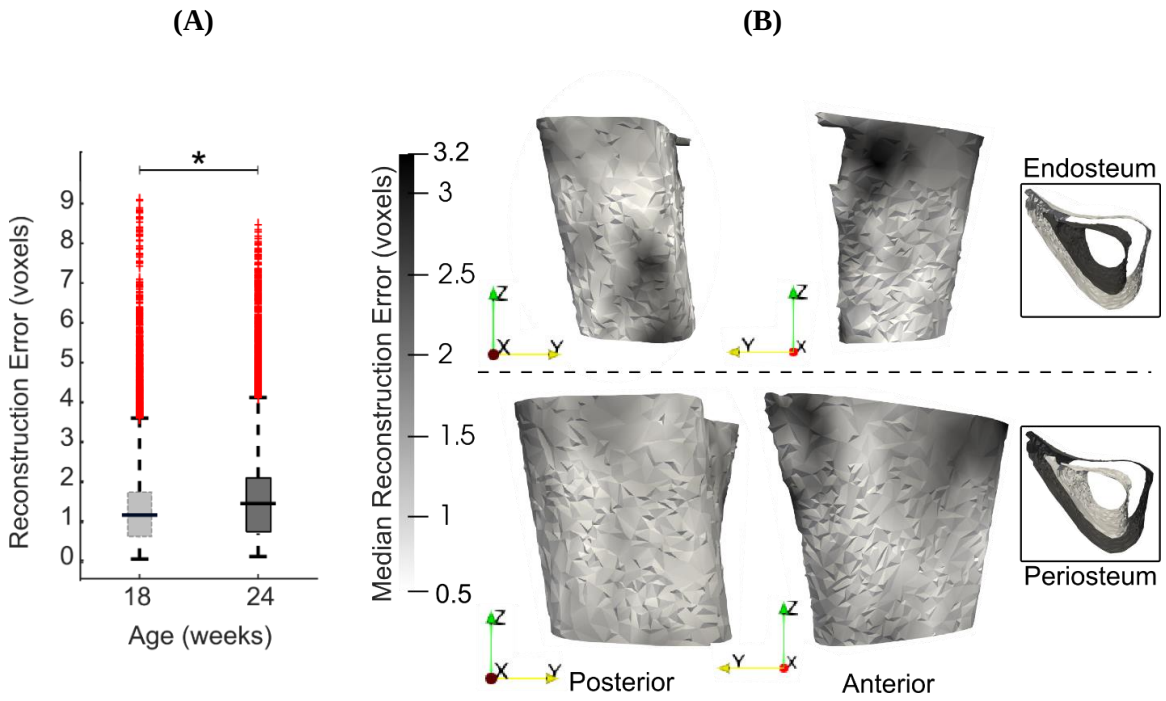


Figure 3.10: (A) Boxplots of the leave-one-out errors in reconstructing the bone geometry of six mice in the "OVX+ML" group at weeks 18 and 24. The overall median error for week 24 is statistically significantly higher ($p<0.05$) than for week 18. (B) Two different views (posterior-to-anterior and anterior-to-posterior) of the endosteal and periosteal surfaces (mean bone shape for week 24) overlaid with contours levels indicating magnitude of median error at each bone surface location for week 24.

3.3.3 Temporal variations and treatment effects categorization

Mouse-specific changes in mode scores, going from week 18 to 24 of age, were statistically significantly different from zero only for Modes 1, 2, 5 and 6 in the "OVX+ML" ($p<0.05$) group (Figure 3.11). As such, Modes 1, 2, 5 and 6 describe geometric features that are associated with a temporal change, and this change is an effect only due to treatment. This could be also visually indicated by the positive trend of score changes for "OVX+ML" group in Figure 3.11, in contrast with score changes for "OVX" group which do not follow a specific pattern. Modes 1, 2, 5 and 6 were found to have Cohen's d effect sizes of 2.0, 0.60, 0.54 and 2.4 respectively.

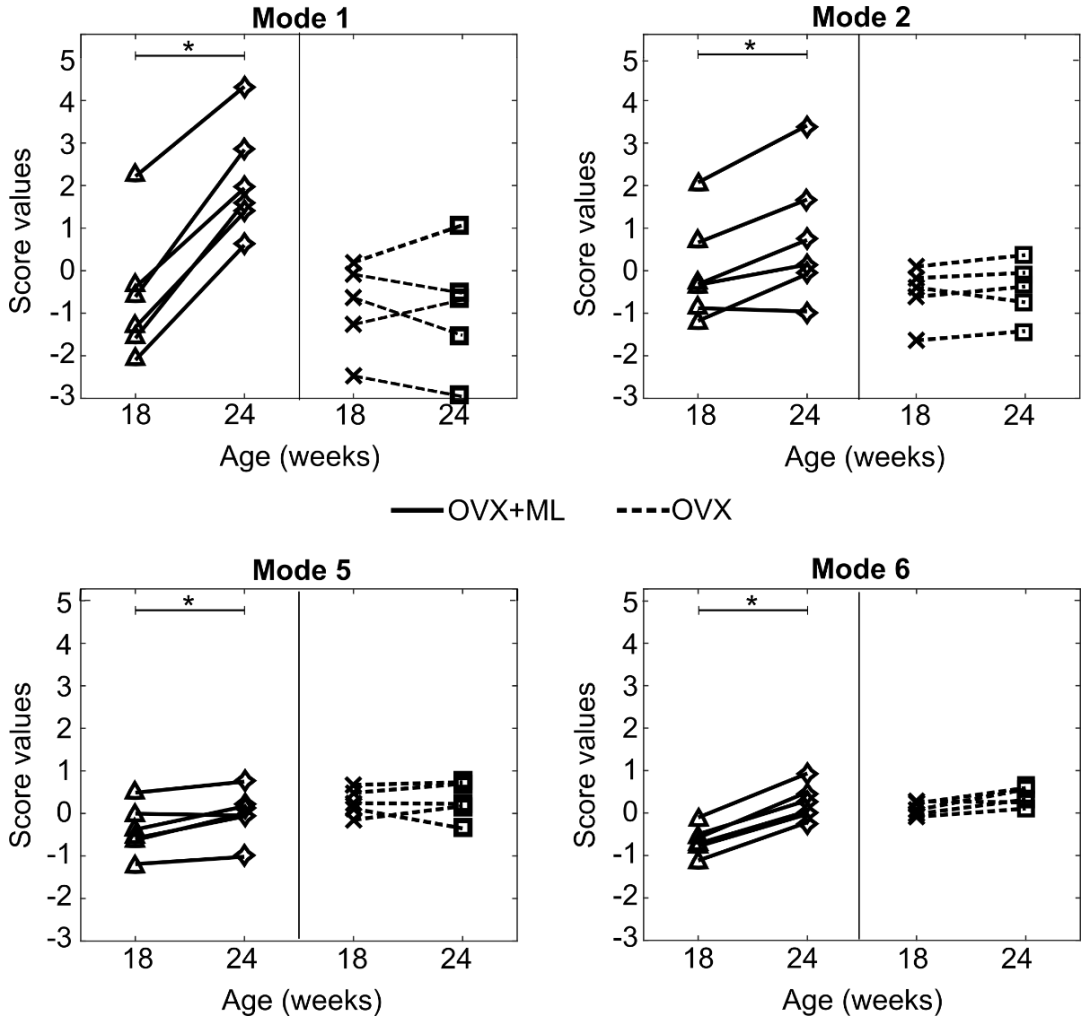


Figure 3.11: Treatment categorization of Modes 1, 2, 5 and 6. Mode score values of individual mice at are shown at week 18 ("OVX+ML", △; "OVX", ╳) and at week 24 ("OVX+ML", ◇; "OVX", ▢). Lines ("OVX+ML", solid; "OVX", dashed) connect mode scores of individual mice between the two time points. Asterisks (*) highlight median changes with time of mode scores in a group that are statistically significantly ($p<0.05$) different from zero.

For these modes, the profiles of the mode-specific temporal surface changes between week 18 and week 24 are shown in Figure 3.12. This figure illustrates the change map on top of the median mode-specific bone profile at week 18. The arrows and magnitudes show that Mode 1 is associated with prominent endosteal deletion of 0.054 mm and periosteal formation of 0.1 mm at the anterior crest along the whole length of the midshaft and less prominent concurrent endosteum and periosteal formation (0.027 mm) at the lateral-posterior aspect. A similar interpretation revealed that Mode 2 describes endosteum deletion and periosteal formation in the range of 0.013–0.032 mm along the medial side. Mode 5 primarily captured periosteal deletion at the distal end of the anterior crest with a magnitude of 0.011 mm. Although this opposed the effect described by Mode 1, the summative effect of both Mode 1 and

Mode 5 is still periosteal expansion of the anterior crest by 0.09 mm. Finally, Mode 6 exhibited endosteal formation at the posterior-lateral aspect of 0.03 mm and periosteal formation along the posterior side of similar magnitudes (0.03 mm).

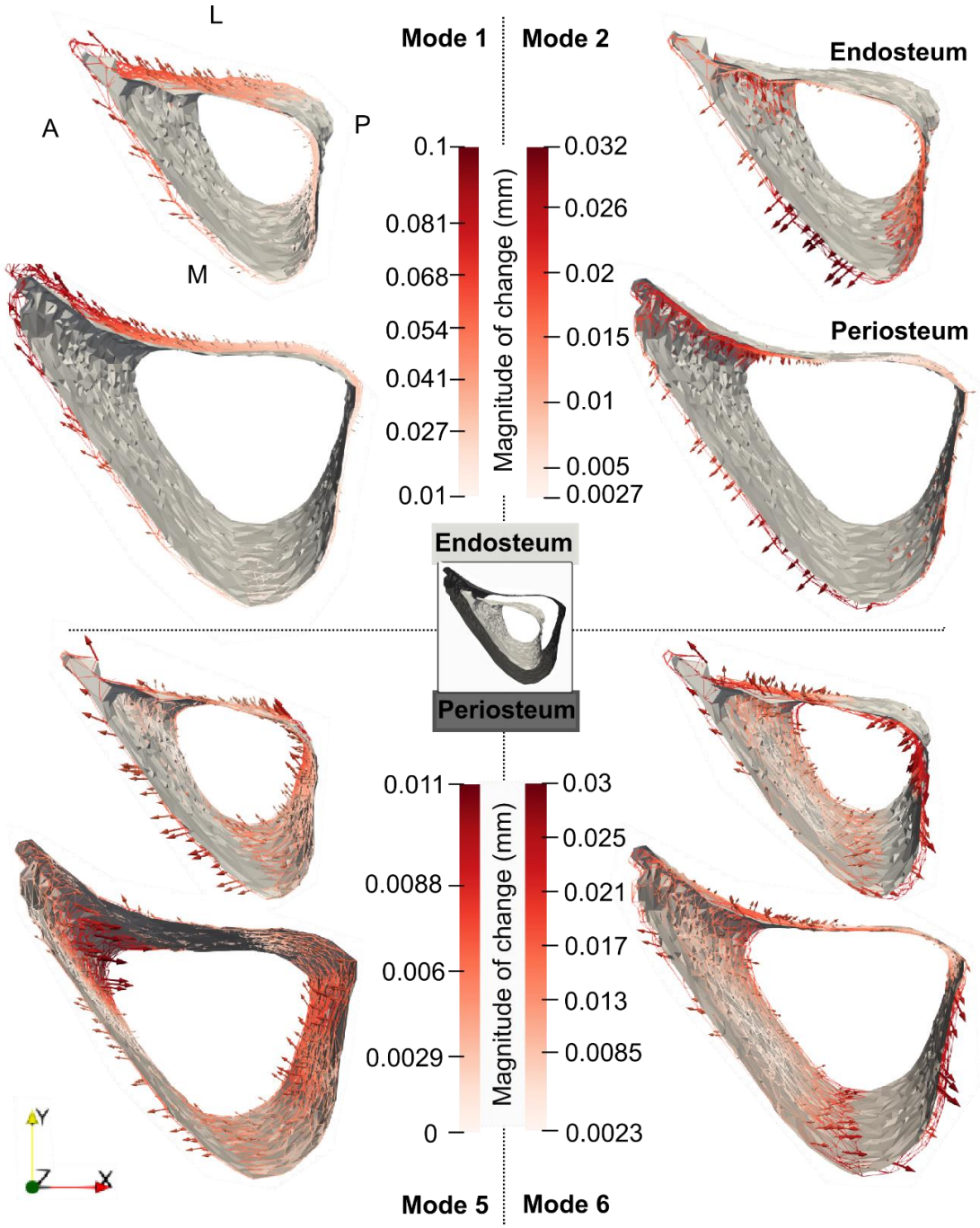


Figure 3.12: Median changes in endosteum and periosteum shapes due to Modes 1, 2, 5 and 6 between weeks 18 and 24 in the "OVX+ML" group. The directions of change at different locations are denoted by the arrows, and a redder arrow indicates a relatively larger change. The median bone profile at week 18 is shown as a solid grey surface, whilst the median profile at week 24 is given as a coloured wireframe. The profiles share the same anatomical orientation indicated by A (anterior), L (lateral), P (posterior) and M (medial).

3.4 Discussion

The main objectives of this study are to demonstrate that: (a) the variations in mouse bone geometry obtained by the novel framework are robust to uncertainty sources contained therein; (b) to demonstrate that this framework enables an assessment of 3D geometric variations induced by treatment.

3.4.1 Robustness of framework including mapping bone images

The robustness of the framework was assessed through case studies focusing on its distinct steps. These analyses were conducted for all mice in the "OVX+ML" group at the age of 18 weeks. At this age, the groups "OVX+ML" and "OVX" were assumed to be biologically indistinguishable as the mechanical loading interventions start only when mice are 19 weeks old. This assertion is further supported by a *posteriori* finding that no statistically significant differences ($p<0.05$, Wilcoxon signed rank test) between these groups were found when comparing the scores of any PCA mode.

All images were rigidly registered to a reference structure at the starting time point (Viceconti & Dall'Ara, 2019), to eliminate spatial errors (Campbell et al., 2014). Since the comparison of the longitudinal images is very sensitive to the registration of each set of microCT scans, a validated registration procedure with errors of less than 3.5% was used (Lu et al., 2016), demonstrating reproducibility in similar mouse tibia studies (Oliviero et al. 2022). Finally, as this study focused on one segment within the entire tibia volume, further local registration at the tibia midshaft was successfully implemented aided by its simple geometry (compared with more heterogeneous trabecular regions).

The first evaluation study (Supplementary Material 1) demonstrated that cortical pores and trabeculae occupy a very small portion of the total bone section volume, and that excluding these features had negligible influence on the mechanical strain distribution. As such, this simplification was considered acceptable. It is not a limitation of the framework because the state-of-the-art method of standard morphometric analysis at midshaft also excludes such features. The second evaluation (Supplementary Material 2) indicated that bone surface topology was preserved across various levels of discretization, with errors converging quickly and remaining small compared to voxel size.

The third evaluation study assessed the accuracy and precision of the deformable registration algorithm in the prediction of the several surface displacements (uniform, non-uniform, global and local). Mean errors in mapping such surfaces were always below 1 voxel. SD of errors were almost always lower than 1 voxel, apart from local and high magnitude deformations which seemed to challenge the algorithm. In those cases ('Study #5: Anterior, Half, Tc=1.2'), a linear relationship between errors and NS was revealed and errors associated with precision were found higher than 1 voxel for NS>35 voxels. Since the distribution of errors in the bone surface was skewed resulting in high errors in the most deformed bone areas, NS was chosen to be 5 voxels to ensure that all important local differences within the examined population of shapes are captured. For NS=5, errors remained smaller than 1 voxel in the presence of existing and simulated image noise and when binarization and geometry correction steps were also applied. Taken together, this evaluation study suggests that the mapping of the reference bone surface to "anatomically similar" locations on the 21 mouse tibial surfaces, achieved using deformable registration with NS=5 voxels, contains errors associated with accuracy and precision smaller than 1 voxel.

Past evaluation studies have found comparable error magnitudes and similar relationships between error magnitudes and NS of deformable registration. For example, Dall'Ara et al., (2014) showed that when predicting uniform displacements of 2 voxels using NS=5 voxels, the mean and standard deviation of error magnitudes were of the order of 10^{-4} voxels and 10^{-2} voxels, respectively. Dall'Ara et al. (2017) simulated uniform virtual displacement fields of similar magnitude in murine tibia samples and found good performance in precision. In the present evaluation study #1, where uniform or non-uniform virtual displacements (i.e. homogeneous but anisotropic) of integer magnitudes (up to 6 voxels) was used, mean and standard deviation of error magnitudes were of the order of 10^{-5} voxels and 10^{-4} voxels respectively, and relatively independent of NS, and similar to those reported in the two studies mentioned above. However, whilst the cortical midshaft section of murine tibia was analysed in the present study, Dall'Ara et al. (2014) focused on localized cubes extracted from trabecular or cortical regions of bovine femur. The lower error magnitudes found in the present study are expected because the topology of the cortical region of the murine midshaft is much simpler than of the trabecular bovine

femur regions.

A non-uniform translation equal to a non-integer number of voxels results in significantly larger mean errors, as also demonstrated previously (Dall'Ara et al., 2014). It is attributable to interpolation, but with less impact on grayscale images. This is expected because interpolation of the displacement on the bone surface is more gradual when applied to grayscale images. Local deformations led to non-uniform errors on the bone surface, with error magnitudes being larger in areas of larger local deformation (Figure 3.7). Therefore, the dependence of the errors on NS is modulated by the deformed features and the magnitude of deformation. For example, a flatter shape of the posterior part in 'Study #5: Posterior, Half, Tc=0.95' resulted in a relatively low dependence of error magnitudes on NS, similarly with applying local deformation in a small fraction of the image. In contrast, the sharp feature of the anterior part explains the linear increase of errors with NS in the case of 'Study #5: Anterior, Half, Tc=1.2'. Image noise slightly increased registration errors. In this study, artificial noise was added to already noisy images, while Dall'Ara et al. (2014) used noise from repeated scans. Despite the difference in noise sources, error magnitudes were similar to those reported previously. The strong dependence of registration errors on NS for complex simulated displacement fields suggests the need to reassess this dependence for new applications. Such reassessment should ideally design tests representative of the length scale, image resolution, alignment, features and noise that are present in such new applications.

Simulated displacement fields, such as those described in the present evaluation study, have the advantage that different sources of complexity can be separately analysed. However, not all past PCA models have analysed registrations errors in similar ways, which makes it challenging to directly compare the magnitude of errors reported here with earlier work. Unlike the present study, where images were mapped using elastic registration, Bryan et al. (2010) registered the surface meshes. The defined registration error as the distance between the registered and target surface meshes were reported with mean and maximum errors of 0.60 mm and 3 mm, respectively. Their error magnitudes cannot be directly compared (i.e. in dimensional units) to the error magnitudes reported in the present study, due to differences in bone sites, scales and imaging modalities. As the image resolution in the study of

Bryan et al. (2010) was 0.78 mm×0.78 mm×2 mm, the errors can be inferred to be in the range of 0.3–3.8 voxels. A similar range of mean registration errors (0.42–3 voxels) was reported by Brown et al. (2014) and Brown et al. (2017) who imaged hind limbs of female C57Bl/6 (similar bone size as the present study) with somewhat lower resolution (14 µm). Registration errors in the present study are slightly smaller than those reported in the above studies. This could be due to differences in elastic registration approach between current and past work and errors in initial positioning / alignment that existed in these past studies, but these effects cannot be further separated.

3.4.2 Decomposition of variations and treatment effects categorization

Among all groups, the "OVX+ML" week 24 group contributed most to the total variation. This indicates that it is furthest from the mean shape of the whole dataset and suggests a strong effect of treatment. The relatively small variation within the "OVX" week 24 group indicates that the effect of untreated ovariectomy from the period from 18 to 24 weeks cannot be reliably distinguished from the natural variation between mice in this study. A posteriori analysis on the standard morphometric parameters of the same mice suggest that the variation in cortical area (both week 18: 19%; "OVX" week 24: 13%, "OVX+ML" week 24: 68%) and in cortical thickness (both week 18: 18%; "OVX" week 24: 11%, "OVX+ML" week 24: 71%) are distributed similarly (Roberts et al., 2019; Roberts et al., 2020). However, while the novel framework presented here can be used to assess the variation in the full 3D bone geometry, standard morphometric analysis can be used to assess only the part of variation in geometry that is captured by the morphometric parameters.

The geometry variation across the 22 mouse image specimens was compact, with 90% of it being explained by the variation of only six principal components (i.e. shape modes). This is interesting because several different factors, including natural variability between individuals, growth, disease, treatment and artifacts of image processing were present in the dataset. The full 3D assessment enabled by PCA allows a concise description of separate findings of previous studies. Specifically, the PCA mode shapes related with specific score changes over time revealed 3D geometric changes describes as a vector field on the 3D bone surfaces. Mode 1 describes simultaneous formation and deletion localized

at the anterior crest, which agrees with the previous findings of Cheong et al. (2020). At this site, the mode vectors on the formation surface are longer than those on the deletion resorption surface indicating higher magnitude of change in periosteum than endosteum (0.1 mm vs 0.054 mm), and this agrees with previous findings of Birkhold et al (2017). This localized cortical thickening of 0.046 mm also agrees with the averaged cortical thickening of 0.064 mm of similar bone region reported by Roberts et al. (2020). Roberts et al. (2020) attributed this change to treatment, and this agrees with the present finding that Mode 1 is characterized as a pure treatment effect (by the statistical tests of the scores).

Mode 2 describes periosteal formation and endosteal deletion of the medial aspect which could potentially explain the increased cortical area in the study of Roberts et al., (2020). The localized changes in the anterior–medial aspect agree with the predictions of Razi et al. (2015) and Javaheri et al. (2020) that indicate higher strain magnitudes at these locations. Both Mode 1 and 6 also describe localized thickening of the posterior-lateral edge, consistent with other loading murine studies which highlight the dependency of the bone remodelling to the strain distribution (Cheong et al., 2021; Sugiyama et al., 2008; Rooney et al., 2023). Overall, posterior and anterior bone response as shown by Modes 1 and 6 are also in agreement with the local thickness changes as measured cross sectionally in the midshaft slice of tibia bone in similar loading models (Miller et al., 2021). The periosteal deletion of the distal anterior crest as described by Mode 5 has not been previously reported elsewhere but this effect is much smaller compared to the opposite effect of Mode 1. Overall, the current framework provides a more specific, precise (local PCA mode vectors) and quantitative assessment (PCA scores) than what could be obtained by previous approaches.

The used sample size (N=6 for each examined cohort and age) was relatively small but similar to other longitudinal studies that quantify bone changes over time and space (Li et al., 2019; Lu et al., 2017; Roberts et al., 2019). The overall median reconstruction errors in the “OVX+ML” groups in weeks 18 and 24 were comparable to the standard deviation of cortical thickness (0.5 voxels and 1.7 voxels respectively) as reported previously in Table 2 of Roberts et al., (2019, 2020) using the same images. This implies that the difference in cortical thickness between an original and a reconstructed geometry

(using all PCA modes) of a mouse is of the same order of magnitude as the difference in cortical thickness between two randomly selected mice in the population. This suggests that a mouse not belonging to the examined population cannot be reconstructed using the discovered PCA modes sufficiently accurately. This highlights that inferences drawn based on the somewhat small dataset must be extrapolated with caution to a larger population of mice undergoing the same course of disease/treatment.

Notwithstanding this limitation, some assertions regarding modes corresponding to treatment effects can still be made reliably. The median reconstruction error across all "OVX+ML" week 24 samples is the highest (~3 voxels) at the anterior crest. The median change due to Mode 1 is also highest (~9 voxels) at a similar location. This suggests that the small dataset used in the present study does not limit the assertion that Mode 1 is an effect of mechanical loading treatment. A similar assertion can be made for Mode 6, as the median change corresponding to this mode is the highest at the endosteum (~3 voxels) which is nearly twice the highest median reconstruction error (found at a similar anatomical position). However, such assertions cannot be made for either Mode 2 or Mode 5, as the maximum median change (3 and 1 voxels, respectively) is not distinguishable from the magnitude of median reconstruction errors in the medial aspect and distal anterior crest, respectively.

The higher confidence in interpreting Modes 1 and 6 as treatment effects is supported by their large effect sizes. The contrast in fractions of total variation explained between Modes 1 and 6 highlights that Mode 1 describes a global change (a large area around the anterior crest) whereas Mode 6 describes a local change (a small area on the posterior-lateral aspect). Modes 2 and 5 explain larger fractions of total variance than Mode 6 but approximately a quarter of the effect size of Mode 6. This highlights that Modes 2 and 5 do not capture treatment effects as reliably as Mode 6, even if these modes separately explain larger fractions of total variance than Mode 6.

The orthogonality of PCA modes guarantees that the treatment effects described by Modes 1, 2, 5 and 6 are uncorrelated to each other. In contrast, the standardized set of morphometric parameters, such as cortical thickness and area, are dependent on each other. This makes it challenging to separate the

fractions of total variation explained by each parameter. In the present study, the treatment-related modes described most of the total variance, indicating the role treatment plays in modifying bone shape. Orthogonality of modes also allows attributing Modes 3, 4, 7–21 to sources of variation other than treatment, such as natural variability including random errors and noise but not growth and disease which have systematic temporal effect. The lack of a temporal effect was supported by our analysis of the original data in Roberts et al., (2019). This analysis showed that neither cortical thickness or cortical area were statistically significantly different (Wilcoxon signed-rank test, $p>0.05$) between 18 and 24 weeks in the "OVX" group. This is expected because the impact of ovariectomy in murine bone models is prominent in the first couple of weeks post-surgery but stabilizes after that period (Roberts et al., 2019).

A major advancement due to the proposed framework is that the experimental design (and corresponding imaging data) drives the discovery of geometric features that automatically separate into independent effects due to treatment, unpacked from disease, growth, random variations and combinations of these, whilst capturing the full 3D variation in the data. This contrasts with the state-of-the-art approach of fixing the morphometric parameters to characterize bone geometries. The study in this chapter successfully demonstrated that there are important local treatment effects that the scalar parameters cannot describe. This contrasts with prevailing approaches that discard a part of total shape variation and moderately decompose the effects of treatment, disease, growth and/or random variations along these parameters, with no guarantee that the decomposition is mutually independent.

3.4.3 Limitations

The 90% coarsening of the surface meshes showed that important geometrical features across the reference bone section are preserved (Supplementary Material 2). Similar levels of coarsening were also suggested previously in the study of Brown et al. (2017). However, a concern is whether the coarse reference surface mesh is representative of all bone samples across the tibia length. Specifically, if there are important features that are present in different regions of bone surfaces for other bones, then coarsening the reference surface mesh cannot guarantee that these features are described by the mapped

reference meshes following deformable registration. In the present study, the size of the cropped section of the tibia was proportional to its length, which ensured that these sections corresponded to similar anatomical regions and thus reduced the possibility of including disparate anatomical landmarks.

This framework is designed to investigate bone variations within an examined population. Although this is based on a PCA model, the latter should be used with care when attempting to approximate new geometries not observed in this study. It is also important to note that individual PCA mode shapes and scores, even those with highest explanatory power or effect size (Modes 1 and 6), cannot predict the full effect of treatment for individual mice, and should not be used in a predictive modelling setting. Therefore, the number of samples used in this study are insufficient for the purpose of building *in silico* physiologic or pathologic cohorts. However, this study can contribute to an ongoing process of data sharing within the scientific community with the long-term aim of reducing the use of animals in bone research (Vicenconti and Dall'Ara, 2019).

This study assumed negligible densitometric variations within the bone samples. Roberts et al. (2020) found 1% change in the tissue mineral density (TMD) in a similar bone segment driven by a double course of mechanical loading. Oliviero et al., (2021) analysed homogeneous (Young's modulus E=14.8 GPa) and heterogenous (E derived from local TMD) FE models of mouse tibia at different ages and cohorts (healthy, diseased and treated) but did not find any significant differences in bone mechanics. In future applications where TMD is expected to change significantly (above levels reported in Roberts et al. (2020)) or be more heterogeneous compared to (Oliviero et al., 2021), further development of the present framework is necessary to describe local TMD variations and separate it into effects of treatment, disease, growth, natural variability and other relevant sources.

This application only focused on the midshaft section, and further development is needed to analyse 3D shape variation in the entire tibia. It is expected that removing the fibula (Cheong, et al., 2020) will be necessary to preserve the topological similarity. The high dimensionality of the problem and the computational demand of each step of the framework should be carefully considered. Here, controlling the surface mesh density using coarsening could be a crucial step. Sophisticated job submission

techniques may be needed for efficient compute time and data management. The application of conventional PCA in the framework presented here inherently assumes that modes can be linearly combined. This assumption might occlude important effects, and the use of several alternative models, such as Probabilistic PCA (Kim & Lee, 2003) and Gaussian processes (Lawrence & Hyvärinen, 2005) should be explored in the future.

The presented PCA-based framework is limited to analysis of topologically equivalent shapes. As such, only the variations in cortical bone regions can be examined and the framework excludes the possibility of analysing trabecular regions. The image processing protocol for removing cortical pores and trabeculae was based on previous evidence of bone remodelling for this specific bone system. However, if the framework is applied to a different volume, site of bone or different bone, evidence of bone remodelling and influence on bone mechanics should be reviewed in the new context.

The selected mouse genetic strain was considered skeletally mature at week 14 of their age and an appropriate experimental model to quantify bone variations in osteoporosis and treatment cohorts. However, bone aging can affect bone response to interventions. Ageing-related effects were not distinguished in this analysis. Particularly, the PCA database did not include any healthy specimens, but it considered ovariectomized mice as the control group. Additionally, some sources of uncertainties within the examined population that could not be eliminated are the exact date of birth related to their age (± some days), genetic phenotypes, success of ovariectomy (complete removal of ovaries with/without additional soft tissue) and recovery.

3.5 Conclusions

The proposed new framework based on longitudinal microCT imaging, image processing and PCA on discretized bone shapes showed potential of improving the preclinical treatment investigation in murine bone models. The framework was demonstrated to accurately describe bone shapes up to ~1 voxel accuracy in the presence of several error sources in the processing pipeline. Application of PCA on discretized tibial midshaft shapes taken from a population of diseased and treated mice identified for the first time six mutually independent geometrical features that explained a significant fraction of the

total variation in the 3D imaging data. Four of these geometric features (modes) were found to be purely an effect of the mechanical loading treatment and described changes over the course of treatment at the anterior crest, medial aspect, posterior area and some specific localized features. Due to the small dataset size, only two of these features could be reliably asserted as being treatment effects. Nevertheless, these features offer a compact description of several changes found in previous studies, and contain new information not discovered until now. The imaging data used here enabled the demonstration of the various methodological aspects of the developed framework, which was the primary focus of this study. However, the application of the framework is not limited to the experimental set-up from where the images are sourced. It has the potential to be used as a more precise strategy for investigating the effect of different treatment strategies on bone structure.

3.6 References

Akhter, M. P., & Recker, R. R. (2021). High resolution imaging in bone tissue research-review. *Bone, 143*, 115620. https://doi.org/10.1016/j.bone.2020.115620

Aldieri, A., Terzini, M., Audenino, A. L., Bignardi, C., & Morbiducci, U. (2020). Combining shape and intensity dxa-based statistical approaches for osteoporotic HIP fracture risk assessment. *Computers in Biology and Medicine*, *127*, 104093. https://doi.org/10.1016/j.compbiomed.2020.104093

Barber, D. C., & Hose, D. R. (2005). Automatic segmentation of medical images using image registration: diagnostic and simulation applications. *Journal of Medical Engineering & Technology*, *29*(2), 53–63. https://doi.org/10.1080/03091900412331289889

Barratt, D. C., Chan, C. S. K., Edwards, P. J., Penney, G. P., Slomczykowski, M., Carter, T. J., & Hawkes, D. J. (2008). Instantiation and registration of statistical shape models of the femur and pelvis using 3D ultrasound imaging. *Medical Image Analysis*, *12*(3), 358–374. https://doi.org/10.1016/j.media.2007.12.006

Birkhold, A.I., Razi, H., Duda, G.N., Checa, S. & Willie, B.M. (2017). Tomography-based quantification of regional differences in cortical bone surface remodeling and mechano-response. *Calcified tissue international, 100*, 255-270. https://doi.org/10.1007/s00223-016-

0217-4

Birkhold, A. I., Razi, H., Weinkamer, R., Duda, G. N., Checa, S., & Willie, B. M. (2015). Monitoring in vivo (re) modeling: a computational approach using 4D microCT data to quantify bone surface movements. *Bone*, *75*, 210–221. https://doi.org/10.1016/j.bone.2015.02.027

Bouxsein, M. L., Boyd, S. K., Christiansen, B. A., Guldberg, R. E., Jepsen, K. J., & Müller, R. (2010). Guidelines for assessment of bone microstructure in rodents using micro--computed tomography. *Journal of Bone and Mineral Research*, *25*(7), 1468–1486. https://doi.org/10.1002/jbmr.141

Bouxsein, M. L., Myers, K. S., Shultz, K. L., Donahue, L. R., Rosen, C. J., & Beamer, W. G. (2005). Ovariectomy-induced bone loss varies among inbred strains of mice. *Journal of Bone and Mineral Research*, *20*(7), 1085–1092. https://doi.org/10.1359/JBMR.050307

Brown, J. M., Naylor, A., Buckley, C., Filer, A., & Claridge, E. (2014). 3D articulated registration of the mouse hind limb for bone morphometric analysis in rheumatoid arthritis. *Biomedical Image Registration: 6th International Workshop, WBIR 2014, London, UK, July 7-8, 2014. Proceedings 6*, 41–50. https://doi.org/10.1007/978-3-319-08554-8_5

Brown, J. M., Ross, E., Desanti, G., Saghir, A., Clark, A., Buckley, C., Filer, A., Naylor, A., & Claridge, E. (2017). Detection and characterisation of bone destruction in murine rheumatoid arthritis using statistical shape models. *Medical Image Analysis*, *40*, 30–43. https://doi.org/10.1016/j.media.2017.05.006

Bryan, R., Surya Mohan, P., Hopkins, A., Galloway, F., Taylor, M., & Nair, P. B. (2010). Statistical modelling of the whole human femur incorporating geometric and material properties. *Medical Engineering & Physics*, *32*(1), 57–65. https://doi.org/10.1016/j.medengphy.2009.10.008

Campbell, G. M., Tiwari, S., Grundmann, F., Purcz, N., Schem, C., & Glüer, C.-C. (2014). Three-dimensional image registration improves the long-term precision of in vivo micro-computed tomographic measurements in anabolic and catabolic mouse models. *Calcified Tissue International*, *94*(3), 282–292. https://doi.org/10.1007/s00223-013-9809-4

Castro-Mateos, I., Pozo, J. M., Cootes, T. F., Wilkinson, J. M., Eastell, R., & Frangi, A. F. (2014). Statistical shape and appearance models in osteoporosis. *Current Osteoporosis Reports*, *12*,

163–173. https://doi.org/10.1007/s11914-014-0206-3

Chan, E. F., Harjanto, R., Asahara, H., Inoue, N., Masuda, K., Bugbee, W. D., Firestein, G. S., Hosalkar, H. S., Lotz, M. K., & Sah, R. L. (2012). Structural and functional maturation of distal femoral cartilage and bone during postnatal development and growth in humans and mice. *The Orthopedic Clinics of North America*, *43*(2), 173–185. https://doi.org/10.1016/j.ocl.2012.01.005

Cheong, V. S., Campos Marin, A., Lacroix, D., & Dall'Ara, E. (2020). A novel algorithm to predict bone changes in the mouse tibia properties under physiological conditions. *Biomechanics and Modeling in Mechanobiology*, *19*, 985–1001. https://doi.org/10.1007/s10237-019-01266-7

Cheong, V. S., Roberts, B. C., Kadirkamanathan, V., & Dall'Ara, E. (2020). Bone remodelling in the mouse tibia is spatio-temporally modulated by oestrogen deficiency and external mechanical loading: A combined in vivo/in silico study. *Acta Biomaterialia*, *116*, 302–317. https://doi.org/10.1016/j.actbio.2020.09.011

Cheong, V. S., Roberts, B. C., Kadirkamanathan, V., & Dall'Ara, E. (2021). Positive interactions of mechanical loading and PTH treatments on spatio-temporal bone remodelling. *Acta Biomaterialia*, *136*, 291–305. https://doi.org/10.1016/j.actbio.2021.09.035

Dall'Ara, E., Barber, D., & Viceconti, M. (2014). About the inevitable compromise between spatial resolution and accuracy of strain measurement for bone tissue: A 3D zero-strain study. *Journal of Biomechanics*, *47*(12), 2956–2963. https://doi.org/10.1016/j.jbiomech.2014.07.019

Dall'Ara, E., Boudiffa, M., Taylor, C., Schug, D., Fiegle, E., Kennerley, A. J., Damianou, C., Tozer, G. M., Kiessling, F., & Müller, R. (2016). Longitudinal imaging of the ageing mouse. *Mechanisms of Ageing and Development*, *160*, 93–116. https://doi.org/10.1016/j.mad.2016.08.001

Dall'Ara, E., Peña-Fernández, M., Palanca, M., Giorgi, M., Cristofolini, L., & Tozzi, G. (2017). Precision of digital volume correlation approaches for strain analysis in bone imaged with micro-computed tomography at different dimensional levels. *Frontiers in Materials*, *4*, 31. https://doi.org/10.3389/fmats.2017.00031

Davies, R., Twining, C., & Taylor, C. (2008). *Statistical models of shape: Optimisation and evaluation*. Springer Science \& Business Media. doi:10.1007/978-1-84800-138-1

Grassi, L., Väänänen, S. P., & Isaksson, H. (2021). Statistical shape and appearance models: development towards improved osteoporosis care. *Current Osteoporosis Reports*, *19*, 676–687. https://doi.org/10.1007/s11914-021-00711-w

Hoshino, Y., Takechi, M., Moazen, M., Steacy, M., Koyabu, D., Furutera, T., Ninomiya, Y., Nuri, T., Pauws, E., & Iseki, S. (2023). Synchondrosis fusion contributes to the progression of postnatal craniofacial dysmorphology in syndromic craniosynostosis. *Journal of Anatomy*, *242*(3), 387–401. https://doi.org/10.1111/joa.13790

Iida-Klein, A., Lu, S. S., Cosman, F., Lindsay, R., & Dempster, D. W. (2007). Effects of cyclic vs. daily treatment with human parathyroid hormone (1–34) on murine bone structure and cellular activity. *Bone*, *40*(2), 391–398. https://doi.org/10.1016/j.bone.2006.09.010

Javaheri, B., Razi, H., Gohin, S., Wylie, S., Chang, Y.-M., Salmon, P., Lee, P. D., & Pitsillides, A. A. (2020). Lasting organ-level bone mechanoadaptation is unrelated to local strain. *Science Advances*, *6*(10), eaax8301. doi: 10.1126/sciadv.aax8301

Jepsen, K. J., Silva, M. J., Vashishth, D., Guo, X. E., & Van Der Meulen, M. C. H. (2015). Establishing biomechanical mechanisms in mouse models: practical guidelines for systematically evaluating phenotypic changes in the diaphyses of long bones. *Journal of Bone and Mineral Research*, *30*(6), 951–966. https://doi.org/10.1002/jbmr.2539

Jilka, R. L. (2013). The relevance of mouse models for investigating age-related bone loss in humans. *The Journals of Gerontology: Series A*, *68*(10), 1209–1217. https://doi.org/10.1093/gerona/glt046

Killian, M. L., Locke, R. C., James, M. G., Atkins, P. R., Anderson, A. E., & Clohisy, J. C. (2019). Novel model for the induction of postnatal murine hip deformity. *Journal of Orthopaedic Research*, *37*(1), 151–160.

Kim, D., & Lee, I.-B. (2003). Process monitoring based on probabilistic PCA. *Chemometrics and Intelligent Laboratory Systems*, *67*(2), 109–123. https://doi.org/10.1016/S0169-7439(03)00063-7

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lawrence, N., & Hyvärinen, A. (2005). Probabilistic non-linear principal component analysis with Gaussian process latent variable models. *Journal of Machine Learning Research*, *6*(11), 1783–1816.

Levchuk, A., Zwahlen, A., Weigt, C., Lambers, F. M., Badilatti, S. D., Schulte, F. A., Kuhn, G., & Müller, R. (2014). The Clinical Biomechanics Award 2012—presented by the European Society of Biomechanics: Large scale simulations of trabecular bone adaptation to loading and treatment. *Clinical Biomechanics*, *29*(4), 355–362. doi: 10.1016/j.clinbiomech.2013.12.019

Li, Z., Betts, D., Kuhn, G., Schirmer, M., Müller, R., & Ruffoni, D. (2019). Mechanical regulation of bone formation and resorption around implants in a mouse model of osteopenic bone. *Journal of the Royal Society, Interface*, *16*(152), 20180667. https://doi.org/10.1098/rsif.2018.0667

Lu, Y., Boudiffa, M., Dall'Ara, E., Bellantuono, I., & Viceconti, M. (2016). Development of a protocol to quantify local bone adaptation over space and time: Quantification of reproducibility. *Journal of Biomechanics*, *49*(10), 2095–2099. https://doi.org/10.1016/j.jbiomech.2016.05.022

Lu, Y., Boudiffa, M., Dall'Ara, E., Liu, Y., Bellantuono, I., & Viceconti, M. (2017). Longitudinal effects of Parathyroid Hormone treatment on morphological, densitometric and mechanical properties of mouse tibia. *Journal of the Mechanical Behavior of Biomedical Materials*, *75*, 244–251. https://doi.org/10.1016/j.jmbbm.2017.07.034

Miller, C. J., Trichilo, S., Pickering, E., Martelli, S., Delisser, P., Meakin, L. B., & Pivonka, P. (2021). Cortical thickness adaptive response to mechanical loading depends on periosteal position and varies linearly with loading magnitude. *Frontiers in Bioengineering and Biotechnology*, *9*, 671606. https://doi.org/10.3389/fbioe.2021.671606

Nepal, A. K., van Essen, H. W., de Jongh, R. T., van Schoor, N. M., Otten, R. H. J., Vanderschueren, D., Lips, P., & Bravenboer, N. (2023). Methodological aspects of in vivo axial loading in rodents: a systematic review. *Journal of Musculoskeletal \& Neuronal Interactions*, *23*(2), 236.

Oliviero, S., Roberts, M., Owen, R., Reilly, G. C., Bellantuono, I., & Dall'Ara, E. (2021). Non-invasive prediction of the mouse tibia mechanical properties from microCT images: comparison between different finite element models. *Biomechanics and Modeling in Mechanobiology*, *20*, 941–955. https://doi.org/10.1007/s10237-021-01422-y

Oliviero, S., Cheong, V. S., Roberts, B. C., Orozco Diaz, C. A., Griffiths, W., Bellantuono, I., & Dall'Ara, E. (2022). Reproducibility of Densitometric and Biomechanical Assessment of the Mouse Tibia From In Vivo Micro-CT Images. *Frontiers in Endocrinology*, *13*, 915938. https://doi.org/10.3389/fendo.2022.915938

Oliviero, S., Giorgi, M., Laud, P. J., & Dall'Ara, E. (2019). Effect of repeated in vivo microCT imaging on the properties of the mouse tibia. *PloS One*, *14*(11), e0225127. https://doi.org/10.1371/journal.pone.0225127

Rajamani, K. T., Styner, M. A., Talib, H., Zheng, G., Nolte, L. P., & Ballester, M. A. G. (2007). Statistical deformable bone models for robust 3D surface extrapolation from sparse data. *Medical Image Analysis*, *11*(2), 99–109. https://doi.org/10.1016/j.media.2006.05.001

Razi, H., Birkhold, A. I., Zaslansky, P., Weinkamer, R., Duda, G. N., Willie, B. M., & Checa, S. (2015). Skeletal maturity leads to a reduction in the strain magnitudes induced within the bone: A murine tibia study. *Acta Biomaterialia*, *13*, 301–310. https://doi.org/10.1016/j.actbio.2014.11.021

Roberts, B. C., Arredondo Carrera, H. M., Zanjani-pour, S., Boudiffa, M., Wang, N., Gartland, A., & Dall'Ara, E. (2020). PTH (1–34) treatment and/or mechanical loading have different osteogenic effects on the trabecular and cortical bone in the ovariectomized C57BL/6 mouse. *Scientific Reports*, *10*(1), 8889. https://doi.org/10.1038/s41598-020-65921-1

Roberts, B. C., Giorgi, M., Oliviero, S., Wang, N., Boudiffa, M., & Dall'Ara, E. (2019). The longitudinal effects of ovariectomy on the morphometric, densitometric and mechanical properties in the murine tibia: A comparison between two mouse strains. *Bone*, *127*, 260–270. https://doi.org/10.1016/j.bone.2019.06.024

Rodan, G. A., & Martin, T. J. (2000). Therapeutic Approaches to Bone Diseases. *Science*, *289*(5484), 1508–1514. https://doi.org/10.1126/science.289.5484.1508

Rooney, A. M., McNeill, T. J., Ross, F. P., Bostrom, M. P. G., & van der Meulen, M. C. H. (2023). PTH treatment increases cortical bone mass more in response to compression than tension in mice. *Journal of Bone and Mineral Research*, *38*(1), 59–69. doi: 10.1002/jbmr.4728

Sugiyama, T., Saxon, L. K., Zaman, G., Moustafa, A., Sunters, A., Price, J. S., & Lanyon, L. E.

(2008). Mechanical loading enhances the anabolic effects of intermittent parathyroid hormone (1–34) on trabecular and cortical bone in mice. *Bone*, *43*(2), 238–248. https://doi.org/10.1016/j.bone.2008.04.012

van't Hof, R. J., & Dall'Ara, E. (2019). Analysis of Bone Architecture in Rodents Using Micro-Computed Tomography. *Methods in Molecular Biology (Clifton, N.J.)*, *1914*, 507–531. https://doi.org/10.1007/978-1-4939-8997-3_28

Viceconti, M., & Dall'Ara, E. (2019). From bed to bench: How in silico medicine can help ageing research. *Mechanisms of Ageing and Development*, *177*, 103–108. https://doi.org/10.1016/j.mad.2018.07.001

Zhang, Y., Dall'Ara, E., Viceconti, M., & Kadirkamanathan, V. (2019). A new method to monitor bone geometry changes at different spatial scales in the longitudinal in vivo μCT studies of mice bones. *PLOS ONE*, *14*(7), 1–21. https://doi.org/10.1371/journal.pone.0219404

Zhou, H., Iida-Klein, A., Lu, S. S., Ducayen-Knowles, M., Levine, L. R., Dempster, D. W., & Lindsay, R. (2003). Anabolic action of parathyroid hormone on cortical and cancellous bone differs between axial and appendicular skeletal sites in mice. *Bone*, *32*(5), 513–520. https://doi.org/10.1016/S8756-3282(03)00057-7

3.7 Supplementary material

Supplementary Material 1: Mechanical importance of cortical pores and trabeculae in the midshaft section

Methods

The second step of the proposed framework, i.e. binarization and geometry correction (Section 2.4), ensures the topological equivalence of the shape observations to an annular cylinder. This step locally "corrected" the geometry, such as by removing trabeculae and cortical pores. The aim of this study is to quantify the influence of these local geometry corrections on the mechanical response. To do so, this study introduced five different metrices, referred to as M1, M2, M3, M4, and M5, to facilitate the comparison of geometry and mechanics between the original and processed bone section.

Two micro-Finite Element (microFE) models of each midshaft tibia section were created by directly converting all bone voxels from the 3D images (original and processed) into linear hexahedral elements (Cheong et al., 2020). The ratio of integration points (IPs), present in either model but not both, to the total number of the IPs in the original model was computed as a measure (denoted M1) of difference in overall volume due to geometry correction. For both models, homogeneous isotropic material properties were assumed, $E = 14.8$ GPa, $\nu = 0.3$ (Oliviero et al., 2018). All nodes on the top surface were constrained to have the same displacement as a control node to which a 1 N static load was applied in the axial direction. All translation degrees of freedom of all nodes on the bottom surface were fixed. The above microFE modelling pipeline has been previously validated against experimental measurement of stiffness (errors of 14% ± 11%) and failure load (errors of 9% ± 9%) (Oliviero et al., 2018). MicroFE analyses were performed using Abaqus 2018. At the IPs common to both models, the absolute difference in minimum principal strain was computed, normalized by the maximum absolute minimum principal strain in the original model and denoted as measure M2. This measure indicated the overall difference in mechanical strain between the models. For both models, "highly compressed" IPs were identified as those with the 10% smallest minimum principal strain (Oliviero et al., 2021). Of these, only those IPs that are common to both models were counted, the ratio of this to the total number of the highly compressed IPs in the original model was determined and denoted as measure M3. This

measure indicates the degree to which the critically compressed region remains unchanged after geometry correction. For the highly compressed IPs common to both models, the absolute difference in minimum principal strain was determined, normalized by the maximum absolute minimum principal strain in the original model and denoted as measure M4. This measure describes the effect of feature removal on the failure load prediction. The area around the corrected features (trabeculae/pores) were identified as follows. Subtracting the 3D processed image from the original yielded the volumes of the removed trabeculae. Subtracting the original image from the processed images yields the volumes of cortical pores (filled). Image dilation by using a disk of 3 voxels was used to expand the volumes of these features, so that the analysis includes the surrounding area around the feature. The connectivity function was used to separate the features within the same bone geometry. The minimum and maximum *x*, *y*, *z* coordinates were determined for each feature separately. This defined the 3D bounding boxes that enclosed areas around each corrected feature. The number of the highly compressed IPs falling inside the bounding boxes enclosing filled cortical pores were counted. The ratio of this to the total number of highly compressed IPs (measure M5-cortical pores) indicates the portion of the fracture locations that is close to this specific type of corrected feature. M5-trabeculae was obtained similarly by considering bounding boxes enclosing deleted trabeculae. In the case of multiple features, the fraction with the highest value is retained.

Results

The values of each measure, with mean and the standard deviation taken over six mouse samples of "OVX+ML" group at 18 weeks, are summarized in the table below. For measures M2 and M4, the average was taken over all relevant integration points and all mice. For measure M5, the mean value comes from the one specimen that exhibited a trabecula and as such no standard deviation can be computed.

All measures are indicative of differences between the two models except M3 is a measure of similarity. Mean values of all difference measures are very low (<1.4%) and the mean of the similarity measure M3 is very high (95%). The mean values of measures M1–4 suggest that the effect of geometry correction on overall volume, overall strain, highly compressed volumes and critical strain values is

negligible. The mean values of M5 measures suggest that the highly compressed volumes have negligible overlap with the region surrounding either type of corrected feature (cortical pore and trabecula). The standard deviations of all measures are very low (<2%) suggesting that the above inferences can be reliably asserted for all mouse samples. In conclusion, geometry correction has a negligible impact on predicting the global and local mechanical response of the bone section.

Supplementary Table 1: Statistics of measures of difference and similarity between original and geometry corrected model. Explanation of measures (M1, M2, …) are given in the main text.

Metric	Mean	Std
M1	0.40%	0.15%
M2	0.23%	0.82%
M3	95%	2.0%
M4	1.4%	0.36%
M5-cortical pores	0.5%	0.94%
M5-trabeculae	6.6×10^{-3}%	–

Supplementary Material 2: Role of mesh coarsening on surface geometry description

Methods

For each mouse, fifteen successively coarse surface meshes were obtained by reducing the number of faces by 20% at each step. Table 1 shows how the number of faces (actual output) and mean edge length depend on the desired number of faces (input) for the reference mouse bone. The coarsest mesh (Mesh-15) was created by 96% reduction of the faces in the original mesh and comprised 5802 faces with a mean edge length of 0.068 mm.

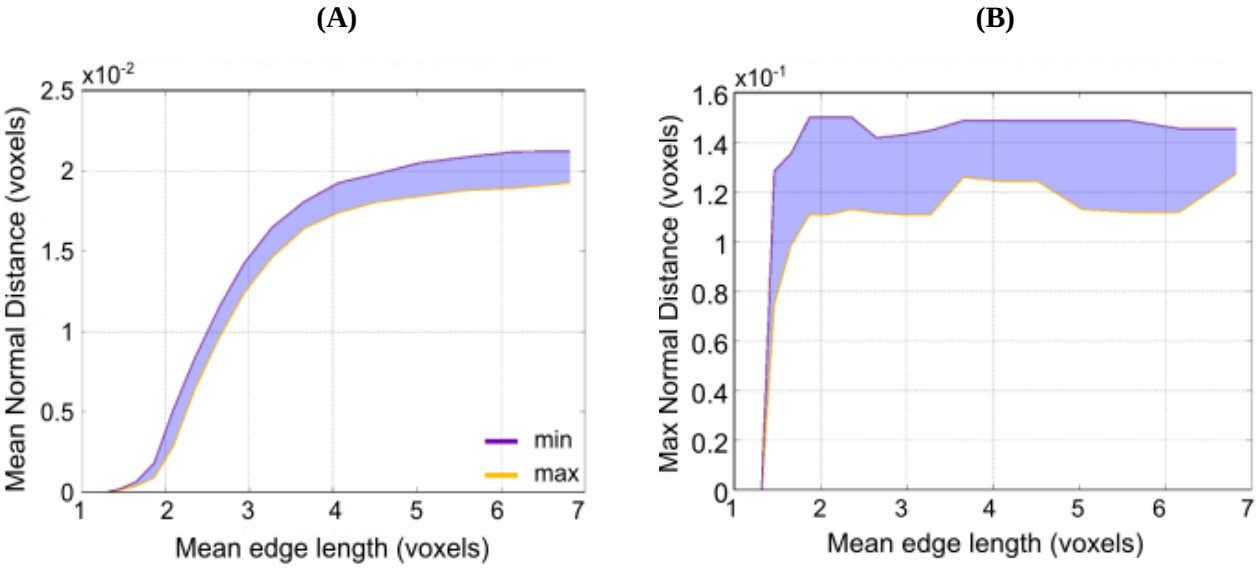
Two distance metrics were defined to quantify the effect of coarsening on shape representation accuracy, these are referred to as the maximum normal distance and the mean normal distance (Pottmann et al., 2001). For the reference mouse, d_{ij}^{v} is defined as the normal distance between vertex v of Mesh-i and its closest face on Mesh-j (i,j=0…15). The maximum and mean forward normal distances between Mesh-i and Mesh-0 are, respectively, the maximum and mean of d_{i0}^{v} taken over all vertices v of Mesh-i. The maximum and mean backward normal distances between Mesh-0 and Mesh-i are, respectively, the maximum and mean of d_{0i}^{v} taken over all vertices v of Mesh-0. Between a pair of meshes, the maximum normal distance is the larger of the maximum forward and maximum backward normal distances, and the mean normal distance is the larger of the mean forward and mean backward normal distances. This approach was used to evaluate both distance metrics for each mouse in the "OVX+ML" group at 18 weeks old at different levels of surface coarseness. A level of surface coarseness was considered admissible if both distance metrics were below a threshold of 1 voxel ($0.0104\ mm$) for all six mice. All levels of coarsening (up to 15) were found admissible according to the above criterion. Therefore, the coarsest mesh was chosen to reduce the computational demand of the subsequent steps of the framework.

Supplementary Table 2: Number of faces and mean edge length for each coarseness level

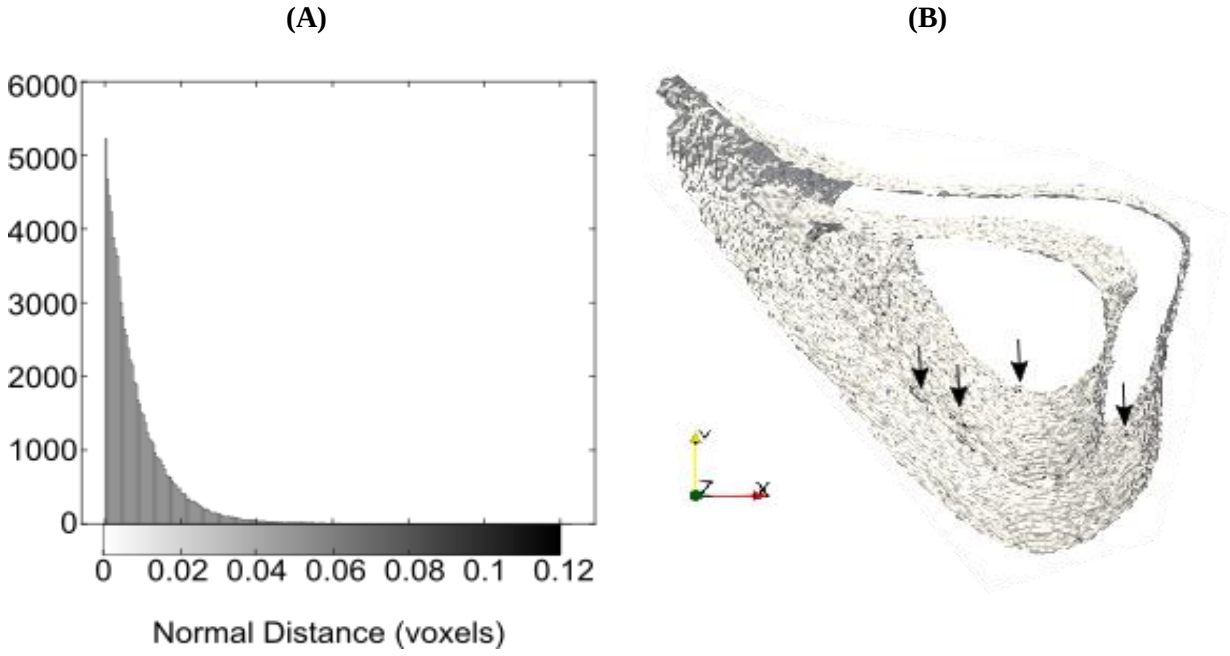
Coarseness step i	Number of faces	Mean edge length (mm)
0	166639	0.013
1	133181	0.015
2	106542	0.017
3	85230	0.019
4	68174	0.021
5	54537	0.024
6	43587	0.027
7	34821	0.030
8	27834	0.033
9	22250	0.037
10	17779	0.041
11	14208	0.045
12	11358	0.050
13	9080	0.056
14	7259	0.062
15	5802	0.068

Results

The mean normal distance was found to rapidly increase in the first coarsening steps, but it gradually flattened after 7 coarseness levels (Supplementary Figure 1). After 13 coarseness levels the mean errors were only slightly increased to 1%. This trend revealed that the simplification algorithm converges into a certain level of surface description error and subsequently no coarser meshes were further tested. At any coarsening level, the mean and the maximum normal distance were smaller than the voxel size and did not exceed 0.22 μm and 1.5 μm for any mice (Supplementary Figure 2). Therefore, all coarsening levels met the admissibility criterion.

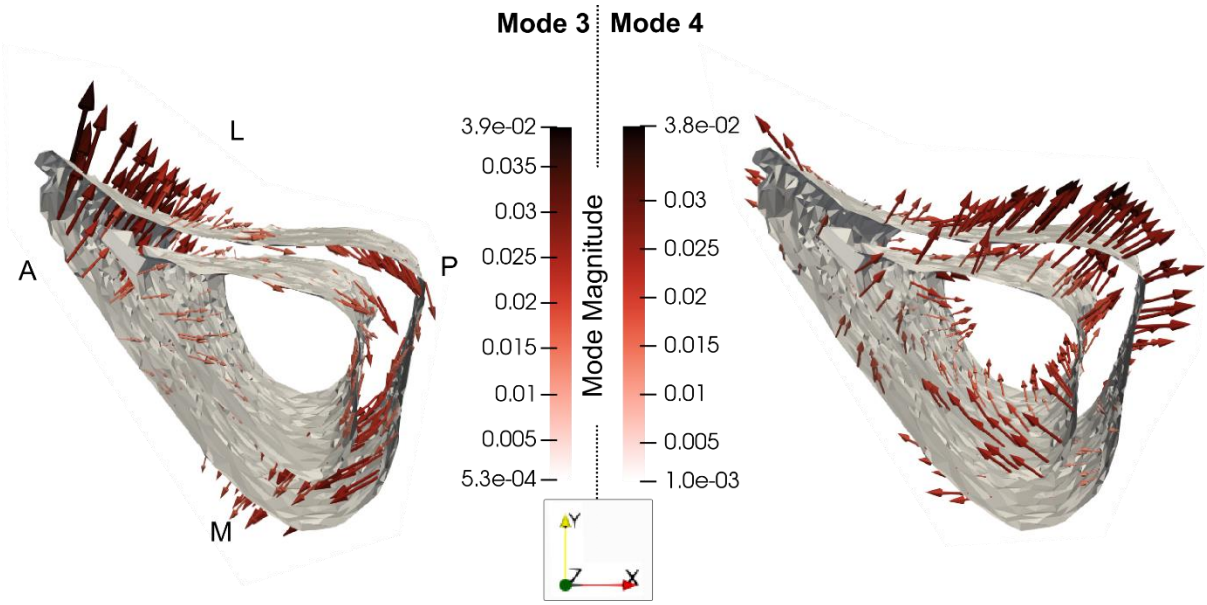


Supplementary Figure 1: Dependence of surface discretization errors on the mean edge length derived from the sequential coarseness levels (A) Mean Normal Distance and (B) Maximum Normal Distance. Quartiles defined by the minimum and maximum values among 6 mouse sample are shown by solid purple and orange lines in both plots.



Supplementary Figure 2: (A) Histogram of the surface discretization errors and (B) their spatial distribution on the bone surfaces. The error profile illustrates the forward normal distance from the finest to the coarsest mesh (Mesh-0 to Mesh-15). This histogram and profile describe only one specimen in the "OVX+ML" group and 18 weeks of age and is representative of other samples. The contour colours give the error magnitude and correspond to specific locations on the 3D shape. Black arrows highlight some areas with the largest errors.

Supplementary Material 3: PCA modes explaining a major fraction of total variation but not related to a treatment effect



Supplementary Figure 3: The 3D profiles of the Modes 3 and 4 illustrated as vectors plotted on the mean shape. The vectors are scaled and colored by mode magnitude at each point of the mesh. The darker and longer the arrow, the higher the variance on that area.

References

Cheong, V. S., Campos Marin, A., Lacroix, D., & Dall'Ara, E. (2020). A novel algorithm to predict bone changes in the mouse tibia properties under physiological conditions. *Biomechanics and Modeling in Mechanobiology, 19*, 985–1001. https://doi.org/10.1007/s10237-019-01266-7

Oliviero, S., Owen, R., Reilly, G. C., Bellantuono, I., & Dall'Ara, E. (2021). Optimization of the failure criterion in micro-Finite Element models of the mouse tibia for the non-invasive prediction of its failure load in preclinical applications. *Journal of the Mechanical Behavior of Biomedical Materials, 113*, 104190. https://doi.org/10.1016/j.jmbbm.2020.104190

Oliviero, S., Giorgi, M., & Dall'Ara, E. (2018). Validation of finite element models of the mouse tibia using digital volume correlation. *Journal of the Mechanical Behavior of Biomedical Materials, 86*, 172–184. https://doi.org/10.1016/j.jmbbm.2018.06.022

Pottmann, H. and Wallner, J. (2001). Computational line geometry. *Berlin: Springer*, 6. https://doi.org/10.1007/978-3-642-04018-4

Chapter 4

Principal Component Analysis for complete assessment of 3-dimensional treatment effects on mouse tibia cortex in an osteoporotic bone model

ABSTRACT

Preclinical *ex vivo* and *in vivo* studies of murine bone models have demonstrated the anabolic effects of the combination of pharmacologic, i.e., Parathyroid Hormone PTH (1-34) and mechanical treatment for osteoporosis. However, the standardised method for reporting the geometric changes over time due to treatment interventions uses a predefined set of scalar quantifiers that fail to explain important localised changes in the entire bone volume. Chapter 2 introduced a novel robust and accurate statistical framework, combining longitudinal imaging, deformable registration, and Principal Component Analysis for assessing the treatment effects in the 3D bone geometry of mouse tibia. This framework was firstly applied to a limited dataset of tibia midshaft received *in vivo* mechanical loading. In the current chapter, Chapter 3, the model considered almost the entire tibia cortex and evaluated the geometric differences between mice treated with either monotherapy of *in vivo* mechanical loading, and PTH or a combination of both. The treatment-related modes captured prominent changes in the proximal region, with the thickening of the anterior crest and lateral aspect and expansion of the medial and posterior compartments and highlighted higher in magnitude changes in the group which received combination of treatments. Additionally, the first two modes described axial variation in the length direction. The first six modes describing growth changes and treatment effects explained 91% of the total variance in geometry. Finally, the model demonstrated rigid body movements as shape variations possibly indicating instance misalignment in the shape data and underscoring the importance of future model development in this direction. This proposed model improves the investigation of bone geometry alterations to different treatment strategies.

4.1 Introduction

The efficiency of long-term pharmacologic monotherapies for osteoporosis such as Parathyroid Hormone (PTH)-based drugs, is debated due to potential side effects such as atypical fractures, bone cancer and stroke (Khosla & Hofbauer, 2017) After the anabolic window of two years of PTH-based drugs use, resorption tends to increase (Aslan et al., 2012). Another suggested approach is exercise but loading a highly porous and weak osteoporotic bone could increase the fracture risk. Therefore, novel treatment strategies including combinations of drugs and exercise have being investigated in clinical and preclinical research (Jee & Tian, 2005; Zhang et al., 2014; Monzem et al., 2023). The current preclinical study aims to provide a compact assessment of the combined treatment effects on bone geometry in osteoporotic mouse bone models.

A large number of preclinical studies have analysed the effects of PTH (1-34) and *in vivo* mechanical loading on rodent long bones, simulating the administration of anabolic drugs based on a synthetic form of PTH and exercise, respectively. Studies have examined the combined effects of pharmacological and mechanical treatment by conducting a cross-sectional analysis of specific sections of the mouse tibia. For instance, Sugiyama et al. (2008) demonstrated that exercise could increase the anabolic effects of drugs beyond what each therapy would achieve individually. However, such *ex vivo* analyses are highly invasive and often focus on small regions in metaphysis or/and diaphysis of the bones, limiting the ability to track the progression of the treatments and inter-subject differences.

On the other hand, longitudinal studies enable the comprehensive monitoring of bone adaptation in 4 dimensions, space and time, with *in vivo* longitudinal micro–Computed Tomography imaging (van 't Hof & Dall'Ara, 2019). This imaging protocol has been used for assessing localised densitometric changes over time across the entire tibia bone (Roberts et al., 2020). However, as elaborated in Chapter 2, a complete and localised assessment of the geometric changes over time considering intra-variations remains a challenge. The standardised method involves reporting differences between time points or treated/untreated cohorts in a predefined set of geometric scalar quantifiers extracted from standard morphometric analysis (Bouxsein et al., 2010). Such studies aiming at geometric assessments of cortical

bone response to concurrent treatments (drug and loading) are either concentrated on limited Volumes of Interest in the diaphysis (usually midshaft) (Gardinier et al., 2015; Birkhold et al., 2017; Roberts et al., 2020; Rooney et al., 2023) or analyse a larger segment of the tibia cortex (Monzem et al., 2023).

Studies that focused on a limited section by Rooney et al. (2023) and Birkhold et al. (2017) provided more detailed information than the traditional standard morphometric measures; by splitting the section into different compartments based on the anatomical directions. Both examined different loading cross-sectional compartments to investigate localised bone adaptation to localised strain distribution in the midshaft. Rooney et al. (2023) measured changes in area and thickness extracted from binary images of bones, while Birkhold et al. (2017) reported volume and thickness of formation and deletion by superimposing the images and measuring voxels of new and deleted bone on the surfaces. Nevertheless, they were still limited to a small part of the entire bone structure.

Such models based on the scalar geometric properties have shown bone adaptation to disease and/or treatment interventions over time in mouse tibia, in terms of increase or decrease of these parameters. In a study by Roberts et al., (2019), the relative increase in cortical thickness in mouse tibia from 16 to 24 weeks of age was 3.7–5.3% lower in ovariectomized mice compared with healthy mice. Similar findings were reported for other bone sites, such as femoral midshaft (Klinck & Boyd, 2008). Roberts et al. (2020) also demonstrated that the combination of the Parathyroid hormone PTH (1-34) with *in vivo* mechanical loading had an increased anabolic effect than either treatment alone, similar to *ex vivo* histomorphometric measures in both murine and rat bones (Sugiyama et al., 2008; Chow et al., 1998). Both the monotherapy of mechanical loading and the combination with PTH caused a consistent cortical thickening between the ages of 18, 20 and 22 weeks equal to a 6−20% increase/week. Additionally, Rooney et al., (2023) demonstrated that bone turnover is higher in compression than tension cross-sectional compartments of the bone. For the pharmacological treatment, mice treated with PTH alone exhibited a 17 % increase in tibial cortical thickness but only in the first administration period between weeks 18 and 20 (Roberts et al., 2020). Monzem et al., (2023) also showed that the daily use of PTH in 21 days increases the cortical area and alters the ellipticity changing the shape in 11-13 weeks-old mice.

These studies have provided a brief description of simple geometric alterations in terms of thickness and area. These quantifiers are scalar variables and are typically extracted from and averaged over pseudo two-dimensional regions. Therefore, they are only partially informative about the osteogenic response of the whole bone to treatment interventions.

Predictive models, utilising micro-Finite Element Modelling and mechanoregulation algorithms have expanded the investigation across larger tibia volumes and supported experimental findings of bone responses. Pereira et al., (2015) measured 3D cortical thickness adaptation in an *in vivo* loading murine tibia model between the proximal and distal tibia-fibula junctions. Another similar model predicted the 3D bone adaptation across nearly the entire tibia volume, allowing a temporal analysis of different treatment strategies (Cheong et al., 2021). Both studies consistently indicated that mechanical loading has high anabolic effects on the posterior-lateral diaphysis with increased formation and decreased resorption, showing prominent bone response at the periosteum. Moreover, combined treatments have shown benefits over monotherapy, with prominent apposition and low resorption due to the combination of treatments (Cheong et al., 2021).

Being a localised approach, mechanoregulation-based models provide new insights into investigating bone adaptation in 3D, by measuring bone formation and deletion frequencies on the active surfaces. While considering mechanical stimulus only, such models need further development to accurately simulate both bone formation and deletion when pharmacological stimulus is administrated (Cheong et al., 2021). Additionally, since these models are sample-specific, it remains unclear which 3D geometrical features consistently change for all samples due to external interventions.

To improve the precision and accuracy of the geometric assessment, we previously proposed a 3D PCA-based bone model using *in vivo* microCT to investigate spatiotemporal changes due to interventions in osteoporotic mouse tibia models. This model offers several advantages. First and foremost, it does not rely on any predefined geometry description model (compared with standard morphometry), but it is solely defined by the input image data and the shapes extracted from them. This approach was thoroughly examined in a previous study in Chapter 2, where it was tested on a midshaft section of the

bone, and the reliability and robustness of the entire framework were evaluated. Briefly, we combined *in vivo* longitudinal imaging, deformable registration and Principal Component Analysis to precisely quantify the geometric variations of the active murine tibia midshaft surfaces due to *in vivo* mechanical loading.

The current study in this chapter aims to expand the application of this novel model in a higher complexity considering the entire mouse tibia cortex and combinations of both pharmacological and mechanical treatment. It is expected that a comprehensive 3D assessment of bone adaptation patterns to interventions will be possible within the examined population of mice tibia, with these patterns better explaining the shape variation of the entire structure and fully capturing all important localised changes.

4.2 Materials and methods

4.2.1 Animals and interventions

Previous longitudinal *in vivo* murine studies generated all experimental data analysed in this study (Roberts et al., 2019, 2020). Specifically, this study examines all treatment groups and the osteoporotic group, as shown in the schematic illustration in Figure 4.1. The experimental protocol is detailed in section 2.7.

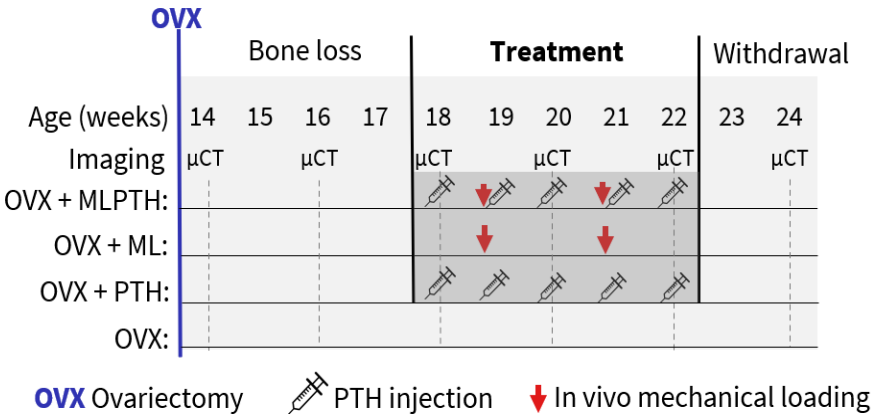


Figure 4.1: Examined groups of mice and the different interventions.

4.2.2 *In vivo* imaging and processing

In those studies, the right tibiae were scanned every two weeks using *in vivo* microCT (Roberts et al., 2019, 2020. The 3D grey-scaled images were reviously rigidly registered by Cheong et al. (2021). The rigid registration was accomplished using the protocol detailed by Lu et al. (2016) (also Chapter 2,

section 2.7). The growth plate, fibula and condyles were removed from the images (Cheong et al. 2020), including 80% of the total tibia length, starting from under the growth plate. All remaining 3D image samples were finally rigidly registered to the reference.

The current study analyses the registered image on bones derived from weeks 18, 20, 22 and 24 for all groups.

4.2.3 Preprocessing and Principal Component Analysis Implementation

The preprocessing framework and the variation decomposition are detailed in Chapter 3. In the present chapter, the author highlights the adaptations of each district step of the framework to enable its implementation to a more complex database, analysing the entire cortex of the tibia bone rather than the midshaft region, as in Chapter 3.

4.2.4 Further cropping

Firstly, the 3D grayscale images of the mouse tibia bones previously cropped to encompass 80% of the total length (Cheong et al., 2021), were further cropped to ensure topological equivalence across all bone samples. Specifically, 10% of the original length was cropped to exclude the proximal trabecular region, where the cortex could not be effectively separated from the trabeculae. Additionally, automatic cropping at the distal end was performed to remove the distal separation of the tibia and fibula by detecting the number of separate objects in the binary images slice by slice. As a result, the total Volume of Interest in this study represented approximately 65% of the total length. Previous evidence showed that the bone turnover is low in the distal sections (Cheong et al., 2021), and, therefore, we expected that after the above cropping scheme, the most critical bone remodelling regions of the mouse tibia were still included in the analysis.

4.2.5 Rigid registration

Next, the yielded cropped 3D greyscale images underwent further rigid registration. That was needed because there were noticeable positional misalignments between the bone samples in the original images. Since PCA would mistreat this position and orientation differences as form differences, the framework applied further registration to eliminate these effects in the data. A similar approach was

used in Chapter 3, where additional local registration of the midshaft was used to reduce the local position mismatch between groups and within groups.

Firstly, global rigid registrations of all image samples to a single reference at 18 weeks were visually inspected for spurious misalignments and they were indicated ineffective. That could be a result of important structural differences between all cohorts, and it could also potentially explain the positional misalignments observed in the original images. To address this issue and with the consideration that bone shapes of similar age might be geometrically 'closer to each other', this study proposed using mouse age-specific references, which were initially co-registered with each other. Specifically, advanced ages were registered to a reference mouse of week 18 and the resulting registered image of the reference mouse of each age was onwards called 'reference week 18/20/22/24'. The remaining mice within each age group were registered to the reference of the same age. For example, images of mice at 20 weeks across all intervention groups were registered to reference week 20. Rigid registration was performed in Amira 2022.2 using Normalized Mutual Information as a similarity metric and Lanczos interpolation for image resampling.

4.2.6 Binarization, geometry correction and extraction of the reference mesh

After rigid registration, the greyscale images were binarized and further manipulated to remove cortical pores and trabeculae following the procedures outlined in Chapter 3, subsection 3.2.4. The manual manipulation of the image data took approximately 1 hour/bone image. Additionally, the binarized images of reference week 18 were processed to extract the reference surface mesh which involved triangulating the free surfaces of the periosteum and endosteum of the tibia cortex. To facilitate faster management of surface data and further processing, the mesh was simplified by 50% of faces, resulting in 1.5 million vertices.

4.2.7 Mapping via deformable registration

Following that, the binarized reference week 18 image was elastically registered to all remaining bone images using Nodal Spacing equal to 5 (optimisation of this parameter in Chapter 3, subsections 3.2.6 and 3.3.1). Due to the high dimensionality of the structure (~1200 image slices/mouse) and the large

number of image samples ((6 mice + 5 mice) x 4 groups x 4 ages = 92 observations), adjustments were made in the deformable registration algorithm, referred to as the Sheffield Image Registration Toolkit to enable task parallelisation on a High-Performance Computing (HPC) system. For efficient job and data management, input and output data folders were named after the mouse, group and age-specific label. These jobs were then executed automatically and in parallel on the Stanage HPC system, with an approximate execution time of 6 hours and a memory requirement of 10.53GB RAM each.

In the next step, the displacement field obtained from the elastic registration of the reference to each image data was tri-linearly interpolated to the reference surface mesh. The output of this step comprised the registered reference tibia mesh to all other bone samples, constituting all bone samples discretised by a fixed number of nodes. The observations on coordinates of the nodes formed the rows in the PCA input.

4.2.8 Investigating systematic shape variations

As in Chapter 3, PCA decomposed the variation in the nodes into mode shapes to describe orthogonal shape variations associated with the different interventions in the tibia system (Equation 2.1, section 2.8.2).

The examined population included longitudinal mouse tibiae observations from three different treatment cohorts and one disease cohort. The PCA modes described multiple sources of variations in the tibia shapes of this examined population. To investigate systematic variations described by a specific mode shape, the scores samples of each mode were grouped according to cohorts and ages. Statistical tests were utilised to investigate differences between these subgroups and ages.

Pairwise analysis (one-sided and 2-sided Wilcoxon signed rank tests) was conducted for each cohort between ages 18 and 20, 20 and 22, 22 and 24 and 18 and 24. One-sided tests were performed to detect monotonic (or not) temporal changes in scores. Monotonic score changes of a specific mode were interpreted as systematic shape alterations due to growth, disease or treatments. This categorisation into growth, disease and treatment-related modes was further expanded from the one described in Chapter 3, in order to account for multiple groups and ages. Mann Whitney U tests were performed to identify

inter-cohort differences between disease and monotherapies/combined treatment. Normalised score changes over age were calculated with respect to the median score value in week 18.

For the modes that describe temporal changes in shape over age, the 3D surface changes, described by these modes, were reconstructed using the median score changes between scores at two ages (when statistically significant differences of the median score change over time with zero were found). The reconstruction was defined as scaled mode vectors with median score changes, and it was accomplished for both cases of one mode and a combination of modes to investigate solo or combined geometric changes over time.

4.3 Results

The first 2 and 6 shape modes captured 82% and 91% of the total shape variance, respectively, as shown in Figure 4.2. The shape profiles of Mode 1 and 2 describe uniaxial variations in the proximal and distal end, respectively (Figure 4.3), as shown by the darkest (almost black) and longest arrows allocated at those areas.

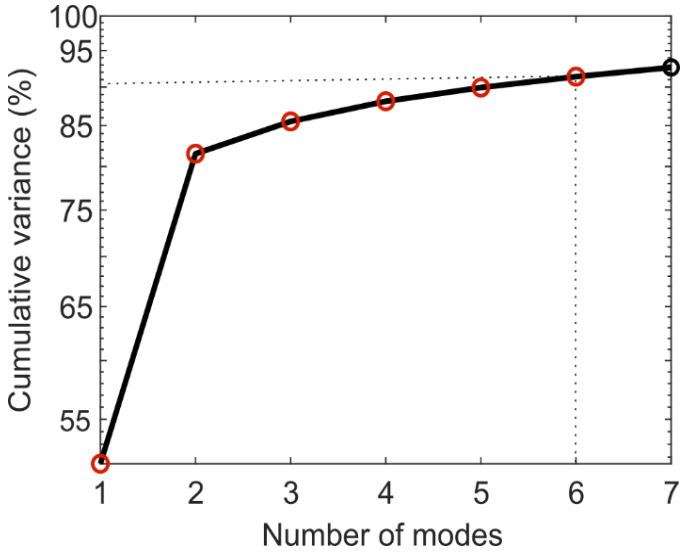


Figure 4.2: *Percentage of the total shape variation captured by PCA modes. Red coloured circles correspond to the first 6 modes which capture 91% of the total variance.*

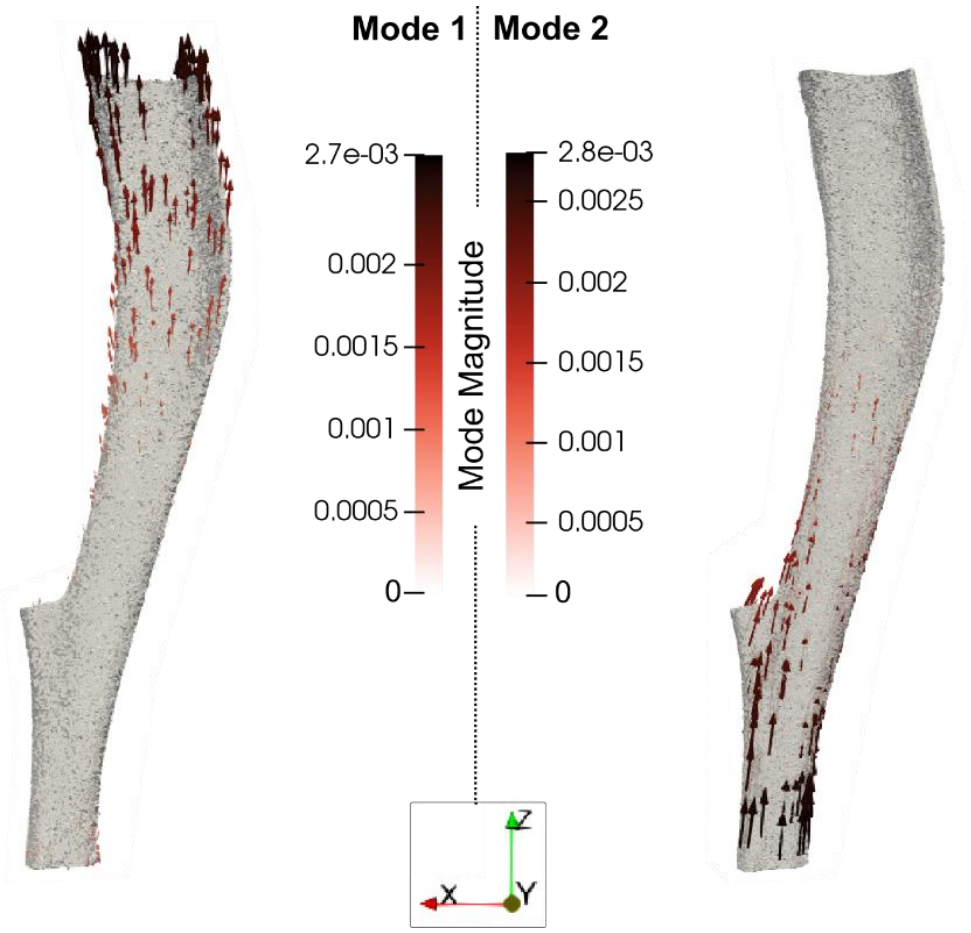


Figure 4.3: The 3D profiles of the first 2 modes are illustrated as vectors plotted on the mean shape. The vectors are scaled and coloured by mode magnitude at each node. Lateral viewpoint is illustrated.

In the following figures, Figures 4.4 and 4.5, the corresponding scores of Mode 1 were grouped with respect to the cohort groups and used to quantify the temporal changes, test for statistically significant changes over time and therefore determine whether Mode 1 represent a temporal shape change. These figures illustrate the median normalized score changes (and the min-max range) over time in pairs of 18-20, 20-22, 22-24 and 18-24 for all cohort groups in subfigures A and B and the individual score changes in subfigures C, D, E, and F corresponding to each age pair. The results of the statistical tests shown in subfigures A and B indicate whether the corresponding mode captures a systematic temporal change between the age pairs.

Specifically, for Mode 1, the temporal score changes between 18-20, 20-22, 22-24 and 18-24 were not statistically significant different (s.s.d.) than zero for all groups (Figures 4.4A, 4.4B, 4.5A, 4.5B). The median normalised score changes over all time points are approximately zero with non-systematic fluctuations, and ranges of [-13−0.45] % for “OVX+ML”, [-0.027−0.78] % for “OVX”, [2.59−4.29] %

for “OVX+PTHML”, and [-52 62] % for “OVX+PTH” group. This is visually indicated in Figure 4.4 (“OVX+ML”, “OVX”) and Figure 4.5. (“OVX+PTHML”, “OVX+PTH”). Additionally, there was large variability in the distribution of score changes, with groups “OVX+ML” and “OVX+PTH” exhibiting the largest ones for 18-24 weeks and min-max ranges of normalised score changes up to 100% and 800% units, respectively (Figures 4.4A, 4.5B). The variability in the score changes was visually complemented by the plots of the individual score changes in subfigures C, D, E and F in Figures 4.4 and 5.4, which do not exhibit a consistent pattern. Thus, non-significant changes over time for Mode 1 suggested that the captured shape variations in the proximal end (3D profile of Mode 1 in Figure 4.3) described natural variability (random) across the examined population, and it could not be attributed to a systematic temporal effect of growth, disease, or treatment.

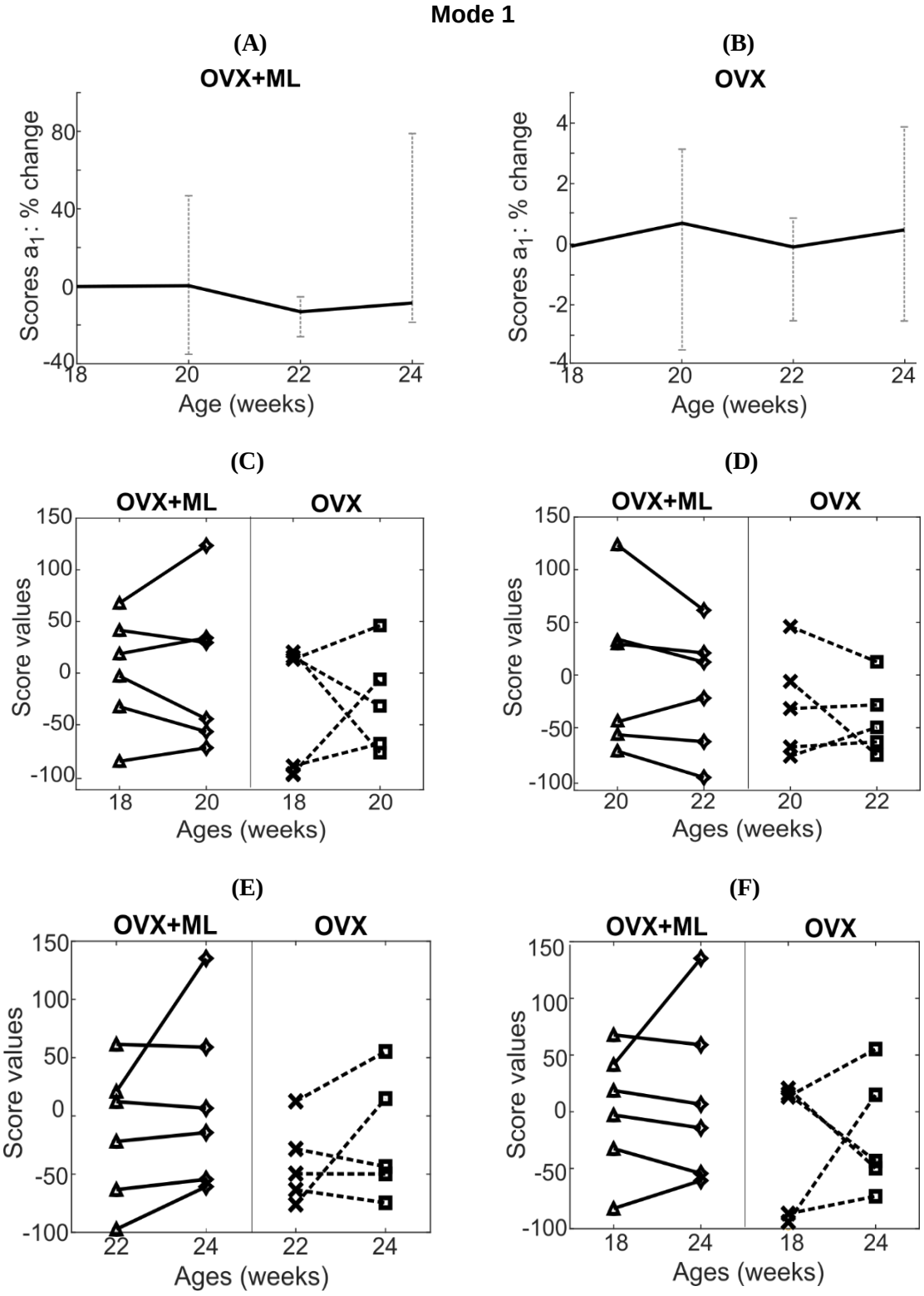


Figure 4.4: Temporal changes of scores for Mode 1 in "OVX+ML" and "OVX". (A-B): Normalised score changes values in 18-24 weeks for "OVX+ML" and "OVX". (C-F): Individual changes between ages, i.e., 18-20, 20-22, 22-24 and 18-24 for "OVX+ML" (solid) and "OVX" (dashed line).

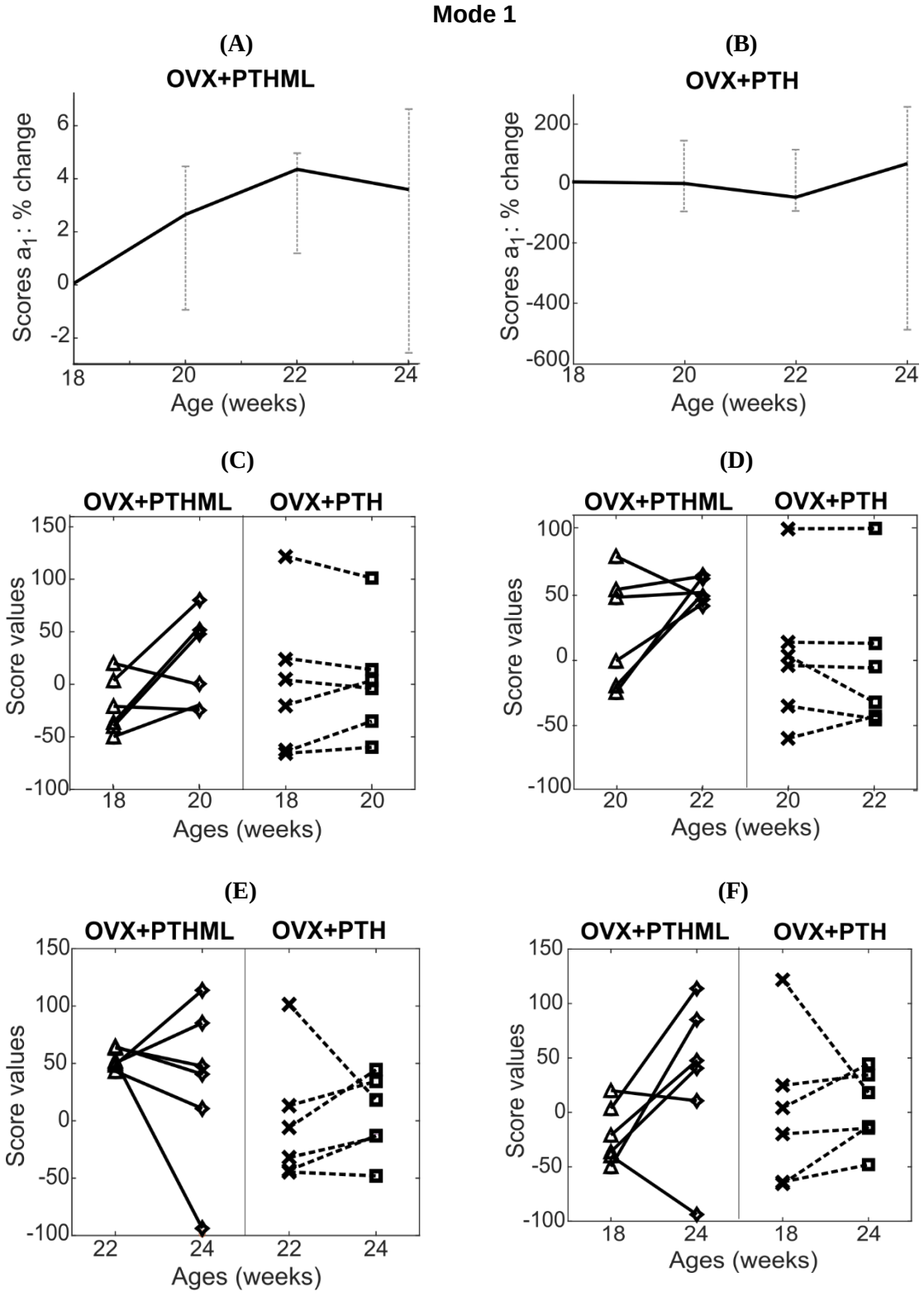


Figure 4.5: Temporal changes of scores for Mode 1 in "OVX+PTHML" and "OVX+PTH". Normalised score changes values from week 18 to weeks 20, 22, and 24 *are plotted on top (A: "OVX+PTHML" and B: "OVX+PTH"). Pairwise changes between ages, i.e., 18-20, 20-22, 22-24 and 18-24 are shown for both the "OVX+PTHML" (solid line) and "OVX+PTH" (dashed line) group* in C-F.

Similarly, as in Mode 1, Figures 4.6 and 4.7 correspond to the scores of Mode 2. In this case, some statistical significance differences ($p<0.05$) between ages were found (Figure 4.6A, 4.6B and 4.7A and 4.7B). For "OVX+ML", the median score change was negative (1-sided, $p<0.05$) between 18, 20 and 22 weeks. The median normalised score changes were -0.40%, -1.74%, and -0.83% for age pairs of 18-20, 18-22, and 18-24 weeks, respectively. The overall change of -0.83% between weeks 18 and 24 was found statistically significant (2−sided).

For the "OVX" group, no statistically significant changes between ages were found, a statement that could be complemented by Figures 4.6C-F, which exhibit inconsistent, i.e., positive and negative, changes of the individuals in the group. Although the overall changes between weeks 18-24 seemed monotonically negative with a gradual median normalised score increase (in magnitude) from -0.49% to -1.25%, they were not significant ($p>0.05$) (Figure 4.7B).

For "OVX+PTHML", changes are significant for the pairs 20-22, and 22-24 weeks, but not for 18-20 weeks of age. The median normalised change is equal to -2.12%, and -13.18% for the periods 18-20 and 18-22, but it decreases in magnitude resulting in an overall median normalised score change of -0.5% from week 18 to week 24 (Figure 4.7A). The difference in the gradient of change from negative to positive is highlighted in Figure 4.7A, but also in the individual score changes in Figures 4.7C-E.

Lastly, for "OVX+PTH", there is a monotonic negative change between all ages, with statistical differences between 18-20, and 20-22 weeks, but it slightly flattens between the period 22-24 weeks ($p>0.05$) (Figure 4.7B). The median normalised score changes are -0.84%, -1.48% and -1.62% from week 18 to week 20, 22 and 24, respectively. The overall change of -1.62% between week 18 and week 24 is statistically significant as shown by the 2−sided test.

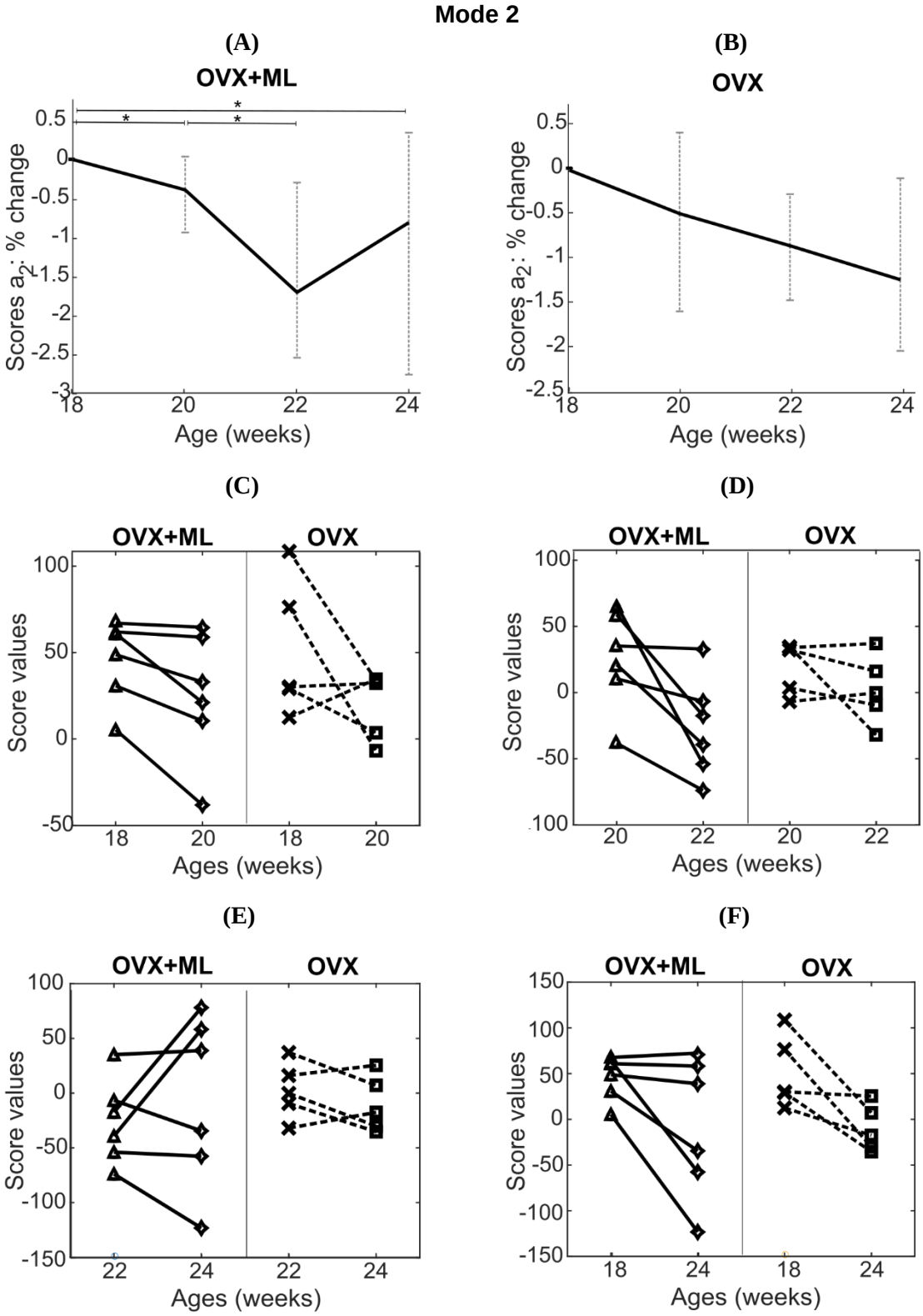


Figure 4.6: Temporal changes of scores for Mode 2 in "OVX+ML" and "OVX". Normalised score changes values from week 18 to weeks 20, 22, and 24 are plotted on top (A: "OVX+ML" and B: "OVX"). Pairwise changes between ages, i.e., 18-20, 20-22, 22-24 and 18-24 are shown for both the "OVX+ML" (solid line) and "OVX" (dashed line) group in C-F.

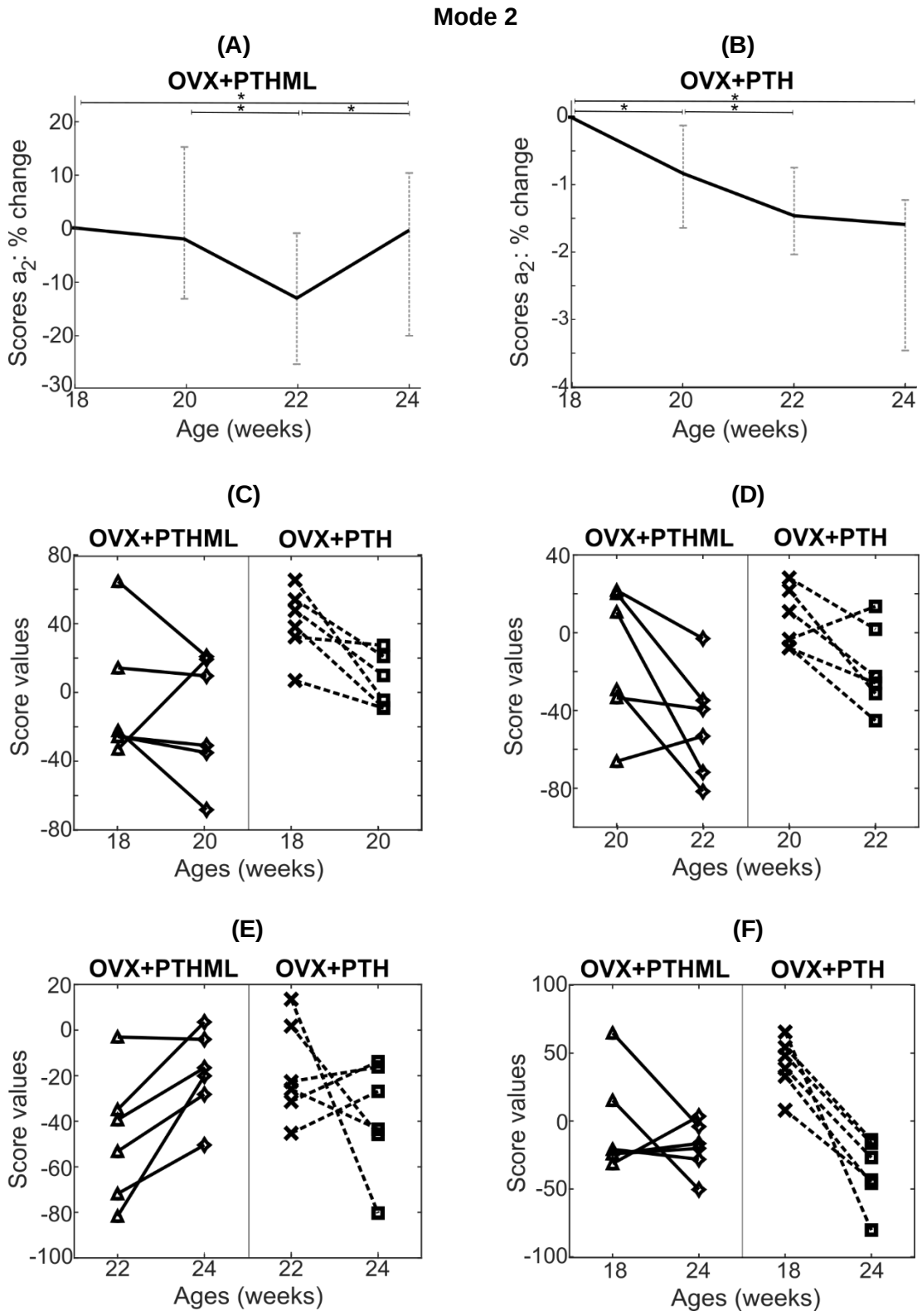


Figure 4.7: Temporal changes of scores for Mode 2 in "OVX+PTHML" and "OVX+PTH". Normalised score changes values from week 18 to weeks 20, 22, and 24 are plotted on top (A: "OVX+PTHML" and B: "OVX+PTH"). Pairwise changes between ages, i.e., 18-20, 20-22, 22-24 and 18-24 are shown for both the "OVX+PTHML" (solid line) and "OVX+PTH" (dashed line) group in C-F.

For the statistically significant changes between the start (week 18) and the end (week 24) of the intervention period for “OVX+ML”, “OVX+PTHML” and “OVX+PTH” groups, the 3D surface changes were calculated, by scaling up the mode vectors using the median score change. Specifically, Mode 2, the median score change between 18 and 24 weeks in the “OVX+ML” group was equal to -52.09 units, associated with at most 0.14 mm movement of the proximal end of the mouse tibia over 18-24 weeks (Figure 4.8). The median score change in the “OVX+PTHML” group was equal to -15.1 and it was translated into at most 0.042 mm movement of proximal end. For the “OVX+PTH” group, the median score change between 18 and 24 weeks was measured to be -79.02, corresponding to at most 0.22 mm uniaxial proximal movement. Looking at the sign of the score changes (negative) and the direction of the mode vector (shrinkage, Figure 4.3), the surface changes over time suggest an overall expansion in the axial direction in all groups going from week 18 to 24 at the distal end. The shape profile of the “OVX+ML” group in Figure 4.8 is representative of the surface changes in the other treatment groups.

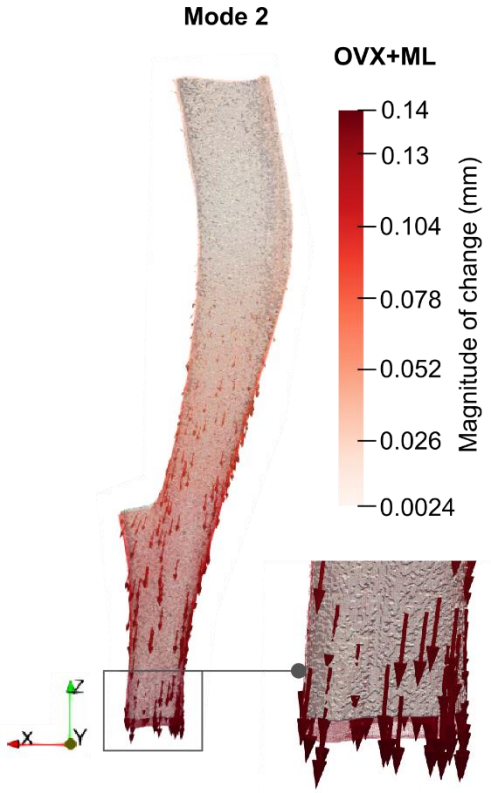


Figure 4.8: Surface changes between week 18 and week 24 as described in Mode 2 for the “OVX+ML”. The directions of change at different locations are denoted by the arrows, and a redder arrow indicates a relatively larger change. The median bone profile at week 18 is shown as a solid grey surface, whilst the median profile at week 24 is given as a coloured wireframe. Zoomed-in distal section is given.

The present study also investigated score distributions in each group for Mode 1 and 2 to further investigate the inter-group differences (Figure 4.9). The score distributions of Mode 1 in the "OVX" group described differences from the mean shape (zero axis) with median equal to -37.65 and 65% of the score samples being negative. Conversely, the score distribution in the "OVX+PTHML" group has median value equal to 41.91 and 67% of the samples positive (Figure 4.9A). Figure 4.9A also highlights the inter-group difference between "OVX+PTHML" and "OVX", with a statistical significance of $p<0.01$.

The score distribution of Mode 2 described similar trends of variations from the mean shape, as for Mode 1, but with opposite signs (Figure 4.9B), and similar systematic differences between cohort groups, as for Mode 1, with most importantly the ones between the "OVX+PTHML" and "OVX" groups.

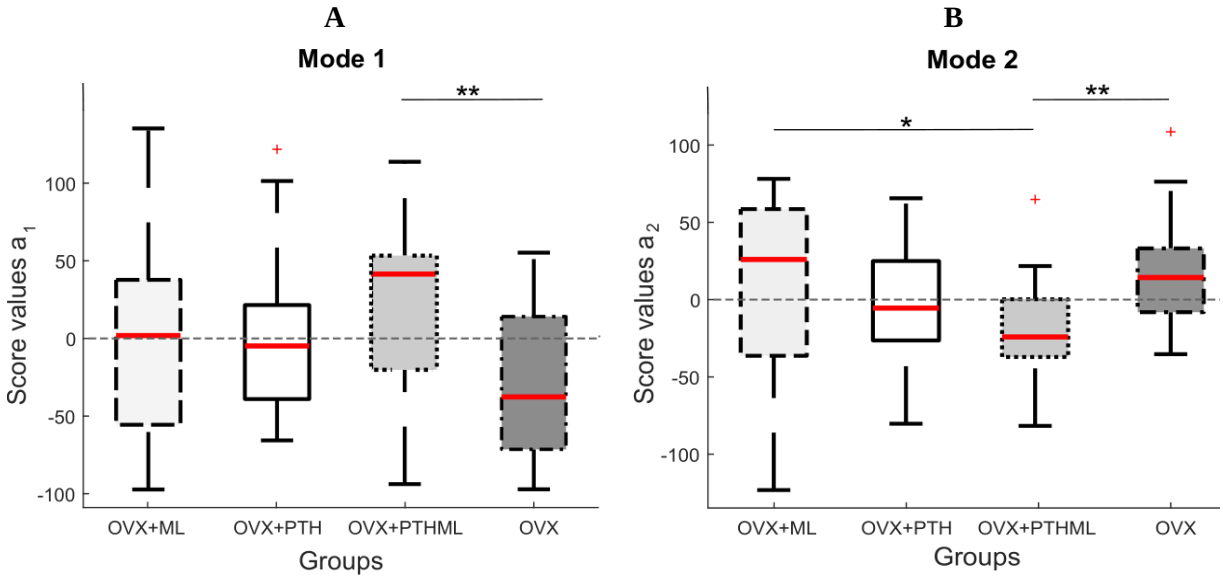


Figure 4.9: Boxplots of the score distributions grouped concerning the cohort groups. Different colours and line styles correspond to the subgroups of the examined population. (A) Mode 1, and (B) Mode 2. (*) statistically significant difference between groups, $p<0.05$, (**) $p<0.01$.

When investigating the combined temporal effects of Mode 1 and Mode 2, the profiles of the 3D surface changes over time revealed uniaxial movements for the "OVX+ML", "OVX" and "OVX+PTH" groups in 18-22 weeks. Their characteristic was the almost constant orientation, going downwards in the z-axis, and similar magnitude of the vectors over the 3D surface (small ranges of 0.022-0.029 mm for "OVX+ML", 0.11-0.14 mm for "OVX" and 0.012-0.015 mm for "OVX+PTH") across the entire bone

shape (Figure 4.10). The "OVX+PTHML" group exhibited concurrent outwards movement of both ends, resulting in expansion of the bone shape and increase in length.

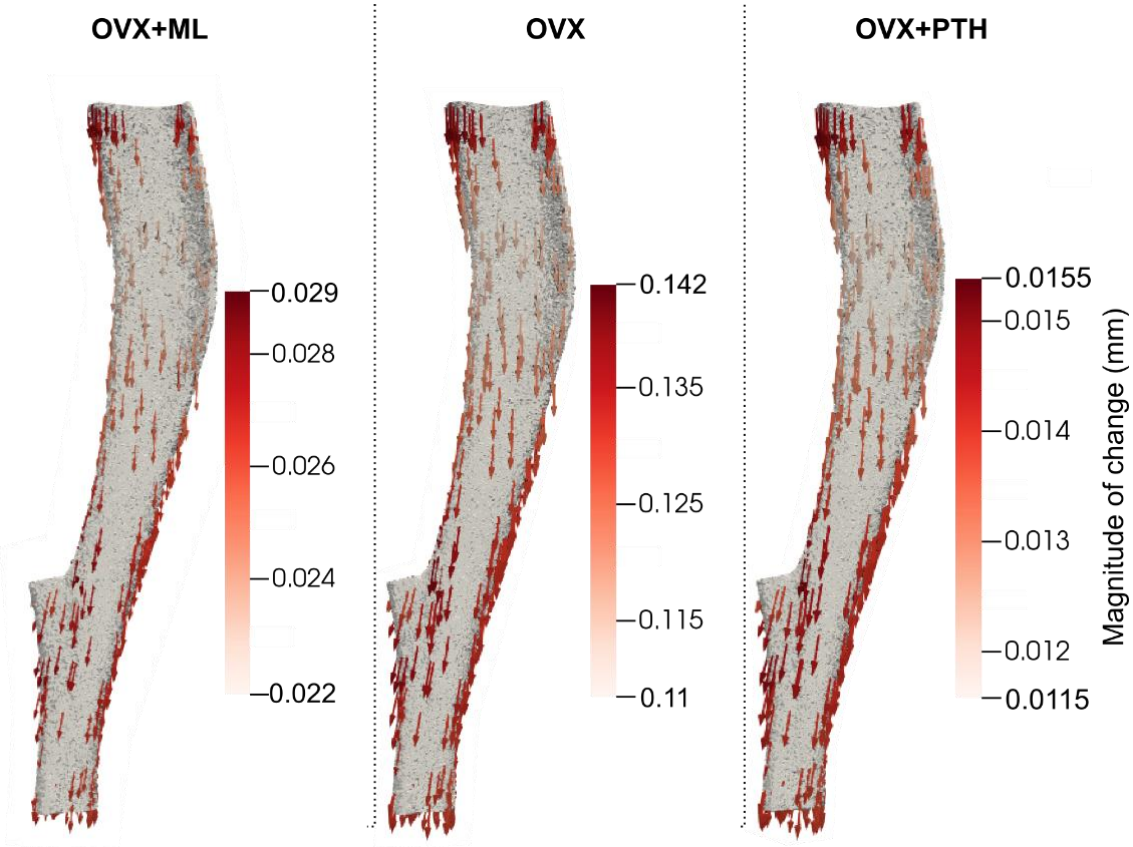


Figure 4.10: Surface changes between week 18 and week 22 as described by the combined Mode 1 and Mode 2 for the "OVX+ML", "OVX" and "OVX+PTH" groups. The vector field is plotted on top of the median profile of week 18. Lateral view is exhibited.

The profiles of the next four modes are shown in Figure 4.11. As in Figure 4.3 for Mode1 and Mode 2, the darker and longer arrows indicating higher magnitudes of the modes reveal the variant 3D features the Mode 3, 4, 5 and 6 capture. Specifically, these four modes prominently described variations at the proximal region but in different localised areas and different directions. They also captured variations of smaller magnitude in regions of the diaphysis at the posterior and anterior compartments. Specifically, Mode 3 captured localised variation at the posterior-lateral edge of the proximal end, where the arrows with the highest magnitude are allocated. Variations of smaller magnitude were also found at the anterior crest on the diaphysis. The darkest and longest arrows of Mode 4 was interpreted that this mode described variations at the anterior crest of the proximal area. Mode 5 described variations at the posterior compartment. Finally, Mode 6 revealed variations in the medial aspect of the proximal end (Figure 4.12). Along with the proximal variations, Mode 6 captured cross-sectional

variation in the diaphysis.

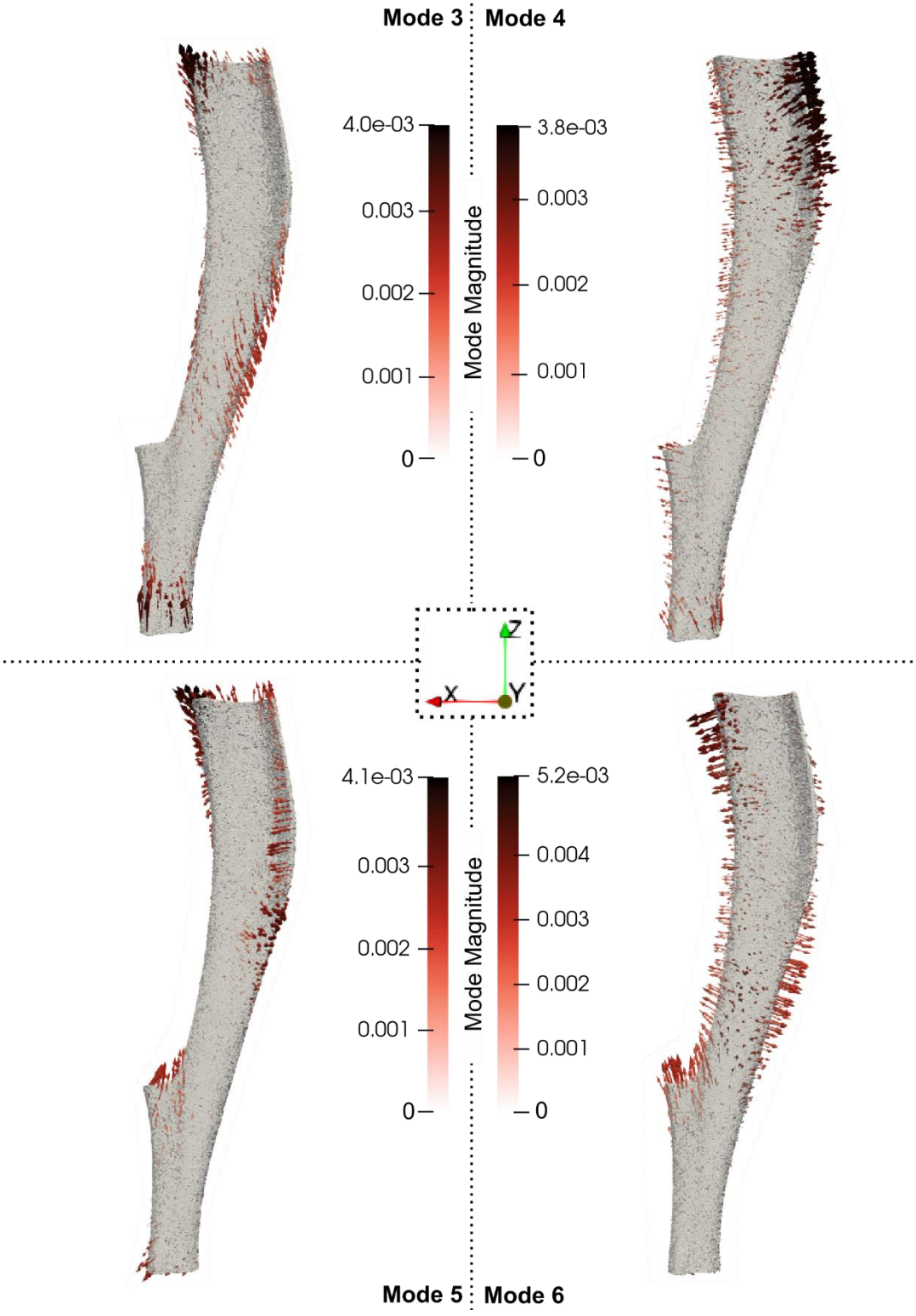


Figure 4.11: The 3D profiles of Modes 3, 4, 5 and 6 are illustrated as vectors plotted on the mean shape. The vectors are scaled and coloured by mode magnitude at each node. Lateral viewpoint is exhibited.

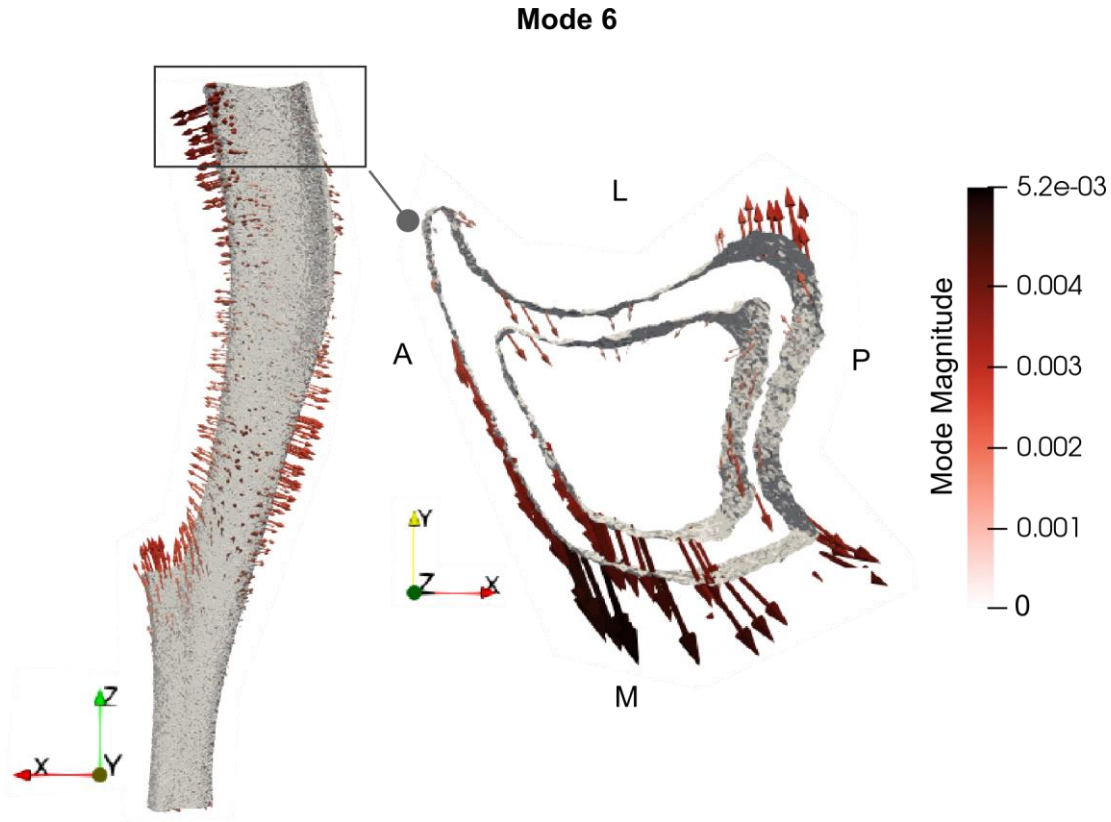


Figure 4.12: Zoom in the proximal region of Mode 6. Zoomed-in view exhibits anatomical orientation as indicated by A (anterior), L (lateral), P (posterior) and M (medial).

To investigate whether Mode 3, 4, 5 and 6 describe any temporal changes in the 3D shape, the scores were grouped with respect of the cohort and the age and the score change were calculated and tests for statistically significance. As in the case of Mode 1 and Mode 2, the median normalized score changes and the results of the statistical tests for all groups and ages and from Mode 3, 4, 5 and 6 are illustrated in Figures 4.13−4.16 The score changes for Mode 3 were not statistically significant for all cohort groups and all time points. The only statistically significant changes over time were found in the "OVX+PTHML" group for 18-24 weeks, underscoring the summative treatment effect with a gradual increase from week 18 to week 24 (Figure 4.13). The median normalised score changes for this group are equal to 0.13%, 0.64% and 2.64% for 18-20, 18-22 and 18-24 weeks, respectively. Notably, "OVX+ML" also exhibit an increase in scores, but with a lower magnitude of 1.9% between week 18 and week 24. "OVX" and "OVX+PTH" exhibit almost stable score values over age.

The score changes of Mode 4 were found statistically significant positive in the "OVX+PTHML" group for the periods of 20-22 weeks and 22-24 weeks and the overall time period from the beginning till the

end of the examination period, week 18-24 (Figure 4.14). The scores for this treatment group gradually increased over time, with median normalised score changes from week 18 being 5.53%, 12.18% and 17.16% to weeks 20, 22 and 24, respectively. Although no other statistically significant changes for the other groups were found, monotonic increases in scores were revealed in the "OVX+ML" group, while scores were almost stable for the "OVX+PTH" group and decreased in the "OVX" group. The positive sign of the score changes in the "OVX+PTHML" group was associated with the 3D geometric change of prominent periosteal bone formation at the anterior ridge (arrows looking outwards in Figure 4.11) between periods 20-22 weeks and 22-24 weeks.

The score changes of Mode 5 were significantly positive for the overall time period of 18-24 weeks for all treatment groups and stable for the disease group (Figure 4.15). The median normalised score changes between weeks 18 and 24 were equal to 1.6% for "OVX+ML", 0.26% for "OVX", 1.3 % for "OVX+PTHML" and 0.85% for "OVX+PTH". The positive score changes in the treatment groups corresponded to captured expansion of the posterior compartment in the proximal area for all treatment groups. That is because the positive sigh of the score changes agrees with the direction of the mode vectors as in Figure 11.

Finally, the score changes of Mode 6 were found statistically significant positive for all treatment groups and statistically significant negative for the disease group (Figure 4.16). The statistical tests showed significant monotonic positive changes for the "OVX+PTHML" and the "OVX+ML" groups in 18-20 and 22-24 weeks. The changes were less pronounced in 20-22 weeks, with no statistical significance. In the "OVX+PTH" group, the score increased dramatically in 18-20 weeks but flattened with time. In the "OVX" group, the overall score change was found negative (1-sided, $p<0.05$). The median normalised score changes between weeks 18 and 24 were found to be equal to 4.8% for "OVX+ML", 2.16% for "OVX+PTHML", 36% for "OVX+PTH", and -1.35% for "OVX". The positive score change over age for the treatment groups agreed with the outward direction of Mode 6 in Figure 4.12 and, therefore, it was associated with the temporal cross-sectional expansion of the medial compartment of the proximal end and the diaphysis over time (18-20, 20-22 and 22-24 weeks) for all treatment groups.

The opposite surface change explained by the negative score changes is detected for the disease group.

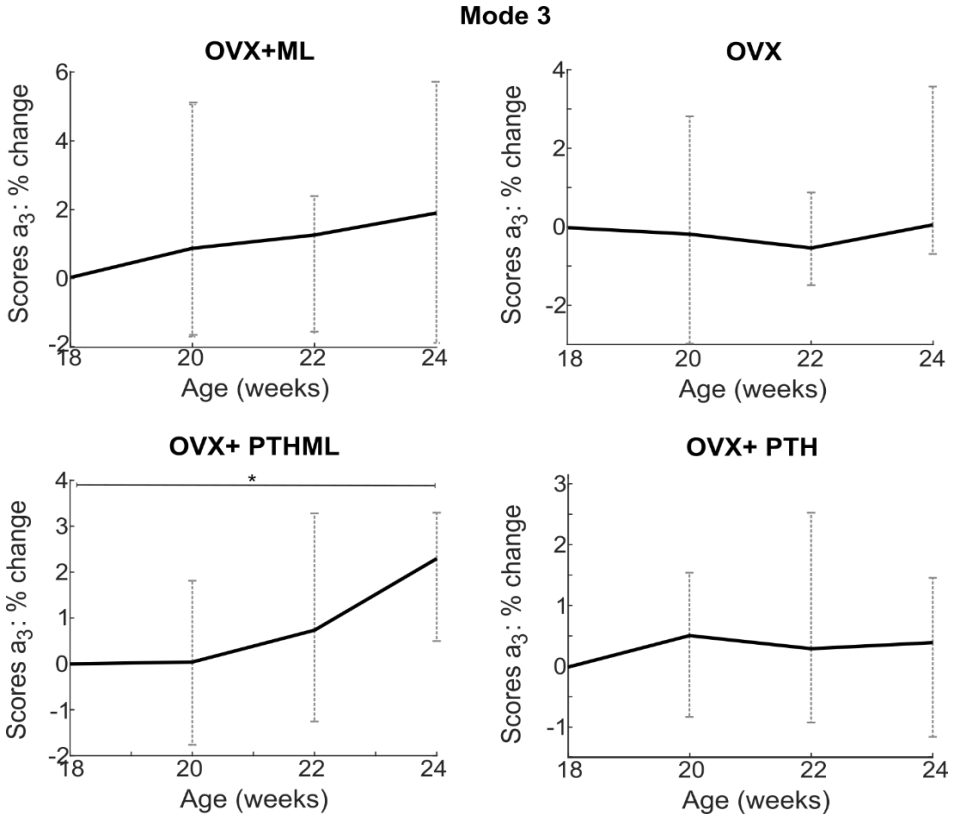


Figure 4.13: Normalised score changes for Mode 3 in all groups.

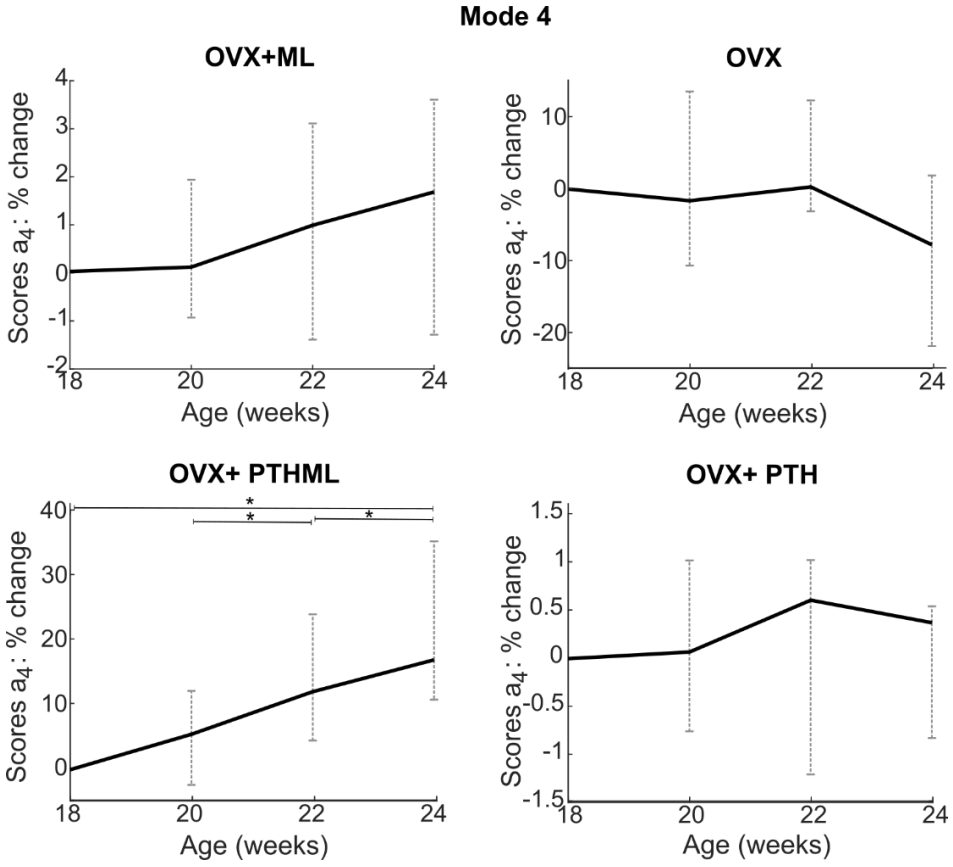


Figure 4.14: Normalised score changes for Mode 4 in all groups.

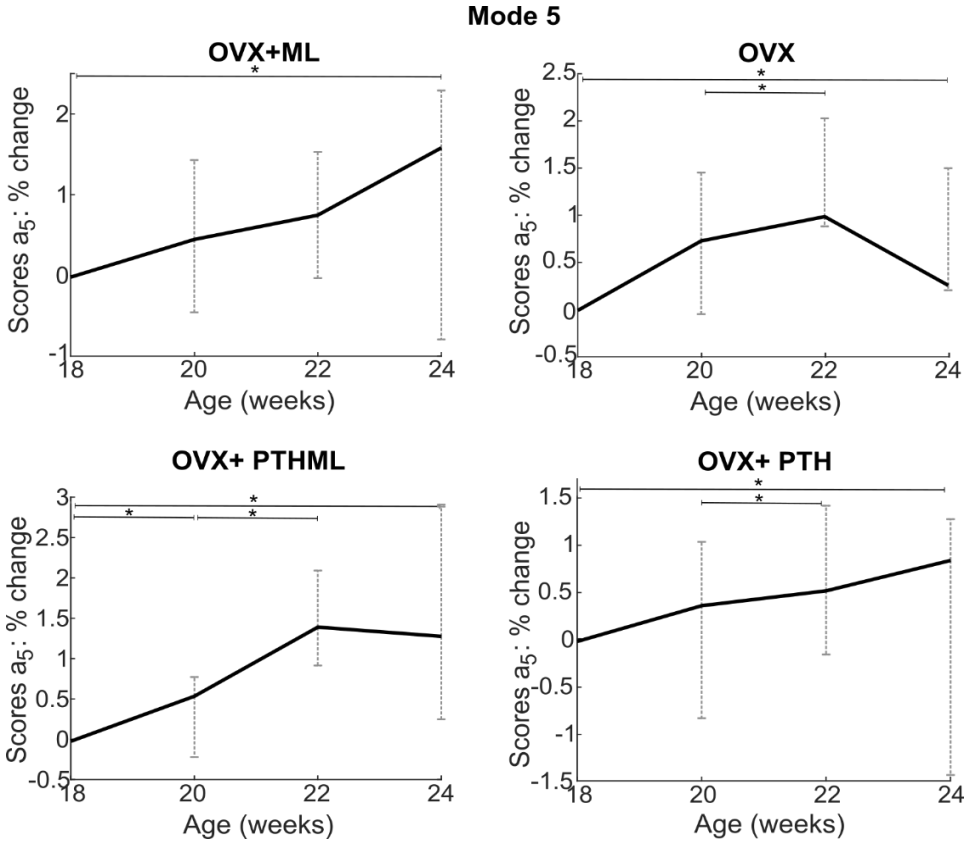


Figure 4.15: Normalised score changes for Mode 5 in all groups.

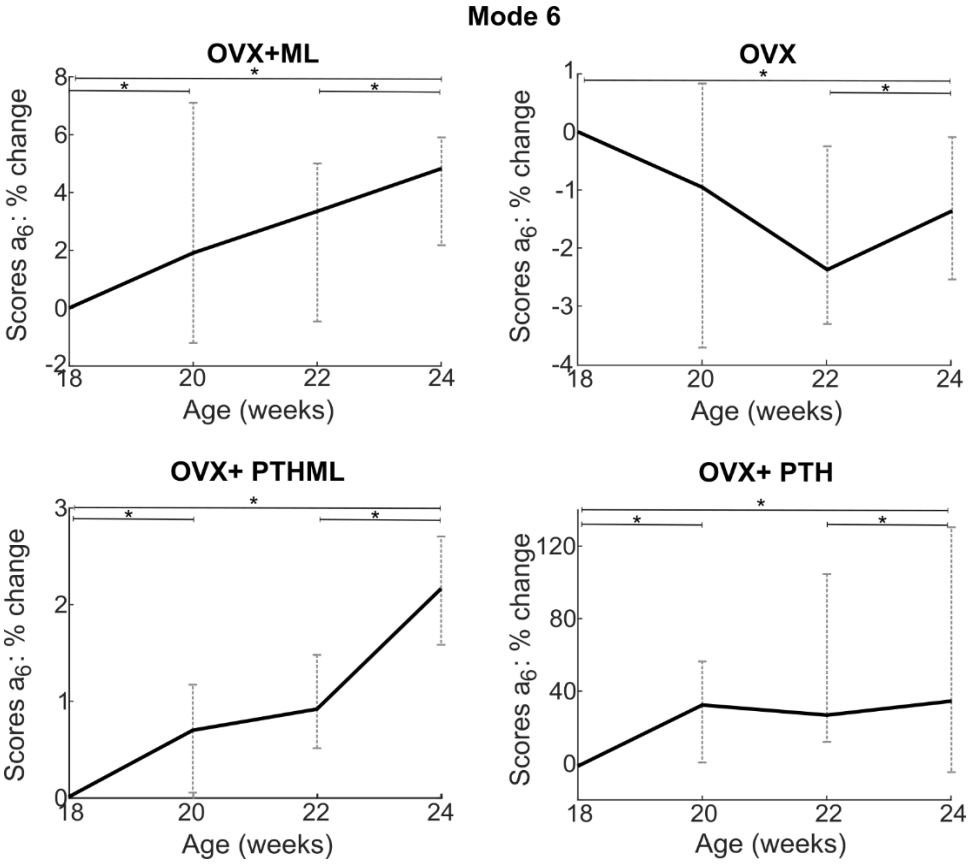


Figure 4.16: Normalised score changes for Mode 6 in all groups.

The 3D mode-specific profiles of the surface changes for Modes 3-6 between week 18 and week 24 were used to quantify the overall 3D geometric changes associated with the treatment and the disease (Figures 4.17, 4.18, 4.19 and 4.20). For "OVX+PTHML", Mode 3 revealed significant (statistical test on scores for 18-24 weeks, $p<0.05$ only for "OVX+PTHML", Figure 4.13) thickening in the posterior-lateral compartment in the proximal end with maximum bone formation at the periosteum equal to 0.1 mm (Figure 4.17). For the same group, with the only significant temporal score changes (Figure 4.14), Mode 4 described the thickening of the anterior crest in a large area of the proximity with 0.065 mm bone formation at the periosteum and concurrent 0.03mm bone deletion in the endosteum (Figure 4.17).

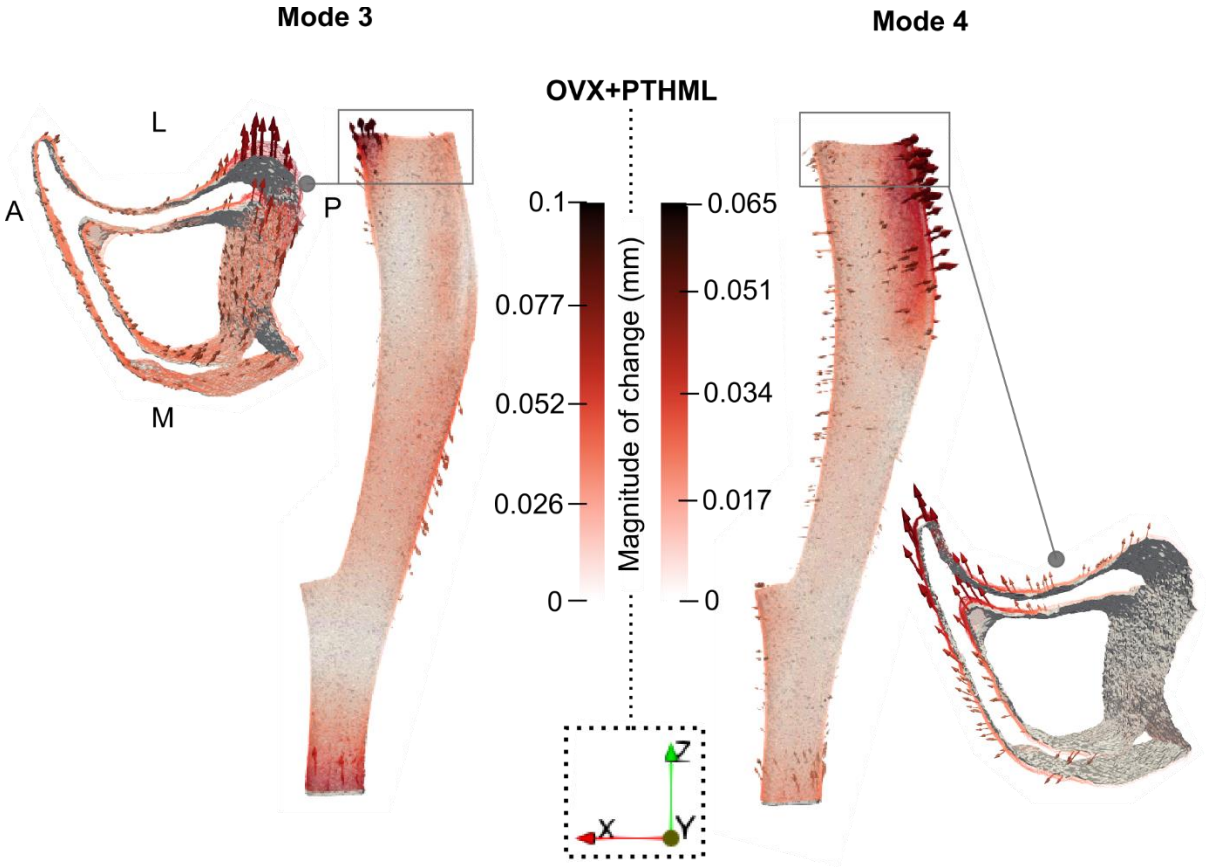


Figure 4.17: Median changes in endosteum and periosteum shapes due to Modes 3 and 4 between weeks 18 and 24 in the "OVX+PTHML" group. The directions of change at different locations are denoted by the arrows, and a redder arrow indicates a relatively larger change. The median bone profile at week 18 is shown as a solid grey surface, whilst the median profile at week 24 is given as a coloured wireframe. The zoomed-in proximal sections share the same orientation denoted as A (anterior), L (lateral), P (posterior) and M (medial). These 3D surfaces changes are defined by the mode vector direction shown in Figure 4.11 and the score changes in Figure 4.13 and 4.14.

Mode 5 described posterior expansion between weeks 18 and 24 for all groups (representative 3D profile in Figure 4.18). However, the magnitude of this change is different for the different groups, with "OVX+ML" exhibiting the largest periosteal expansion of 0.06mm, followed by "OVX+PTHML" with

0.048mm and “OVX+PTH” with 0.026mm. Finally, the “OVX” group holds the smallest periosteal bone formation equal to 0.0075mm.

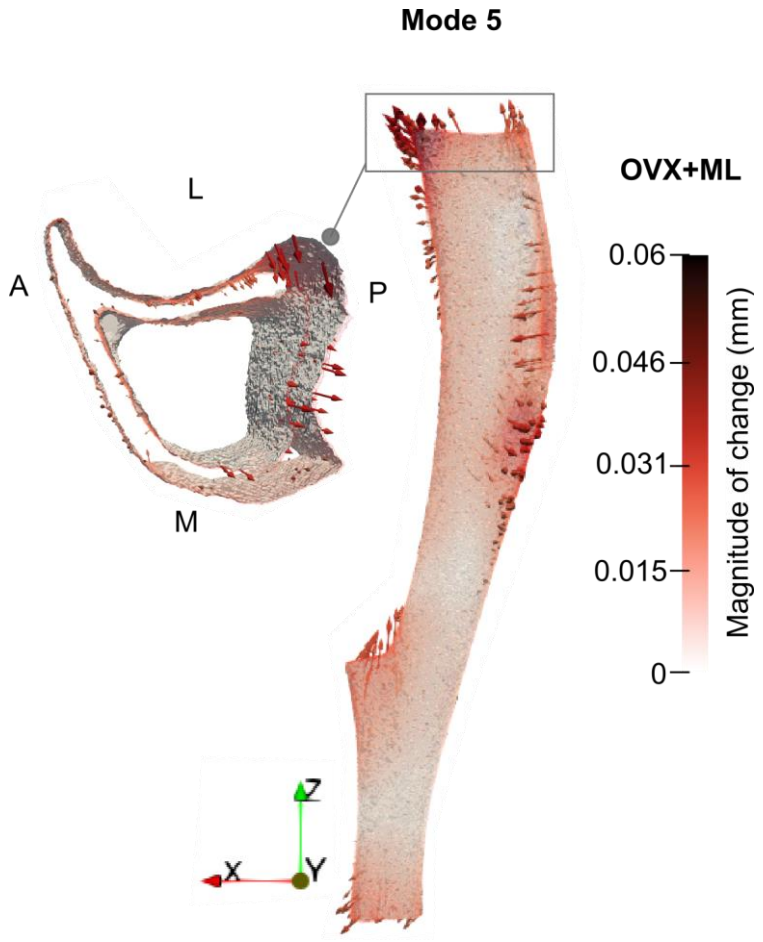


Figure 4.18: Median changes in endosteum and periosteum shapes due to Mode 5 between weeks 18 and 24 in all groups. The directions of change at different locations are denoted by the arrows, and a redder arrow indicates a relatively larger change. The mean bone profile at week 18 is shown as a solid grey surface, whilst the mean profile at week 24 is given as a coloured wireframe.

Mode 6 indicated significant geometric changes for all groups, as found from the statistical tests of the scores between weeks 18 and 24 (Figure 4.16). These changes were the medial expansion at the proximal end for all treatment groups, as shown by the representative profile of mode changes in the “OVX+ML” group (Figure 4.19). However, the negative score change over the age in the “OVX” group was translated to the shrinkage of this feature in the 3D bone profile (Figure 4.20). The combination of treatments in the “OVX+PTHML” indicated the highest magnitude of change up to 0.068mm. Then, monotherapy of mechanical loading exhibited 0.066mm medial periosteal bone formation. Lastly, PTH showed the lowest magnitude of change across all treatment groups with 0.015mm. The disease group exhibited the opposite change with medial shrinkage up to 0.023mm.

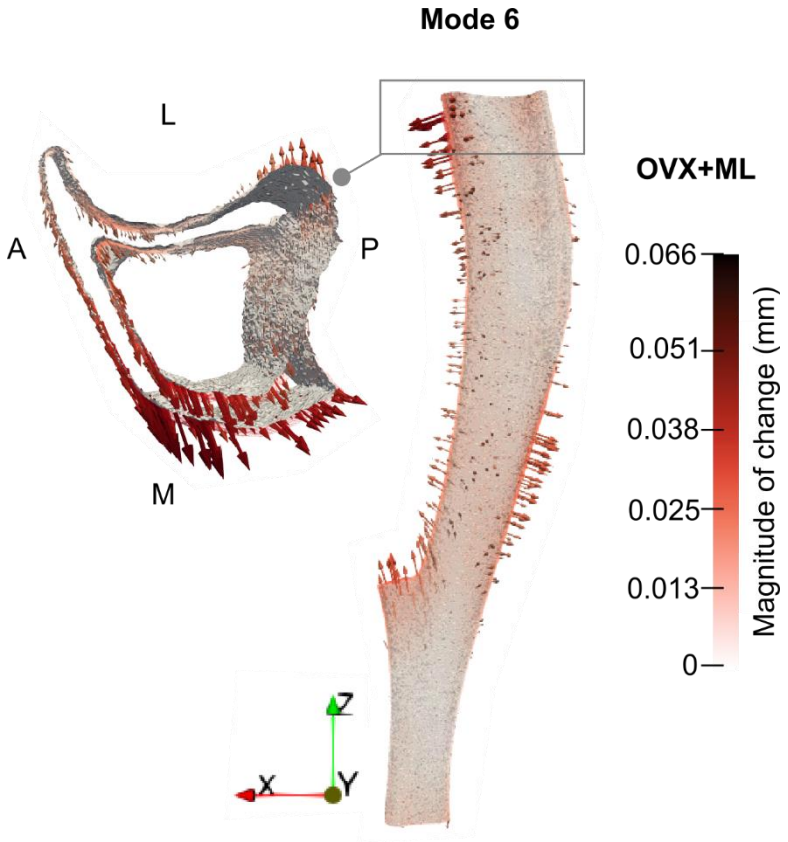


Figure 4.19: Median changes in endosteum and periosteum shapes due to Mode 6 between weeks 18 and 24 in all treatment groups. The directions of change at different locations are denoted by the arrows, and a redder arrow indicates a relatively larger change. The mean bone profile at week 18 is shown as a solid grey surface, whilst the mean profile at week 24 is given as a coloured wireframe.

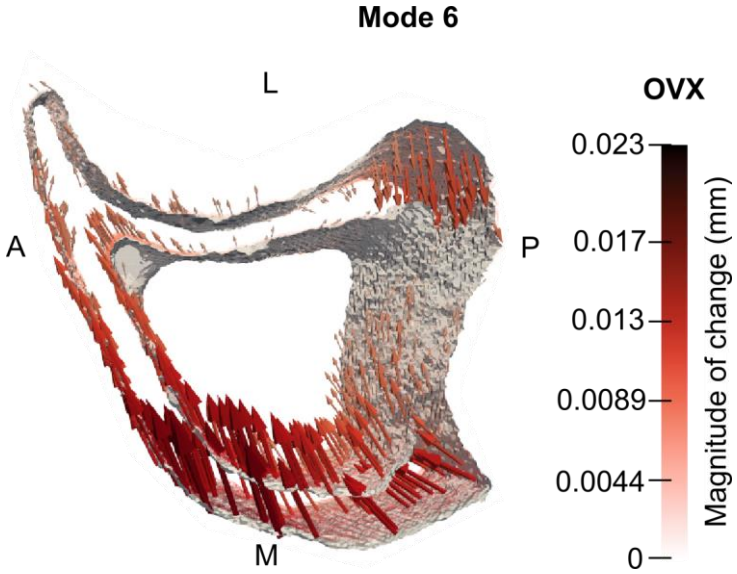


Figure 4.20: Median changes in endosteum and periosteum shapes due to Mode 6 between weeks 18 and 24 in the "OVX" group.

4.4 Discussion

In vivo longitudinal imaging has provided a valuable means to monitor bone turnover over time and investigate how disease and treatment affect bone structure. However, the state-of-the-art approaches used for the geometric assessment to quantify the bone shape changes are underdeveloped. Based on scalar variables, they hold major assumptions about the bone geometry that reduce the precision and accuracy of the findings. The present study implements a novel PCA-based model to extract a complete statistical representation of the entire tibia cortex and provide bone remodelling patterns in 4D (space and time), testing different osteoporosis treatment strategies.

4.4.1 Investigating systematic variations

This study considered one disease cohort (OVX) and three treatment cohorts including administration of PTH (1-34), *in vivo* mechanical loading and combined pharmacological and mechanical treatment. Longitudinal microCT images from these cohorts were processed using the PCA framework developed previously (detailed in Chapter 3) for a small section of the bone. The pipeline was applied as designed and only one major adjustment to the rigid registration protocol was needed. PCA of the extracted shape data did not present significant computational challenges (computation time in seconds), but the geometry correction and the deformable registration were the most time-consuming steps. In the current study, task parallelisation and usage of the HPC system for deformable registration substantially aided the preprocessing of the data.

The PCA modes showed that the first six independent variant features described 91% of the shape variations (Figure 4.2). Mode 1 and 2 described more globalised variations across a large portion of the bone shape and explained 82% of the total variance already (Figure 4.3). In contrast, the subsequent 4 modes described more localised volumes in the entire cortex (Figure 4.11), justifying the dramatic decrease in the portion of variations explained (Figure 4.2).

Mode 1 and 2 primarily described variations in length, as indicated by the direction of the mode vectors, which predominantly followed the z-axis in the length direction. The post-processing of the score distribution did not reveal any statistically significant changes over time for Mode 1, leading to the

conclusion that Mode 1 describes random variance among groups and ages. However, previous studies have pinpointed the changes in length over time (Roberts et al., 2024) and therefore the interpretation of Mode 1 to be only a natural variability effect is not expected. That could possibly be a result of the large range of the score values (Figure 4.9) and subsequently of the score changes (Figure 4.4, 4.5). In contrast, Mode 2 described a temporal shape change at the distal end with significant score changes for the treatment groups and almost all age pairs. Notably, the score changes were monotonic for all treatment groups, with only the last period of 22-24 weeks exhibiting stability or a slightly catabolic trend. This difference in the gradient of change in all treatment groups might be an effect of the withdrawal period, where no treatment was applied. The surface changes described by Mode 2 indicated a distal expansion describing the increase in length (Figure 4.8). However, the distal expansion varied across the treatment groups, being equal to 0.14 mm, 0.042 mm and 0.22 mm for "OVX+ML", "OVX+PTHML" and "OVX+PTH". These differences in magnitude could be possibly due to the variation in the position of the distal tibia-fibula separation, where the tibia samples were cropped. For the disease group, the score changes underscore a monotonic change suggesting distal expansion as in the treatment groups, but with no statistical significance. This result highlighted the disease effect in the length increase, similar to a previous study which showed smaller tibia lengths in ovariectomised mice over treated ones (Roberts et al., 2024).

Mode 3, 4, 5, and 6 described localised variations prominently around the proximal volume of the tibia (Figure 4.11). These variations explain previous findings which report higher bone turnover in proximal sections compared with the distal volume of the tibia including the tibia-fibula fusion (Monzem et al., 2023). The high turnover at the proximal compartments might be somewhat associated with the existence of the growth plate in rodent long bones, as also shown in (Fan et al., 2008), and the integration of trabeculae into the cortical shell as a longitudinal growth change mechanism (Cadet et al., 2003).

In this study, the statistical tests of the temporal score changes of Mode 3 have shown that the proximal variation at the posterior-lateral aspect is a time-dependent change only for the combination of the pharmacological and physical treatment. However, considering the mean normalised score change,

monotonical changes over time were also found in OVX+ML groups. That could possibly suggest that mechanical stimulus provokes bone formation at the posterior-lateral aspect. Moreover, when it is combined with pharmacological treatment this change is exaggerated. However, solo administration of PTH seemed to have almost no effect in the posterior-lateral compartment (small score changes. Figure 4.13), whereas osteoporosis ("OVX" group) leads to its slight thinning as shown by the small decrease in scores (Figure 4.13). Similar geometric changes and increased anabolic effects due to the combination of treatments were also previously found by the mechanoregulation-based model applied to the same bone data but younger, 18−22 weeks old (Cheong et al., 2021). Specifically, no important effect of solo PTH and bone deletion due to OVX leading to thinning of the lateral part was shown in the 3D illustration of bone formation and deletion in Cheong et al., (2021).

Another variant feature that the PCA model described is the anterior crest at the proximal area, as captured by Mode 4 and shown in Figure 4.11. Analysis of scores across groups and ages revealed that expansion and thickening of this feature occurred only when the mice received the concurrent treatment between the periods 20-22 and 22-24 weeks, signifying that the monotherapy does not have similar efficiency. This could be potentially attributed to the increase in Bone Mineral Density (BMC) values in a similar bone section during the 20–22-week period, as previously measured (Roberts et al., 2020). Although the mechanical loading in mice showed increased BMC in the proximal anterior compartment (Cheong et al., 2021), this was not translated in a significant change of this feature in the "OVX+ML" group, as described by Mode 4.

Mode 5 described temporal variations in the proximity of the bone at the posterior compartment in all treatment groups with substantially higher magnitude than the one in the disease. Similarly, Cheong et al., (2021) have previously shown increased localised turnover due to treatments. The posterior expansion described by Mode 5 could also explain the increased proximal cross-sectional area as in Monzem et al., (2023) for similar treatments and similar age of mice. Notably, Mode 5 exhibited slight thickening and rotation of the anterior crest, but the latter has not been previously reported elsewhere.

Finally, Mode 6 exhibited the medial endosteal deletion and medial periosteal formation, resulting in

expansion in the medial compartment over time in all treated mice, consistent with similar findings in Sugiyama et al., (2008). Mode 6 also described thickening in the anterior crest and the lateral compartment in the diaphysis, with the anterior crest thickening being prominent around the midshaft. Previous studies have shown remodelling in both the lateral and anterior of the midshaft section when tibiae are loaded, showing the bone responses in both tension and compression (Sugiyama et al., 2008; Rooney et al., 2023). The study in Chapter 3 also confirms these findings. Notably, all surface changes described by Mode 6 are the highest in magnitude in the "OVX+PTHML", followed by the "OVX+ML" and "OVX+PTH" groups. In contrast, the "OVX" group exhibits the opposite changes with a low magnitude.

Overall, the current framework presented the increased anabolic effects of the combined PTH and mechanical loading, which is in line with previous experimental murine studies (Sugiyama et al., 2008). Although there are big ranges in the score changes, leading to no statistically significant results, the model still gives some monotonic changes over time. The main advancement of the proposed model is that it provides specific, compact (local PCA mode vectors) and quantitative assessment of the 3D geometric changes of the entire 3D bone cortex. Additionally in contrast with the standard morphometric analysis, due to the orthogonality of the modes, the model ensures that the geometric changes are independent of each other and describe different portions of shape variations.

4.4.2 Limitations

This study indicated high variations in the score distributions and temporal score changes, especially for Mode 1 in the "OVX+ML" and "OVX+PTH" groups (Figure 4.4A, 4.5A). Figure 4.9 showed a high ranges in the score distributions for Mode 1 and 2 for all cohort groups. Possible reasons for the high variations in scores of those modes that describe axial variation in the length direction could be the differences in distal cropping and 3D positioning of the tibia samples, derived from the image rigid registration.

Firstly, the bones were cropped to exclude the tibia-fibula separation due to its topological heterogeneity. The cropping was accomplished automatically by detecting the first evidence of the distal

tibia-fibula separation. However, the longitudinal location of this feature could vary among observations, resulting in conservative or more extensive cropping between 25-40 slices (0.16 mm discrepancy). Consequently, modes that describe longitudinal variations in the distal end, such as Mode 2, encompass a combination of length variation and positional variation of the distal tibia-fibula separation.

Another reason for the large score variability could be the 3D positioning of the bone samples. Figure 4.10 highlighted that combined Mode 1 and Mode 2 described uniform vector fields with constant orientation in the uniaxial direction for "OVX+ML", "OVX" and "OVX+PTH" groups. The uniformity of the vector field and the small ranges in magnitude made the combined Mode 1 and Mode 2 challenging to biologically interpret. Therefore, they were considered to describe rigid body movements in the length direction, possibly suggesting systematic positional misalignments in the z-axis between groups. That explains the inter-group score differences highlighted in Figure 4.9. Positional misalignment could potentially arise from poor rigid registrations, challenged by inherent prominent differences in length between groups and ages. This analysis showed that the misalignment in the z-axis was mistreated as a longitudinal variation and coupled with the length variations, generating these large SDs in the score distributions. Finally, the effect of moderate alignment of the bone samples on the PCA model is possibly high due to the relatively small sample size (6 for each group and age).

Applying PCA-based modelling to an entire bone structure while considering multiple sources of variations (disease and various treatments) presents challenges. The major challenge arises from the prominent structural differences between the bone samples of different ages and cohort groups that provoke moderate positional alignment between them. This is because of the increased heterogeneity across the entire volume (vs bone segment) and the high dimensionality of the data with increased differences between subgroups of the examined population. This attribute was indicated by the rigid movements described by the combined Mode 1 and Mode 2. Since these misalignments are typically blended in the mesh data, PCA mistreats them as form differences (Baab, 2013). With image rigid registration used as a superposition approach to align all tibia samples in the 3D space, its errors and

the impact of different protocols on the PCA model should be evaluated in future studies.

Although this PCA model didn't describe length changes as expected, it described cross-sectional localised variations compactly, summarising findings separately reported by previous studies. That showed the potential of the 3D statistical models of bones to provide a new more precise strategy in quantifying localised shape changes due to external interventions for improving osteoporosis treatments (Grassi et al., 2021). Nevertheless, this study showed the need for such PCA models to evaluate their robustness to the instance alignment.

4.5 Conclusions

In conclusion, this study introduced a novel PCA-based approach to assess the shape variations of the entire mouse tibia cortex among a population of mice subjected to ovariectomy and various treatment strategies. While the statistical shape model somewhat struggled to capture axial variations, associated with growth changes, it successfully described localised and cross-sectional variations associated with treatment and disease to some extent. The latter were mostly detected on the proximal end with prominent posterior-lateral thickening and medial expansion due to the treatments. It also showed the increased anabolic effect of the combined PTH and mechanical loading leading to higher magnitudes of surface changes. However, due to the large standard deviations of the modal scores, the reported geometric changes illustrate moderate reliability, underscoring the need to evaluate the instance alignment effect on the assessment of the 3D shape variations in future work.

4.6 References

Aslan, D., Andersen, M. D., Gede, L. B., de Franca, T. K., Jørgensen, S. R., Schwarz, P., & Jørgensen, N. R. (2012). Mechanisms for the bone anabolic effect of parathyroid hormone treatment in humans. *Scandinavian Journal of Clinical and Laboratory Investigation*, *72*(1), 14–22. https://doi.org/10.3109/00365513.2011.624631

Baab, K. L. (2013). The impact of superimposition choice in geometric morphometric approaches to morphological integration. *Journal of Human Evolution*, *65*(5), 689–692. https://doi.org/10.1016/j.jhevol.2013.07.004

Birkhold, A. I., Razi, H., Duda, G. N., Checa, S., & Willie, B. M. (2017). Tomography-based quantification of regional differences in cortical bone surface remodeling and mechano-response. *Calcified Tissue International*, *100*, 255–270. https://doi.org/10.1007/s00223-016-0217-4

Bouxsein, M. L., Boyd, S. K., Christiansen, B. A., Guldberg, R. E., Jepsen, K. J., & Müller, R. (2010). Guidelines for assessment of bone microstructure in rodents using micro--computed tomography. *Journal of Bone and Mineral Research*, *25*(7), 1468–1486. https://doi.org/10.1002/jbmr.141

Cadet, E. R., Gafni, R. I., McCarthy, E. F., McCray, D. R., Bacher, J. D., Barnes, K. M., & Baron, J. (2003). Mechanisms responsible for longitudinal growth of the cortex: coalescence of trabecular bone into cortical bone. *JBJS*, *85*(9), 1739–1748.

Cheong, V. S., Campos Marin, A., Lacroix, D., & Dall'Ara, E. (2020). A novel algorithm to predict bone changes in the mouse tibia properties under physiological conditions. *Biomechanics and Modeling in Mechanobiology*, *19*, 985–1001. https://doi.org/10.1007/s10237-019-01266-7

Cheong, V. S., Roberts, B. C., Kadirkamanathan, V., & Dall'Ara, E. (2021). Positive interactions of mechanical loading and PTH treatments on spatio-temporal bone remodelling. *Acta Biomaterialia*, *136*, 291–305. https://doi.org/10.1016/j.actbio.2021.09.035

Chow, J. W. M., Fox, S., Jagger, C. J., & Chambers, T. J. (1998). Role for parathyroid hormone in mechanical responsiveness of rat bone. *American Journal of Physiology-Endocrinology and Metabolism*, *274*(1), E146--E154. https://doi.org/10.1152/ajpendo.1998.274.1.E146

Fan, W., Crawford, R., & Xiao, Y. (2008). Structural and cellular differences between metaphyseal and diaphyseal periosteum in different aged rats. *Bone*, *42*(1), 81–89. https://doi.org/10.1002/jbmr.2432

Gardinier, J. D., Mohamed, F., & Kohn, D. H. (2015). PTH signaling during exercise contributes to bone adaptation. *Journal of Bone and Mineral Research*, *30*(6), 1053–1063. https://doi.org/10.1002/jbmr.2432

Grassi, L., Väänänen, S. P., & Isaksson, H. (2021). Statistical shape and appearance models: development towards improved osteoporosis care. *Current Osteoporosis Reports*, *19*, 676–687.

https://doi.org/10.1007/s11914-021-00711-w

Jee, W. S. S., & Tian, X. Y. (2005). The benefit of combining non-mechanical agents with mechanical loading: a perspective based on the Utah Paradigm of Skeletal Physiology. *Journal of Musculoskeletal and Neuronal Interactions*, *5*(2), 110.

Khosla, S., & Hofbauer, L. C. (2017). Osteoporosis treatment: recent developments and ongoing challenges. *The Lancet Diabetes & Endocrinology*, *5*(11), 898–907. https://doi.org/10.1016/S2213-8587(17)30188-2

Klinck, J., & Boyd, S. K. (2008). The magnitude and rate of bone loss in ovariectomized mice differs among inbred strains as determined by longitudinal in vivo micro-computed tomography. *Calcified Tissue International*, *83*, 70–79. https://doi.org/10.1007/s00223-008-9150-5

Lu, Y., Boudiffa, M., Dall'Ara, E., Bellantuono, I., & Viceconti, M. (2016). Development of a protocol to quantify local bone adaptation over space and time: Quantification of reproducibility. *Journal of Biomechanics*, *49*(10), 2095–2099. https://doi.org/10.1016/j.jbiomech.2016.05.022

Monzem, S., Valkani, D., Evans, L. A. E., Chang, Y.-M., & Pitsillides, A. A. (2023). Regional modular responses in different bone compartments to the anabolic effect of PTH (1-34) and axial loading in mice. *Bone*, *170*, 116720. https://doi.org/10.1016/j.bone.2023.116720

Oliviero, S., Cheong, V. S., Roberts, B. C., Orozco Diaz, C. A., Griffiths, W., Bellantuono, I., & Dall'Ara, E. (2022). Reproducibility of Densitometric and Biomechanical Assessment of the Mouse Tibia From In Vivo Micro-CT Images. *Frontiers in Endocrinology*, *13*, 915938. https://doi.org/10.3389/FENDO.2022.915938/BIBTEX

Pereira, A. F., Javaheri, B., Pitsillides, A. A., & Shefelbine, S. J. (2015). Predicting cortical bone adaptation to axial loading in the mouse tibia. *Journal of the Royal Society Interface*, *12*(110), 20150590. https://doi.org/10.1098/rsif.2015.0590

Roberts, B.C., Cheong, V.S., Oliviero, S., Arredondo Carrera, H.M., Wang, N., Gartland, A. and Dall'Ara, E. (2024). Combining PTH (1-34) and mechanical loading has increased benefit to tibia bone mechanics in ovariectomised mice. *Journal of Orthopaedic Research*, *42*(6), 1254-1266. https://doi.org/10.1002/jor.25777

Roberts, B. C., Arredondo Carrera, H. M., Zanjani-pour, S., Boudiffa, M., Wang, N., Gartland, A., &

Dall'Ara, E. (2020). PTH (1–34) treatment and/or mechanical loading have different osteogenic effects on the trabecular and cortical bone in the ovariectomized C57BL/6 mouse. *Scientific Reports*, *10*(1), 8889. https://doi.org/10.1038/s41598-020-65921-1

Roberts, B. C., Giorgi, M., Oliviero, S., Wang, N., Boudiffa, M., & Dall'Ara, E. (2019). The longitudinal effects of ovariectomy on the morphometric, densitometric and mechanical properties in the murine tibia: A comparison between two mouse strains. *Bone*, *127*, 260–270. https://doi.org/10.1016/j.bone.2019.06.024

Rooney, A. M., McNeill, T. J., Ross, F. P., Bostrom, M. P. G., & van der Meulen, M. C. H. (2023). PTH treatment increases cortical bone mass more in response to compression than tension in mice. *Journal of Bone and Mineral Research*, *38*(1), 59–69. https://doi.org/10.1002/jbmr.4728

Sugiyama, T., Saxon, L. K., Zaman, G., Moustafa, A., Sunters, A., Price, J. S., & Lanyon, L. E. (2008). Mechanical loading enhances the anabolic effects of intermittent parathyroid hormone (1–34) on trabecular and cortical bone in mice. *Bone*, *43*(2), 238–248. https://doi.org/10.1016/j.bone.2008.04.012

van 't Hof, R. J., & Dall'Ara, E. (2019). Analysis of Bone Architecture in Rodents Using Micro-Computed Tomography. *Methods in Molecular Biology (Clifton, N.J.)*, *1914*, 507–531. https://doi.org/10.1007/978-1-4939-8997-3_28

Zhang, J., Gao, R., Cao, P. and Yuan, W. (2014). Additive effects of antiresorptive agents and exercise on lumbar spine bone mineral density in adults with low bone mass: a meta-analysis. *Osteoporosis International*, *25*, 1585-1594. https://doi.org/10.1007/s00198-014-2644-2

Chapter 5

Rigid registration protocol influences the assessment of the 3D geometric variations in mouse tibia

ABSTRACT

The outcome of image-based bone shape analyses, such as PCA and Statistical Shape Modelling, depends on the alignment of instances. As regards the previously introduced PCA-based framework of mouse tibia geometry assessment starting from microCT images, there is currently no gold standard for instance alignment technique (rigid registration). The study presented in this chapter evaluates the influence of two rigid image registration protocols on changes in mouse tibia geometry due to OP treatment effect between two ages as assessed by the above framework. In the first protocol, a pair of co-registered age-specific references are defined which are used to register the images at different ages; in the second protocol, samples at both ages are registered to a global reference, followed by a mouse-specific registration correction. Comparative analyses are conducted with regards to three aspects, image positional disagreement, differences in individual and mean shape discretisation, differences in PCA mode shapes and their relation to treatment effects. The highest difference between the two registration protocols was found in determining the axial position of the bone of an average magnitude equal to 0.16 mm. The first two PCA modes for the first registration protocol, but not the second, estimated negligible proximal expansion in the diseased group and negligible distal expansion in the treated group associated with moderate changes in length, in disagreement with past studies. Additionally, one mode in the first protocol described a rotation effect which is not biologically interpretable. Overall, the choice of rigid registration protocol leading to different instance alignments significantly influences the inference drawn from the PCA model. The main conclusion is that the registration protocol including the mouse-specific registration correction leads to better match with biomechanically meaningful inferences.

5.1 Introduction

Murine models are widely used to study how bone remodels and adapts to external interventions over time and space. In the last few decades, *in vivo* imaging techniques have considerably advanced such that it is now possible to monitor spatiotemporal changes in bone structures longitudinally (Dall'Ara et al., 2016). To assess such changes or differences in the full three-dimensional bone shapes, bone images acquired at different time points and/or on different animals need to be compared. It is necessary to exclude from shape comparisons differences such as of position, orientation and ensure mutual alignment of image instances (Campbell et al., 2014). Such differences can arise due to the irregular shape of bones and the practical limitations of imaging such as positioning of the specimens in the imaging equipment. These are typically excluded using alignment techniques including rigid image registration. Therefore, rigid image registration is an important and often the first step in image-based analysis of bone shape.

An image-based framework for analysis of mouse tibia shapes was previously presented in Chapter 2. That framework, starting from *in vivo* microCT images, identified Principal Component Analysis (PCA) modes that are separately related to changes with time due to growth, disease progression or treatment effects. In Chapter 3 the framework was applied to assess changes in the whole mouse tibia. Although it is expected that the outcomes of such pipelines would depend on the rigid image registration protocol (Baab, 2013), a gold standard to rigidly register mouse bone images does not exist. The framework presented in Chapters 3 and 4 implemented a variation of the protocol proposed by (Lu et al., 2016a) to align *in vivo* microCT image instances of mouse tibia. However, the first PCA modes detailed in Chapter 4 described a rigid movement of the tibia along its length, which is not easily explained biomechanically. The individual mouse scores of these modes a suggest combined effect of instance misalignments in the axial direction, which could be a spurious outcome of the rigid image registration protocol.

The aim of the present chapter is to evaluate the influence of rigid image registration protocols on the PCA framework described previously. This will aid the standardisation of rigid registrations protocols

generally. It will also address a specific gap that exists in the use of *in vivo* images of whole murine bone. Past studies performing image-based bone shape analysis (Campbell et al., 2014) are mostly limited to the analysis of specific regions of proximal or midshaft tibia and/or used *ex vivo* images (cross-sectional studies), where the challenge of positioning is relatively lower. As such it is not clear whether image registration protocols used in previous studies are fit for purpose when analysing *in vivo* images of whole murine bones.

The study presented in this chapter considers two different image rigid image registration protocols for aligning the *in vivo* mouse tibia images. The protocols differ in the choice of sequences of reference shape that are used to align the images to these using rigid image registration. The use of reference shape is common in the literature for PCA-based bone shape analysis (Castro-Mateos et al., 2014; Grassi et al., 2021; Li et al., 2009) but their choice and/or sequence of application is not standardised. The framework for whole mouse tibia analysis demonstrated in Chapter 3 is performed separately for both registration protocols. This leads to two PCA models of bone changes following concurrent pharmacological and physical treatment in an osteoporotic mouse model. The outcomes at distinct steps of the PCA-based framework are compared, thus describing the influence of the two protocols on the changes in the 3D surfaces of mouse tibia and on the corresponding PCA modes and scores.

5.2 Materials and methods

This section describes the image database, two rigid registration protocols investigated and the comparative analysis approach. Particularly, the two registration protocols, denoted as “Registration 1” and “Registration 2”, are performed at the first step of the pipeline (Figure 5.1) and two databases processed differently occurred in the end of the pipeline. The two databases were compared at three distinct steps of the framework, i.e., after step ‘1. Image Alignment’, step ‘4. Mapping bone images via deformable registration’ and extraction of the mesh samples and finally PCA model construction, referred to henceforth as Objectives 1, 2 and 3.

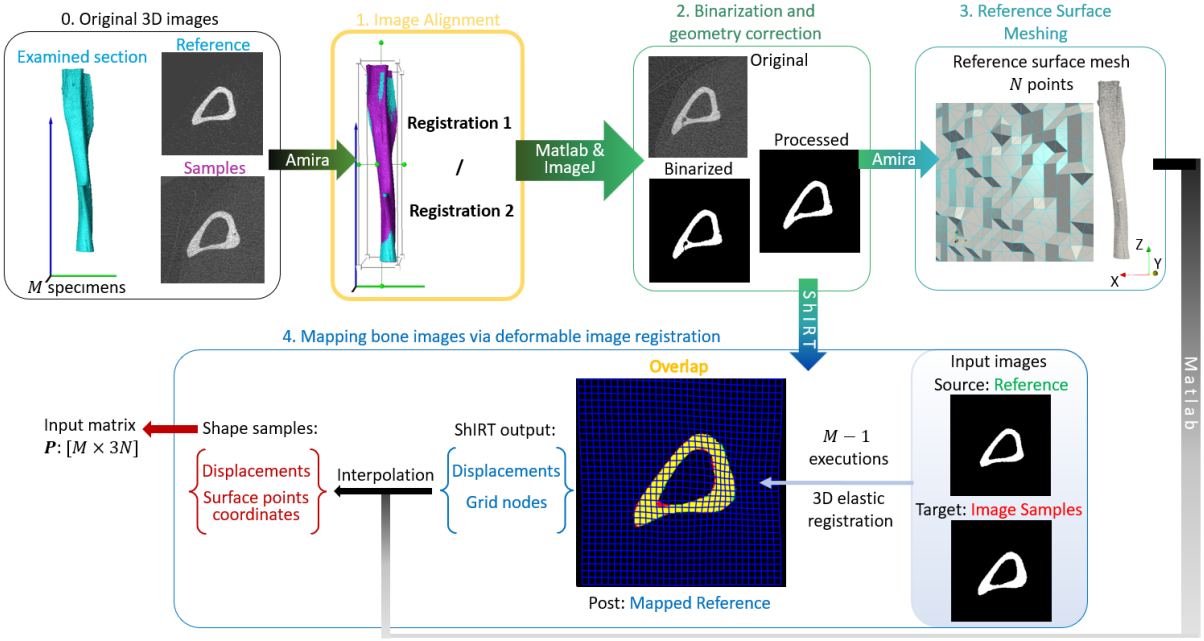


Figure 5.1: Flowchart of the framework incorporating the use of two different rigid registration protocols, i.e., "Registration 1" and "Registration 2" in step 1. Image Alignment (highlighted with bold and yellow).

5.2.1 Animals and interventions

Data used in this study were collected from previous longitudinal murine studies (Roberts et al., 2019;Roberts et al., 2020). The current study focused on two groups of mice: the "OVX" group, which underwent ovariectomy at the age of 14 weeks and received no treatment; the "OVX+PTHML" group, which underwent ovariectomy at the age of 14 weeks, followed by a combination of pharmacological (PTH injections) and biomechanical (external *in vivo* mechanical loading) treatments in between 18 and 22 weeks of age (Figure 5.2).

In this study two time points were analysed: the beginning of the treatment at 18 weeks of age, and the end of the study, i.e. 24 weeks of age.

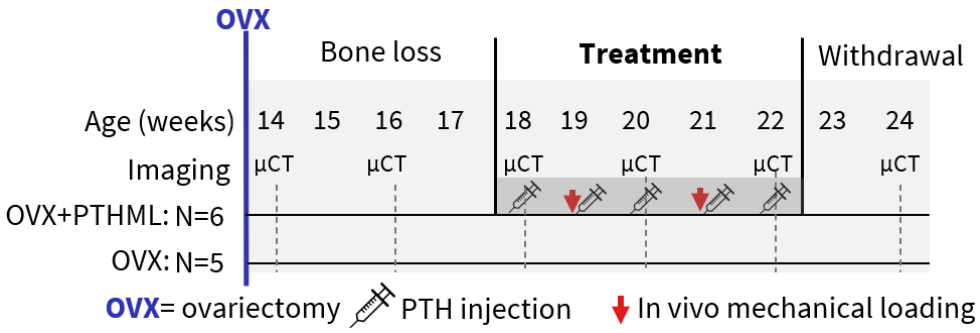


Figure 5.2: Examined groups of mice and the different interventions.

5.2.2 Tested registration protocols

The steps of the PCA-based framework are shown in Figure 5.. The first step of the framework is the alignment of the images using rigid registration. In the current study, either of two rigid registration protocols are used in this first step. Application of the PCA-based framework on the image database described in the previous section leads to (i) two differently registered images of each bone sample, (ii) two different discretisations (meshes) of each bone sample surface, and finally, (iii) two differently processed PCA databases, PCA modes and scores. Later sections 5.2.4–5.2.6 detail how these outcomes of the framework are compared.

The two registration protocols are modifications of the registration protocol as described in Lu et al., (2016). Therein, a reference image was defined and manually aligned to the anatomical coordinate system. All image samples (all mice, all groups, and all ages) were registered to this reference. The two modifications of this registration protocol are described below and illustrated in Figure 5.3.

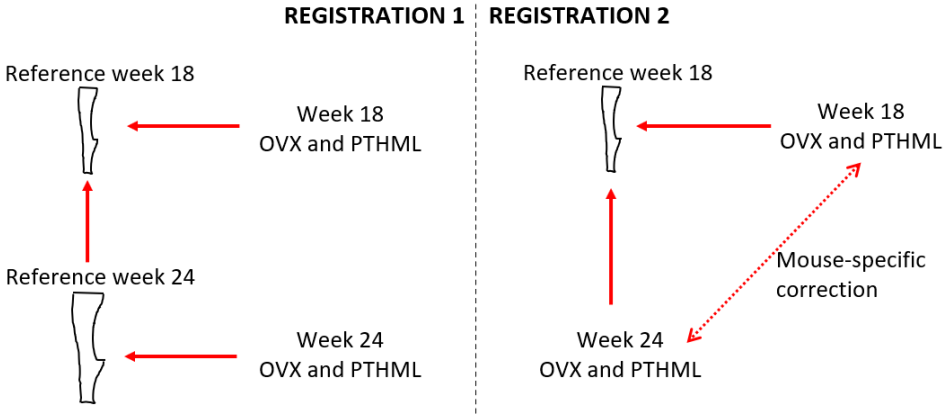


Figure 5.3: The two registration protocols, "Registration 1" and "Registration 2", used in this study.

The first examined protocol, "Registration 1", used one mouse-specific reference for each age. The choice of age-wise references was motivated by the presence of prominent age-related differences, which could challenge the registration between bone image samples from different ages. This registration protocol has been used and previously detailed in Chapter 4. Briefly, the image of one mouse at weeks 24 was registered to its image at 18 weeks of age. The registered image and the reference image (week 18) of this reference mouse are referred to here onward as "Reference week 24" and "Reference week 18", respectively. The images of the remaining mice (in both "OVX" and "OVX+PTHML groups) at any age were registered to the corresponding reference of the same age.

In the second protocol, referred to as “Registration 2”, all bone sample images were first visually inspected to check that the slices at the top and bottom extents contained similar anatomical features. This was necessary because a large variation in the z location of the distal tibia–fibula separation was reported previously in Chapter 4, and the fully automated cropping algorithm led to a large variation in the number of slices cropped. Therefore, extensive cropping could lead to exclude different volumes of tibia. For this reason, the outcome of the algorithm was inspected and at times of automatic extensive cropping that the algorithm suggested, manual cropping was instead used. Next, the image of the mouse at week 18 referred to as “Reference week 18” in the first protocol was considered here as the reference image, thus ensuring the same image origin and dimensionality and same bone geometry across registration protocols. Next, all remaining images (from all mice and ages) were rigidly registered to the reference image. The registered images at 24 weeks were compared with the registered image at 18 weeks of age in a mouse-specific manner, visually checking against spurious misalignments (typically in the axial direction). This step was also confirmed by re-registering the image at 24 weeks to the mouse-specific registered image at 18 weeks. When the transformation coefficients of such registration were found to be large, it indicated a spurious misalignment, and the new re-registered image at 24 weeks was instead used in the following analysis.

The primary reference image was taken from an independent group of mice (“OVX+ML” group). In this group, mice were subjected to external mechanical loading at weeks 19 and 21, after ovariectomy at week 14, but without injecting PTH. The choice of an independent group avoided bias towards any cohort of the examined population.

Prior to any registration, manual translation and rotation were performed to align the image sample as closely as possible (confirmed visually) to the reference image. All registrations were performed in Amira (Amira 3D 2022.2) using an optimization algorithm and with the following settings:

- Normalized Mutual Information as the similarity metric,
- Lanczos method for interpolation, rigid transform with rotation enabled,
- registration in 3D without realignment,

- extensive direction as optimizer type,
- initial and final optimizer steps set to 50 and 0.098, respectively,
- gradient optimizer with finest level equal to 1,
- tolerance -0.0009,
- threshold outside 0.2,
- 256 and 257 histogram bins for reference and model (image to be registered), respectively,
- 'Bell' resampling interpolation used for both model and reference,
- coarsest resampling of 1, -1 and 1 for 1 for x, y and z respectively, and
- the option to ignore finest resolutions enabled.

5.2.3 3D shape analysis for assessing treatment effects

More details about the following steps of the proposed PCA framework can be found in Chapters 3 and 4. Briefly, the registered images (Cheong et al., 2021) were binarized and corrected to enforce topological equivalence to annual cylinder. The binarized reference image was mapped to each image sample, and the yielded displacement field was interpolated to the reference surface mesh to provide the individual meshes of each remaining bone sample. Location coordinates and displacement field are expressed in a Cartesian coordinate system with x, y and z axis identified with the anterior-posterior, medial-lateral and distal-proximal directions, and alternatively with the horizontal, vertical and axial directions in the 3D images (see Chapter 2, Figure 2.5 and Chapter 3, Figure 3.4). The displaced point of the reference surface mesh to any other bone observation was then decomposed into orthogonal modes applying Principal Component Analysis. Following this, the PCA scores were tested for statistically significant differences between cohorts and ages and revealed differences across the inter-group and time domain. If a statistically significant score change was found, and with similar magnitude for both cohort groups, then the corresponding mode was characterised to describe growth change. If the score changes were different between groups, then coupled disease and treatment effects were found. In this case, if the magnitude of the score change in the treatment group was higher than the one in the "OVX" group, then the treatment was defined to exaggerate the disease. In the opposite case, when the change magnitude was smaller in the treatment than in the disease group, then the treatment

opposed the disease progression. Finally, if there was no significant change in the disease group but there was in the treatment, then the mode was considered to describe a pure treatment effect.

5.2.4 Objective 1: 3D position disagreement of the registered images

The discrepancy in the positioning of the same bone image sample after undergoing the different rigid registration protocols (Figure 5.1, after 1) was quantified. The two differently registered images of each sample were co-registered, and the corresponding transformation matrix was computed. This matrix provides the translation in x, y and z directions and the rotation along the x, y and z axes required to align the geometry of "Registration 2" to geometry of "Registration 1" (Figure 5.4), and therefore quantifies their relative 3D position disagreement, describing differences in the location on the three axes and in the orientation along the three axes.

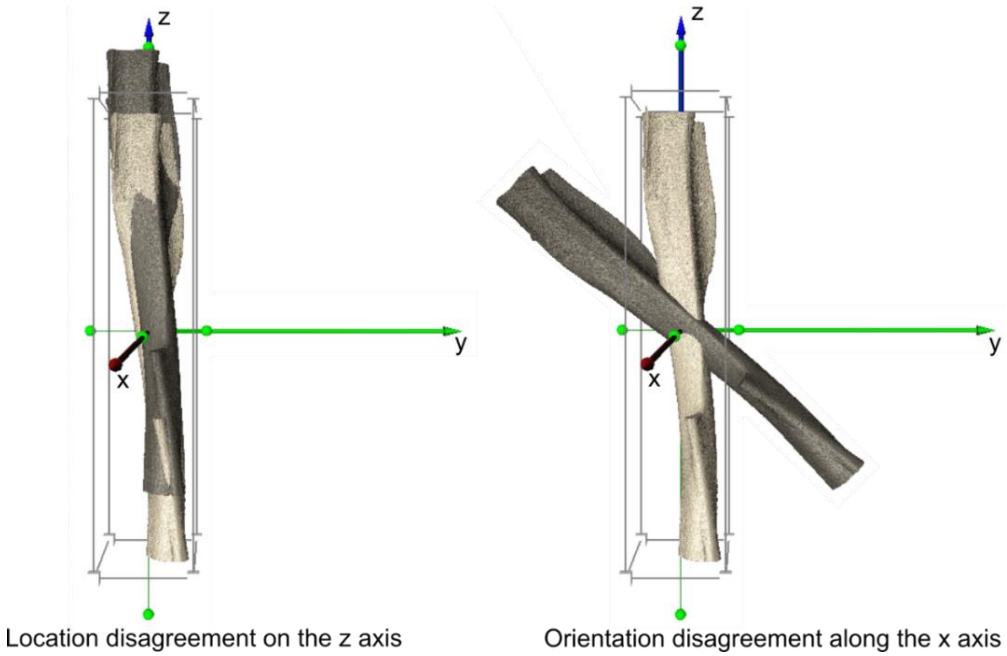


Figure 5.4: Schematic of differences in the 3D positions of differently registered images, defined as disagreement in the uniaxial location on the axes and in the orientation along the axes. Two examples are illustrated. Location disagreement on the z-axis, meaning that the centre of one image bounding box of the light grey bone- aligned with the centre of the coordinate system- has a different z-coordinate from the centre of the image bounding box of the dark grey bone. Orientation disagreement along the x-axis, meaning that the bounding box of the one image is oriented differently along the x-axis, with this example showing 45^{o} degree rotation anticlockwise (negative sing) on the zy plane.

5.2.5 Objective 2: 3D shape differences in meshes

Differences in rigid registration protocol resulted in a differently registered image for the same bone sample, thus altering its mapping to the reference bone image and a different bone sample surface mesh

(Figure 5.1, after step 4). As the bone sample surface mesh was obtained by deforming the reference bone surface mesh, all meshes had the identical number of nodes and face–node connectivity. Thus, differences between the surface meshes of the same bone sample were quantified by computing the displacements of each node from a mesh in "Registration 2" to the other in "Registration 1". The x, y and z displacement components were examined separately to identify the direction of the highest discrepancy.

5.2.6 Objective 3: Differences in mean shapes, PCA mode shapes and scores

Differences in individual bone sample surface meshes led to the mean discretised shape being dependent on the registration protocol used. The registration protocol also influences the deviation of individual surface meshes from the mean surface mesh, leading to different PCA mode shapes and scores (referred to together as the PCA model). The difference between the two mean discretized surfaces was quantified by evaluating the displacement of each node from the mean shape of "Registration 2" to "Registration 1". The compactness of the two PCA models was compared by evaluating the fraction of the total shape variance explained when PCA modes are sequentially added. The important PCA modes of the two PCA models were compared visually and in the score distributions as grouped by age and groups for the temporal analysis. It compared the results of the statistical tests of the scores in both models to characterise the systematic bone variations over time. For the temporal modes, when the score changes were significant in either one of the "OVX+PTHML" and "OVX" groups or both, the median surface change was calculated for both PCA models. That was quantified by scaling the mode vectors with the median score changes corresponding to a cohort group. A comparison of the 3D surface changes described by each mode was accomplished.

5.3 Results

5.3.1 Objective 1: 3D position disagreement of the registered images

The axial location disagreement between the centres of bounding boxes of the differently registered images – i.e. the magnitude of z translation ("Registration 2" to "Registration 1") – in the "OVX+PTHML" group at week 18 was on average equal to -0.14 mm (Figure 5.5, 'Translation (mm)'). All z translations, ranging from -0.2 mm to -0.028 mm, were negative for the samples in the "OVX+PTHML" group at week 18. This indicates that an image obtained using "Registration 2" had to be always translated in the negative z direction to meet the corresponding image obtained using "Registration 1". At week 24, the mean translation in z was -0.11 mm (range, -0.25 to 0.018 mm), and except for one specimen, all z translations were negative in sign. The average x or y translations at either week 18 or week 24 of age did not exceed 0.0094 mm in magnitude. These translations in individual mice in the "OVX+PTHML" group were inconsistent in sign and were typically much smaller in magnitude than the z translation.

The orientation disagreement – i.e. the magnitude of rotation – over the z-axis between the registered images for the "OVX+PTHML" group at week 18 was on average equal to -0.1º (range -0.29º to 0.20º) (Figure 5.5, 'Rotation (degrees)'). This was an order of magnitude larger than the average orientation disagreement over the x and y axes at week 18. The z-axis orientation disagreement in the samples of the "OVX+PTHML" group at week 24 (mean, -2.7º; range -8.35º to -0.97º) was larger (in magnitude) than that in the week 18 samples, consistently negative in sign, and demonstrating a large variability among the samples. It is also interesting to note that for the bone sample at week 24 with the largest magnitude of z-rotation (-8.35º), the difference in z-alignment between the two registrations was also the highest (0.25 mm). The x- and y-axes orientation disagreement in the samples of the "OVX+PTHML" group at week 24 were relatively small in magnitude (ranges 0.085º to 1.3º and -0.22º to 2.2º respectively).

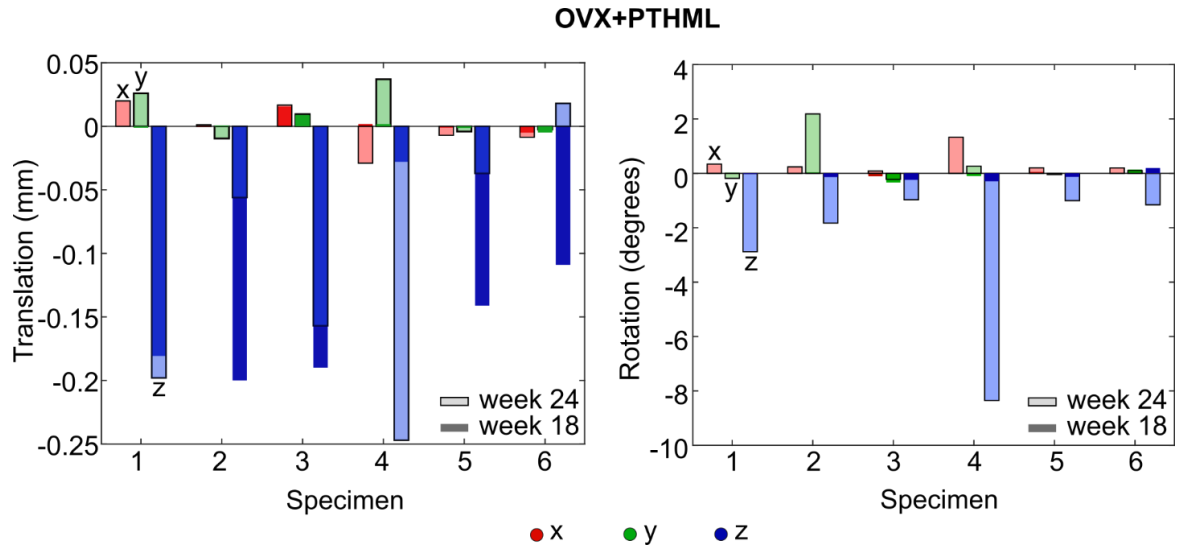


Figure 5.5: Location disagreement (described as translation) and orientation disagreement (described as rotation) between the two differently registered images in x (red), y (green) and z (blue), for each mouse in the "OVX+PTHML" groups at both ages, week 18 (dense colour) and week 24 (transparent colour with black border). For instance, the z translation for specimens 2, 3 and 5 at week 24 is less in magnitude than the one at week 18.

The mean magnitude of translation in the z-direction in the "OVX" at week 18 was measured to be -0.018 mm (range -0.12 mm to 0.092 mm) (Figure 5.6). The mean values for the translation in x and y were equal to 0.0023 mm and 0.0008 mm, respectively, much smaller in magnitude than the one in z. Unlike at week 18, the z-translation for the image samples at week 24 was always negative (mean -0.16 mm, range -0.23 mm to -0.08 mm). A larger magnitude of the mean value of translation and consistency in the sign of translation were also reported in the x and y directions (mean values 0.012 mm and 0.01 mm, respectively) at week 24 in the "OVX" group.

The mean rotations along the x, y and z axes for the "OVX" group at week 18 were found to be 0.033º, -0.034º and -0.14º, respectively, with the rotation along the z-axis being the highest (Figure 5.6). At week 24, the z-rotations were much higher in magnitude (mean -1.31º) and, unlike at week 18, were mostly negative in sign. For the rotations along the x and y axes, the magnitudes remained relatively small at week 24.

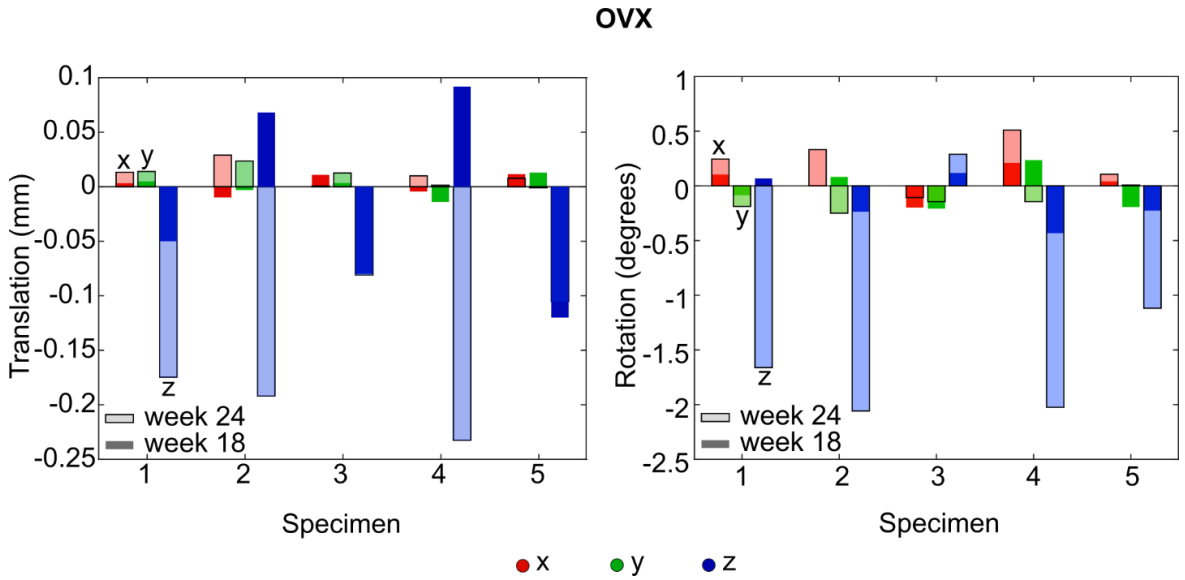


Figure 5.6: Location disagreement (described as translation) and orientation disagreement (described as rotation) between the two differently registered images in x (red), y (green) and z (blue), for each mouse in the "OVX" groups at both ages, week 18 (dense colour) and week 24 (transparent colour with black border). For instance, the z translation for specimen 5 at week 24 is less in magnitude than the one at week 18.

5.3.2 Objective 2: 3D shape differences of meshes

The median surface mesh nodal displacements over all mesh nodes and all samples in the "OVX+PTHML" group at week 18 were $2.4 \cdot 10^{-4}$ mm, $2.1 \cdot 10^{-4}$ mm and -0.14 mm respectively in the x, y and z directions. The 75% of the nodal displacements (Q3-Q1) ranges [-0.023 – 0.024] mm, [-0.021 – 0.021] mm and [-0.22 – 0.022] mm for x, y and z, respectively. For the "OVX+PTHML" group at week 24, the surface meshes derived from the two registration protocols differed with each other with median surface distance in x, y and z equal to 0.0057 mm, 0.0024 mm and -0.11 mm. The Q3-Q1 ranges were measured to be [-0.048 – 0.063] mm, [-0.04 – 0.051] mm and [-0.40 – 0.08] mm for x, y, and z, respectively. The larger median displacement magnitude in z (and its negative sign) found for almost all specimens in the group (Figure 5.7) agrees with the previously observed larger z-translation between the two registrations (also with a negative sign). Figure 5.8 illustrates the boxplots of the displacement distribution in x, y, and z for the one specimen in the group at week 24, highlighting the highest magnitude in z. The 3D profile of the displacements ("Registration 2" to "Registration 1") for this specimen is representative of all samples in "OVX+PTHML", with negative displacement in z.

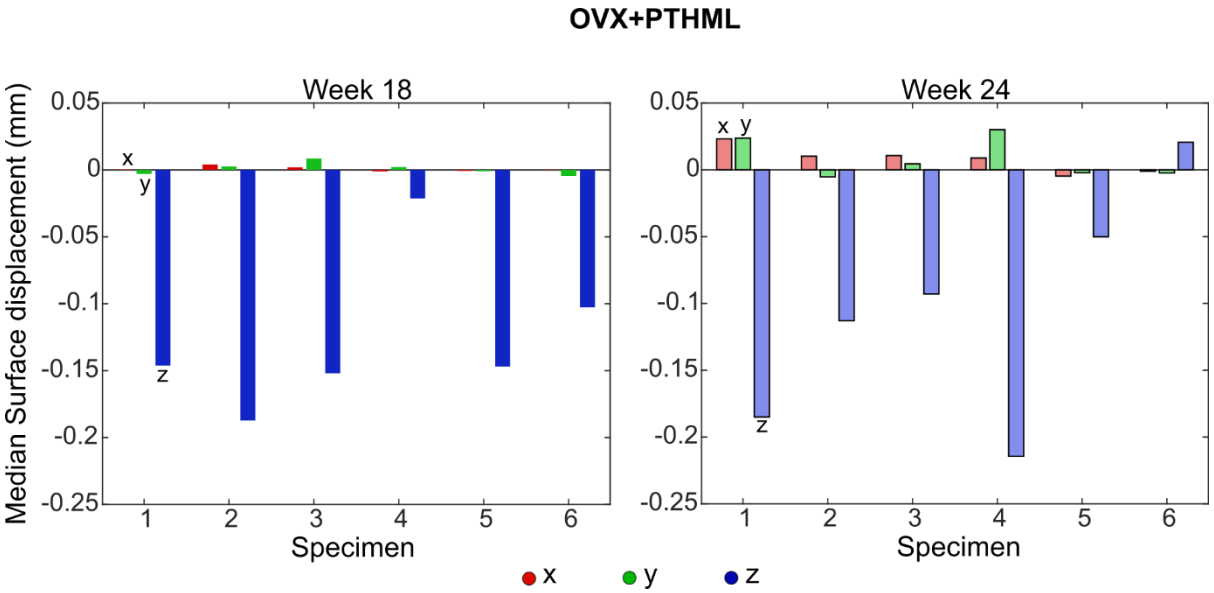


Figure 5.7: Median displacement in x(red), y (green), and z (blue) between the surface meshes corresponding to the two differently registered images of each bone sample. Displacement components are plotted for each bone sample in the "OVX+PTHML" group and both ages (week 18: solid colors in the left figure, and week 24: transparent with black line in the right figure).

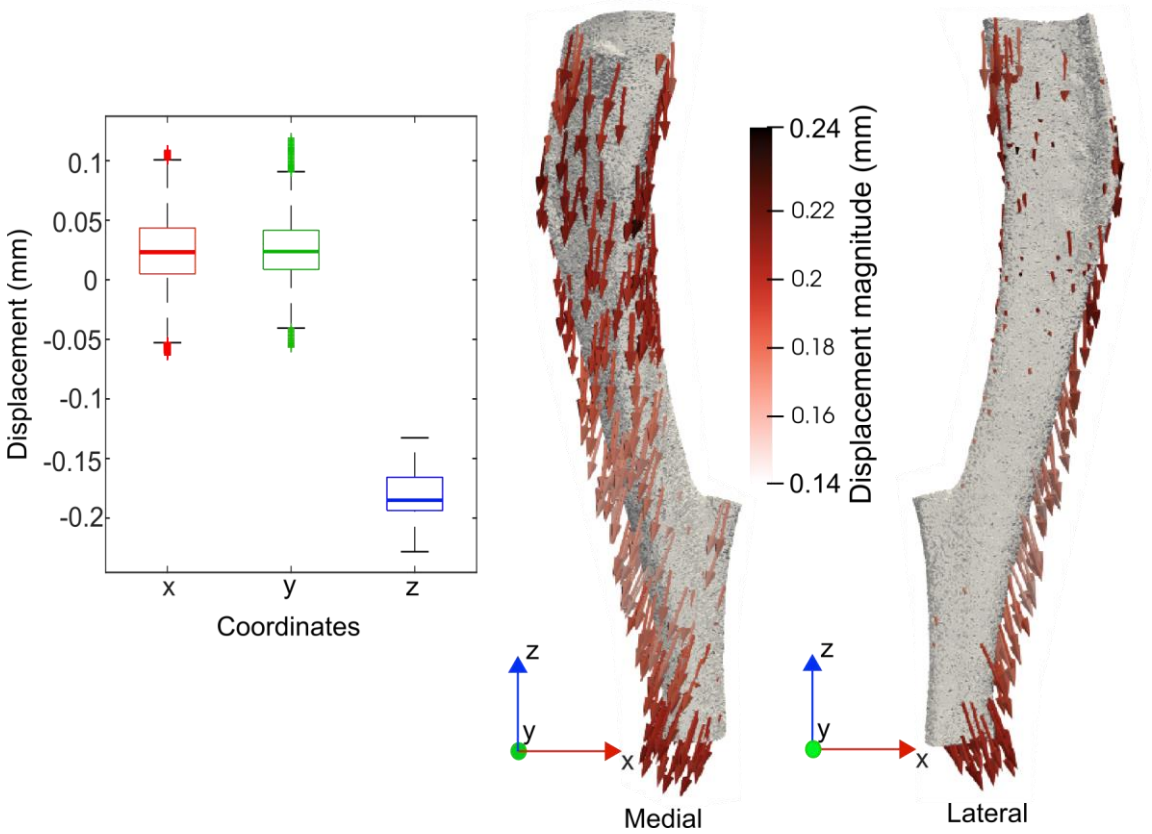


Figure 5.8: Boxplot and 3d profile for one specimen in "OVX+PTHML" and week 24. Two viewpoint are given: media-to-lateral (denoted as medial) and lateral-to-medial (denoted as lateral).

For the meshes in the "OVX" group at week 18, the median surface mesh nodal displacements in x, y and z between the meshes derived from the two registrations was found equal to $1.4 \cdot 10^{-4}$ mm, -0.0019 mm and -0.013 mm. The Q3-Q1 ranges were [-0.025 - 0.024], [-0.035 – 0.031] and [-0.15 – 0.18] mm for x, y and z, respectively. At week 24, the mesh differences in the "OVX" group increased in magnitude, resulting in median displacement equal to 0.0096, 0.0077 and -0.12 mm in x, y and z coordinates, respectively. The Q3-Q1 ranges were [-0.044 – 0.064], [-0.034 – 0.048] and [-0.26 – 0.043] mm. As shown in Figure 5.9, the median magnitude of the displacement was consistently increased for all samples in the "OVX" group, as also indicated by "Objective 1" (Figure 5.6).

Notably, the Q3-Q1 ranges in the "OVX+PTHML" group was higher than the ones in the "OVX" group (~0.048 mm vs 0.032 mm).

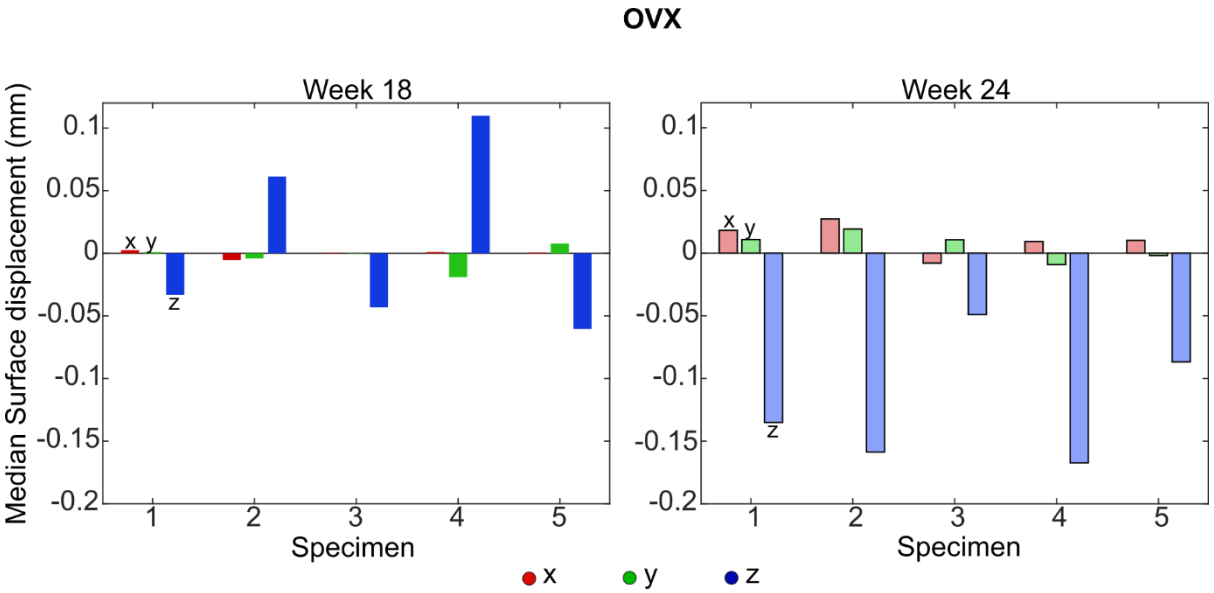


Figure 5.9: Mean and standard deviation of the components of the displacement (x: red, y: green, z: blue) between the surface meshes corresponding to the two differently registered images of each bone sample. Displacement components are plotted for each bone sample in the "OVX" group and both ages (week 18: solid, and week 24: transparent).

5.3.3 Objective 3: Differences in mean shapes, mode shapes and scores

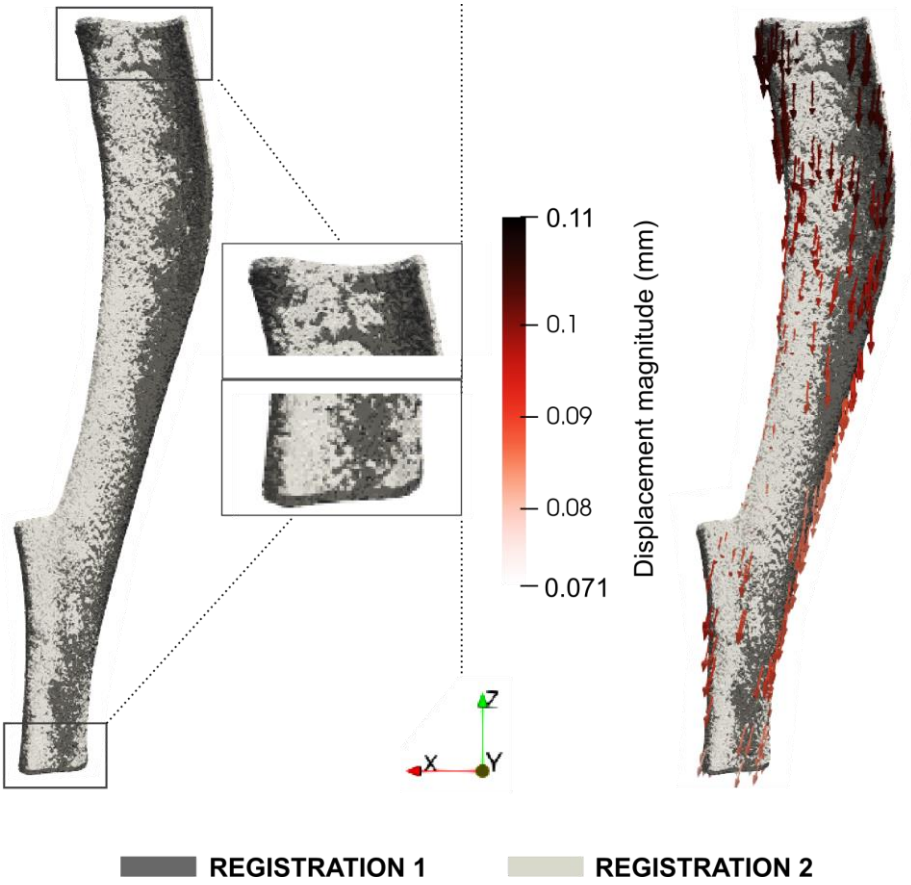


Figure 5.10: Mean shapes of the databases processed with "Registration 1" (dark grey) or "Registration 2" (light grey). The 3D displacement profile from the "Registration 2" profile to "Registration 1" is illustrated as a vector field plotted on top of the "Registration 2" profile. Lateral view is exhibited.

The difference in the mean shapes is shown in Figure 5.10. The displacement profile illustrated as vector field highlights that the prominent difference is the axial alignment with magnitudes [0.071 -0.11] mm as also suggested by Objective 1 and 2.

The geometric variance in the z coordinates present in the mesh databases of "Registration 1" and "Registration 2" and averaged across all nodes in the mesh are respectively 0.011 mm and 0.0088 mm. The geometric variance in the x and y coordinates are of a similar magnitude in both databases. PCA of these databases shows that Modes 1–6 capture, in that order, 54%, 26%, 5%, 4%, 3%, and 2% of the total variance in the case of "Registration 1" and 69%, 10%, 7%, 4%, 3% and 2% of the total variance in the case of "Registration 2". Together the first six modes explain ~94% of the total variance (Figure 5.11).

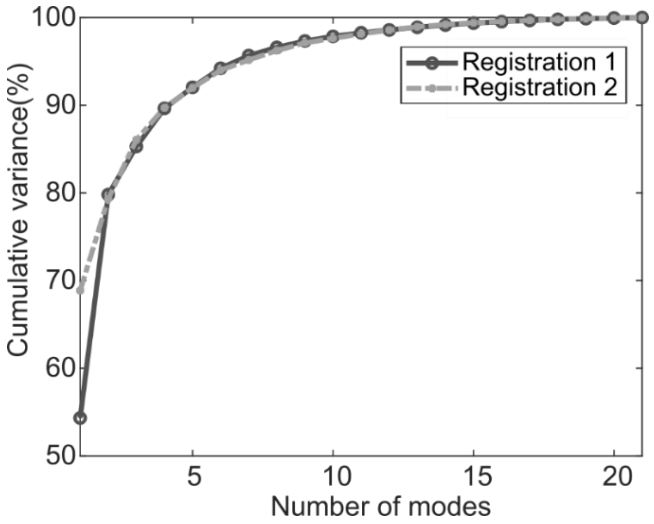


Figure 5.11: Percentage of the total variation in geometry captured by PCA modes. The black solid line corresponds to the PCA results of the database processed with "Registration 1". The Grey dashed line corresponds to the PCA results of the database processed with "Registration 2".

Figures 5.12, 5.15 and 5.16 exhibit the first six mode shapes given by the two PCA models, illustrated next to each other. Modes 1 and 2 described similar global variations in the proximal and distal epiphysis in both databases (Figure 5.12). They also shared similar distributions of magnitudes (note that the range of the colour bar was the same for a mode extracted from both databases to highlight the possible magnitude differences in the 3D shapes).

However, these two modes also described some different features. Firstly, the noticeable difference in Mode 1 between the two registrations was the distal expansion, shown in "Registration 1", but not in "Registration 2" (Figure 5.12). Secondly, in "Registration 2", Mode 2 exhibited higher magnitudes in the x and y coordinates of the mode vectors in the diaphysis resulting in more prominent cross-sectional expansion than the ones in "Registration 1", as highlighted in Figure 5.13. Additionally, it captured medial expansion at the proximal end, which is not present in "Registration 1" (Figure 5.14).

Modes 3 captured similar variations in both databases but with different magnitudes (brighter red arrows in "Registration 1", Figure 5.15). Specifically, Mode 3 in both databases described some uniaxial distal variations, but the magnitude in "Registration 1" is 0.004, lower than 0.0045 in "Registration 2".

Mode 4 described different variations for the differently registered data (Figure 5.15). Specifically, for "Registration 1", Mode 4 captured variations at the anterior ridge at proximity and a global cross-

sectional expansion of the diaphysis and a shrinkage of the distal end. For “Registration 2”, Mode 4 described variations at the anterior crest in the midshaft.

Mode 5 also described different variations in the two databases (Figure 5.16). Specifically, Mode 5 demonstrated the expansion of the anterior crest in midshaft when data were processed with “Registration 1”. Additionally, it described a moderate expansion of the posterior compartment at the proximal epiphysis. In “Registration 2”, Mode 5 captured the expansion of the anterior ridge at the proximity of the bone.

Visual inspection of shape profiles in Figures 5.15 and 5.16 suggest that Mode 4 in “Registration 1” is similar to Mode 5 in “Registration 2”, and Mode 5 in “Registration 1” is similar to Mode 4 in “Registration 2”. Notably, Mode 4 in “Registration 1” describes 4% of the total in shape, 1% higher than the percentage that Mode 5 describes in “Registration 2”.

Finally, Mode 6 captured posterior expansion and anterior ridge expansion in the proximal epiphysis in both registrations (Figure 5.16). However, the mode magnitude is smaller in “Registration 2” than in “Registration 1” (0.003 vs 0.0042 mm). Mode 6 also described some different features in the two databases. Specifically, for “Registration 1”, the mode exhibited an expansion of the anterior crest in the midshaft and a distal expansion; but this variation was not captured by Mode 6 in “Registration 2”.

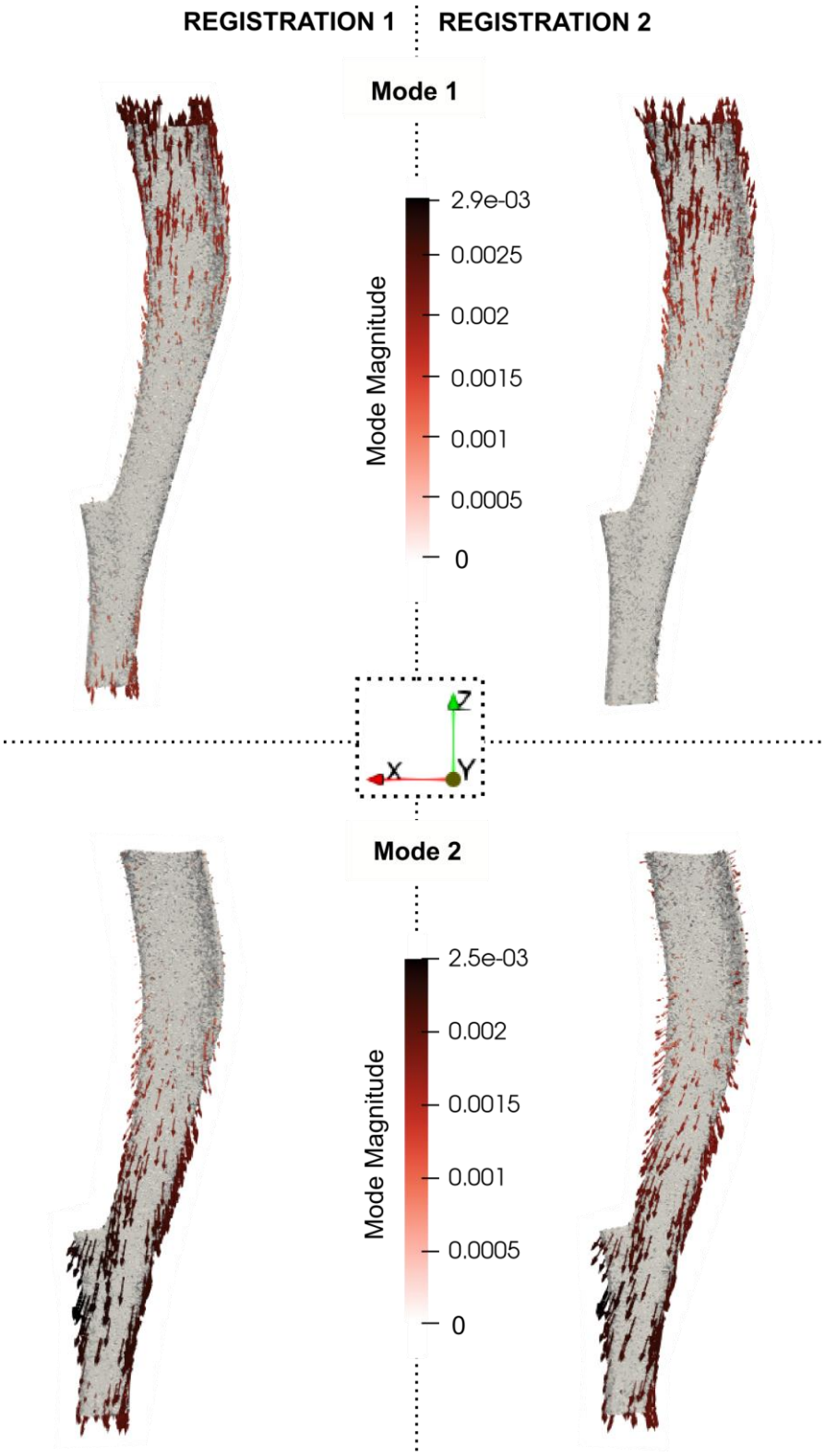


Figure 5.12: The 3D profiles of Modes 1 and 2 for both databases were processed with "Registration 1" or "Registration 2". The vector fields are plotted on top of the mean shape. The vectors are scaled and coloured by mode magnitude at each node. Same colorbar and ranges are used to enable comparison of the contour colours on the 3D shape.

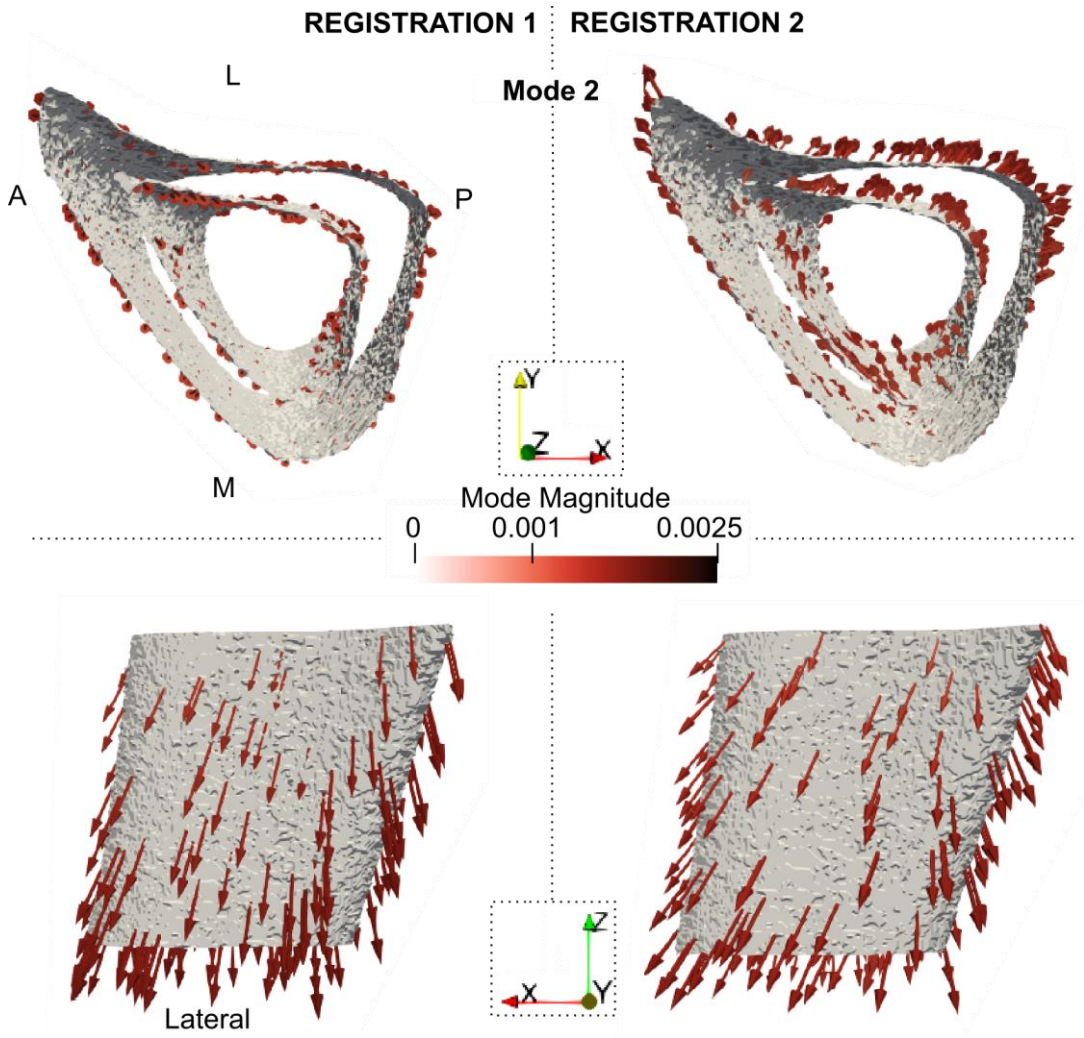


Figure 5.13: 3D profile of Mode 2 for "Registration 1" and "Registration 2"- zoomed-in view on the midshaft provided with two different views: proximal to distal and lateral to medial.

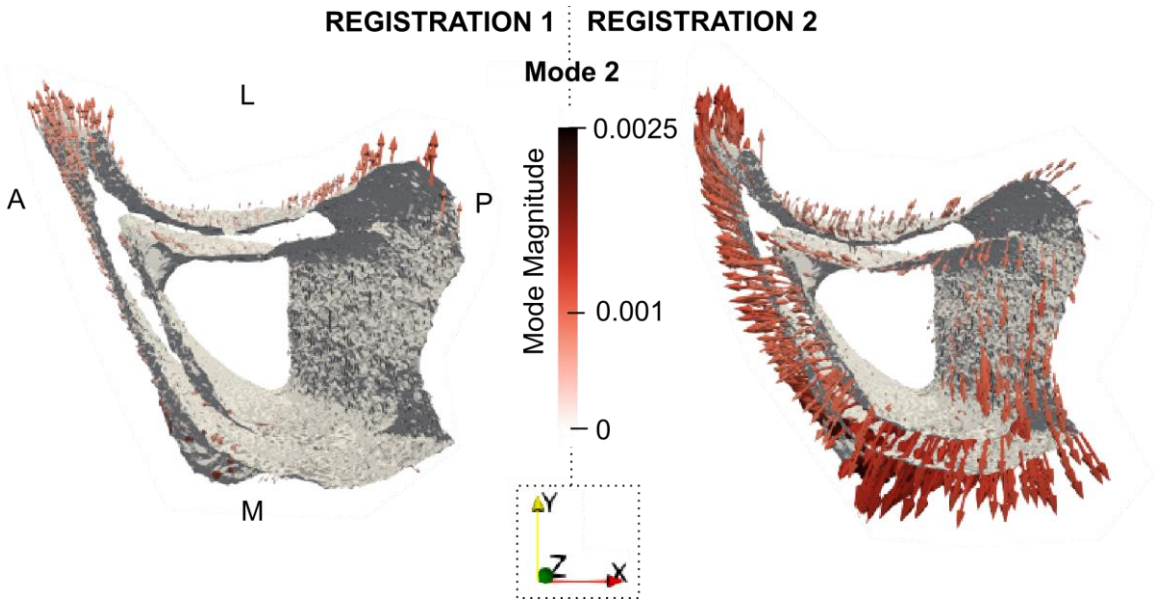


Figure 5.14: 3D profile of Mode 2 for "Registration 1" and "Registration 2", zoom in on the proximal end. The sections share the same orientation denoted as A (anterior), L (lateral), P(posterior) and M (medial)

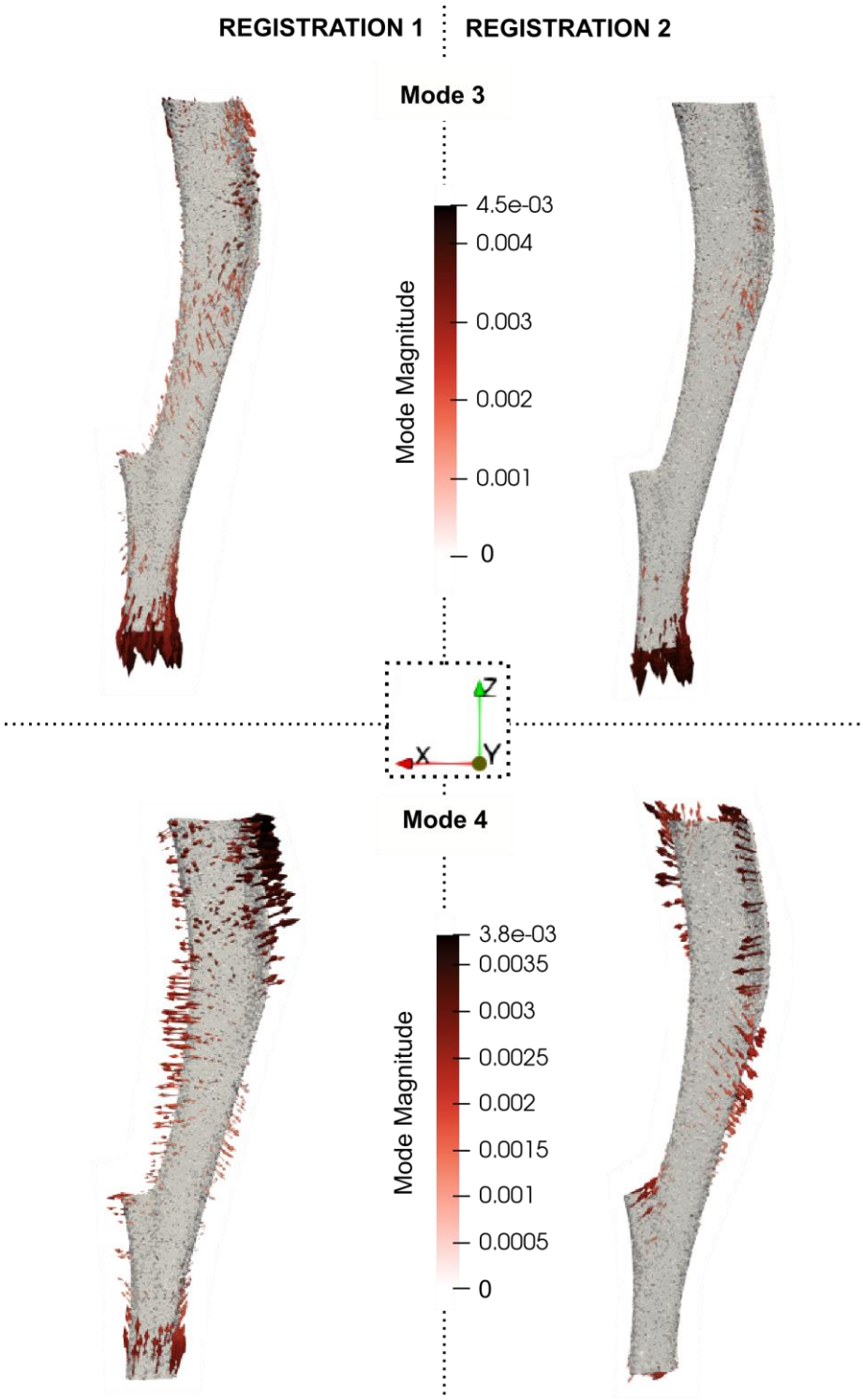


Figure 5.15: The 3D profiles of Modes 3 and 4 for both databases were processed with "Registration 1" or "Registration 2". The vector fields are plotted on top of the mean shape. The vectors are scaled and coloured by mode magnitude at each node.

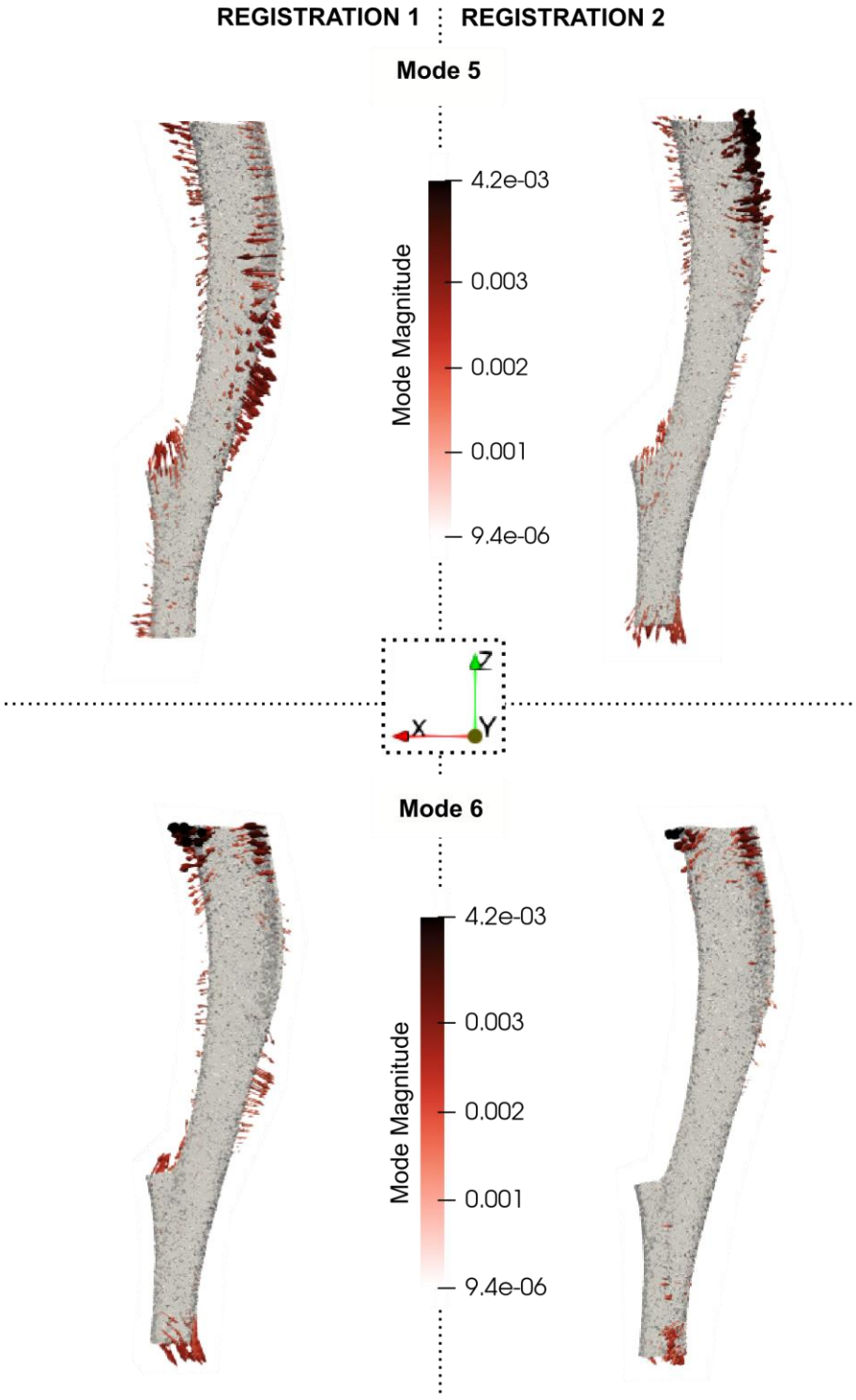


Figure 5.16: The 3D profiles of Modes 5 and 6 for both databases were processed with "Registration 1" or "Registration 2". The vector fields are plotted on top of the mean shape. The vectors are scaled and coloured by mode magnitude at each node.

Figures 5.17 and 5.18 illustrate the individual score differences between the two registration approaches, highlighting how these differences affect the temporal analysis of the scores. Based on the statistical analysis of the score changes, for “Registration 1”, the treatment-related modes were determined to be Modes 1, 2, 3, 4, 5 and 11, with statistically significant score changes over time in one of the cohort groups (Figures 5.17, 5.18). For “Registration 2”, the treatment-related modes were Modes 3, 4, 5, 6, 9, 11 and 13. For the second protocol, Mode 1 was characterised to be a growth-related mode, with significant and of similar magnitude score changes in both “OVX+PTHML” and “OVX” groups ($p<0.05$). Additionally, Mode 2 exhibited almost consistent score changes of similar positive patterns for both disease and treatment groups, indicating growth change, however, the statistical tests did not reveal significant changes ($p>0.05$).

For “Registration 1”, the score changes for Mode 1 were statistically significant in the “OVX+PTHML” group ($p<0.05$, but $p>0.05$ for “OVX”) (Figure 5.17). In Registration 2, statistically significant score changes between the two ages were found in both disease and treatment groups.

For “Registration 1”, Mode 2 described temporal variation only in the “OVX” group ($p<0.05$) (Figure 5.17). However, when the data were processed using “Registration 2”, Mode 2 exhibited not significant score changes over time but with a consistent trend of change for almost all samples in both “OVX” and “OVX+PTHML” groups.

For Mode 3, when data were processed using “Registration 1”, significant score changes were found only in the “OVX+PTHML”, although some consistent trend of change could be seen in the “OVX” group too (Figure 5.17). In contrast, for “Registration 2”, significant score changes in Mode 3 were found in “OVX” and not in the “OVX+PTHML” group.

For Mode 4, temporal positive score changes were revealed in the “OVX+PTHML” for both registrations ($p<0.05$), with the “OVX” having the opposite effect, but not significant (Figure 5.18). This score change was translated to anterior crest thickening and cross-sectional diaphysis expansion in the treatment group. Conversely, the negligible score changes in the “OVX” group meant no change

in the anterior crest and diaphysis. Similar score changes were found in Mode 5 in “Registration 2” (Mode 4 in “Registration 1” and Mode 5 in “Registration 2” share similar shape profiles). Specifically, for Mode 5, a significant temporal positive score change was found in the “OVX+PTHML” ($p<0.05$), but a statistically zero score change in the “OVX” group ($p>0.05$). Similarly, the score changes in Mode 5 for “Registration 1” and Mode 4 for “Registration 2” share similar patterns with statistical significance only in the “OVX+PTHML” group (Figure 5.18).

Finally, Mode 6 describes a treatment effect only when the data were processed with “Registration 2”, with the “OVX+PTHML” exhibiting significant positive score change ($p<0.05$, but $p>0.05$ in “Registration 1”) (Figure 5.18).

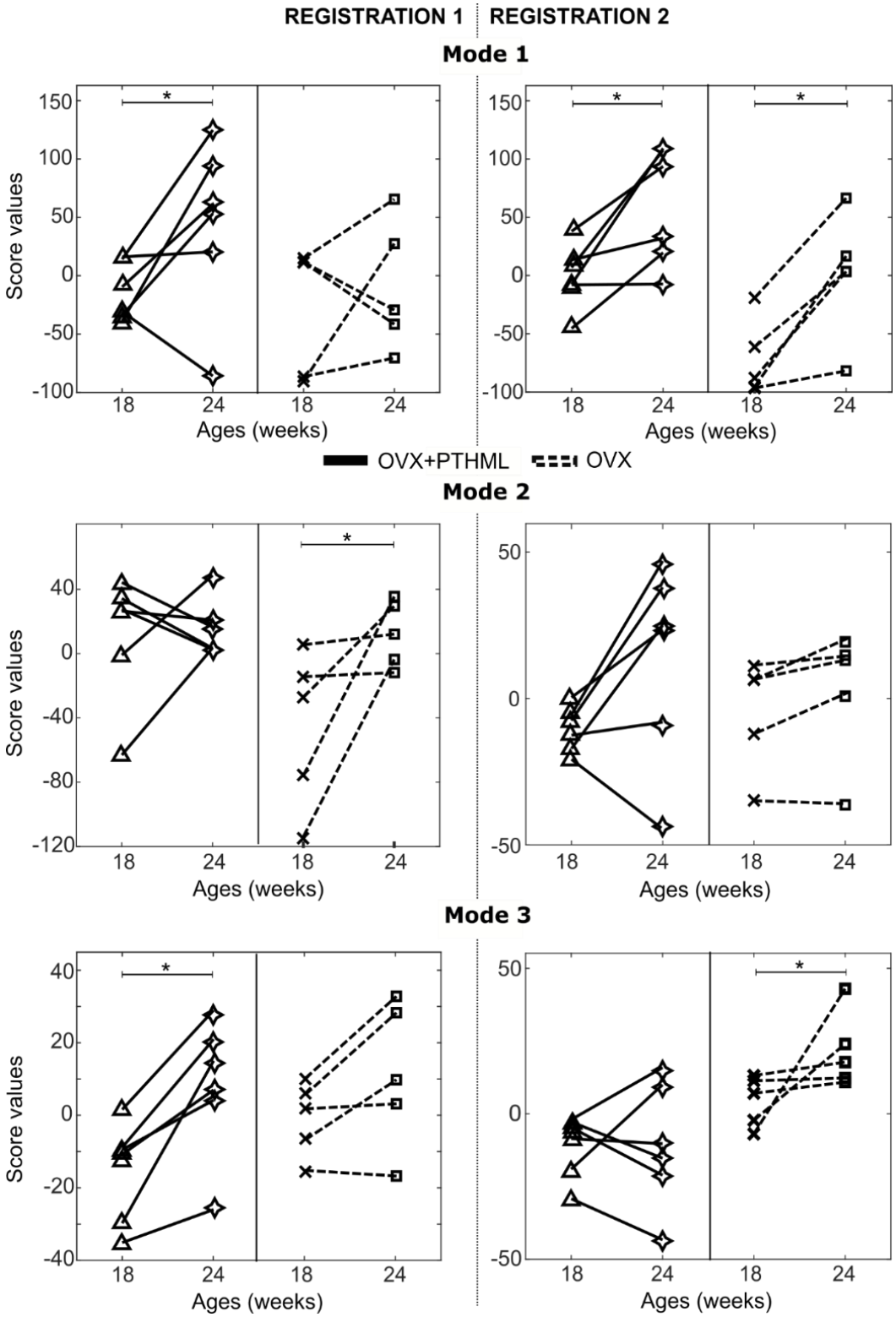


Figure 5.17: Mode categorization of the first 3 modes. Score values of each specimen in "OVX" (black and dashed line) and "OVX+PTHML" (black, solid line) group at the two ages, 18 and 24 weeks. The trend of change over time is illustrated by the connecting lines. Statistically significant results with $p<0.05$ are indicated with (*).

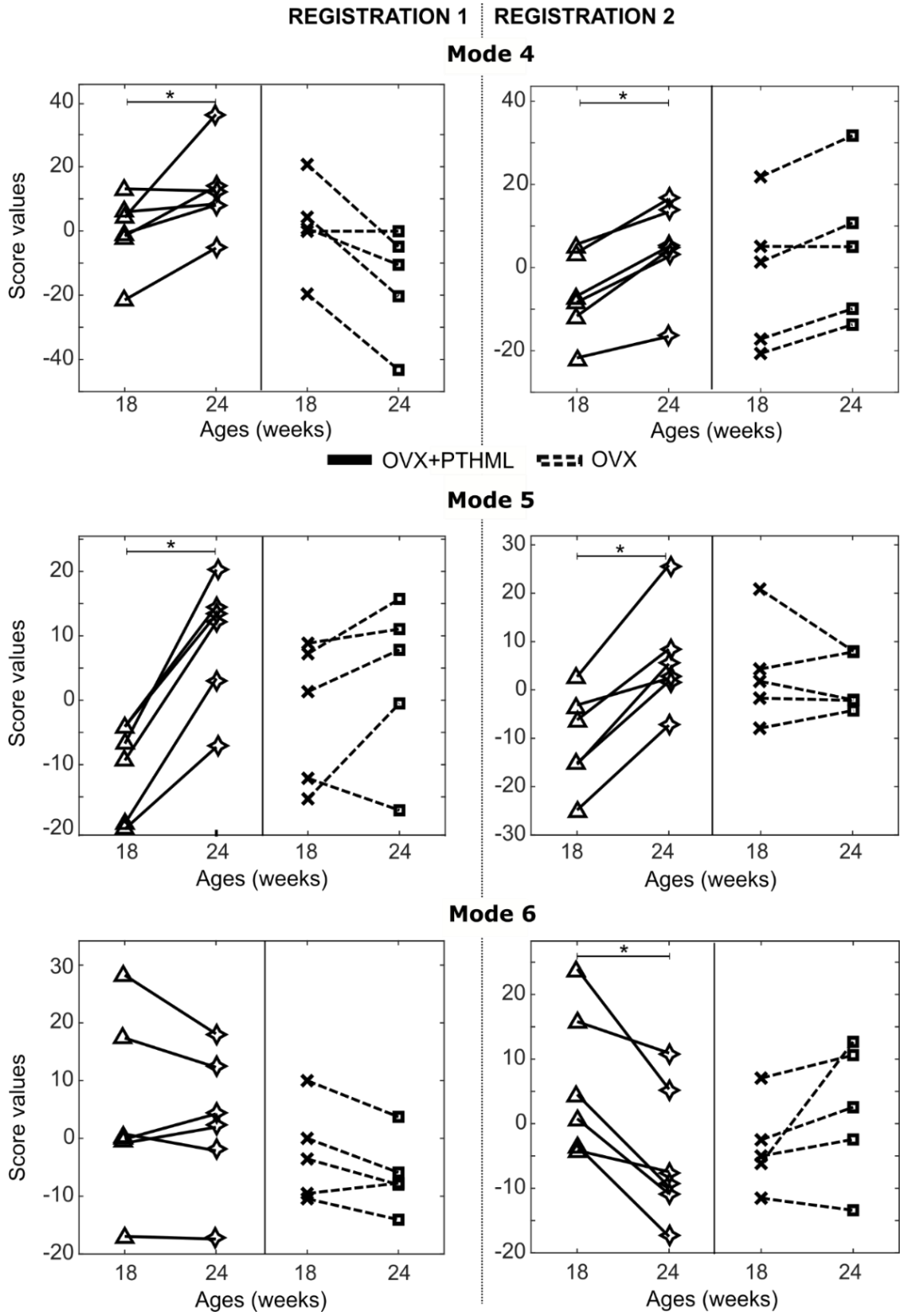


Figure 5.18: Mode categorisation of Modes 4, 5, 6. Score values of each specimen in "OVX" (black and dashed line) and "OVX+PTHML" (black, solid line) groups at the two ages, 18 and 24 weeks. The trend of change over time is illustrated by the connecting lines. Statistically significant results with $p<0.05$ are indicated with (*).

In both protocols, Mode 1 described proximal expansion in the “OVX+PTHML” group but with different magnitudes of 0.22 mm and 0.19 mm for “Registration 1” and “Registration 2”, respectively (Figure 5.19). Mode 1 also described significant temporal expansion of the proximal end in the “OVX” group ($p < 0.05$), but only in “Registration 2”, as suggested by the statistical results of the score changes in that group (Figure 5.17).

Mode 2 captured distal expansion with a maximum magnitude of 0.086 mm and 0.040 mm in “Registration 2” and “Registration 1”, respectively (Figure 5.19). Notably, although the statistical test of the score changes for Mode 2 and “Registration 2” did not reveal any significance (Figure 5.17), their almost consistent score change pattern is associated with 8 pixel length increase. Additionally, Mode 2 in “Registration 2” described cross-sectional expansion in the diaphysis with periosteal formation equal to 0.04 mm and endosteal deletion of a smaller magnitude.

For “Registration 1”, the positive score changes in Mode 3 in the “OVX+PTHML” group (Figure 5.17) were associated with distal expansion and proximal posterior shrinkage, as shown in Figure 5.20 and 5.21. However, for “Registration 2”, Mode 3 described proximal posterior thickening in the “OVX+PTHML” group. Specifically, the mode described 0.027 mm posterior thickening at the proximal epiphysis, but also 0.025 mm shrinkage at the distal end in the treatment group. Notably, mode’s vector field in “Registration 1” seems to include a rotation effect that could be possibly related to orientation misalignment, as highlighted in Figure 5.22. Another difference between the protocols is that Mode 3 for “Registration 2” exhibited a moderate cross-sectional expansion at the midshaft with local thickening at the posterior compartment and the anterior crest (~0.02 mm), not evident in “Registration 1” (Figure 5.19).

For Mode 4, the maximum expansion of the anterior ridge was 0.033 mm in “Registration 1” (Figure 5.20). A similar expansion of the anterior ridge was described by Mode 5 in “Registration 2”, with 0.061 mm periosteal bone formation (Figure 5.23). Moreover, the surface change described by Mode 5 in “Registration 1” was detected at the anterior crest with more prominent periosteal expansion equal to 0.068 mm and lower endosteal resorption of ~0.04 mm, resulting in 0.028 mm thickening of the

cortical bone. In contrast, for “Registration 2”, thickening of the anterior crest in the midshaft was described with the combination of Mode 2 (as described before), Mode 3 (0.025 mm periosteal apposition and 0.007 mm endosteal resorption) and Mode 4 (0.023 mm periosteal apposition and 0.011 mm endosteal resorption).

Mode 6 describes the posterior thickening of the proximal epiphysis for both databases (Figure 5.23). The periosteum formation was found to have higher magnitudes than the endosteal deletion for both registrations. However, the magnitudes of the surface change maps describing the two databases are different: local periosteal formation of 0.054 mm (vs 0.015 mm in “Registration 1”) and endosteal deletion of 0.036 mm for “Registration 2” (vs 0.013 mm in “Registration 1”) (Figure 5.23). Notably, the temporal analysis of the scores showed non-significant temporal changes in the “OVX+PTHML” group in “Registration 1”.

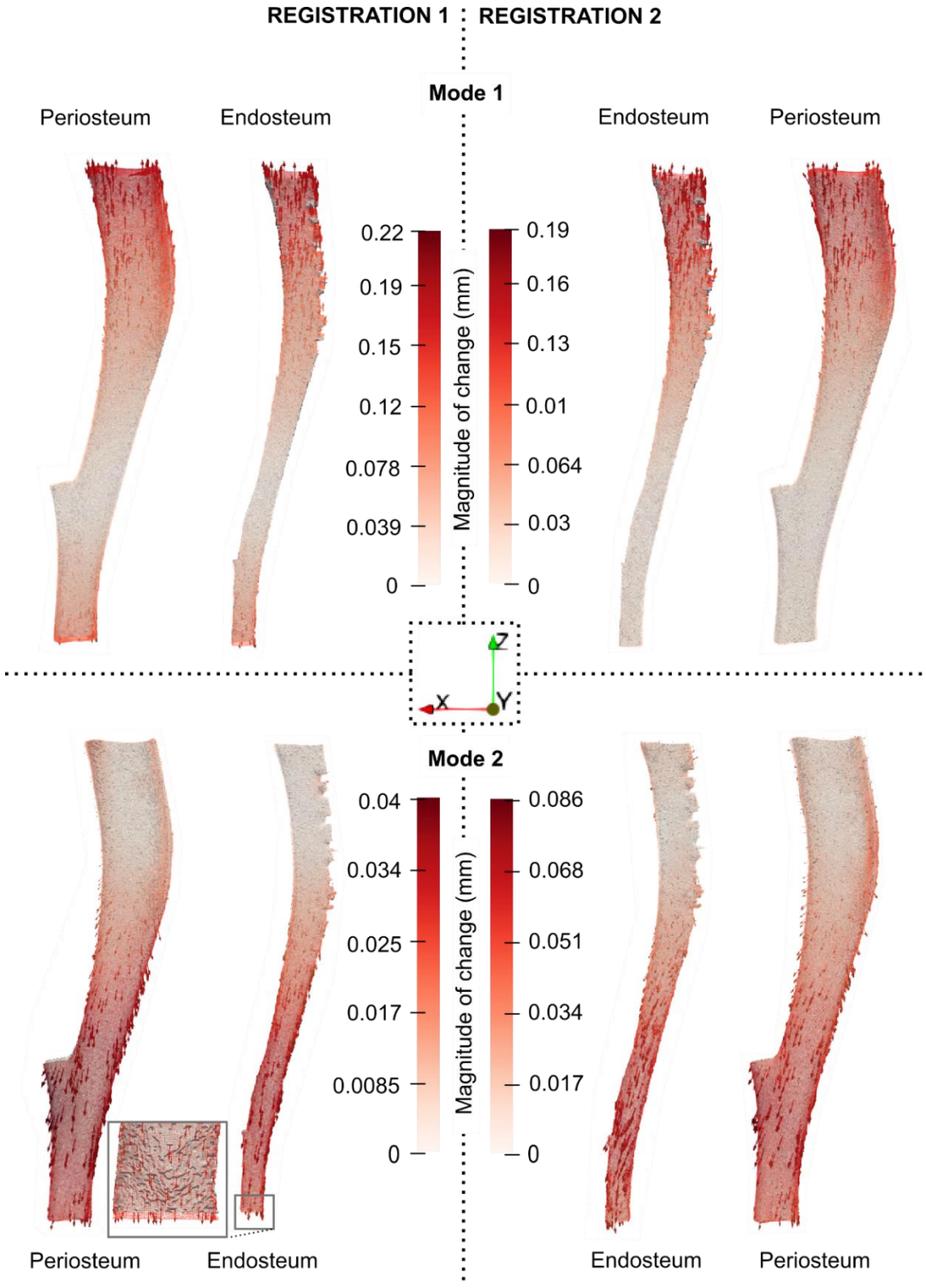


Figure 5.19: The median geometric changes described by Modes 1 and 2 for the "OVX+PTHML" group in both databases derived from "Registration 1" and "Registration 2". Endosteum and periosteum profiles are given for both ages and reconstructed using the PCA formula. Specifically, they were reconstructed using the mean shape and adding the scaled mode vector using the median score value in the "OVX+PTHML" group at week 18 and week 24. The bone profile of week 18 is represented as a solid grey surface, whilst the profile at week 24 is given as a coloured wireframe. The colours represent the magnitude of the geometrical change over time as described by each Mode. The vectors describe the direction of the change, and they are scaled and coloured with respect to the magnitude. The arrows' starting point is the bone surfaces corresponding to week 18 of age. The redder the colour, the higher the magnitude is. Zoomed-in view for Mode 2 in "Registration 1".

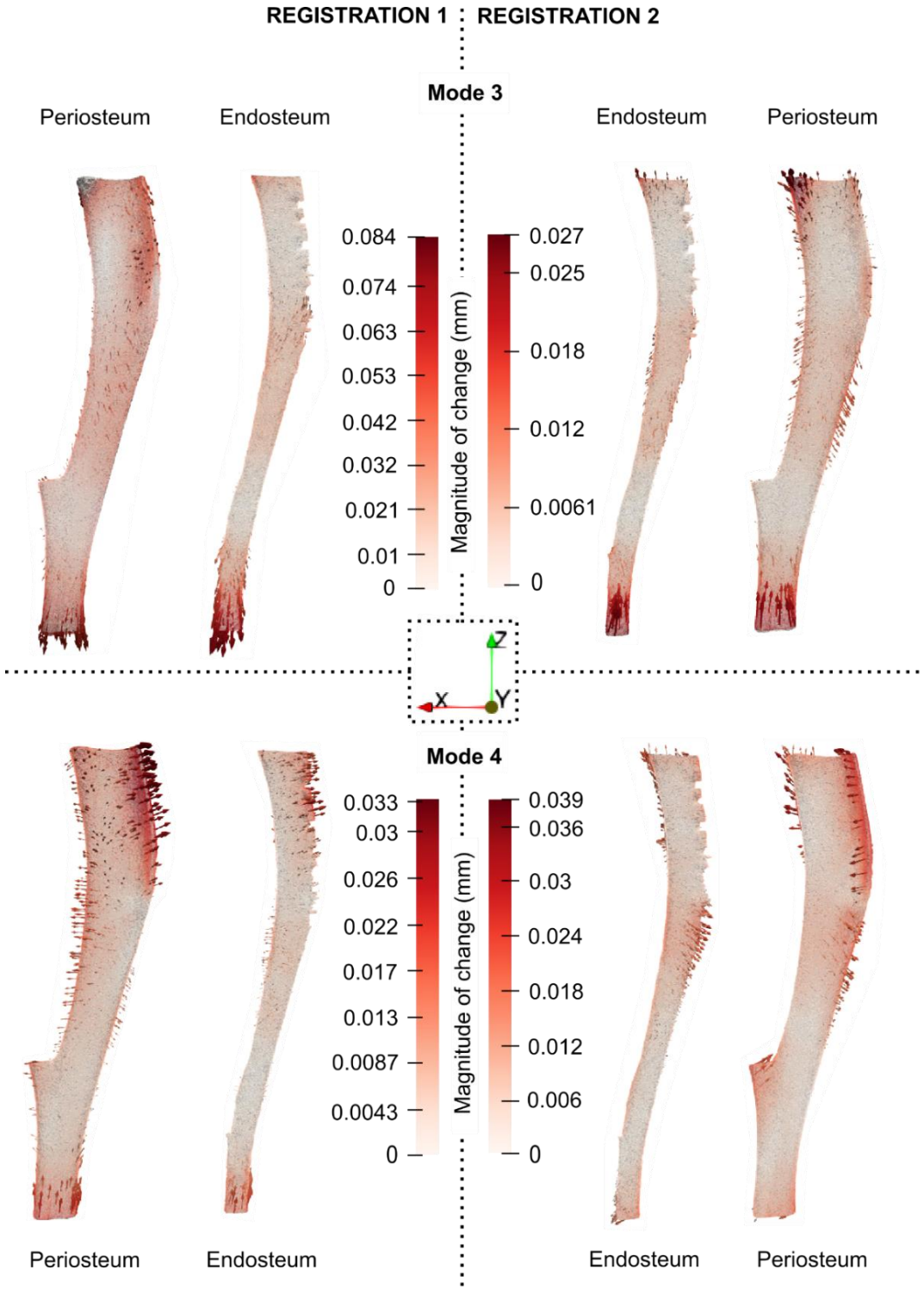


Figure 5.20: The median geometric changes described by Modes 3 and 4 for the "OVX+PTHML" group in both databases derived from "Registration 1" and "Registration 2". Endosteum and periosteum profiles are given for all modes and both ages and reconstructed using the PCA formula. Specifically, they were reconstructed using the mean shape and adding the scaled mode vector using the median score value in the "OVX+PTHML" group at week 18 and week 24. The bone profile of week 18 is represented as a solid grey surface, whilst the profile at week 24 is given as a coloured wireframe. The colours represent the magnitude of the geometrical change over time as described by each Mode. The vectors describe the direction of the change, and they are scaled and coloured with respect to the magnitude. The arrows' starting point is the bone surfaces corresponding to week 18 of age. The redder the colour, the higher the magnitude is.

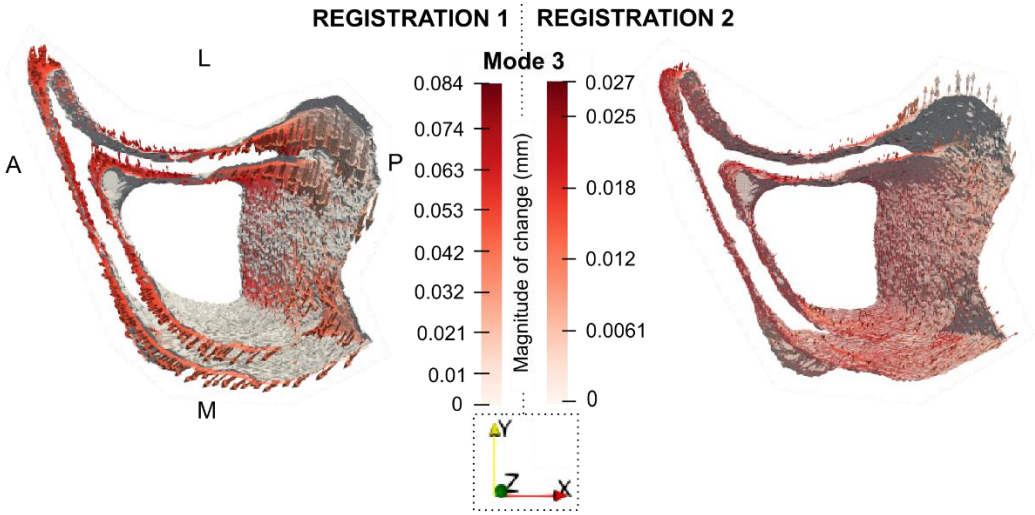


Figure 5.21: 3D profiles of median surface change described by Mode 3 for "Registration 1" and "Registration 2", zoomed-in view on the proximal end and anatomical orientation denoted with A (anterior), L (lateral), P (posterior), M (medial).

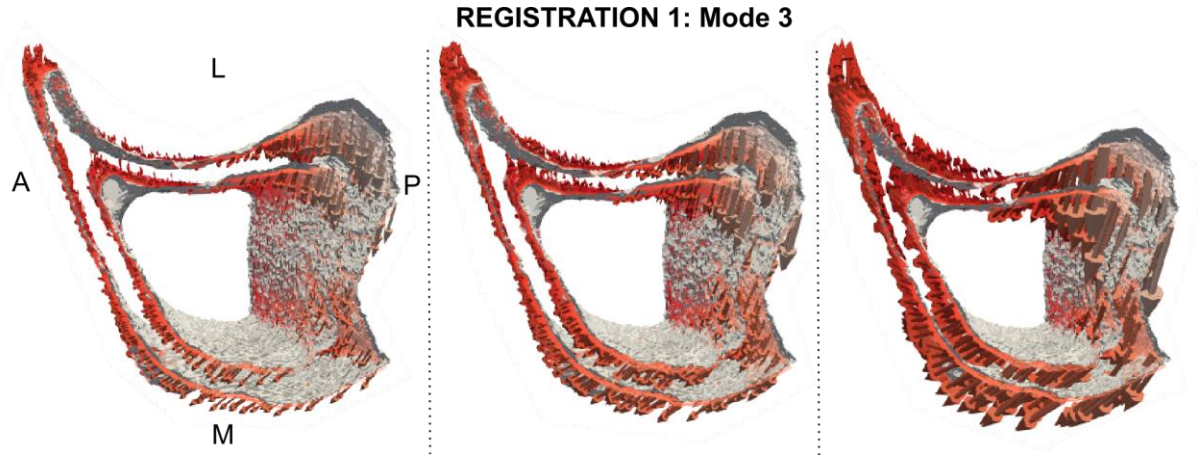


Figure 5.22: 3D profiles of median surface change described by Mode 3 for "Registration 1"- three different scales of vectors.

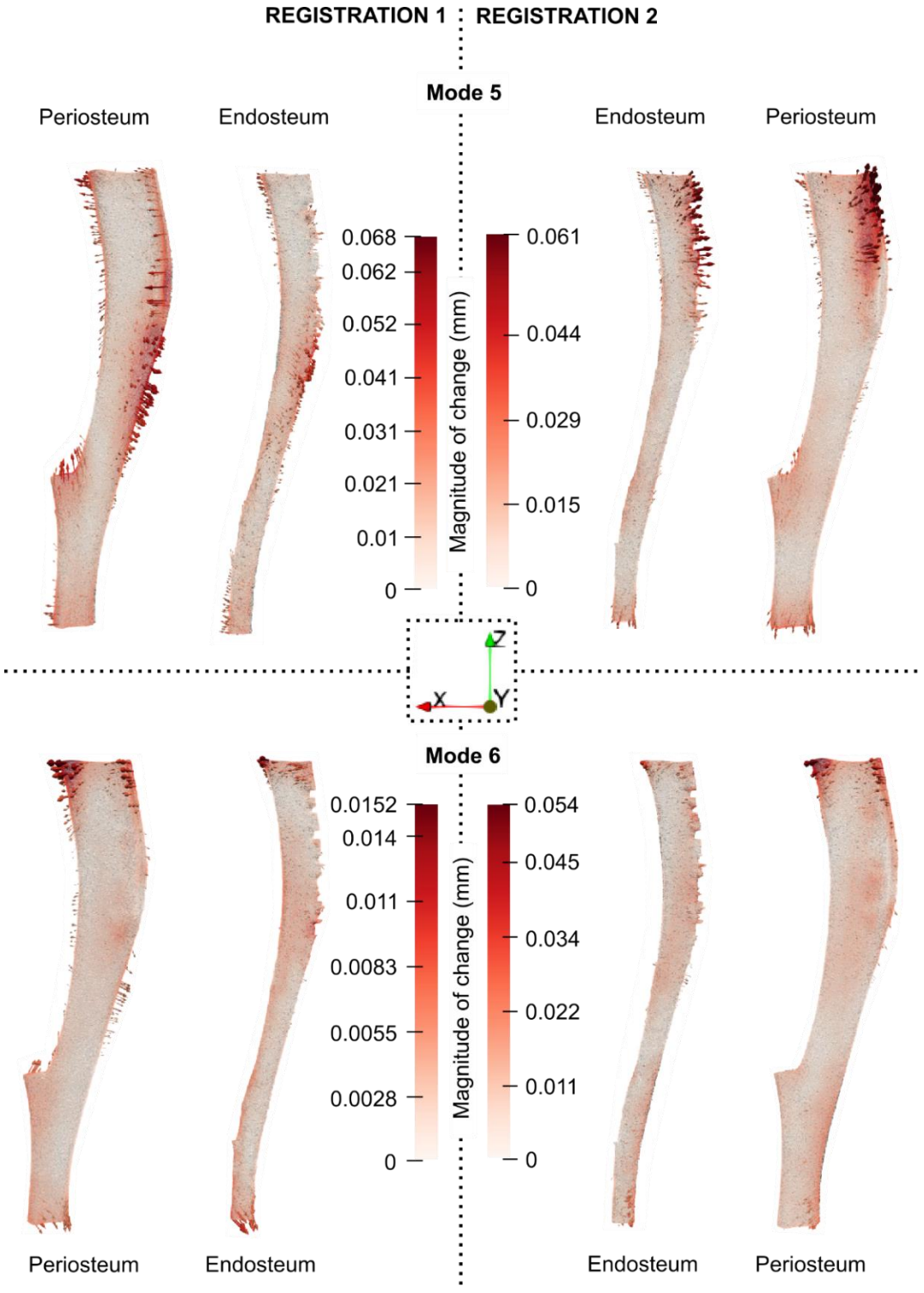


Figure 5.23: The median geometric changes described by Modes 5 and 6 for the "OVX+PTHML" group in both databases derived from "Registration 1" and "Registration 2". Endosteum and periosteum profiles are given for all modes and both ages and reconstructed using the PCA formula. Specifically, they were reconstructed using the mean shape and adding the scaled mode vector using the median score value in the "OVX+PTHML" group at week 18 and week 24. The bone profile of week 18 is represented as a solid grey surface, whilst the profile at week 24 is given as a coloured wireframe. The colours represent the magnitude of the geometrical change over time as described by each Mode. The vectors describe the direction of the change, and they are scaled and coloured with respect to the magnitude. The arrows' starting point is the bone surfaces corresponding to week 18 of age. The redder the colour, the higher the magnitude is.

5.4 Discussion

5.4.1 Comparison of the two rigid registration protocols

The current study aimed to evaluate how the rigid registration protocol as an instance alignment technique affects the assessment of the shape variations of a population of treated and diseased mouse tibia. This aim was achieved by assessing the differences between two registration protocols in distinct steps of the PCA-based framework.

The study used longitudinal data of treated and untreated mice to test the influence of registration in the assessment of treatment effects on bone geometry. An image database of ovariectomized mice and the same treated with combined pharmacological and physical treatment at week 18 and 24 was used. The motivation for using this database was that large geometric changes were induced by the combined treatment over time (Roberts, et al., 2020). The challenge of such large differences in bone shapes present to the registration process is representative of image-based shape analyses.

The two first studies ("Objective 1" and "Objective 2") demonstrated the disagreement in uniaxial location and orientation and how these are propagated into differences between surface meshes. The first study highlighted that the location in the axial direction was the most sensitive to the choice of rigid registration protocol. It also demonstrated that the axial location disagreement between the registrations for the mouse observations at week 24 was almost always of the same sign. This signifies that the two rigid registrations hold systematic differences for all samples at week 24, possibly because the first protocol uses age-wise references. The differences in the axial direction may be more prominent than any other because of the length difference over age. This could also be an indicator that in a sequence of registrations, as in the first protocol, the initial co-registration of the reference mouse at the two ages to create "Reference week 18" and "Reference week 24" will have an effect on later registrations of the other samples. The second study confirmed the relatively larger axial location disagreement, as the displacements between the meshes extracted from the two differently processed images had larger z-component values on average. The mesh differences were larger for the mesh samples at week 24, confirming the findings in "Objective 1".

The third study highlighted how the differences in the registration outputs were propagated to the PCA model. The mean shapes derived from the differently processed databases were found to have different 3D locations, with more prominent differences in the z-direction. This result was consistent with the prominent axial location disagreement that 'Objective 1' measured. The differences in the registration propagated differences in the shape variance especially along the z axis, with higher variance for the mesh database of "Registration 1". This led to differences in the proportions of total variation explained by each mode. For example, the first two modes together explained ~80% of the total geometry variation in both PCA models, but the individual contributions of these modes substantially differed between the PCA models. The 3D shape profiles of Modes 1, 2 and 3 showed similar global features, including proximal and distal expansion. However, one difference was that in "Registration 2" a local variation in the medial aspect of the proximal end was captured in Mode 2 (Figure 5.14). Mode 2 also described a cross-sectional expansion in diaphysis more prominent in "Registration 2" than in "Registration 1". Additionally, the ordering of the modes was also different, with Modes 4 and 5 flipped between the two registrations. Overall, similar variant features detected in the first modes in both PCA models suggested that the variations in features that occupy a large area in the bone were larger than the differences in surface meshes induced by registration. In contrast, differences in the modes describing more localised variations and the ordering of later modes were sensitive to the choice of registration, possibly because the smaller variations contained in these modes were comparable to registration-induced differences.

The different score distributions between the two rigid registration protocols substantially affect the mode categorisation to growth, treatment or disease-related modes. Particularly, differences in the individual score values affected the score changes and the statistical post-processing of the scores led to different categorisation of temporal modes. Figures 5.17 and 5.18 signify this aspect. A striking example of these differences was the score differences in the first two modes. Specifically, for "Registration 1", Mode 1 gave significant temporal changes in scores only in the treatment group, and conversely, Mode 2 gave significant changes only in the diseased group. Based on these results, the analysis suggests that Mode 1 describes a pure treatment effect, and Mode 2 describes a treatment effect as opposed to the disease progression. Therefore, the shape profiles of these modes demonstrate uniaxial

variations leading to length increase with proximal expansion for only “OVX+PTHML” and distal expansion only in the “OVX” group. In contrast, previous studies have shown length increase in both cohorts “OVX” and “OVX+PTHML” (Roberts et al., 2023). The PCA results for “Registration 2” better agree with the literature showing that Mode 1 and Mode 2 gave length increase in both groups and, therefore, they were successfully characterised as growth-related modes. Looking at the unexpected temporal changes of length for “Registration 1”, one could argue that the important differences in the score values, that lead to completely different biological interpretations of the PCA modes, could be strongly linked with the translational misalignment as shown in “Objective 1” and “Objective 2”.

Additionally, the differences in scores also lead to different median surface change profiles (Figures 5.19, 5.20 and 5.23). One striking example includes the direction of changes as described by Mode 3. Specifically, as highlighted in Figure 5.21, for “Registration 1”, the mode describes posterior shrinkage at the proximity, whereas, for “Registration 2”, it describes the thickening of the same feature. According to previous loading experimental and computational studies, as the bone remodelling is strain-driven, the posterior side exhibits high turnover leading to local thickening (Monzem et al., 2023; Cheong et al., 2021). Therefore, Mode 3 for “Registration 2” better explains expected anabolic treatment effects in the posterior compartment whereas it describes unexpected posterior shrinkage for “Registration 1. Although this mode for “Registration 2” also describes some distal shrinkage, which is not expected, this is of a low magnitude, and it is defeated by the distal uniaxial expansion of Mode 2 with higher magnitude. This variant feature might also be related to the distal tibia-fibula separation landmark which was considered for the further cropping of the bone volume.

Other shape changes such as diaphyseal cross-sectional expansion, thickening and increase of sharpness of the anterior crest, diaphyseal posterior thickening and proximal expansion of the anterior crest are described in both Registrations but in different order of importance (i.e., Modes 4 and 5) or with differences in magnitudes. Expansion of the proximal medial compartment was only detected when the database was processed with “Registration 2”, consistent with previous findings (Cheong et al., 2021). Differences in magnitudes of change in the anterior crest between the current study and Chapter 3 might

be due to the different examined bone volumes and how local and global rigid registration could lead to different alignments.

5.4.2 Evaluating advantages and disadvantages of the two registration protocols tested in this study

The rigid registration protocol "Registration 1" was utilised in Chapter 3. The major characteristic of this registration protocol was the definition of time-wise sub-references. The possible advantage coming from this choice of references is that mice of the same age might be more geometrically "close" to the reference possibly leading to better positional alignments than using one single reference A disadvantage, however, is the systematic bias to the reference geometry and translation of initial registration errors (registering week 24 to the reference week 18) to the subsequent ones (registering samples at week 24 to the sub-reference week 24). That was also indicated by 'Objective 1', which found systematic differences in the 3D axial position of the images at week 24 between the two registrations.

The rigid registration protocol "Registration 2" was applied using two references, i.e., (a) the randomly chosen mouse sample at week 18 for the registrations, and (b) the registered mouse-specific image at week 18 for registration corrections. If the first registration using the original reference of a randomly chosen mouse at week 18 resulted in a registered image, misaligned from the registered mouse-specific image at week 18, then the registration output of (b) was used. The advantage is that the rigid registration protocol would possibly lead to random errors (noise) as it doesn't follow a sequence of registrations, with the output of the first affecting the next. Thus, it is expected that PCA would have the ability to throw the registration noise away at the very last modes. The disadvantage, however, is that the structural differences between ages challenged the registrations, resulting in spurious misalignments (local minima in the optimisation). Therefore, the usage of mouse-specific reference was sometimes necessary, with almost half of the samples using reference (b), meaning that the registrations used multiple references within the examined population.

5.4.3 Limitations and future work

One limitation of the current study is that it only tested two different image rigid registration protocols, but other alternatives could be examined in the future. One possible protocol could define the mean image over the mice at week 18 and use this as the reference. Additionally, future work should explore the effect of other techniques of instance alignment that use meshes rather than images, such as Procrustes Analysis and other extensions of it (Gower, 1975). For instance, some relevant mouse bone applications with similar bone scale and image resolutions have applied mesh instance alignment (Brown et al., 2014; Hoshino et al., 2023).

A future study could compare the shape variations described by the PCA model with other predictive models. For example, effects over time, such as due to ageing, disease and treatment can be predicted by mechanoregulation algorithms and expressed as local bone formation and resorption (Pereira et al., 2015;Cheong et al., 2020). However, some challenges in such comparison can be anticipated. Firstly, mechanoregulation models are predictive, and have their own set of assumptions and accuracy considerations, separate to those of a data projection method such as PCA. Secondly, the PCA approach used here discretised the bone surface geometry using triangulated meshes such that changes to this geometry is described as displacement of mesh nodes. In contrast, the changes occurring on the bone surface due to mechanoregulation activity are typically expressed in computational models as new (formed) and/or deleted (resorbed) voxels (Cheong et al., 2021). Therefore, a common basis for comparison of the results of both models needs to be developed.

5.5 Conclusions

This study highlighted the non-negligible effect of image rigid registration in the PCA-based model for localised assessment of the bone shape variations. It approached this by conducting a comprehensive comparative analysis of two image registration protocols, and it signified important disagreement in the uniaxial translation of the bone images with higher differences in advanced ages, maximum of ~24 slices. These positional differences were propagated into the PCA models, with “Registration 1” not being able to accurately describe uniaxial variations of bone shapes happening due to growth change. In contrast, PCA detected expected length increases in both treatment and disease groups when data

were processed using “Registration 2”. Less location disagreement between the registration outputs was found in the translation in the x and y axes and the orientation of the bones. However, variation in the medial aspect in the proximal end was only described by “Registration 2”. Additionally, Mode 3 in “Registration 1” described a rotation along the z axis, mostly noticeable in the proximal end. This variation was not biologically interpretable. Differences in the score values also affected the mode categorisation into treatment/disease and growth-related variations and the biological interpretation of the modes, leading to different calculated surface changes over time. Considering all the above, this study signified the better performance of “Registration 2”. It suggested that the PCA framework better detects expected geometric changes in mouse tibia over time when the image bone samples were registered and corrected using the mouse -specific images at the start of the examination period at week 18. This study proposed a comprehensive approach for providing a better understanding in evaluating how the instance alignment affects image-based shape analysis including shape comparisons.

5.6 References

Baab, K. L. (2013). The impact of superimposition choice in geometric morphometric approaches to morphological integration. *Journal of Human Evolution*, *65*(5), 689–692. https://doi.org/10.1016/j.jhevol.2013.07.004

Brown, J. M., Naylor, A., Buckley, C., Filer, A., & Claridge, E. (2014). 3D articulated registration of the mouse hind limb for bone morphometric analysis in rheumatoid arthritis. *Biomedical Image Registration: 6th International Workshop, WBIR 2014, London, UK, July 7-8, 2014. Proceedings 6*, 41–50. https://doi.org/10.1007/978-3-319-08554-8_5

Campbell, G. M., Tiwari, S., Grundmann, F., Purcz, N., Schem, C., & Glüer, C.-C. (2014). Three-dimensional image registration improves the long-term precision of in vivo micro-computed tomographic measurements in anabolic and catabolic mouse models. *Calcified Tissue International*, *94*(3), 282–292. https://doi.org/10.1007/s00223-013-9809-4

Castro-Mateos, I., Pozo, J. M., Cootes, T. F., Wilkinson, J. M., Eastell, R., & Frangi, A. F. (2014). Statistical shape and appearance models in osteoporosis. *Current Osteoporosis Reports*, *12*,

163–173. https://doi.org/10.1007/s11914-014-0206-3

Chan, E. F., Harjanto, R., Asahara, H., Inoue, N., Masuda, K., Bugbee, W. D., Firestein, G. S., Hosalkar, H. S., Lotz, M. K., & Sah, R. L. (2012). Structural and functional maturation of distal femoral cartilage and bone during postnatal development and growth in humans and mice. *The Orthopedic Clinics of North America*, *43*(2), 173–185, v. https://doi.org/10.1016/j.ocl.2012.01.005

Cheong, V. S., Campos Marin, A., Lacroix, D., & Dall'Ara, E. (2020). A novel algorithm to predict bone changes in the mouse tibia properties under physiological conditions. *Biomechanics and Modeling in Mechanobiology*, *19*, 985–1001. https://doi.org/10.1007/s10237-019-01266-7

Cheong, V. S., Roberts, B. C., Kadirkamanathan, V., & Dall'Ara, E. (2021). Positive interactions of mechanical loading and PTH treatments on spatio-temporal bone remodelling. *Acta Biomaterialia*, *136*, 291–305. https://doi.org/https://doi.org/10.1016/j.actbio.2021.09.035

Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*, *40*, 33–51.

Grassi, L., Väänänen, S. P., & Isaksson, H. (2021). Statistical shape and appearance models: development towards improved osteoporosis care. *Current Osteoporosis Reports*, *19*, 676–687.

Hoshino, Y., Takechi, M., Moazen, M., Steacy, M., Koyabu, D., Furutera, T., Ninomiya, Y., Nuri, T., Pauws, E., & Iseki, S. (2023). Synchondrosis fusion contributes to the progression of postnatal craniofacial dysmorphology in syndromic craniosynostosis. *Journal of Anatomy*, *242*(3), 387–401. https://doi.org/10.1111/joa.13790

Li, W., Kornak, J., Harris, T., Lu, Y., Cheng, X., & Lang, T. (2009). Hip fracture risk estimation based on principal component analysis of QCT atlas: a preliminary study. *Medical Imaging 2009: Biomedical Applications in Molecular, Structural, and Functional Imaging*, *7262*, 470–477.

Lu, Y., Boudiffa, M., Dall'Ara, E., Bellantuono, I., & Viceconti, M. (2016). Development of a protocol to quantify local bone adaptation over space and time: Quantification of reproducibility. *Journal of Biomechanics*, *49*(10), 2095–2099.

https://doi.org/https://doi.org/10.1016/j.jbiomech.2016.05.022

Monzem, S., Valkani, D., Evans, L. A. E., Chang, Y.-M., & Pitsillides, A. A. (2023). Regional modular responses in different bone compartments to the anabolic effect of PTH (1-34) and axial loading in mice. *Bone*, *170*, 116720. https://doi.org/https://doi.org/10.1016/j.bone.2023.116720

Pereira, A. F., Javaheri, B., Pitsillides, A. A., & Shefelbine, S. J. (2015). Predicting cortical bone adaptation to axial loading in the mouse tibia. *Journal of the Royal Society Interface*, *12*(110), 20150590.

Roberts, B. C., Arredondo Carrera, H. M., Zanjani-pour, S., Boudiffa, M., Wang, N., Gartland, A., & Dall'Ara, E. (2020). PTH(1–34) treatment and/or mechanical loading have different osteogenic effects on the trabecular and cortical bone in the ovariectomized C57BL/6 mouse. *Scientific Reports*, *10*(1), 8889. https://doi.org/10.1038/s41598-020-65921-1

Roberts, B. C., Giorgi, M., Oliviero, S., Wang, N., Boudiffa, M., & Dall'Ara, E. (2019). The longitudinal effects of ovariectomy on the morphometric, densitometric and mechanical properties in the murine tibia: A comparison between two mouse strains. *Bone*, *127*, 260–270. https://doi.org/https://doi.org/10.1016/j.bone.2019.06.024

Chapter 6

Assessment of treatment effects on osteoporotic mouse tibia reveals novel 3D geometric features with important influence on bone strength

ABSTRACT

Monotherapy of Parathyroid hormone is acknowledged as an anabolic treatment of osteoporosis; however, the side effects of this treatment have sparked concerns regarding its long-term efficacy. In the last few decades, research has shifted its focus towards combining pharmacological treatments with exercise, demonstrating increased anabolic effects on bone geometry and strength. In preclinical studies, murine tibia models are extensively utilised for rapid testing of these treatment strategies, involving the administration of PTH and concurrent *in vivo* compressive mechanical loading. However, the geometric changes over time resulting from these interventions are typically quantified using predefined scalar parameters which cannot describe important localised geometric alterations. Moreover, limited research has been conducted to identify the structural determinants of bone mechanics and its reliability depends on how the geometry is parameterised. The present study aims to identify the 3D geometric features whose variation due to growth, disease or treatment exerts influence on bone strength. Particularly, this study extended the framework developed previously to extract shape changes in osteoporotic and treated mouse tibia and employed the Partial Least Squares to determine the geometric modes that contribute to bone strength changes over time. The treatment effects were found to be absent in the distal section but pronounced in the diaphysis and proximity of the tibia. These were characterised by an increase in the cross-sectional area globally and a thickening of the posterior-lateral segment. This local thickening was associated with a prominent periosteal apposition. These findings offer a novel and compact description of the changes in the entire tibia cortex geometry due to treatment and lead to new insights into how these changes regulate its mechanics. The framework developed here can be used to evaluate new treatment plans' efficacy in terms of changes in both geometry and bone strength in preclinical mouse studies.

6.1 Introduction

Numerous preclinical and clinical studies aim to quantify the disease and treatment effects in both bone geometry and mechanics to understand the mechanisms of osteoporosis and the anabolic impact of novel treatment plans. This focus is motivated by the fact that osteoporotic bones exhibit weak structural and mechanical properties compared to healthy bone, leading to an elevated risk of fracture. Given that the primary objective of drug development for osteoporosis is to decrease fracture risk (Cummings et al., 2013), research should focus on exploring the treatment effects on both geometry and mechanics.

An example of the current FDA-approved osteoporosis treatment includes anabolic drugs based on a synthetic form of PTH. As elaborated in the literature review in Chapter 4, the PTH effects are characterised by an anabolic window of two years with increased osteogenic response. However, after this period, this anabolic response decreases and the continuation of this treatment is not recommended. Additionally, it has been shown that PTH-based drugs come with side effects such as atypical fractures, bone cancer and stroke. Therefore, there is an increased interest in exploring the alternative of combined pharmacological treatment with physical exercise (Jee & Tian, 2005).

In preclinical studies, longitudinal murine models have been extensively developed to rapidly test pharmacological or/and physical treatments (Vandamme, 2014). *In vivo* micro-Computing Tomography is the gold standard to generate longitudinal image data to monitor the bone alterations within the intervention period (Viceconti & Dall'Ara, 2019). The standard methods for assessing geometric changes include a pre-defined set of scalar parameters to parameterise the bone geometry (Bouxsein et al., 2010). These are typically extracted from small volumes of interest such as the midshaft (Lu et al., 2016). More details on the literature review on this aspect are provided in Chapter 3.

Traditionally, bone mechanical properties are assessed experimentally (Jepsen et al., 2015). However, such an invasive approach allows only cross-sectional analysis. Relatively non-invasive microCT-based Finite Element (microFE) modelling of the mouse tibia was validated and accurately measured bone strength and stiffness, and local deformation under uniaxial compression (Oliviero et al., 2018). In a later study, the failure criterion was optimised to accurately predict bone strength (Oliviero, et al., 2021).

This model was very recently used to investigate the longitudinal effects of monotherapies and combination of PTH and mechanical loading on tibia mechanics in an ovariectomized mouse model (Roberts et al., 2024). Therein, the correlation between morphometric parameters and bone strength and stiffness was assessed to investigate the structural determinants of bone mechanics.

The current state-of-the-art methods (morphometric parameters and microFE models) in longitudinal murine tibia models have shown the effects of disease and treatments in both geometry and mechanics. Osteoporosis was shown to result in stable cortical thickness and area of C57BL/6 mice and no statistically significant changes in bone strength over time (weeks 18–24 of age) (Roberts et al., 2019). In contrast, treatments were found to have anabolic effects on both geometry and strength. For example, the combination of parathyroid hormone and *in vivo* mechanical loading was shown to have a higher anabolic effect than any monotherapy (Meakin et al., 2017; Cheong et al., 2021; Rooney et al., 2023). Roberts et al. (2020) showed the increased effect as described by an increase in cortical standard morphometric properties and in a later study (Roberts et al., 2024) also showed a 29% average increase in bone strength between 18 and 24 weeks. Additionally, correlations of both cortical area and thickness with bone strength were found to be high.

Major limitations of the current state-of-the-art primarily arise from the scalar nature of the geometric parameterisation. While temporal differences in structural quantifiers such as cortical thickness and cortical area provide an approximate measure of temporal bone turnover, they lack preciseness. They are scalar parameters averaged over limited regions of interest and subsequently give a globalised approximation of the bone structure, assuming that it can be modelled as a hollow cylinder. Therefore, they fail to describe important regional or more localised bone adaptation to treatments (Cheong et al., 2021; Monzem et al., 2023), which may also be mechanically important leading to anabolic effects in mechanical properties (Roberts et al., 2023). Additionally, inter-dependent geometric features such as thickness and area could possibly result in misleading high correlations with mechanical properties. Considering the simplified nature of the geometric properties such as cortical thickness and area, correlations between these and the mechanical properties to understand the changes in strength through

changes in geometry would need to be treated very carefully. Hence, it is still a challenge to identify the 3D geometric features in bone structure that are mechanically important, changes of which due to disease or treatments will generate important changes in bone mechanical properties.

Considering the aforementioned, the present study suggests using the localised parameterisation of the bone tibia structure as described in Chapter 3. Taking advantage of this complete 3D parameterisation of the shapes, it proposes an alternative statistical model to investigate the shape variations as projected in the mechanical properties' domain. Particularly, motivated by the PCA approach, described in Chapters 3 and 4, which answers the question 'which geometrical features exert the maximum variation in the tibia structure', the current study in this chapter determines 'which geometric features exert the maximum influence on bone strength'. Specifically, the implementation of a regression model, called Partial Least Squares (PLS), is proposed.

PLS is a statistical linear regression model that could extract patterns of bone geometry variations that capture important portion of strength variation. Mathematically speaking, PLS identifies the orthogonal vectors of a predictor variable X (here geometry) by maximizing the covariance between predictor X and response Y (here strength)(Rosipal & Krämer, 2005). There are several PLS applications in bone research. It has been combined with bone imaging of femurs and Finite Element models to improve risk assessment (Aldieri et al., 2020). It has also been extensively used for developing prediction *in silico* bone models (Lekadir et al., 2015; Zhang et al., 2016; Oura et al., 2017). To the best of the author's knowledge, Partial Least Squares is yet to be applied on longitudinal mouse data to investigate how disease/treatment affects bone geometry and mechanics.

The aim of the current study is to provide a complete map of the treatment impact on the entire mouse tibia cortex that exerts change in bone strength in osteoporotic murine models. Motivated by prior findings that treatment increases bone strength, the development of a novel PLS-based shape model has the potential to unlock hidden information about systematic geometric alterations induced by treatments and provide them in a compact manner describing their effects in bone mechanics.

6.2 Materials and methods

6.2.1 Input images

The examined population of mouse tibia constituted the same image data considered in Chapter 5. The examined treatment intervention was the concurrent administration of Parathyroid Hormone PTH (1-34) and *in vivo* mechanical loading. Both subgroups, labelled as "OVX" and "OVX+PTHML" consist of mouse tibia observation of age 18 and 24 weeks. The two ages of ovariectomised and treated mice were included to investigate the disease progression and the treatment effects over time. The input images of these bone samples were extracted from Cheong et al., (2021)

The 3D grayscale images underwent the PCA preprocessing framework as described in Chapter 3 to consistently discretise the surfaces of the tibia samples and produce the matrix of the predictors X for the PLS model. Briefly, the images were registered, binarized, corrected and mapped and surface meshes with a fixed number of nodes were extracted from all bone samples. The matrix of the coordinates of the nodes for all samples constituted the predictor matrix X. This study applies the rigid registration protocol 2, "Registration 2", as suggested in Chapter 5.

6.2.2 Assessment of bone strength

Bone strength values were extracted by micro-FE models of the bone samples (Roberts et al., 2023), and formed the response Y used in the PLS. The input images analysed therein included 80% of the tibia length. Particularly, the fibula, the volume proximally to the growth plate and the distal end where the condyles are exhibited were removed (Cheong et al., 2020b) and further cropping of the trabecular region and the distal fibula-tibia separation was herein applied, as detailed in Chapter 5.

The finite element models of the tibia were created from the bone images used in Cheong et al., (2021)– same input image for both geometry discretization and bone strength assessment. The pipeline which converted the images to micro Finite Element Models converted the image voxels to hexahedral 8-noded elements (Cheong et al., 2020a). Homogenous, isotropic and linear material properties, with Young's modulus E = 14.8 GPa and Poisson's ratio v = 0.3, characterized the bone material. The distal end was fixed and the proximal was compressed applying an axial displacement equal to 0.1mm. The

failure load was estimated as the force required for 10% of the nodes to reach the yield strain of –14420 microstrain for the minimum principal strain (Oliviero et al., 2021). To reduce the boundary effects, 10% of the length at each extremity was excluded from the strength calculation, thus analysing 60% of the total length. Hence, the volume of interest in analysing bone strength (Roberts et al., 2024) and geometric variations (the present study) is primarily identical. The strength predictions derived from the micro-FE models were validated against experiments using Digital Volume Correlation with a mean error of 9% (Oliviero et al., 2018).

6.2.3 Partial Least Squares

PLS regression was employed to detect the most important geometric features that concurrently 1) alter over time and 2) exert the maximum change in bone strength under compression. As a statistical learning tool, PLS investigated the relationship between bone geometry and strength, with the former being the predictor (denoted **X**) and the latter the response (denoted **Y**).

Let's denote the input matrices, **X** and **Y**. **X**: [M x 3N] is the geometry matrix where M the number of mice and N the number of nodes on the surface mesh (details in Chapter 3). **Y**: [M x 1] is the strength vector which summarized all bone strength samples of the examined population (section 6.2.2). Firstly, PLS centred both **X** and **Y** matrices. For simplicity, the following description assumes that **X** and **Y** are already centred. Then, it calculated the cross-covariance matrix between X and Y and performs singular value decomposition to calculate the modes for **X** and **Y** (De Jong, 1993) (section 2.8.3).

PLS was applied in Matlab, using the function 'plsregress'. Alongside, it provided the coefficients for the linear regression model that links shape modes with strength. From now onwards, the following modes and scores always refer to those obtained using PLS, and further analysis uses the shape modes and shape scores representing the **X**.

6.2.4 Temporal geometric variations and treatment categorisation

Similarly to the study elaborated in Chapter 3, the current data used longitudinal images from two ages of treated and diseased mice. This study, however, investigated the temporal variations in bone strength

as mediated through geometry. A shape mode was characterized as treatment-related if the median shape score change (MSSC), i.e., median of the differences of score values over time, of the shape mode in the “PTHML” group is statistically significant different (s.s.d.) than the MSSC in the “OVX” group. This condition could be further split into the following scenarios: 1) pure treatment effect: MSSC is s.s.d. from 0 only in the treatment group, 2) opposed effects: MSSCs of both groups are s.s.d. from 0 but have opposite signs, or MSSC in the “OVX+PTHML” is 0 but not in the “OVX” group, or finally the MSSCs are s.s.d form 0, but it is lower in the “OVX+PTHML” group than the one in the “OVX”, 3) summative effects: MSSCs of both groups are s.s.d. from 0 and the MSSC in the “OVX+PTHML” group has higher magnitude than the one in the “OVX” group. Specifically, Wilcoxon signed rank test was performed for testing differences in shape scores between ages in each group and the Mann-Witney U test for testing differences in temporal shape score changes between groups. A one-sided Wilcoxon signed rank test was also used to reveal the sign of the temporal score change.

Next, the surface changes described by each treatment-related shape mode were computed by assessing the temporal difference in median shape scores between ages within both “OVX” and “OVX+PTHML” groups. The shape score changes was then used to scale the shape mode vectors. It is important to note that the surface change map is represented by a 3D vector for each node of the surface mesh, measured in length units. To visualise these changes, shape mode-specific median profiles were generated by scaling the shape mode vectors with the median score value of a specific group at week 18. Subsequently, the surface change map was plotted over these profiles .

6.2.5 Mode effects on bone strength and its change over time

To assess the contribution of different shape modes on bone strength, the linear regression model was created to reconstruct the bone strength by using the linear coefficients that PLS provided. Firstly, geometry was reconstructed by the shape modes using linear composition. The linear composition used the mean shape and the shape modes, scaled using the individual shape scores for each mouse sample. The reconstructed geometry was considered the predictor **X** in the linear regression model. Following this, the strength samples for each mouse sample (response **Y**) were reconstructed via the linear

regression model that linked the strength with the reconstructed geometry.

Combinations of growth and treatment-related modes were sequentially inserted in the regression model. A similar model was constructed for analysing the bone strength change using the temporal differences of the score samples instead of their values. This approach allowed for the examination of how changes in geometric features influenced alterations in bone strength over time.

To evaluate the effectiveness of each geometric mode in describing bone strength, normalized residuals between the reconstructed strength and the real samples were calculated, indicating the contribution of each mode's to describing the bone strength. The absolute residuals were also calculated from the singed residuals, and the mean value was also reported as the mean absolute residual (MAR). The R-squared was also provided with MAR to further assess the model's fit to the real strength data.

6.3 Results

6.3.1 Assessment of bone strength

The media strength for the "OVX+PTHML" and "OVX" groups at week 18 is 42N and 37N, respectively. The strength increased over time in the treatment group ($p<0.05$), with the median value of 52N. In contrast, in the "OVX" group, the strength remained stable ($p>0.05$) at 36N. A significant difference was found between the groups at week 18 as shown in Figure 6.1.

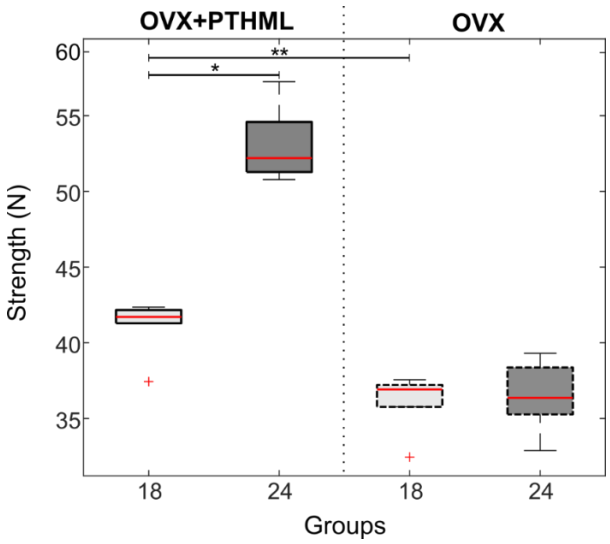


Figure 6.1: Boxplots of strength values for the "OVX+PTHML" (solid box) and "OVX" (dashed box) groups and both ages at week 18 (light grey) and week 24 (dark grey).

6.3.2 Partial Least Squares

The first four PLS modes describe 87% of the total geometry variance and 98% of variance in strength, with the first two occupying a substantial portion (77% and 83%) in both geometry and strength (Figure 6.2) within the examined population of all “OVX+PTHML” and “OVX” mice. As shown in Figure 6.3, Mode 1 describes uniaxial variation at the proximal end, capturing 68% of the total variance in geometry and 58% in strength. Mode 2 describes a global cross-sectional variation but more prominently in the posterior compartment along the proximal end and diaphysis with 9.3% of the total variance in geometry and 25% in strength. Finally, Mode 3 and Mode 4 describe uniaxial variation at the distal end along with some more localized variations detected at the proximal end (Figure 6.4). Mode 3 captures 6.2% and 8.6% of the total variance in geometry and strength, respectively. Mode 4 occupies 3.7% in geometry and 6.2% in strength.

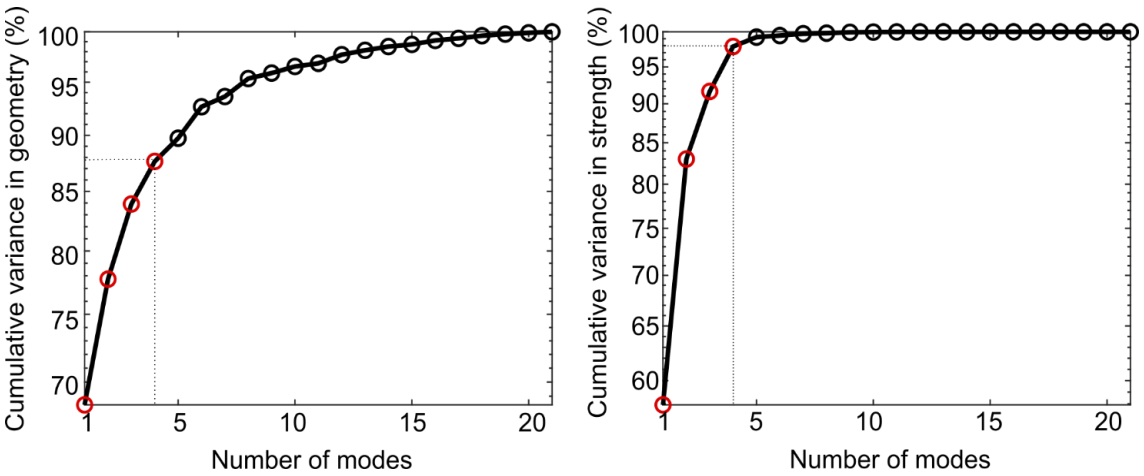


Figure 6.2: Percentage of the total variation in geometry and strength captured by PLS modes. Red coloured circles correspond to the first 4 modes.

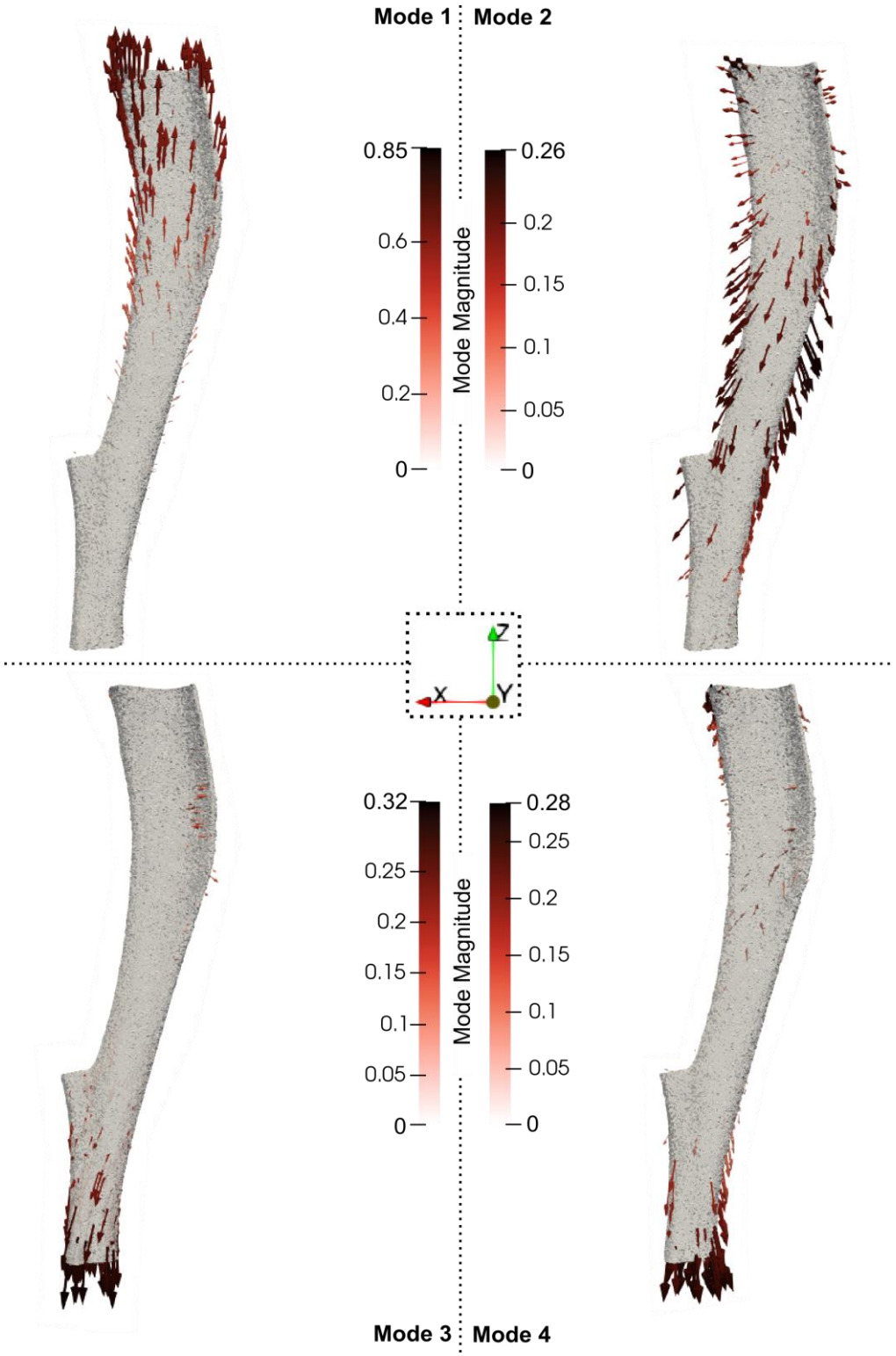


Figure 6.3: The 3D profiles of the first 4 PLS modes illustrated as vectors plotted on the mean shape. The vectors are scaled and coloured by mode magnitude at each node. Different ranges in the colorbars were used.

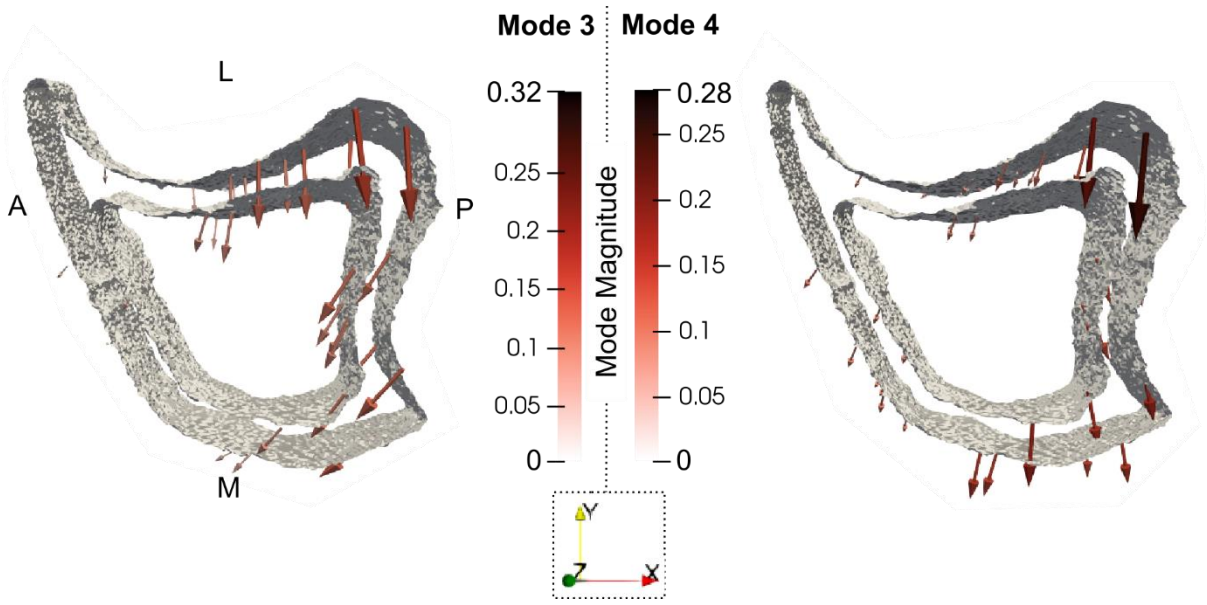


Figure 6.4: Mode 3 and 4, zoom at the proximal end.

6.3.3 Temporal geometric variations and treatment categorisation

The statistical tests of the temporal geometry score changes showed that Mode 1 describes growth change with median shape score change (MSSC) being statistically significant different (s.s.d.) from 0 in both groups and of similar magnitude (Figure 6.5). Modes 2 and 3 were shown to exhibit MSSC s.s.d. from 0 in the "OVX" group only, while Mode 4 was identified to characterize temporal changes in the "OVX+PTHML" group only.

The 3D profiles of the mode-specific geometric temporal changes as described by Mode 1, 2, 3 and 4 in both groups are exhibited in Figures 6.6 ("OVX+PTHML") and 6.7 ("OVX"). For the treatment group "OVX+PTHML", Mode 1 describes the expansion of the proximal end of maximum magnitude equal to 0.22 mm leading to an increase in length. For the "OVX" group, the mode describes the same geometrical change of similar magnitude. Mode 2 describes the global expansion of the cross-sectional area in the "OVX+PTHML" group with an average magnitude around the diaphysis equal to 0.07 mm. As can be seen, by the vectors and the contour colours on the endosteum and periosteum in Figure 6.6 and Figure 6.7, it describes prominent periosteum apposition at the posterior compartment in the proximal epiphysis equal to 0.073 mm and low turnover in the similar location in the endosteum (0.01 mm). In the midshaft, the periosteum apposition at the posterior compartment was equal to 0.065 mm, also higher in magnitude than the endosteum deletion (0.03 mm). In contrast, this mode describes the

opposite shape variation in the "OVX" group, with small magnitudes of resorption lower than 0.018 mm and located at the diaphysis (Figure 6.7).

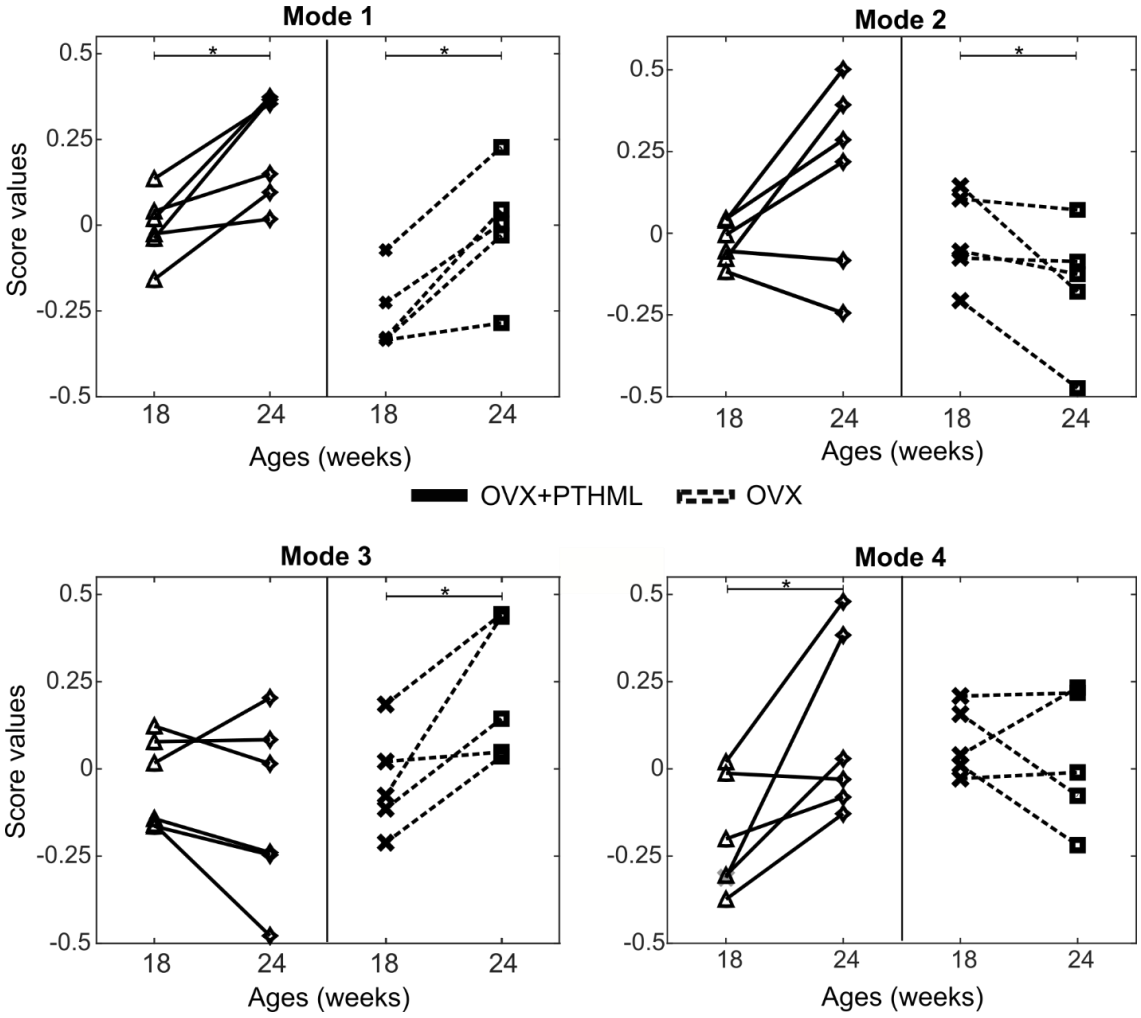


Figure 6.5: Treatment categorization of Modes 1, 2, 3 and 4. Mode score values of individual mice are shown at week 18 ("OVX+ML", △; "OVX", ×) and at week 24 ("OVX+ML", ◇; "OVX", □). Lines ("OVX+ML", solid; "OVX", dashed) connect mode scores of individual mice between the two time points. Asterisks (*) highlight changes with time of median mode scores in a group that are statistically significantly ($p<0.05$) different from zero.

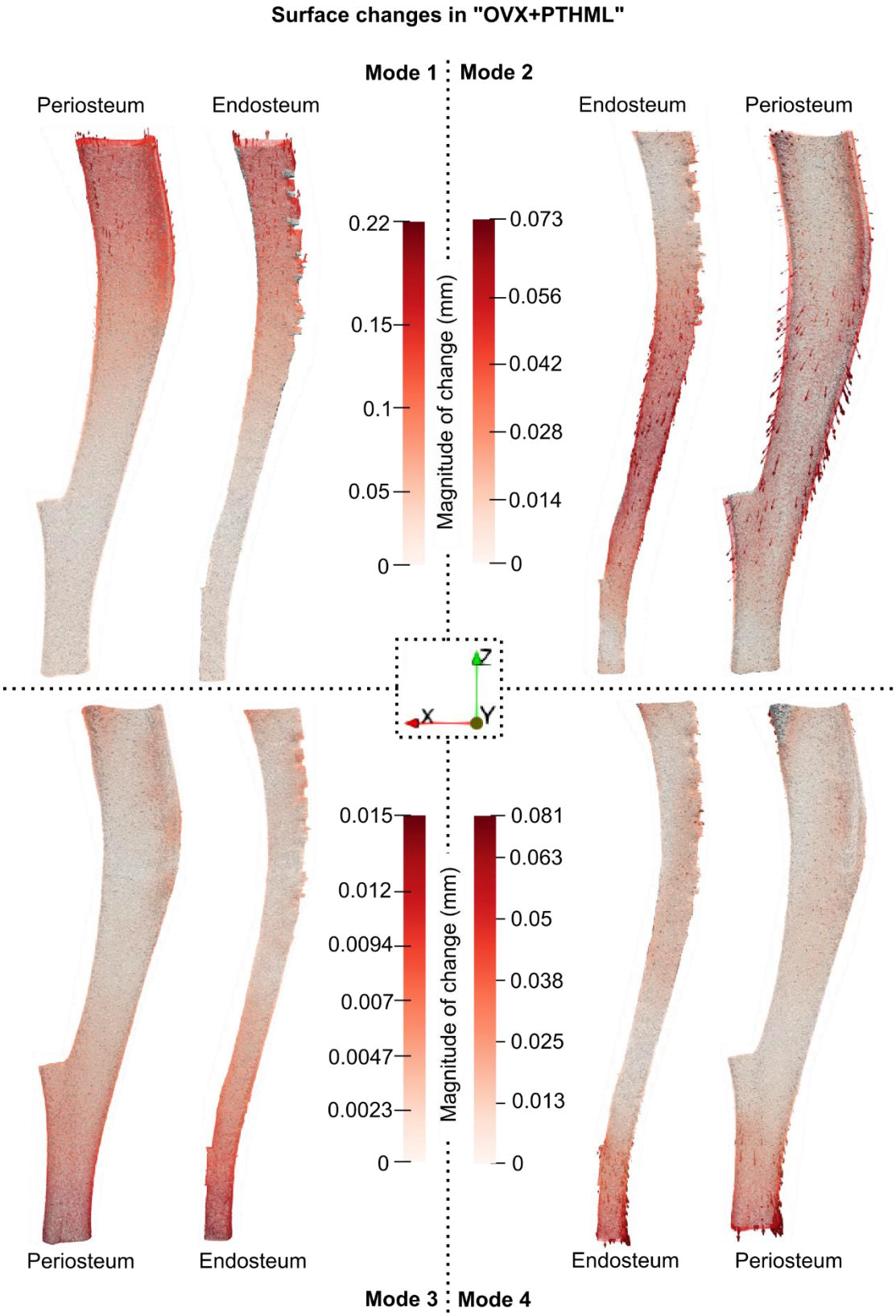


Figure 6.6: Changes in the endosteum and periosteum profiles described by PLS modes 1–4 for the "OVX+PTHML" group. Profiles are obtained by adding to the mean shape the mode vector scaled by the corresponding median scores at week 18 (solid grey surface) and week 24 (coloured wireframe). Arrows indicate the direction of the change, starting from the week 18 bone surface and their length and colour indicate the magnitude of geometrical change over time (a higher magnitude is indicated by a deeper red).

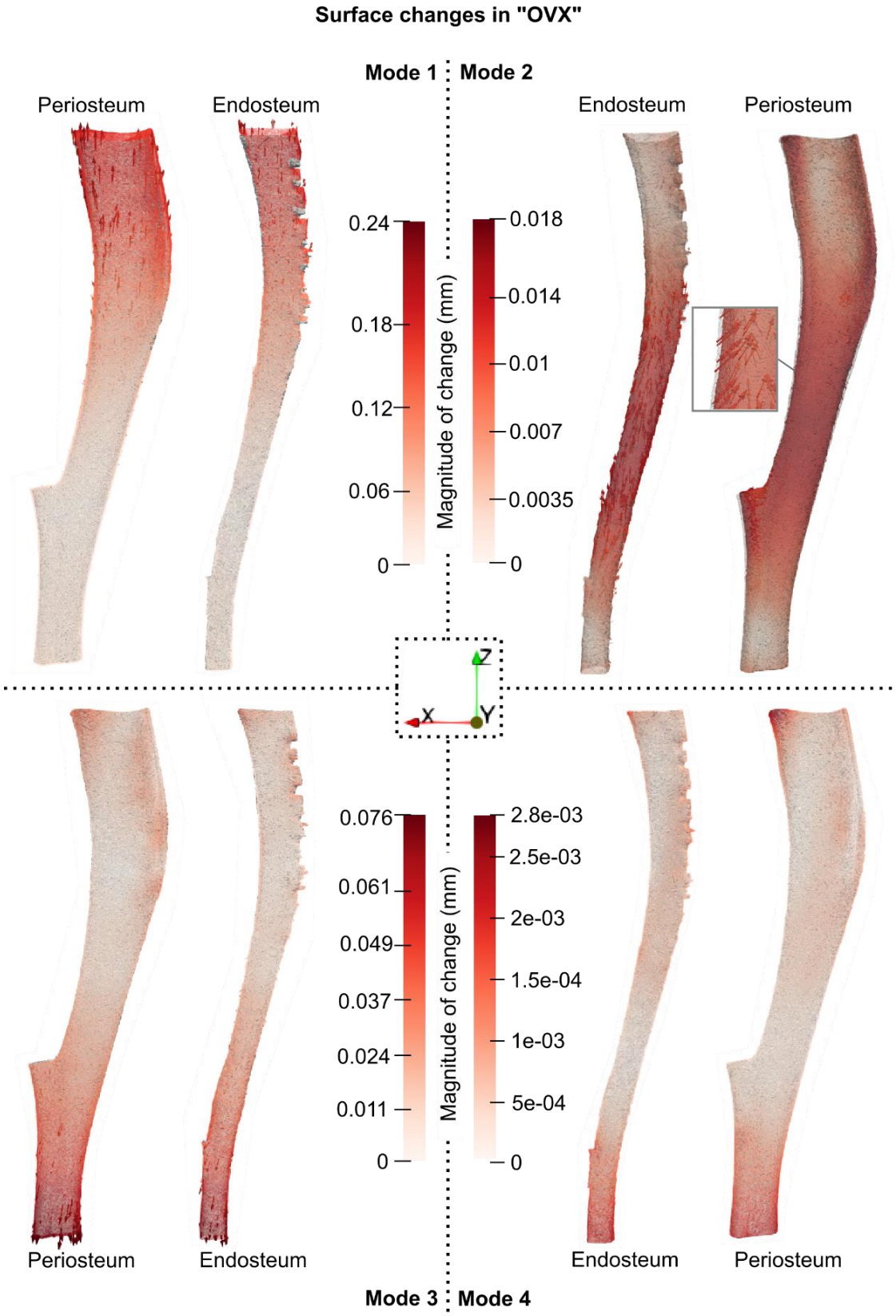


Figure 6.7: Changes in the endosteum and periosteum profiles described by PLS modes 1–4 for the "OVX" group. Profiles are obtained by adding to the mean shape the mode vector scaled by the corresponding median scores at week 18 (solid grey surface) and 24 (coloured wireframe). Arrows indicate the direction of the change, starting from the week 18 bone and their length and colour indicate the magnitude of geometrical change over time (a higher magnitude is indicated by a deeper red). A zoomed-in view of the midshaft region is also shown for Mode 2 indicating the inward vector direction.

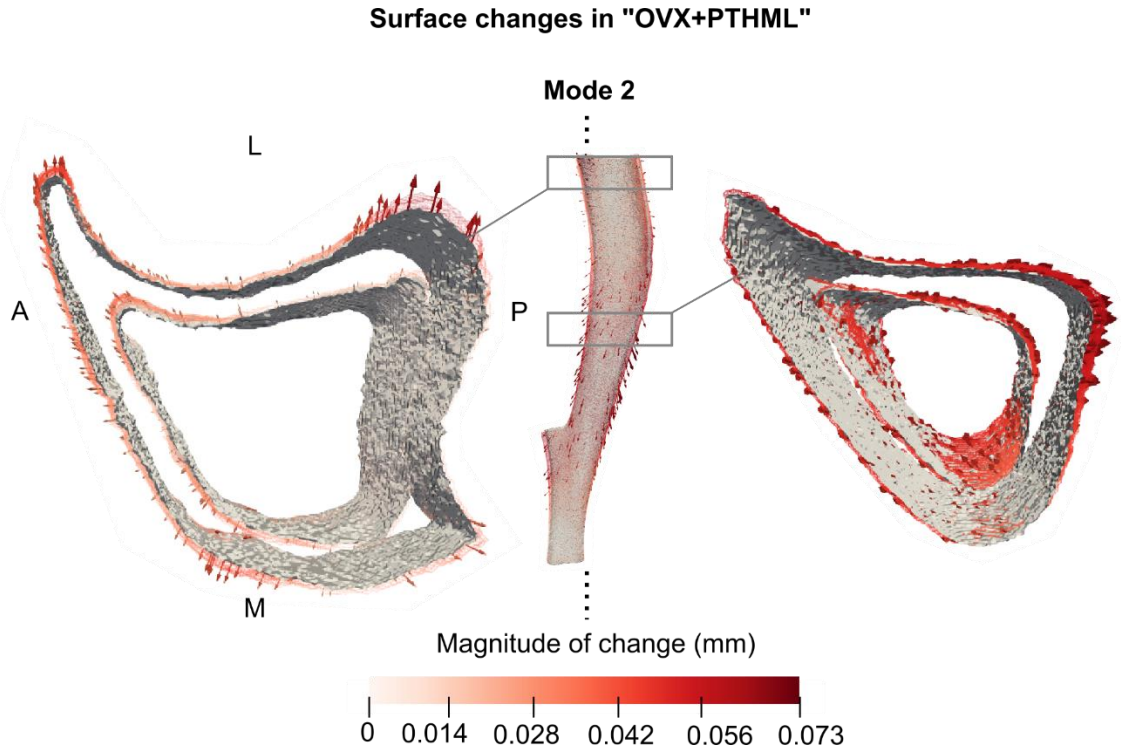


Figure 6.8: Surface changes as described by Mode 2 for the "OVX+PTHML" group. Zoomed-in view of two different regions: proximal and midshaft. The grey surface represents the median mode shape for week 18 and the "OVX+PTHML" group, while the coloured in the median profile for week 24 and same group. The redder the area, the higher the magnitude.

Mode 3 and 4 captured an overall expansion of the distal end together with other less prominent and very localised changes. The combined information of the mode shapes (Figure 6.3) and the score changes (Figure 6.5) suggested that Mode 3 and 4 describe similar temporal geometric changes, i.e., distal expansion, with each one representing each group solely. Together, these described an overall expansion of the distal end equal to 0.066 mm (Mode 4: 0.081 mm expansion, and Mode 3: 0.015 mm shrinkage) in the treatment group (Figure 6.6) and 0.073 mm in the diseased one (Figure 6.7). Along with this growth change, these modes were also shown to exhibit shrinkage of the proximal posterior compartment lower than 0.06 mm in the treatment group (Figure 6.9). This temporal change controverted the expansion of the posterior expansion of higher magnitude in the same location described by Mode 2 (Figure 6.8), leading to an overall proximal posterior expansion of 0.013 mm (= 0.073 – 0.060). For the diseased group, the combined modes describe a shrinkage of 0.015 mm at the same location. Finally, Mode 4 also captured the expansion of the proximal medial aspect equal to ~0.035 mm.

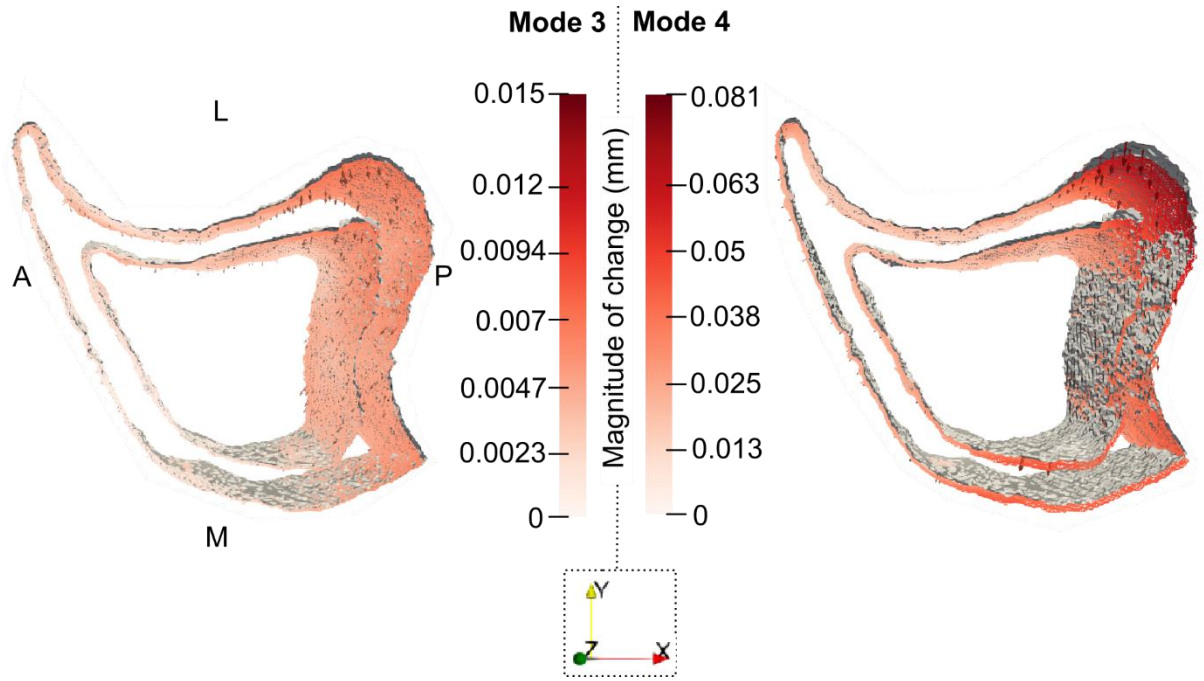


Figure 6.9: Surface changes of Modes 3 and 4 for "OVX+PTHML" at the proximal end. The grey surface represents the median mode shape for week 18 and the "OVX+PTHML" group, while the coloured in the median profile for week 24 and same group. The redder the area, the higher the magnitude.

Other more localised geometric changes only due to treatment (MSSC≠0 only in the "OVX+PTHML" group) were described by Mode 6 and Mode 8, capturing 3% and 2% of the total geometric variance, respectively, and less than 0.3% of the strength variation altogether. The shape mode profiles (Figure 6.10) demonstrate variations at the anterior ridge (both), at the medial-posterior edge (Mode 6) and the posterior aspect (Mode 8) in the proximal epiphysis of the tibia. By calculating the surface changes, Mode 6 was associated with a maximum 0.07 mm uniaxial expansion of the posterior aspect on the proximal end. Mode 8 was associated with 0.03 mm cross-sectional expansion of a similar location, along with the increase of the sharpness of the anterior ridge with the same magnitude.

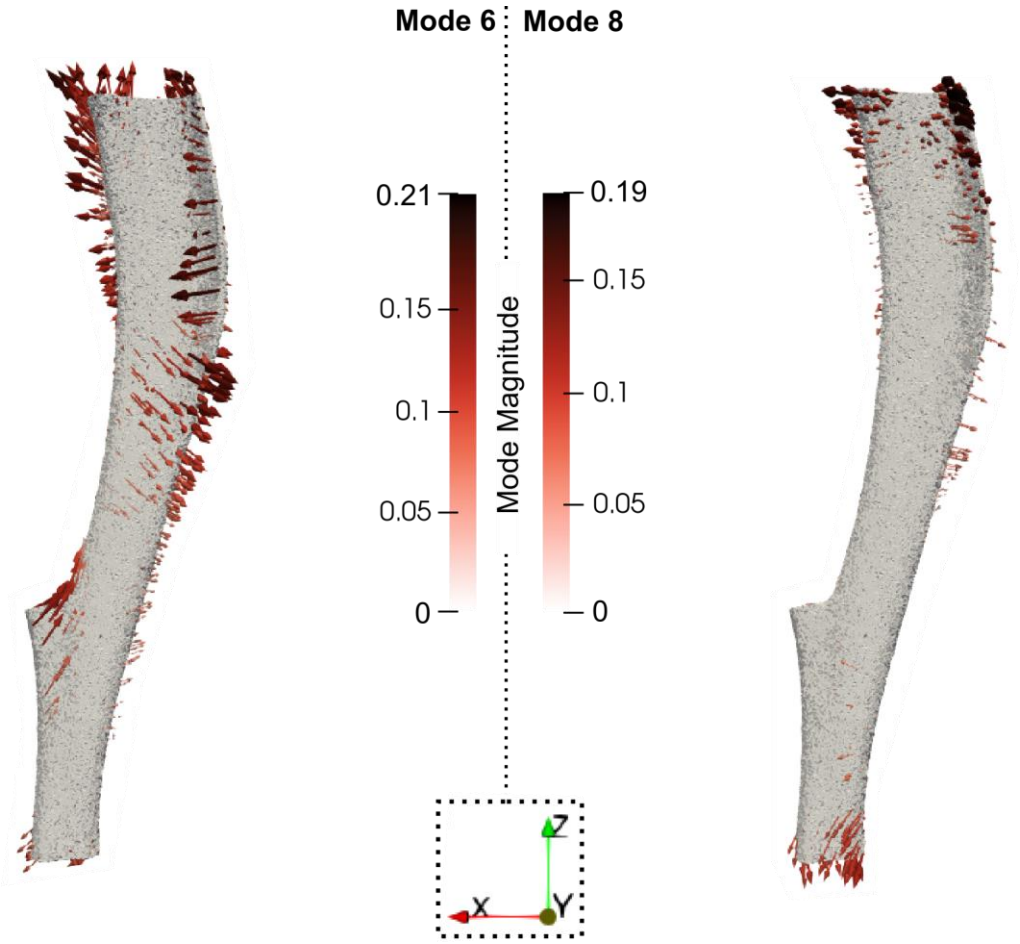


Figure 6.10: Shape profiles of Modes 6 and 8 illustrated as vectors plotted on the mean shape. The vectors are scaled and coloured by mode magnitude at each node.

6.3.4 Mode effects on bone strength and its change over time

The use of the important modes (growth and treatment-related modes: Modes 1, 2, 3, 4, 6, 8) to predict the bone strength in both groups, "OVX+PTHML" and "OVX", has shown a very good fit to the real response with mean absolute relative error 1.7% and 1.9%, and correlations R^2=98% and R^2=90%, respectively. For the bone strength change between weeks 18 and 24, MAR was equal to 9% and 78% for the treatment and the diseased groups, respectively.

The decrease of the errors with the number of modes inserted in the reconstruction model of the geometry and subsequently in the linear regression model of strength is summarized in Table 6.1. Figure 6.11 illustrates the decrease of the individual normalized residuals and the increase of the R^2 for both groups and both ages while gradually adding more shape modes in the strength reconstruction model. Specifically, for the "OVX+PTHML" group, the MAR was reduced from 6.2% to 1.7% when

sequentially adding Mode 1, 2, 3, 4, 6 and 8 to reconstruct the geometry. Similarly, the R^2 was increased from 64% to 98%. For the "OVX" group, the mean absolute normalized residual was reduced from 12% to 1.9%, while the R^2 was increased from 0.14 to 0.90. These trends highlighted the contribution of each mode in the strength prediction in both diseased and treatment groups and both ages.

Table 6.1: Mean absolute normalized residuals (MAR, %) for both strength and strength change prediction in both groups using a combination of Modes in the reconstruction model.

Modes	"OVX+PTHML"		"OVX"
	Strength (N)	Strength Change (N)	Strength (N)
1	6.2	53	12
1, 2	4.5	24	6.5
1–3	3.9	24	3.4
1–4	2.1	13	2.6
1–4, 6	1.6	9.6	2.0
1–4, 6, 8	1.7	9.0	1.9

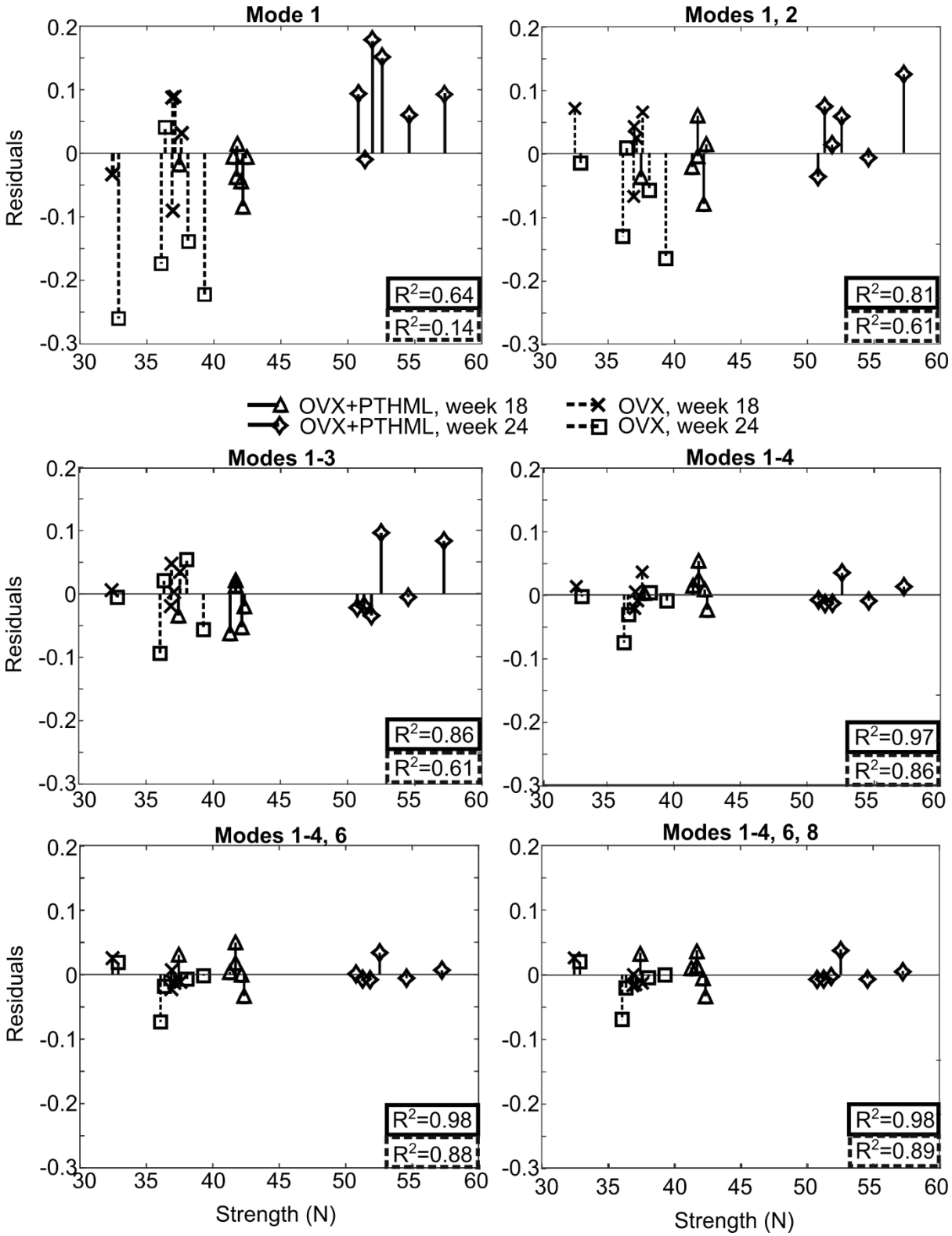


Figure 6.11: Model normalized residuals for bone strength. The model to predict the bone strength for all samples in both groups and both ages are updated adding more modes and residuals and R-squared are calculated for every model and given for each group "OVX+PTHML" group is in black solid line and "OVX" black dashed line. Residual values of individual mice are shown at week 18 ("OVX+ML", △; "OVX", ×) and at week 24 ("OVX+ML", ◇; "OVX", □).

Figure 6.12 shows that the MAR of bone strength changes decreased from 53% to 9% in the treatment

group. This demonstrates that when the model is used to predict the change in bone strength between the two ages in the treated group, its predictability increases with the number of modes used in the model. Although MAR in the "OVX" group decreased from 730% to 78%, the regression model is not useful in this case because there is no significant change in strength in this group and as such the strength change observations are noisy. As all MAR values in the "OVX+PTHML" group are positive, it suggests that Mode 1 describes an average and rough approximation of bone strength increase over age, underestimating the bone strength change in this group. Contrarily, it shows a high overestimation of the bone strength change in the "OVX" group with negative mean residuals across mouse samples. Additionally, Mode 3 which describes the distal expansion between weeks 18 and 24 in the "OVX" group (Figure 6.7) was shown to reduce the prediction errors only in the "OVX" group, while the residuals in the "OVX+PTHML" remained the same. Similarly, as Mode 4 describes similar distal expansion in the treatment group, it was shown that it reduced the residuals in the "OVX+PTHML" group, with a reduction of the MAR from 24% to 13%.

Mode 2 was found to describe a compact treatment effect of changes in geometry and strength due to treatment. While explaining the globalized increase of the cortical area and the local thickening of the posterior-lateral compartment for "OVX+PTHML" group (Figure 6.6), adding Mode 2 in the model to predict bone strength and its change decreased normalized residuals and increased the fraction of variation in strength described in the "OVX+PTHML" group up to 81%. Contrarily, while Mode 2 described the minor cross-sectional shrinkage and thinning in the "OVX" group (< 0.018 mm in magnitude), its contribution to strength prediction resulted in stabilizing the bone strength estimation. This led to an increase of R^2 for the "OVX" group from 14% to 61%.

In contrast with Mode 2, Mode 6 and 8 were shown to influence only a very small area of the proximity with minor influence in describing bone strength and its change (R^2: no change in the treatment group, and 4% increase in the diseased group, and MAR: less than 5% reduction).

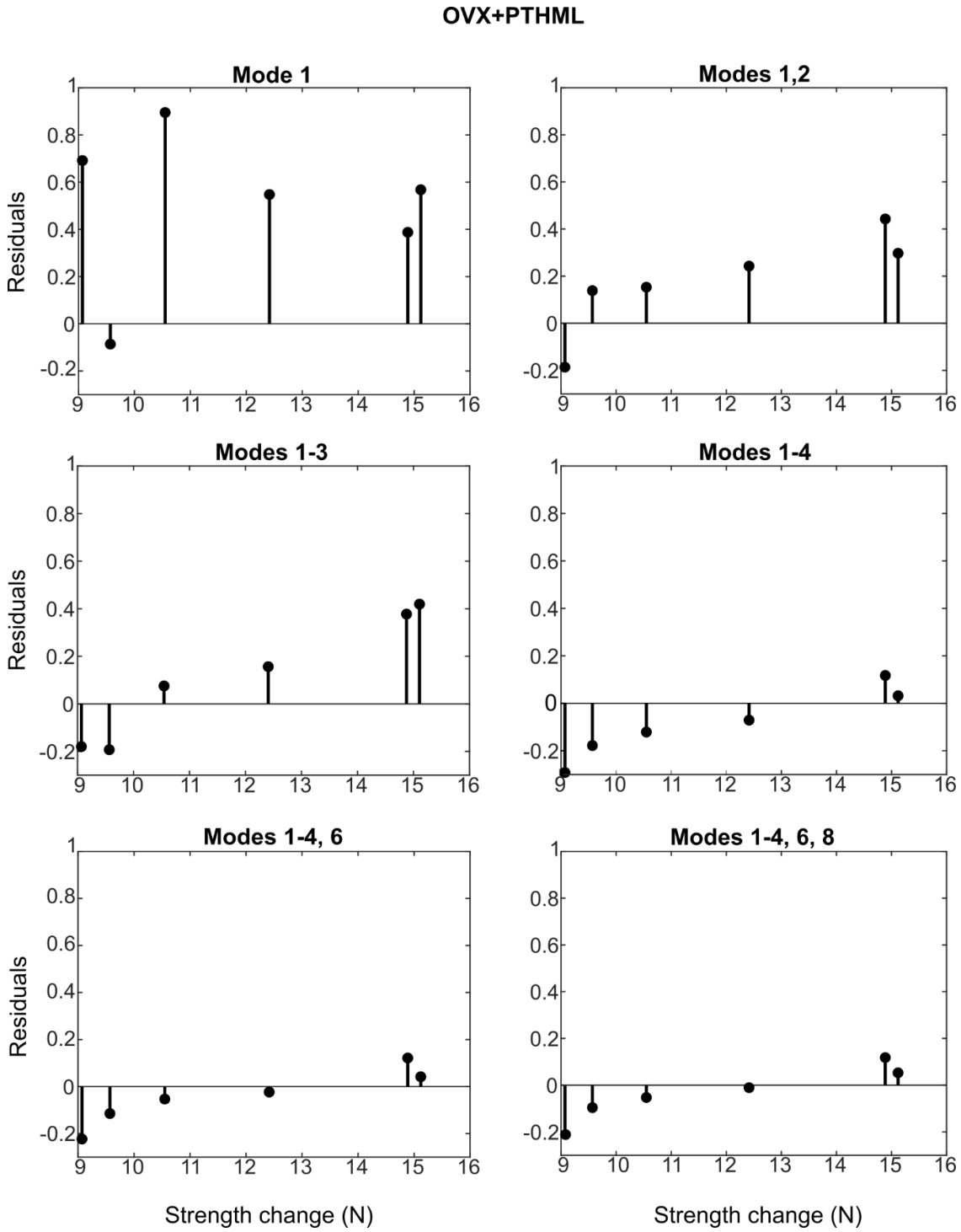


Figure 6.12: Model prediction errors for bone strength changes in the "OVX+PTHML" group. The model to predict the bone strength changes between ages for all samples in both groups is updated adding more modes and residuals and R-squared are calculated for every model and given for each group.

6.4 Discussion

The goal of this study was to identify the 3D geometric mouse tibia features that are affected due to osteoporosis and treatment and exert important influences on bone strength. With the motivation that

treatment efficiency could be expressed by the anabolic effects in bone mechanics, this study identified the structural determinants of bone strength and provided a linear model that links 3D geometry variations due to treatment with bone strength changes.

This study considered the disease cohort of osteoporosis and the treatment cohort of combined PTH administration with exercise. Longitudinal image data of the diseased and treated mice corresponding to two ages of 18 and 24 weeks were processed using the pipeline elaborated in Chapter 2. The localized parameterization of each bone instance constituted the geometric predictor in the PLS model. The PLS provided four important geometric modes that describe substantial portions of variation in both geometry and compressive strength, i.e., >77% (Figure 6.2). Mode 1, 3, and 4 were determined to describe growth changes with proximal and distal expansion (Figure 6.6, Figure 6.7). Mode 2 was determined to be the main mode describing important geometric alteration due to treatment with an increase of the cross-sectional area and local thickening of the posterior-lateral compartment (Figure 6.6). Other minor localized changes at the proximal end due to treatment were captured by Mode 6 and 8 (Figure 6.10). Mode 1, 2, 3, and 4 seemed to explain an important portion of the variation in strength (83%) and they were crucial for strength description with 2.1% and 13% errors for bone strength and its changes over time, respectively. Mode 6 and 8 showed low contribution to describing bone strength.

Mode 1, 3 and 4 primarily captured uniaxial variation detected on the extremities of the bone. The score distributions and score changes revealed that Mode 1 describes the proximal expansion in both groups, whereas Mode 3 describes distal expansion in the disease group and Mode 4 describes similar shape variation in the treatment group. The map of the median surface changes signified the proximal expansion equal to 0.22 mm and 0.24 mm in the treatment and the diseased group and distal expansion of 0.066 mm in the treatment and 0.073 mm in the diseased group. This led to a 0.28 mm and 0.31 mm increase in length in the “OVX+PTHML” and “OVX”, respectively. Although the increase of length is surprisingly higher in the “OVX”, the values are close with only less than 1 voxel difference. A possible explanation for the higher changes in “OVX” is that the two groups belong to different generations of mice processed in subsequent studies. The results of the length measurements published by (Roberts et

al., 2023) suggest that the median length increase is equal to 0.35 mm in the “OVX+PTHML” group and 0.34 mm in the “OVX” group. The difference in the values of length changes might be because the cited study measured the length including 80% of the total tibia length, in comparison with the present study which proceeded with further cropping including 65% of the total length.

Mode 1, which describes proximal expansion, captured almost 60% of the strength variation. Although the “OVX” observations exhibit stable strength over time (Figure 6.1), Mode 1 described temporal variation in length associated with strength differences. This finding could be explained by the fact that both strength (Figure 6.1) and length (data derived from Roberts et al., (2024)) are lower in the “OVX” group than in the “OVX+PTHML” group. Therefore, Mode 1 described the inter-group structural and mechanical differences. This was confirmed by the differences in the score distributions between the “OVX+PTHML” and “OVX” groups at week 18 (Mann Witney test: $p<0.01$, Figure 6.4), signifying that the groups are biologically distinguishable at week 18. This finding was expected as these data were generated from two different generations of mice, and therefore other sources of variations such as genetic factors could have affected the growth of these two populations.

On the other hand, the scores in Mode 2 were not statistically significantly different between the two groups at week 18. With this finding, we could claim that the temporal variations that Mode 2 described were primarily associated with structural changes induced by the treatment or/and the disease. Figures 6.5 and 6.6 illustrate the mode-specific temporal changes in the “OVX+PTHML” and “OVX” groups. The profiles suggested an overall almost cross-sectionally uniform expansion of the diaphysis with a magnitude of 0.07 mm in the “OVX+PTHML” group. More importantly, it showed the local thickening at the posterior-lateral compartment which is possibly a result of the loading of the bone, as previous murine experimental and predictive loading models have shown (Rooney et al., 2023). Particularly, Figures 6.5 and 6.7 signified higher periosteal apposition than endosteum resorption in two segments of the bone, proximal and midshaft, in line with mechanoregulation algorithms (Cheong et al., 2021). This result would also explain the concurrent increase of cortical area and thickness in the standard morphometric parameters (Roberts et al., 2023). However, the current results are 3D localised surface

changes that describe the bone turnover in a compact way regardless of whether it is very localised or global across the entire bone volume.

In the OVX group, Mode 2 showed almost no change in the bone shape (<0.015), which is consistent with the standard morphometric parameters, showing no statistically significant changes in the midshaft (area and thickness) between weeks 18 and 24 (Roberts et al., 2019). Although the statistical tests of the scores herein showed significant results for "OVX" and not for "OVX+PTHML", the surface changes using the median score values suggest the opposite: important treatment effect with high magnitude changes and negligible minor changes due to disease (Figure 6.6, 6.7). The negligible changes in the cortical bone which lead to stable bone strength is a limitation of the current osteoporosis model as other bone sites have demonstrated the catabolic effects of the disease (Monzem et al., 2022). Nevertheless, this study has successfully demonstrated the anabolic effects of the osteoporosis combined treatment.

Other localized changes detected on the proximal end included thickening of the posterior region and expansion of the medial aspect. These changes were distributed among Modes 2, 6 and 8 and they were characterized as treatment effects. These results corroborate the prominent increase in the proximal cortical area as shown in (Monzem et al., 2023). None of the important modes described any cross-sectional variation in the distal epiphysial bone apart from the uniaxial expansion, indicating the absence of bone turnover in that region. Previous studies have confirmed this statement (Cheong et al., 2021; Monzem et al., 2023).

The novelty of the present study arises from the relationship between the geometric modes and the strength. The implementation of the Partial Least Squares established a novel model that considered the geometric modes to describe bone strength in both diseased and treatment groups. This linear model showed a very good fit with the real bone strength when Modes 1, 2, 3 and 4 were included in the geometry reconstruction formula for both "OVX+PTHML" and "OVX" (MAR<0.021 and R^2>86%). That indicates that 1) the increase of the strength over time is growth-related and associated with difference in length, 2) the further increase of strength due to treatment is due to the increase of the

cross-sectional area and local thickening at the posterior aspect, and 3) the decrease of the strength to end up being stable over time (similar values between week 18 and 24) in the OVX group come with the slight catabolic effect described by Mode 2. The last attribute signifies the importance of Mode 2 in concurrently describing the anabolic effects of the treatment but also the small catabolic effects of the disease in geometry and the stable strength over time. This comes in agreement with the high correlations of the cortical area and thickness with bone strength reported by Roberts et al., (2024). Mode 6 and 8 showed a small influence slightly improving the strength description in both groups.

The linear regression model was found to better describe the strength and its change in the "OVX+PTHML" than "OVX". Particularly, the mean error in describing the strength and its change over time in the "OVX+PTHML" were very low equal to 1.7% and 1.9%. On the other hand, it resulted in high errors in predicting the strength change in the "OVX" group (78% mean error) highlighting that this is a model to describe only the treatment effect in strength as mediated through the treatment effects in geometry. That is because the linear regression model facilitated the geometry reconstruction using the growth and treatment-related modes only.

6.4.1 Limitations

The anabolic effects of the combination of treatments in geometry and mechanics were quantified between weeks 18 and 24. This period included two rounds of mechanical loading, performed every two weeks between weeks 18 and 22 and a treatment withdrawal period of another 2 weeks at the end, week 22-24. PTH was uniformly administered across the weeks 18-22. The current model investigated the overall changes observed in bone geometry and mechanics from the beginning till the end of the experimental period, including only the data generated at week 18 and week 24.

Future work should consider all time points to investigate the individual treatment effects of the intermediate stages of the treatment course. The first set of loading at week 19 was shown to have increased benefits in strength higher than ones from the second round at week 21 (Roberts et al., 2024). On the other hand, PTH was indicated to have a strong but small anabolic window, with a notable increase in resorption after week 20 (Cheong et al., 2021). The strong anabolic effect of PTH at the

beginning of the administration was also reported in an osteoporotic vertebra model (Levchuk et al., 2014). Additionally, the current experimental database includes 2 weeks of treatment withdrawal, observing stability or slight but not significant reduction of the structural and mechanical properties during this period (Roberts et al., 2024). Since these effects are time-dependent and dose-dependent, it is important to include all intermediate stages to assess the direct anabolic effects of the treatments at different ages and the following geometric changes after the end of the course.

The current model assumes a linear relationship between geometry and strength. However, nonlinear models are recommended for future work. Moreover, it was only used to identify these effects induced by the treatment by using only the growth and treatment-related modes for the geometry reconstruction. Further development is needed to add the geometric modes related to osteoporosis in order to investigate the impact on mechanics. Additionally, as reported previously in Chapter 3, the sample size is limited and therefore predictions using this PLS model should be treated very consciously. Therefore, it is utilized only to explore the link between the 3D geometric features and the bone strength, rather than developing a predictive model. Specifically, the R-squared and the residuals are used as indicators of how much contribution each geometric mode has in describing the bone strength. Finally, as Chapter 5 elaborated, instance alignment is crucial to accurately describe the shape variations. Future work should focus on optimising the alignment protocol, to reveal more systematic temporal alterations and minimize any noise inserted from the image rigid registration protocol. This could potentially improve the score distributions and reduce the variance of the scores in the subgroups of age.

6.5 Conclusions

In conclusion, this study developed a novel PLS-based shape model, identified the 3D geometric modes that describe geometric and strength changes over time and provided a complete assessment of these changes. The growth-related changes encompassed proximal and distal expansion leading to an approximate 0.30 mm increase of length over the age as described by Modes 1, 3 and 4. The ones categorised as treatment effects were shown to describe both global and localised changes in the mouse tibia shape. The most important treatment effect, described by Mode 2, entailed the uniform expansion of the diaphysis (0.07 mm) and the prominent local thickening of the posterior-lateral compartment in

the diaphysis and proximal epiphysis, exhibiting higher periosteal bone formation than endosteal deletion. This Mode encompasses compact information about a large area of the diaphysis, describing both anabolic treatment effects and slight catabolic disease effects. Other less important treatment effects, captured by Modes 6 and 8, were detected on the proximal end with expansion of the medial and posterior quadrant, indicating higher bone turnover at the proximity of the tibia compared to the distal epiphysis. Modes 1, 2, 3 and 4 effectively describe the bone strength variations of the mouse tibia subjected to PTH and *in vivo* mechanical loading. Specifically, the mean absolute normalized residuals were lower than 2.1% when describing both bone strength at the ages of 18 and 24 weeks by using these modes. This methodology identified the 3D geometric features that are mechanically important and exert influence on bone strength during a course of treatment in osteoporotic mouse tibia bones. Hence, drug development and treatment optimisation should target these 3D regions to enhance bone mechanics and reduce fracture risk. Furthermore, this model holds promise for evaluating the interactions of other treatments and optimising them based on their impact on bone strength.

6.6 References

Aldieri, A., Terzini, M., Audenino, A. L., Bignardi, C., & Morbiducci, U. (2020). Combining shape and intensity dxa-based statistical approaches for osteoporotic HIP fracture risk assessment. Computers in Biology and Medicine, 127, 104093. https://doi.org/10.1016/j.compbiomed.2020.104093

Bouxsein, M. L., Boyd, S. K., Christiansen, B. A., Guldberg, R. E., Jepsen, K. J., & Müller, R. (2010). Guidelines for assessment of bone microstructure in rodents using micro--computed tomography. Journal of Bone and Mineral Research, 25(7), 1468–1486. https://doi.org/10.1002/jbmr.141

Cheong, V.S., Roberts, B.C., Kadirkamanathan, V. and Dall'Ara, E. (2021). Positive interactions of mechanical loading and PTH treatments on spatio-temporal bone remodelling. Acta Biomaterialia, 136, 291-305. https://doi.org/10.1016/j.actbio.2021.09.035

Cheong, V.S., Campos Marin, A., Lacroix, D. and Dall'Ara, E. (2020a). A novel algorithm to predict bone changes in the mouse tibia properties under physiological conditions. Biomechanics and

modeling in mechanobiology, 19, 985-1001. https://doi.org/10.1007/s10237-019-01266-7

Cheong, V.S., Roberts, B.C., Kadirkamanathan, V. and Dall'Ara, E. (2020b). Bone remodelling in the mouse tibia is spatio-temporally modulated by oestrogen deficiency and external mechanical loading: A combined in vivo/in silico study. Acta Biomaterialia, 116, 302-317. https://doi.org/10.1016/j.actbio.2020.09.011Cummings, S. R., Cosman, F., Eastell, R., Reid, I. R., Mehta, M., & Lewiecki, E. M. (2013). Goal-directed treatment of osteoporosis*. Journal of Bone and Mineral Research, 28(3), 433–438. https://doi.org/10.1002/jbmr.1854

Jee, W. S. S., & Tian, X. Y. (2005). The benefit of combining non-mechanical agents with mechanical loading: a perspective based on the Utah Paradigm of Skeletal Physiology. Journal of Musculoskeletal and Neuronal Interactions, 5(2), 110.

Jepsen, K. J., Silva, M. J., Vashishth, D., Guo, X. E., & Van Der Meulen, M. C. H. (2015). Establishing biomechanical mechanisms in mouse models: practical guidelines for systematically evaluating phenotypic changes in the diaphyses of long bones. Journal of Bone and Mineral Research, 30(6), 951–966. https://doi.org/10.1002/jbmr.2539

Lekadir, K., Hazrati-Marangalou, J., Hoogendoorn, C., Taylor, Z., van Rietbergen, B., & Frangi, A. F. (2015). Statistical estimation of femur micro-architecture using optimal shape and density predictors. *Journal of Biomechanics*, *48*(4), 598–603. https://doi.org/10.1016/j.jbiomech.2015.01.002

Levchuk, A., Zwahlen, A., Weigt, C., Lambers, F. M., Badilatti, S. D., Schulte, F. A., Kuhn, G., & Müller, R. (2014). The Clinical Biomechanics Award 2012—presented by the European Society of Biomechanics: large scale simulations of trabecular bone adaptation to loading and treatment. Clinical Biomechanics, 29(4), 355–362. https://doi.org/10.1016/j.clinbiomech.2013.12.019

Lu, Y., Boudiffa, M., Dall'Ara, E., Bellantuono, I., & Viceconti, M. (2016). Development of a protocol to quantify local bone adaptation over space and time: Quantification of reproducibility. Journal of Biomechanics, 49(10), 2095–2099. https://doi.org/10.1016/j.jbiomech.2016.05.022

Meakin, L. B., Todd, H., Delisser, P. J., Galea, G. L., Moustafa, A., Lanyon, L. E., Windahl, S. H., & Price, J. S. (2017). Parathyroid hormone's enhancement of bones' osteogenic response to

loading is affected by ageing in a dose- and time-dependent manner. Bone, 98, 59–67. https://doi.org/10.1016/J.BONE.2017.02.009

Monzem, S., Gohin, S., Yagüe Ballester, R., de Souza, R., Meeson, R., & Pitsillides, A. A. (2022). An examination of two different approaches for the study of femoral neck fracture: Towards a more relevant rodent model. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 236(2), 199–207. https://doi.org/10.1177/09544119211053057

Monzem, S., Valkani, D., Evans, L. A. E., Chang, Y.-M., & Pitsillides, A. A. (2023). Regional modular responses in different bone compartments to the anabolic effect of PTH (1-34) and axial loading in mice. Bone, 170, 116720. https://doi.org/10.1016/j.bone.2023.116720

Oliviero, S., Giorgi, M., & Dall'Ara, E. (2018). Validation of finite element models of the mouse tibia using digital volume correlation. Journal of the Mechanical Behavior of Biomedical Materials, 86, 172–184. https://doi.org/10.1016/j.jmbbm.2018.06.022

Oliviero, S., Owen, R., Reilly, G. C., Bellantuono, I., & Dall'Ara, E. (2021a). Optimization of the failure criterion in micro-Finite Element models of the mouse tibia for the non-invasive prediction of its failure load in preclinical applications. Journal of the Mechanical Behavior of Biomedical Materials, 113, 104190. https://doi.org/10.1016/j.jmbbm.2020.104190

Oura, K., Otake, Y., Shigi, A., Yokota, F., Murase, T., & Sato, Y. (2017). Prediction of forearm bone shape based on partial least squares regression from partial shape. *The International Journal of Medical Robotics and Computer Assisted Surgery*, *13*(3), e1807.

Roberts, B.C., Cheong, V.S., Oliviero, S., Arredondo Carrera, H.M., Wang, N., Gartland, A. and Dall'Ara, E. (2024). Combining PTH (1-34) and mechanical loading has increased benefit to tibia bone mechanics in ovariectomised mice. Journal of Orthopaedic Research, 42(6), 1254-1266. https://doi.org/10.1002/jor.25777

Roberts, B. C., Arredondo Carrera, H. M., Zanjani-pour, S., Boudiffa, M., Wang, N., Gartland, A., & Dall'Ara, E. (2020). PTH(1–34) treatment and/or mechanical loading have different osteogenic effects on the trabecular and cortical bone in the ovariectomized C57BL/6 mouse. Scientific Reports, 10(1), 8889. https://doi.org/10.1038/s41598-020-65921-1

Roberts, B. C., Giorgi, M., Oliviero, S., Wang, N., Boudiffa, M., & Dall'Ara, E. (2019). The

longitudinal effects of ovariectomy on the morphometric, densitometric and mechanical properties in the murine tibia: A comparison between two mouse strains. Bone, 127, 260–270. https://doi.org/10.1016/j.bone.2019.06.024

Rooney, A. M., McNeill, T. J., Ross, F. P., Bostrom, M. P. G., & van der Meulen, M. C. H. (2023). PTH treatment increases cortical bone mass more in response to compression than tension in mice. Journal of Bone and Mineral Research, 38(1), 59–69. https://doi.org/10.1002/jbmr.4728

Rosipal, R., & Krämer, N. (2005). Overview and recent advances in partial least squares. *International Statistical and Optimization Perspectives Workshop" Subspace, Latent Structure and Feature Selection"*, 34–51.

Vandamme, T. F. (2014). Use of rodents as models of human diseases. Journal of Pharmacy & Bioallied Sciences, 6(1), 2. DOI: 10.4103/0975-7406.124301

Viceconti, M., & Dall'Ara, E. (2019). From bed to bench: How in silico medicine can help ageing research. Mechanisms of Ageing and Development, 177, 103–108. https://doi.org/10.1016/j.mad.2018.07.001

Zhang, J., Hislop-Jambrich, J., & Besier, T. F. (2016). Predictive statistical models of baseline variations in 3-D femoral cortex morphology. *Medical Engineering \& Physics*, *38*(5), 450–457.

Chapter 7

Conclusions

7.1 Contributions

In vivo high-resolution imaging is the gold standard for monitoring bone turnover and adaptation to experimental interventions. Despite the advancements in imaging technologies, processing bone images to describe and measure the geometric variations in 4D, over time and space, remains a challenge.

Currently, spatio-temporal changes are measured using a predefined set of scalar quantifiers. These quantifiers, which are averaged over specific bone regions, do not fully capture the local effects of interventions on the entire bone shape. This limited scope and dimensionality reduce the precision of the geometric assessment, posing significant challenges for studies investigating treatment effects on bone health.

The current project suggested a comprehensive, complete and precise assessment of bone surface variations using deformable registration and statistical learning tools. The standardised and localised surface discretisation of a mouse tibia bone population was analysed using Principal Component Analysis to investigate shape variations. By categorising them into systematic or random variations, this study categorised the variation into growth/disease/treatment modes and extracted systematic patterns of shape changes over time. The major advancements of the proposed framework are two:

i. Localized discretization of shapes: This enables detailed analysis of bone surface changes at all anatomical bone features.
ii. Extraction of systematic variation patterns: Using statistical tools and with longitudinal input data, the framework can characterise changes due to growth/disease or treatment over time.

The key contributions accomplished in each chapter are summarised in Figure 7.1 and they are also explained as follows:

1) **Novel statistical framework**: This study introduced a robust statistical framework for quantifying the treatment effects on tibia midshaft of osteoporotic murine bone models. This addresses a significant research gap where previous methods lacked precision in capturing local variations and treatment impacts on bone structures.
2) **Application in various complexities**: This model was tested in high-complexity scenarios,

analysing the entire mouse tibia cortex and various treatment strategies. This demonstrated its versatility across different levels of anatomical and experimental complexity, with some adjustments.

3) **Challenges and evaluation**: This study identified the challenges in the PCA model associated with applying the instance alignment in the entire tibia cortex. It specifically evaluated the influence of the rigid registration protocol on the accuracy of the PCA-based geometric assessment. This is a technical study which highlights a critical point for the creation of similar statistical frameworks. Finally, it suggests a better performance for one rigid registration protocol.
4) **Alternative Statistical Learning Model:** Finally, it proposed an alternative statistical learning model to investigate the relationship between bone geometry and mechanics. This approach provides new insight into understanding how bone shape changes affect bone strength, filling a gap in the current understanding of treatment impact on bone mechanics as mediated through geometric adaptations.

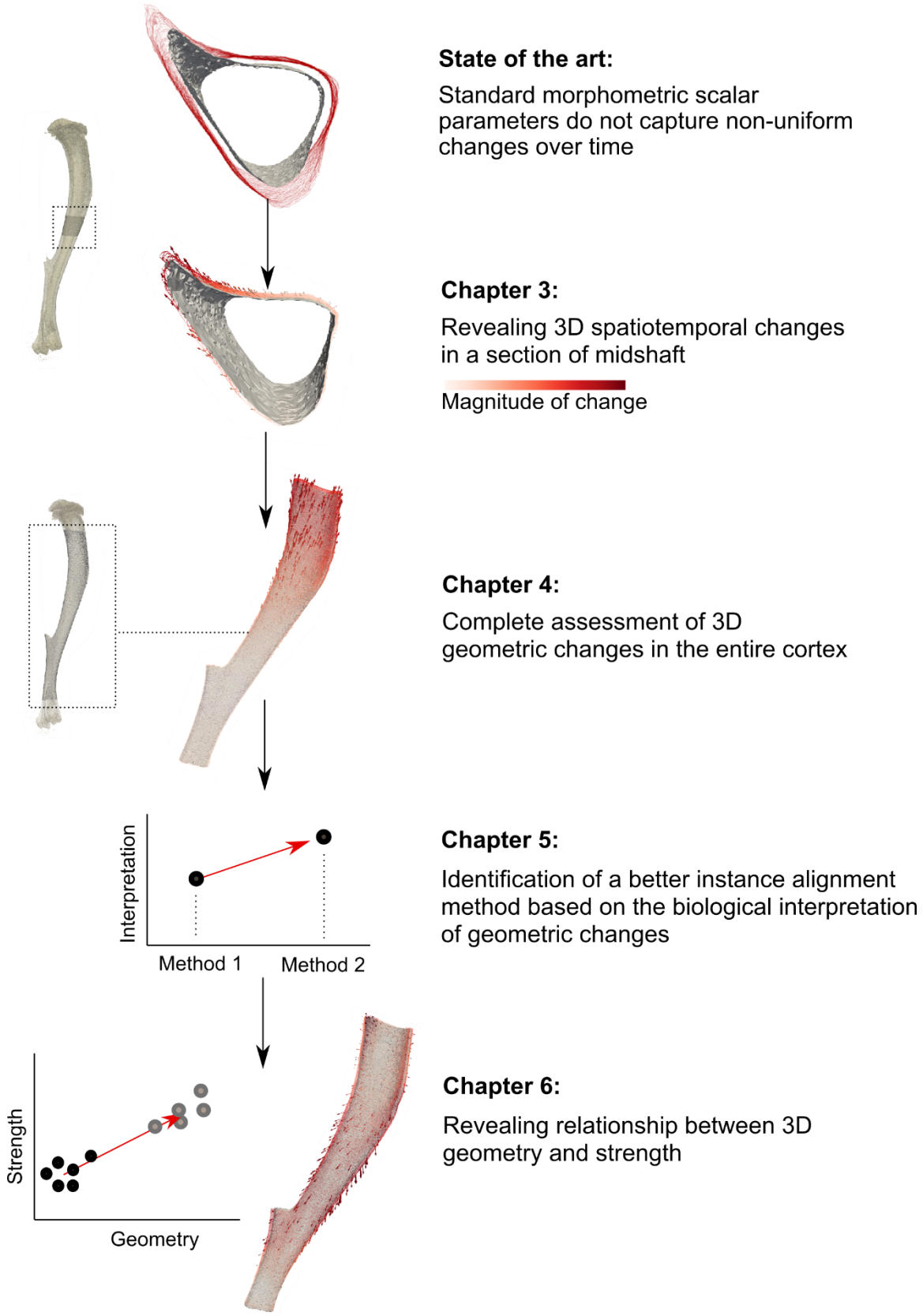


Figure 7.1: Key contribution in this thesis: Moving from simplistic scalar assessments to a full accurate and robust 3D assessment of bone shape changes over time and investigation of their influence on bone strength.

7.2 Thesis summary and conclusions

Chapter 3 elaborated on the first contribution of this project. It detailed the development of the PCA framework, and it served as a proof of concept. This chapter demonstrated how the statistical model was used to detect temporal shape variations related to treatment effects. Briefly, the framework included the following key steps: (1) image rigid registration for instance alignment, (2) binarization and (3) reference meshing. Following this, surface meshes were extracted with a systematic manner using the reference mesh and (4) mapping it to the other bone samples via image deformable registration.

Chapter 3 also included evaluation studies to assess uncertainties and errors at each step of the pipeline. Three robustness analyses were conducted:

Geometry correction:

- This study evaluated the mechanical importance of the cortical pores and trabeculae features, the features that were removed to enforce topological equivalence on the tibia midshaft.

It demonstrated that these features occupied a very small portion of the bone midshaft and that the strain distribution of the original structure and the processed one did not differ.

Mesh density analysis:

- This study focused on discretising the reference bone surface. It evaluated the average and maximum surface errors when coarsening the surface mesh.

It demonstrated that important anatomical features were preserved with coarsening leading to very small errors compared to the finest meshes.

Optimization of Nodal Spacing for deformable registration:

- This study measured the accuracy and precision of the algorithm and testing different sources of complexities in the input images and different spatial resolutions.

It measured the lowest errors, < 1 voxel size when Nodal Spacing was equal to 5 voxels.

The optimisation of certain parameters and the minimisation of errors proved the robustness of the proposed model. At this stage, the framework was applied in a limited bone section and limited dataset,

since the focus was the proof of concept.

Chapter 4 provided the implementation of the framework in more complex applications including a large dataset and the entire bone cortex (the second contribution). It described the required method refinement due to the high complexity and high dimensionality of the data. These refinements included removing the proximal trabecular region and the distal tibia-fibula separation, and adjustment of the registration protocol. Specifically, the further cropping of the extremities was applied to enforce topological equivalence across all bone samples. Additionally, an image rigid registration protocol using age-wise references was suggested to tackle the challenge of failed rigid registrations of some bone samples to a single reference. This PCA model successfully identified cross-sectional shape variations associated with treatment effects in global and localised areas. However, the model also found that some axial shape variations (along the length direction) captured by PCA were not biologically interpretable. For instance, the distribution of the scores for two modes suggested that the modes when combined exhibited consistent orientation of vector fields and small ranges of magnitudes, associated with instance misalignment in the axial axis. This was the initial evidence that rigid registration errors were inaccurately interpreted as shape variations by PCA.

Chapter 5, motivated by the findings in Chapter 4, aimed to comprehensively evaluate the effects of different rigid registration protocols on the assessment of the 3D shape variation on mouse tibia (the third contribution). This chapter conducted a comparative analysis of two rigid registration protocols with three objectives. These objectives investigated the discrepancy in (1) images, (2) surface meshes and (3) mean shapes, PCA modes and scores derived by the same original dataset but processed using different registration protocols. Regarding the tested registration protocols, the first protocol, “Registration 1”, used timewise references which were initially co-registered with each other. Following this, all mouse samples of one age were registered to the corresponding reference of the same age. The second protocol, “Registration 2”, used one reference mouse of age at week 18. All mice were registered to this reference. When registrations were challenged, a mouse -specific at age 18 registered image was used as a secondary reference. This study highlighted the sensitivity of the PCA model to

the rigid registration protocol. It successfully demonstrated that the PCA model detected expected shape variations when image data were processed using mouse-specific references for the instance alignment compared to using age-wise references.

Finally, Chapter 6 investigated alternative statistical models (the fourth contribution). It focused on a linear statistical regression model called Partial Least Squares (PLS) to detect variant geometrical features that exert an important influence on bone strength under uniaxial compression. Here, PLS was used as a type of supervised model using bone strength as a predictor variable. This variable fed the statistical model with more information to provide the shape modes which are mechanically important and explain variation in strength. The model provided compact information about the geometric changes that alter the bone strength highlighting local features in the tibia diaphysis. Thus, this chapter showcased how PLS could provide a more targeted analysis of geometric alterations related to changes in bone strength due to external interventions in osteoporotic mouse bone models.

7.3 Limitations

This project identified several shortcomings, each elaborated in different chapters but also summarized here to highlight the main ones.

The first limitation concerns the instance alignment protocol. As comprehensively demonstrated, choosing the correct protocol is crucial for developing a reliable PCA model. Given the limited sample size, the sensitivity to instance alignment was heightened. This resulted in systematic and random registration errors being mistakenly treated as significant shape variations, coupled with actual structural differences and described by the first modes. Furthermore, the study only compared two protocols and exclusively investigated image rigid registration as an instance alignment technique. Future research should explore additional techniques to address this limitation comprehensively.

The second shortcoming is related to the statistical learning models. This project investigates linear models. PCA assumes that shape variation can be described as a linear combination of orthogonal modes. Additionally, PLS assumes that strength variations are linearly described by the shape

variations. Non-linearities or uncertainties in the data were not thoroughly investigated, representing a potential area for future exploration and refinement.

Last and foremost, a major limitation of the current model is the relatively limited sample size. Despite the overall database appearing substantial, particularly in Chapter 4, specific age and group cohorts contain six samples. This limitation impacted the categorization of modes, when considering age subgroups and cohorts. Consequently, the model was not used for predictive purposes but rather for exploring shape variations within the examined population. Despite this limitation, Chapter 3 demonstrated that significant variations with large effect sizes could still be observed, as indicated by leave-one-out tests.

7.4 Future work

7.4.1 Instance alignment

When comparing shape observations of a system, orientation and/or position differences are not typically considered as shape variations. However, since they are recorded in the coordinates of the shape meshes, statistical models describe combined shape variations and orientation/position differences. Chapter 5 showed that the initial positioning of the bone samples and the alignment can be an important source of error in assessing localised variations. This is associated with bias to the reference structure and inaccuracies of the registration algorithm. This subsection details some suggestions regarding computational approaches for consistently aligning the mouse tibia bone structures.

One suggestion considers the definition of some biological landmarks in the 3D images and the alignment of those. For investigating temporal variations by PCA, landmarks that do not change over time should be selected for aligning similar bone regions. That would potentially result to alignment of bone regions with low turnover and local differences in region with high bone turnover. However, obtaining such invariant landmarks remains challenging with current techniques.

Another approach could suggest the definition of a global coordinate system based on the anatomical orientation with the origin being the centre of mass of a reference bone structure. Following this, alignment of all samples could be accomplished by translating their centre of mass to this origin and rotating until their local axes overlap with the global ones.

Another suggestion could consider defining an average mean of mouse bone at the starting point as a reference for reducing the bias to a mouse sample inserted by the rigid registrations. This mean shape could be extracted by applying the Generalized Procrustes Analysis (GPA) (Gower, 1975) on the mouse samples of the youngest age. This approach works with surface meshes instead of images. The algorithm 1) arbitrarily chooses a reference shape, 2) superimposes all samples to this reference (translation and rotation, no scaling), 3) computes the mean shape of the superimposed shapes, 4) if the distance between the mean shape and the reference is higher than a threshold, the new reference is the mean shape, and the algorithm repeats from step 2. This mean shape could serve as a reference to superimpose the other bone meshes of advanced age using Ordinary Procrustes Analysis (Goodall, 1991).

Alignment using an optimization could be another elaborate computational approach. In this case, each sample mesh could be described as the sum of the transformation matrix derived from its superposition to a reference and the inherent differences in shape. In this model the shape differences are described as residual errors for the superposition of each point of a sample to the reference point. Briefly, this model utilizes Bayes hierarchical modelling to maximize the likelihood estimation of Gaussian Procrustes problem considering the discrepancy between the reference and the sample shapes (Bronstein et al., 2008). More details about the algorithm could be found in Theobald & Wuttke, (2006).

The aforementioned Procrustes techniques have been used in statistical shape modelling applications of bone joints, as seen in studies by Audenaert et al., (2019) and Audenaert et al., (2020). They represent sophisticated methods aimed at improving the robustness of PCA models by enhancing the instance alignment protocols.

7.4.2 Further model development

An increased interest in treatment investigation focuses on including variability in mechanics and reporting effect sizes on mechanical properties. In this direction, some studies have developed and applied stochastic Finite Element Modelling to include variability in structural or compositional properties of bone and measure variability in mechanics (Querol et al., 2006; Pascoletti et al., 2021). The current PCA or PLS-based model of treatment-related bone shapes could be combined with mechanical modelling such as Finite Element Modelling. This model could investigate the variability in mechanics derived from PLS or PCA-based variant features and investigate how different treatment strategies affect bone mechanics as the assessment of effect-sizes of treatment interventions in bone strength.

The ultimate goal of such models is to be used *in silico* and reduce or replace the use of animals for conducting experiments. A key aspect of developing such models is their reliability. As PCA/PLS-based models are data-driven, their reliability depends on the sample size. As shown in Chapter 3, the predictability of the current model is low due to the small sample size. However, future work could potentially focus on increasing the population with artificial image data derived from predictive models. Such models could be based on validated mechanical and biological regulation algorithms that consider bone turnover to rely on a combination of mechanics and biology, so the augmented data are physically meaningful.

This study proposes PCA as one statistical learning model that could detect statistical patterns of shape variations in the examined population by linearly projecting the data to a new hyperspace and preserving the covariance of them. One challenge with traditional PCA is that it is an unsupervised statistical learning technique as it does not use any labels to characterise the data. It is widely used when researching data classification but without prior knowledge regarding existing clusters in the examined population. However, in the current study the mouse population included different cohort groups of treated and untreated mice, and these could be considered as labels.

In the direction of feeding the statistical learning technique with more information, a PLS-based model was tested to detect variation features that exert the maximum variance in bone strength. Since strength

increases due to treatments, the response variable was used to feed more information to the statistical computation in detecting variations induced due to treatment interventions in terms of increase in bone strength. Stiffness could also be another variable of interest.

Another supervised statistical model could be explored further. For example, the PLS Discriminant Analysis (PLS-DA) (Barker & Rayens, 2003) is a supervised data classification model that could find an interesting application in investigating shape variation of bone geometry due to different treatment interventions. In this case, the response variable is not a physical parameter, as in the PLS model developed in this project, but a label variable which drives the supervised classification to separate clusters of data. Additionally, more sophisticated statistical models that consider non-linearity and uncertainty could be also used. These could include non-linear PCA and stochastic PCA (Lawrence & Hyvärinen, 2005).

7.4.3 Other applications

The proposed statistical framework could be potentially applied to any image data of long bone structures in mice, focusing on the following aspects:

1) Investigating variations in the cortical bone of mouse femur. Given the differences in the volume of interest, image topology, and resolution, it is essential to conduct evaluation studies to measure errors and uncertainties at different stages of the framework. Guidelines for such studies are outlined in Chapter 3 and Chapter 5.
2) The framework could be potentially used to assess both natural variability and temporal changes in longitudinal studies, exploring growth changes or the progression of other bone diseases over time. Post-processing the scores can help reveal patterns of shape variations as time progresses.
3) The framework supports the extraction of variations within different subgroups of the examined population. It provides a proof of concept by utilising both unsupervised and supervised statistical learning tools to investigate bone shape variations due to different interventions. Similar methodologies can be applied to inter-group comparisons of bone shapes, such as assessing the effects of various novel treatment strategies to determine which one leads to the greatest increase

in bone strength.

4) When dealing with large sample sizes, the framework facilitates the development of *in silico* models. To evaluate the predictability of these models, leave-out tests, as described in Chapter 3, should be conducted. These tests can also help determine the minimum sample size required for building a predictive PCA model.

5) The proposed framework could potentially be used to improve and refine experimental protocols in murine longitudinal studies. One could develop an *in silico* wild type mouse tibia model by gathering large number of existing experimental data. This could facilitate collaboration among research groups and create an open-source virtual wild type mouse bone model, reducing the need for internally creating such models experimentally. For such applications, mice should meet some specifications: same strain, skeletally mature (14 weeks of age or older), similar age, same gender, similar weight.

6) Another idea to improve the existing experimental protocols could be to use measured mode-specific effect sizes in the power analysis for identifying the sample size for future experiments. Additionally, the intervention period and dose of treatment could potentially be better identified by facilitating PCA-based predictive models for remodelling. This model, however, is data-based and therefore the sample size determines the model predictability. If there were extensive data to proceed with the creation of such longitudinal models, one could investigate how specific modes change over time between multiple rounds of the same treatment and generate a relationship between time and score changes which could potentially show the efficacy of each round of treatment.

Other bone sites of rodent skeleton that could be processed by this framework could be the femur. However, as there is still a challenge to model trabecular regions with PCA, one could focus their analysis on the cortical regions of the bones. In such studies, clinical relevance is a key aspect to be considered. Hence, one should consider the rationale in investigating the shape variations of the femur diaphysis. For instance, PCA-modelling of the cortex of murine femur diaphysis would possibly reveal some importance in understanding the global changes in shapes and subsequently global mechanics.

In addition to different bone sites, bones of different small animals than mice could be investigated. This model could be potentially applied on images of rat long bones for describing shape variations. Considered as small-sized mammals, the rat and mouse skeleton are less complex compared with the ones from larger mammals, exhibiting simpler structural characteristics. Subsequently, the densitometric variation and heterogeneity is not a prominent factor in rats as similar to mice.

On the other hand, the statistical model of large mammalian bones should consider variations in both geometry and bone mineral density, due to high skeletal complexity. However, bone ageing research in large mammals comes with great challenges including size, long lifespans, difficulty to use current longitudinal imaging techniques (due to size or radiation), ethical considerations, limited availability. Therefore, longitudinal image data for large mammals are not easily available to generate such Statistical Shape models for treatment investigation and optimizations. However, this framework could find a good application in the preclinical stage of the drug development for improving bone disease management and treatment.

7.5 References

Audenaert, E. A., Van den Eynde, J., de Almeida, D. F., Steenackers, G., Vandermeulen, D., & Claes, P. (2020). Separating positional noise from neutral alignment in multicomponent statistical shape models. *Bone Reports*, *12*, 100243. https://doi.org/10.1016/j.bonr.2020.100243

Audenaert, E. A., Pattyn, C., Steenackers, G., De Roeck, J., Vandermeulen, D., & Claes, P. (2019). Statistical Shape Modeling of Skeletal Anatomy for Sex Discrimination: Their Training Size, Sexual Dimorphism, and Asymmetry. *Frontiers in Bioengineering and Biotechnology*, *7*, 302. https://doi.org/10.3389/fbioe.2019.00302

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics: A Journal of the Chemometrics Society*, *17*(3), 166–173. https://doi.org/10.1002/cem.785

Bronstein, A. M., Bronstein, M. M., & Kimmel, R. (2008). *Numerical geometry of non-rigid shapes*. Springer Science \& Business Media. DOI: 10.1007/978-0-387-7301-2

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal*

Statistical Society: Series B (Methodological), *53*(2), 285-321. https://doi.org/10.1111/j.2517-6161.1991.tb01825.x

Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*, *40*, 33–51. https://doi.org/10.1007/BF02291478

Lawrence, N., & Hyvärinen, A. (2005). Probabilistic non-linear principal component analysis with Gaussian process latent variable models. *Journal of Machine Learning Research*, *6*(11).

Pascoletti, G., Aldieri, A., Terzini, M., Bhattacharya, P., Cal\`\i, M., & Zanetti, E. M. (2021). Stochastic PCA-based bone models from inverse transform sampling: Proof of concept for mandibles and proximal femurs. *Applied Sciences*, *11*(11), 5204. https://doi.org/10.3390/app11115204

Querol, L. B., Büchler, P., Rueckert, D., Nolte, L. P., & Ballester, M. Á. G. (2006). Statistical finite element model for bone shape and biomechanical properties. *Medical Image Computing and Computer-Assisted Intervention--MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part I 9*, 405–411. https://doi.org/10.1007/11866565_50

Theobald, D. L., & Wuttke, D. S. (2006). Empirical Bayes hierarchical models for regularizing maximum likelihood estimation in the matrix Gaussian Procrustes problem. *Proceedings of the National Academy of Sciences*, *103*(49), 18521–18527. https://doi.org/10.1073/pnas.050844510