Flinders University

Computationally efficient assessment of bone strain for a proximal humeral fracture with a fracture fixation plate

By

Daniela Mini

Thesis
Submitted to Flinders University
for the degree of

Doctor of Philosophy

College of Science and Engineering

November 2024

Table of Contents

Abstract

The treatment of proximal humeral fractures with fracture fixation plates has been associated with a failure rate of up to 35% and the cause of this is not yet fully understood. However, several factors contribute to the failures, such as fracture configuration, bone quality, screw orientation, and screw length. Finite Element (FE) analysis is a commonly used computational approach to investigate the biomechanics of fracture fixation devices. Still, FE techniques are too computationally expensive for complex problems with a high number of variables. The overall aim of this thesis is to develop and assess computationally efficient methodologies in order to investigate the bone deformation of a fractured humerus with a fracture fixation plate, varying different implant parameters and within subjects. The first study of this thesis aimed to develop a computational model to predict the minimum principal strain within the fractured humerus by varying the length of the proximal screws. To achieve this, an FE analysis and an Adaptive Neural Network (ANN) method were combined. A semi-automated FE workflow on a single subject was developed to generate up to a thousand different configurations by varying the length of the proximal screws in the humeral head. The data generated were used to train ANN models, which showed a high level of accuracy (R^2 = 0.96-0.99, RMSE = 0.51-1.27 % strain). After the training process, the best ANN model was applied to predict bone strain for a full factorial scenario. This confirmed that the length of the cortical screw has a significant impact on bone strain. Additionally, using the ANN to make predictions for the full factorial scenario only took a few seconds, whereas performing FE analysis for the same number of configurations would have required 170.6 days.

The second study aimed to develop an efficient computational model to reproduce the minimum principal strain within the fractured humerus with the variation of the orientation of the proximal screws. Similar to the first study, a semi-automated pipeline was developed to generate FE models based on a single subject, with varying orientations of the proximal screws, resulting in up to a thousand simulations. The data from the models created were used to train

ANN models. In particular, two types of ANN models were developed, one to predict the screw collision and the other to predict the strain around the screws when the orientation of the screws varies. The two models combined had a good level of accuracy, showing a percentage of error of 15.60% for the first model, and an R^2 = 0.98-0.99 and an RMSE = 0.65-3.77 % strain for the second model. The two ANN models were then used to make predictions of a full factorial scenario, showing again the high impact of the cortical screw. Moreover, using the ANN for predictions in a full factorial scenario took only a few seconds, compared to the FE analysis which would have taken approximately 5,911.8 days.

Despite the high efficiency of ANN models, they have the limitation of predicting a single value of strain, instead of the distribution of bone strain around the surface of the screws. For this reason, a more advanced DL technique has been introduced in the last two studies.

In the third study, the FE data generated from the first and second analyses were used to train two types of Graph Neural Network (GNN) models. The models were trained using nodal information from the bone surface of proximal screws from the FE data and could predict the distribution of bone strain around the screws, with an R^2 = 0.87-0.95 and RMSE = 2.81-3.86 % strain. The GNN models showed a high level of accuracy, and the main advantage was the considerably smaller training and testing time, consisting of respectively a few hours and a few seconds.

The last study aimed to generate an efficient computational model on a cohort of subjects. A semi-automated FE workflow was developed to generate data from 434 subjects. These data were used to train a GNN model, which could predict the distribution of minimum, middle and maximum principal strain around the proximal screws. The results of the GNN model were satisfactory, with an R^2 = 0.76 and an RMSE = 4.93 % strain for the prediction of the minimal principal strain of the bone, showing again the time efficiency of GNN models.

This thesis has successfully developed methodologies with different levels of complexity that combine FE analysis with DL techniques to enhance and expedite the computational process.

Specifically, it has been shown to be time-efficient without compromising the accuracy of the FE analysis. Integrating DL algorithms into FE setups for evaluating medical device performance has the potential to improve surgical planning for individual patients, ultimately leading to better outcomes in medical procedures.

Declaration

I certify that this thesis:

1. does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university

2. and the research within will not be submitted for any other future degree or diploma without the permission of Flinders University; and

3. to the best of my knowledge and belief, does not contain any material previously published or written by another person except where due reference is made in the text.

Daniela Mini

August 22, 2024

Acknowledgements

Firstly, I would like to thank my supervisors, Mark and Karen. Thank you both for the incredible support I received since the challenging start of this PhD journey in Italy and thank you for the time and dedication you both spent guiding me to the end. Mark, thank you for believing from the start I would complete this journey, for your encouragement and pushing me to challenge myself to do better than I expected. Thank you Karen for all your precious feedback, support and guidance during this time. My achievements would not have been possible without your input and supervision, so I am incredibly grateful to have had both of you as my supervisors.

To Dave and Rami, thank you for all the precious feedback on every stressful milestone of this PhD. To all the members of the MDRI group, thank you for being a great community and support throughout this journey, I am grateful for the encouraging environment that helped me to grow and do my best. To the ARC CMIT group, being part of it was such an amazing opportunity to meet incredibly smart and fun people and feel inspired to improve myself.

To the entire PhD group at Tonsley – the Round Table – I am forever grateful for each one of you being a part of this journey. Thank you for your friendship, support, laughter, cakes, and knock offs. To friends and family here and in Italy, especially to my Paccari, thank you for your amazing friendship during these years. To Mum and Dad, thank you for the unconditional love and support, even though none of you hasn't really understood yet what I'm doing here in Australia.

Last but not least, to Dimi. Thank you for your patience, presence, and support and for putting up with my occasional meltdowns. Your strength and support played an important role in this PhD, so thank you for always encouraging me to do and be my best.

List of Publications

PEER-REVIEWED JOURNALS ARTICLES

Mini D, Reynolds KJ, Taylor M, 2024, *Assessing screw length impact on bone strain in proximal humerus fracture fixation via surrogate modelling*, International Journal for Numerical Methods in Biomedical Engineering, 40(8):e3840. doi:[10.1002/cnm.3840](10.1002/cnm.3840)

Mini D, Reynolds KJ, Taylor M, 2024, *Assessing the influence of screw orientation on fracture fixation of the proximal humerus using finite element informed surrogate modelling*, Submitted to the International Journal for Numerical Methods in Biomedical Engineering.

CONFERENCE ABSTRACTS

Mini D, Zhang Z, Shi JQ, Reynolds KJ, Taylor M, *Bone strain field prediction via graph neural network for a proximal humeral plate*, Australian New Zealand Orthopaedic Research Society, 28th Annual Scientific Meeting. Auckland, New Zealand. Dec 2023. Podium Presentation, PhD Awards Session

Mini D, Reynolds KJ, Taylor M, *Screw length impact on bone strain for a proximal humeral plate via a neural network model*, 28th Congress of the European Society of Biomechanics. Maastricht, The Netherlands. Jul 2023. Podium Presentation

HONOURS AND AWARDS

Travel Grant, Australian and New Zealand Orthopaedic Research Society 28th Annual Conference, Dec 2023, $AU 300

Travel Award, European Society of Biomechanics Research, Jul 2023, €400 (≈$AU 650)

CSE HDR International Conference Support, Flinders University, May 2023, $AU 1500

Higher Degree Scholarship, Australian Research Council Training Centre for Medical Implant Technologies (ARC CMIT), Mar 2021, $AU 32,000 per year for 3.5 years

International Tuition Fee Sponsorship, Flinders University, Mar 2021, $AU 127,200

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Proximal humeral fractures are one of the most common fractures among elderly patients, constituting 6-9% of all diagnosed fractures (Roux *et al.* 2012; Sporer *et al.* 2006). With the growing elderly population, this incidence is expected to rise in the coming years, leading to increased strain on the healthcare system and associated costs (Maravic *et al.* 2014; Palvanen *et al.* 2006). Approximately 15% of these fractures require surgical intervention (Patel *et al.* 2022), with the use of fracture fixation plates being a widely used treatment. However, there is evidence in the literature suggesting a significant failure rate of up to 36% for these implants, and the specific reasons for these failures are not yet fully understood due to the complex nature of the biomechanical problem (Kralinger *et al.* 2014). A variety of parameters, both related to the implant and the patient, are thought to be involved in the problem, yet there is still limited understanding of how these factors individually and collectively contribute to failure. Some of these parameters include screw length, screw orientation, screw configuration, plate position, and bone quality (Lewis *et al.* 2021). Finite Element (FE) analysis is considered the computational gold standard for evaluating implant failure and facilitating the development of new implant designs. One of the primary advantages of FE analysis is its time efficiency

compared to the experimental methods and its capability to accurately compute bone stress and deformation anywhere within the bone structure. These measurements are challenging to obtain through experimental settings, making FE analysis a valuable tool for predicting potential implant failures and offering insights for enhanced implant design. Several studies in the existing literature have conducted FE analysis to address the effects of various implant parameters associated with failure for a proximal humeral fracture within subjects (Fletcher *et al.* 2019a, 2019c, 2019b; Jabran *et al.* 2019b; Mischler *et al.* 2020b; Tilton *et al.* 2020b; Varga *et al.* 2018, 2020). These studies have focused the analysis on a limited number of configurations, performing from less than one hundred to up to a few thousand simulations. Indeed, due to the large number of parameters involved, FE analysis can be a time-consuming process, especially when attempting to assess the impact of numerous variables in a problem with a significant number of simulations and a large solution space. This limitation highlights the need for the development of a more efficient computational technique that can support FE analysis and accurately predict the mechanical environment associated with fracture fixation devices across a wide range of configurations.

At present, some surrogate models have been formulated using FE data to investigate various biomechanical issues and enhance their computational efficiency. For instance, models trained using FE data have been developed to address biomechanical problems with a focus on femoral fracture, hip and knee implants, using techniques such as Kriging, Adaptive Neural Network (ANN), and Gaussian process (GP) methods (Al-Dirini *et al.* 2020; O'Rourke *et al.* 2016; Taylor *et al.* 2017). The development and training of ANN models can be a relatively straightforward and efficient process, making them a valuable computational tool. Despite their success in various applications, ANN models have not been used for addressing the biomechanical behaviour of proximal humeral fractures with fracture fixation plates. Given their potential, it is critical to explore the application of ANN models in this area, and they could be implemented in order to make a prediction of strain with the variation of implant parameters. Additionally,

the literature contains more complex Deep Learning FE-informed models, and they should be explored in an orthopaedic setting, particularly in the context of fracture fixation, as they could make predictions of the distribution of strain of an area of interest. An example is a study conducted by Krokos et al., which demonstrated the effectiveness of a Graph Neural Network (GNN) technique in predicting variations in internal stress within an object with the changes in its porous structure (Krokos *et al.* 2022a). This work highlights the potential of using GNN models combined with FE data to predict the mechanical response of materials, even in complex scenarios. The Krokos study gives confidence to explore GNN models to predict internal bone strain variations in response to changes in implant parameters and within subjects.

In summary, to gain a deeper insight into the biomechanical characteristics of a proximal humeral fracture treated with a fixation plate and eventually reduce its failure rate, more efficient computational methodologies such as ANN and GNN should be implemented to assist the FE analysis, which has limited computational capabilities when exploring a problem with a large solution space.

1.2 Research Aims

The overall aim of this thesis is to develop and assess computationally efficient methodologies in order to investigate the bone deformation of a fractured humerus with a fracture fixation plate, varying different implant parameters and within subjects. Due to the problem's complexity and the extensive solution space, only single variations are made to explore the applicability of the computational methods implemented. More specifically:

- Develop a semi-automated pipeline to efficiently generate an FE analysis varying different implant-related parameters on a single subject.
- Develop a semi-automated pipeline to efficiently generate FE analyses for a large cohort of 100's of subjects.

- Develop a computationally enhanced methodology using FE data and an ANN methodology to efficiently investigate the impact of variations in several implant parameters on humeral strain.

- Develop a computationally enhanced methodology using FE data and a GNN methodology to efficiently investigate the impact of variations in several implant parameters and subject variabilities on humeral strain, and give a strain distribution prediction inside the humerus.

- Apply the developed FE informed ANN and GNN surrogate models to explore the influence of screw length and orientation on the humeral strain.

- Apply the developed FE informed GNN models to explore the influence of subject variability on the humeral strain.

1.3 Thesis Outline

To achieve the research aims set out, the thesis is organised into the following chapters:

Chapter 2 provides a review of the literature. It includes descriptions of the humerus and shoulder and their biomechanics, as well as a review of proximal humeral fractures and fracture fixation plates. A focus on the different investigating approaches has been made, such as in vitro testing, FE models and surrogate techniques. Moreover, it describes various DL techniques available.

Chapter 3 investigates the influence of screw length on the bone strain around the surface of the screws. It involves the use of FE data combined with an ANN method. The FE workflow described the development of the ANN model and its application to make predictions of a full factorial scenario.

Chapter 4 investigates the influence of screw orientation on the bone strain around the surface of the screws. It involves the use of FE data combined with an ANN method. The FE workflow

described the development of the ANN model and its application to make predictions of a full factorial scenario.

Chapter 5 uses the dataset developed in Chapters 3 and 4 to predict the distribution of bone strain around the screws using a GNN technique.

Chapter 6 investigates the influence of variation in subject anatomy on the predicted distribution of bone strain around the screws for the proximal humeral fracture. A semi-automated workflow was developed to generate FE data of 434 subjects, and their data were used to train a GNN model.

Chapter 7 summarises the main findings of this project and its future projection on fracture fixation failure analysis. The limitations of this work and potential future research directions are also discussed.

Chapter 2

Literature Review

2.1 Anatomy of the Shoulder

The shoulder is a complex structure that connects the upper extremity to the axial skeleton. It is composed of four joints and three main bones, and this complex interaction allows movements in six degrees of freedom. Considering the complexity of the shoulder joint structure, there are a variety of injuries or pathologies that can cause shoulder instability, pain, or loss of motion.

2.1.1 Humerus

The longest and largest bone in the upper extremity is the humerus, which is the extension of the shoulder joint that allows arm movements in space (Figure 2.1). Its structure is divided into three different sections such as the proximal end, the shaft and the distal end. The proximal end of the humerus is the portion of the bone that articulates with the scapula adjacent to the glenoid cavity, and together they constitute the shoulder joint (Huri *et al.* 2020). The head has a hemispherical shape, with a diameter between 32 and 58 mm (Boileau and Walch 1997). At the proximal end of the humerus are the greater and lesser tubercles, the first in an upper lateral position and the second in an anterior position, and they act as attachments for the surrounding muscles. The supraspinatus, infraspinatus and teres minor muscles are connected

to the greater tubercle and the subscapularis muscle to the lesser tubercle (Huri and Paschos 2017). Moreover, the anatomical and surgical neck can be identified. The first connects the humeral head to its shaft, the second is the region just below both tuberosities and oriented horizontally in space. This region is the weakest and therefore most commonly at risk of fracture (Drake *et al.* 2010).
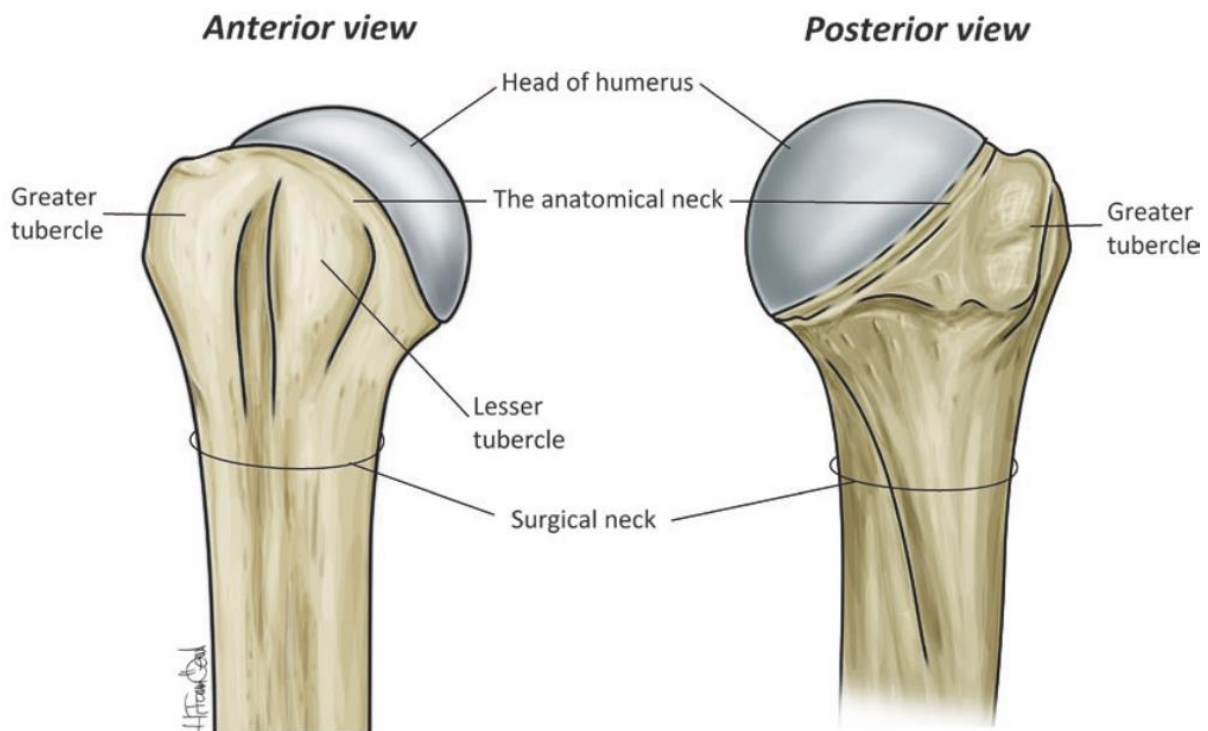
Figure 2.1 - Anterior and posterior view of the proximal humerus (Yılmaz et al. *2020)*

2.1.2   Glenohumeral Joint

The shoulder structure includes four distinct joints, the acromioclavicular, sternoclavicular, scapulothoracic and glenohumeral joints.

The proximal end of the humerus and the glenoid cavity of the scapula are bound by the glenohumeral joint, a synovial ball-and-socket junction. This joint is highly mobile and the diameter of the head of the humerus is larger than the glenohumeral cavity, which makes the joint unstable. To increase its stability, several tissues surrounding the glenohumeral joint

21

serve as dynamic constraints, such as muscles, or static constraints, such as the glenoid labrum, the glenohumeral joint capsule and the glenohumeral ligaments (Huri *et al.* 2020).

2.2 Shoulder Biomechanics

The movements of the shoulder are permitted by the interaction of the sternoclavicular joint, the acromioclavicular joint, the scapulothoracic joint and the glenohumeral joint, where the majority of the movements are performed.

2.2.1   Shoulder Motion

The shoulder is a very flexible structure and is able to make large movements and consequently making it less stable. The movements of the shoulder are enabled by the correct interaction of the four joints together (Lippert 2006)(Figure 2.2).

Extension and flexion occur in the sagittal plane around the frontal axis, allowing rotation up to 180° and hyperextension up to 45° from the anatomical position, i.e. when the arm is aligned to the vertical axis. With regard to flexion, 120° is performed in the glenohumeral joint and the remaining 60° is allowed by the movement of the scapula (Inman and Abbott 1991). The flexion is performed mainly by the pectoralis major and deltoid muscle but also by the coracobrachialis, and biceps brachii, whereas the extension is performed mainly by the latissimus dorsi and teres major, and partially by the pectoralis major, deltoid, triceps brachii (Lippert 2006).

Adduction and abduction take place in the frontal plane around the sagittal axis. The abduction can be done 120° passively when the scapula is constrained by the motion, but it can reach up to 180° only if the shoulder joint is also rotated laterally (Lucas and Francisco 1973). It is mainly enabled by the deltoid and supraspinatus, which also have the role of pulling the humeral head into the glenoid fossa during movement (Lippert 2006). Adduction has a maximum range of 30° and is allowed by pectoralis major, latissimus dorsi, and teres major.

Medial and lateral rotation, also known as internal and external rotation, occur in the transverse plane around the vertical axis. From a neutral position, it is possible to move 90° in any direction, and the muscles involved are subscapularis, teres major, latissimus dorsi, pectoralis major, deltoid, teres minor, infraspinatus (Lippert 2006).
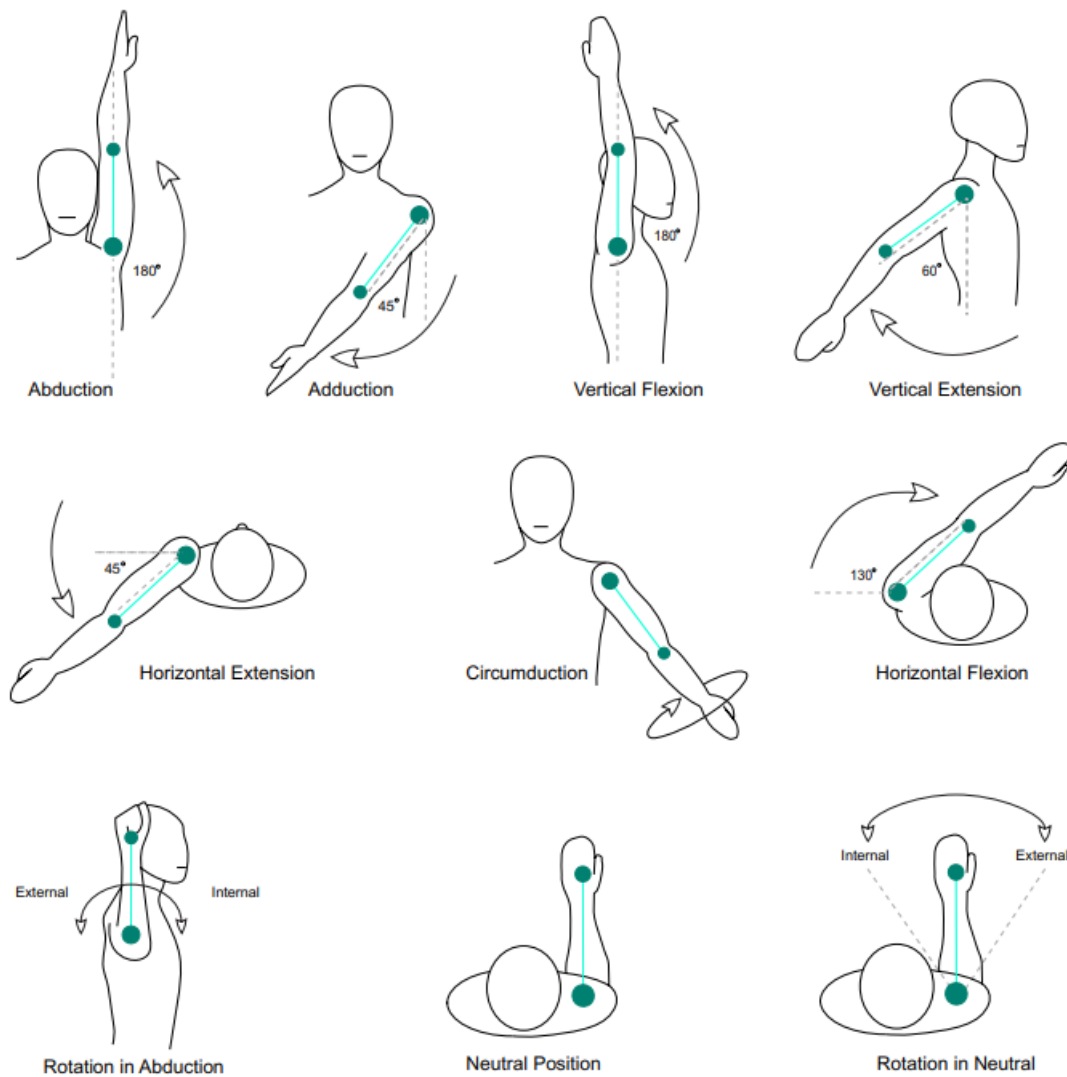
Figure 2.2 - Shoulder Motion (Cardoso and Gasparik 2023)

2.2.2 Humerus Coordinate System

In order to analyse the movement of the glenohumeral joint and the shoulder, a reference system of the humerus is used as a local coordinate system (Figure 2.3). This is defined by three anatomical landmarks: the glenohumeral centre of rotation (GH), the most caudal point on the

lateral epicondyle (EL), and the most caudal point on the medial epicondyle (EM). Once these three points are identified, the origin of the reference system is coincident with GH, and the X axis is the line perpendicular to the plane formed by EL, EM, and GH, pointing forward. The Y axis is the line connecting GH and the midpoint of EL and EM, pointing to GH, and the Z axis is the common line perpendicular to the Y and X axes, pointing to the right (Wu *et al.* 2005).

The displacements are determined as follows once the reference system has been defined:

- Translation along the x-axis = anterior/posterior translation

- Translation along the y-axis = inferior/superior translation

- Translation along the z-axis = joint distraction

- Rotation around the x-axis = abduction/adduction

- Rotation around the y-axis = internal/external rotation

- Rotation around the z-axis = flexion/extension

Motion standardization is only defined for right shoulder joints. When measuring left shoulders, it is advised to mirror the original position data relative to the sagittal plane. (Wu *et al.* 2005).

2.2.3 Forces in the Glenohumeral Joint

In order to reproduce the shoulder joint movements in an experimental or computational model, it is necessary to have a precise knowledge of realistic loads acting in vivo. Although this is not an easy task as the shoulder is considered to be very complex and is less studied than the hip or knee joints, several studies have been found in the literature which attempt to either measure or calculate the contact forces at the interface between the humerus and the shoulder and the forces generated by the surrounding muscles.
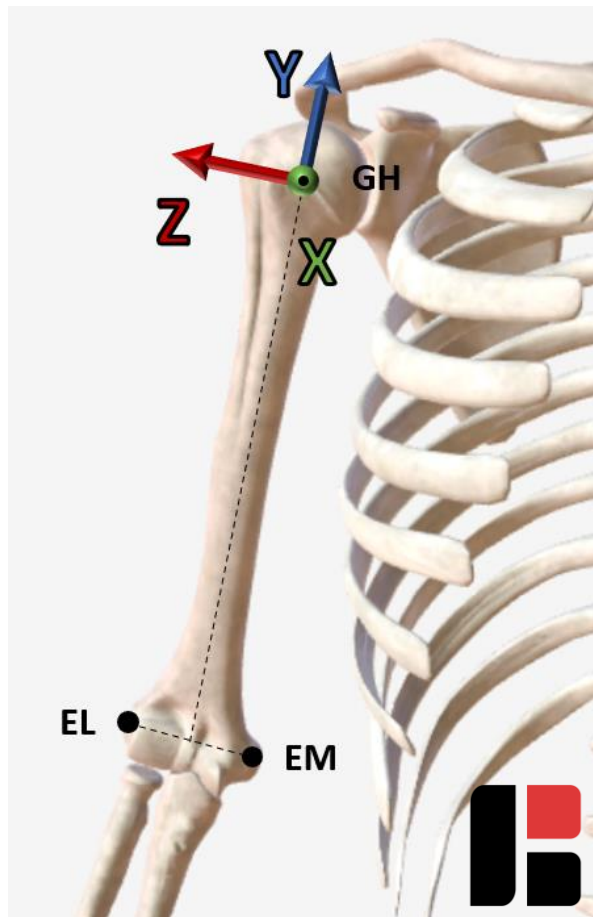
Figure 2.3 - Humerus coordinate system. GH: glenohumeral centre of rotation. EL: lateral epicondyle. EM: medial epicondyle. (Anatomical model powered by BioDigital, image crafted by the author)

2.2.3.1 Contact Forces

The measurement of muscle and joint reaction forces in a healthy intact shoulder joint remains challenging due to the lack of a direct method for such measurements. Additionally, direct measurement of joint reaction forces has only been conducted on a limited group of subjects with a shoulder replacement, limiting the understanding of these forces in healthy individuals. Several studies in the literature have tried to calculate muscle forces and contact loads in the glenohumeral joint using musculoskeletal models, which have produced a wide range of different results (Poppen and Walker 1978; Prinold *et al.* 2013). In contrast to these, Bergmann et al developed an instrumented shoulder implant with telemetric data transmission that made possible the measurement of six components of joint contact forces and moments for the first time ever (Bergmann *et al.* 2007, 2011).

In an initial study, they analysed simple movements of a single patient, such as abduction and flexion, but also more complex movements, such as lifting objects or using crutches (Bergmann *et al*. 2007). Regarding the force values observed, they recorded a value of 40%BW (Body Weight) for the first months of physiotherapy, and up to 120% BW in the following months. However, in terms of the direction of the forces, for almost all the activities performed, the angles of the contact forces with respect to the humerus varied only minimally in the sagittal plane, from 17° to 26°, and slightly more in the frontal plane, from 14° to 31° (Figure 2.4). This finding is in agreement with the further study conducted on seven different patients (Bergmann *et al*. 2011).

Analysing the force values of the second study, for flexion without the use of weights and varying the speed of the movement performed, the mean value of the force ranged from 58%BW to 73%BW, similar for abduction where the mean force recorded was between 65%BW and 81%BW. For both activities, the addition of a 2-kg weight increased the contact force magnitude by 51-75% at both slow and fast speeds, while raising the arm above 90° increased strength by 21-40% (Bergmann *et al*. 2011). Despite the validity and uniqueness of these studies, it must be recognised that these are subjects who have undergone shoulder implant surgery, and therefore unhealthy subjects and this may affect the loads.
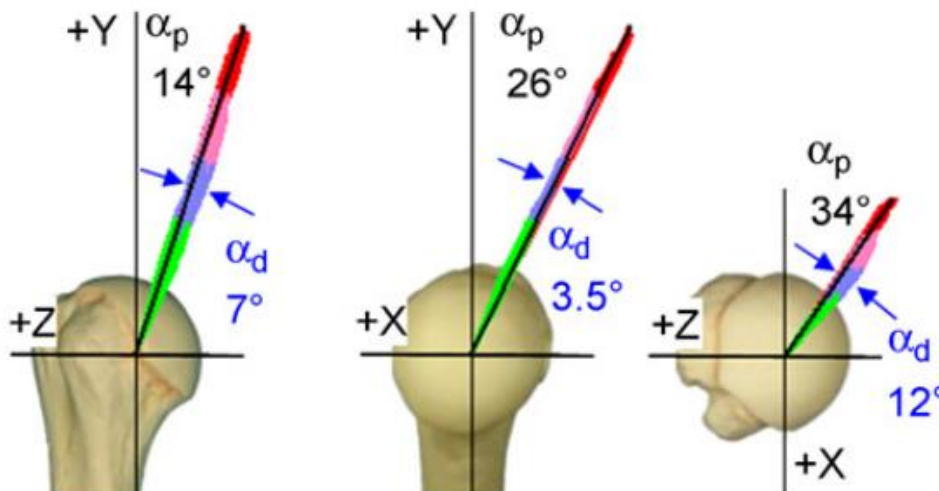
Figure 2.4 - Reaction force at glenohumeral joint during flexion (Bergmann et al. 2007)

2.2.3.2 Muscular Forces

In order to simulate shoulder movements accurately, it would be desirable to know the muscle forces. One technique used to measure them is through EMG signals in vivo, but this method can only reliably detect the activity of superficial muscles, with poor signal strength (Inman and Abbott 1991). Otherwise, a computational approach can be used to develop a musculoskeletal model, and over the last few years, several models have been developed that vary in complexity and the number of muscles considered. The models believed to be most relevant and complete were analysed by Prinold et al. in their review (Prinold *et al.* 2013). The models included are: The Delft shoulder and elbow model (DSEM)(Nikooyan *et al.* 2011; Van Der Helm 1994), the UK National Shoulder Model (UKNSM)(Charlton and Johnson 2006), The Garner and Pandy model (GPM)(Garner and Pandy 2001), The Swedish shoulder model (SSM) (Högfors *et al.* 1995), The Waterloo model (WSM)(Dickerson *et al.* 2007), Stanford-VA model (Holzbaur *et al.* 2005, 2007).

These models used an inverse dynamics simulation analysis, and Hill-type muscle models to describe muscle-tendon behaviour and force generation, which are driven by a series of muscle-tendon parameters. In order to develop a realistic MS model, it is necessary to accurately define these parameters, which for example can be direct inputs to the model such as bone kinematics, landmarks and length of the glenohumeral centre of rotation, muscle insertions and ligament lengths, and assumptions about muscle force characteristics. The studies considered gave similar glenohumeral joint reaction force values for some activities, but it is necessary to underline that for the upper limb there is still a big gap for the experimental validation of these models, as it is difficult to measure muscle forces and joint reaction forces in vivo. In terms of muscular forces, all the models showed that the physiological cross-sectional area (PCSA) was found to be higher for the deltoid muscle compared to other muscles. Generally, the PCSA has been consistently shown to have a linear correlation with the maximum force that a muscle can generate, highlighting the importance of PCSA as a key determinant of muscle strength

27

(Fukunaga *et al.* 2001). Table 2.1 presents the PCSA values defined for the studies considered in the Prinold review, consistently showing a higher value for the deltoid muscle (Table 2.1). Moreover, as shown in the WSM model, the comparison of muscle force prediction patterns during humeral abduction with the DSEM model shows that the deltoid muscle exhibits the highest relative force involvement during the movement (Dickerson *et al.* 2007). These models have limitations and are indeed sensitive to errors in their input parameters, which significantly affect the output results. Furthermore, there is currently a lack of a consistent scaling methodology that provides realistic kinematic reproductions of individual subjects, which would be relevant to support for example the planning of a surgery or to evaluate the performance in a pathological condition (Prinold *et al.* 2013).

Table 2.1 - Comparison of shoulder muscle PCSAs (cm²) of the studies reviewed by Prinold et al. The ratio of deltoid muscle PCSA to the specific shoulder muscle is displayed in the parenthesis. For the SSM and WSM, the normalised values are available. Table adapted from (Prinold et al. 2013).

GH muscles	DSEM	UKNSM	SSM	WSM	Standford-VA	GPM
Total deltoid	33.07 (1)	12.2 (1)	(1)	(1)	25.0 (1)	81.98 (1)
Infraspinatus	14.32 (2.31)	6.0 (2.03)	(1.91)	(1.91)	11.9 (2.09)	33.3 (2.46)
Teres minor	4.97 (6.65)	2.10 (5.81)	(8)	(8)	3.70 (6.76)	6.77 (12.1)
Supraspinatus	6.21 (5.33)	3.0 (4.07)	(3.95)	(3.95)	4.8 (5.23)	20.8 (3.93)
Subscapularis	14.31 (2.31)	7.80 (1.56)	(1.54)	(1.54)	14.1 (1.74)	35.7 (2.30)

2.3 Proximal Humeral Fractures

Proximal humeral fracture is one of the most common fractures for patients over 65 years old and it accounts for 6-9% of the total diagnosed fractures (Iglesias-Rodríguez *et al.* 2021; Roux *et al.* 2012; Sporer *et al.* 2006). In particular, proximal fractures are the most common, about 50% of all humeral fractures and are the third most common type of fracture in the elderly, after the femoral neck and radius (Holloway *et al.* 2015; Maravic *et al.* 2014; Roux *et al.* 2012).

The incidence of proximal humerus fractures increases with age and is higher for female patients between 45 and 85 years of age after a fall from a standing position or an episode of high-energy trauma (Holloway *et al.* 2015; Kim *et al.* 2012; McLean *et al.* 2019). It is therefore well established from multiple studies that proximal humerus fractures are frequent among the elderly, generally with a low BMD, and most cases require hospital treatment (Fleischhacker *et al.* 2021; Maravic *et al.* 2014).
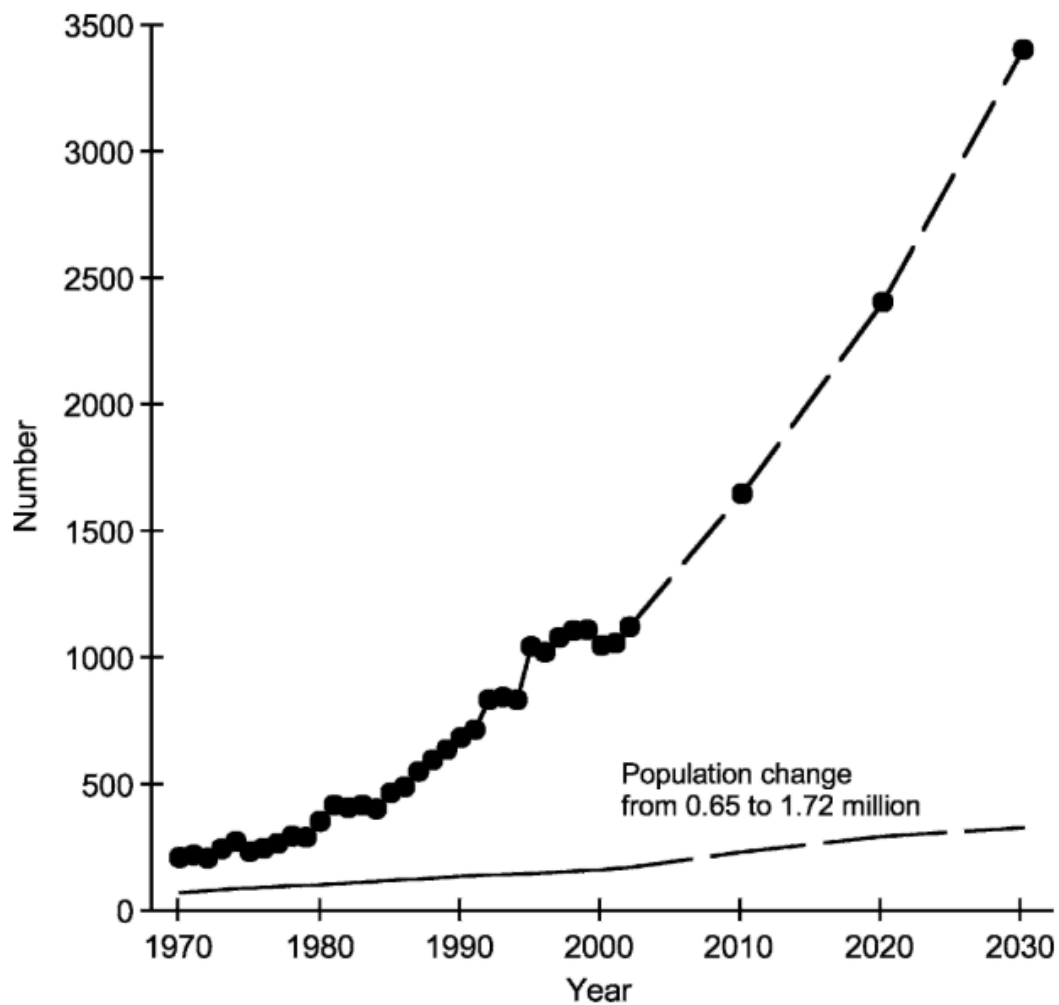
Figure 2.5 - Estimation of proximal humerus fractures in individuals aged 60 and above with osteoporosis in the Finnish population, determined using a regression model (Palvanen et al. 2006)

In addition, the patient's daily activities become more difficult, which can have a great impact on the quality of life after the fracture and the risk of subsequent hospitalisation is high (Maravic *et al.* 2014).

Therefore, since this kind of fracture is typical for elderly osteoporotic patients and considering that the geriatric population is continuously increasing, an increase of this kind of injury is expected in the future and additional demand for hospitalisation, raising the consumption of hospital resources and consequently their cost (Kim *et al.* 2012; Maravic *et al.* 2014; Palvanen *et al.* 2006)(Figure 2.5)

2.3.1 Classification

Several different types of classification of proximal bone fracture are present in the literature and each one focuses on different aspects.

Neer's classification is structured into six groups and divides the fracture patterns by the number of fragments, direction of displacement and articular surface involvement (Carofino and Leopold 2013; Neer 1987)(Figure 2.6). Following this classification system, fractures with displacements less than 1.0 cm or angles less than 45° are included in Group 1. This type of fracture accounts for more than 85% of all fractures of the humerus and is not usually treated surgically. Group 2 refers to fractures with anatomic neck displacement, which normally causes avascular necrosis, and Group 3 identifies fractures with surgical neck displacement and can be angulated, separated, or impacted. The fractures of this group may damage the surrounding arterial structures and the axillary nerve. Displaced fractures involving the greater tuberosity belong to Group 4 and cause longitudinal rotator cuff tears. They may be two-, three-, or four-part fractures such as Group 5, which includes dislocations of the lower tuberosity. Group 6 includes fractures with dislocations; four-part fractures are the most severe and neurovascular symptoms are common (Lasanianos *et al.* 2015).

The AO/ASIF classification system divides proximal humeral fractures into three main groups based on joint involvement and a number of major fracture lines (Foundation AO 2014)(Figure 2.7). The Type A group represents unifocal extraarticular fractures further subdivided based on tuberosity involvement and whether the metaphysis is impacted. Type B fractures are bifocal extra-articular and may be with or without metaphyseal impaction, or with glenohumeral dislocation. The Type C group includes intra-articular fractures, which can be categorized further according to whether the fracture is impacted with slight displacement, marked displacement, or dislocated.
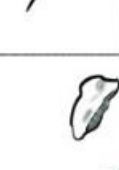
Figure 2.6 - The different types of proximal humeral fractures as per Neer's classification system (Lasanianos et al. 2015)

Overall, it can be seen that the classification of proximal humerus fracture is complicated since fracture can occur in a large number of different patterns.

11 proximal (types according to topography and extent of bone lesion)

| 11-A1 | 11-A2 | 11-A3 | 11-B1 | 11-B2 | 11-B3 | 11-C1 | 11-C2 | 11-C3 |

11-A   extraarticular unifocal fracture
11-A1   tuberosity
11-A2   impacted metaphyseal
11-A3   nonimpacted metaphyseal

11-B   extraarticular bifocal fracture
11-B1   with metaphyseal impaction
11-B2   without metaphyseal impaction
11-B3   with glenohumeral dislocation

11-C   articular fracture
11-C1   with slight displacement
11-C2   impacted with marked displacement
11-C3   dislocated

Figure 2.7 - The types of proximal humeral fractures according to the AO/ASIF classification system. Copyright by AO Foundation, Switzerland (Foundation AO 2014)

2.3.2   Surgical Treatment

In general, most (85%) proximal humeral fractures are minimally displaced or nondisplaced and therefore can be treated non-surgically, usually mobilized with a sling, and passive range-of-motion exercises are performed after rest (Iglesias-Rodríguez *et al.* 2021; Neer 1987). Non-surgical management of displaced fractures is challenging and often results in less than satisfactory outcomes (Neer 1987). Highly displaced fractures and three- and four-part fractures are difficult to treat especially when the patient is osteoporotic and the risk of necrosis caused by humeral head devascularization increases. Fracture instability, i.e. the possibility of the fractured fragments moving during rehabilitation or minimal arm movements, is the main criterion for determining the necessity of surgery (Dahan *et al.* 2019).

The remaining 15% of proximal humeral fractures require operative intervention in order to achieve recovery (Huri and Paschos 2017; Owsley and Gorczyca 2008; Schlegel *et al.* 1994).

There are several techniques for treating the injury, and no one implant is ideal for all fractures. However, the purposes of surgery are the same with all implants: to achieve and maintain satisfactory reduction to allow early shoulder motion, obtain health, and restore functionality (Owsley and Gorczyca 2008).

Surgical options include percutaneous pinning, intramedullary fixation, open reduction and internal plate fixation, hemiarthroplasty, or reverse shoulder arthroplasty (Greiwe 2015)(Figure 2.8).

Percutaneous pinning is a minimally invasive technique and reduces blood vessel compromise. This procedure can be used for two-part neck fractures, three-part fractures and 4-part valgus fractures in patients with good bone quality, minimal metaphyseal comminution and intact medial calcar (Domingue *et al.* 2021). Failures related to this treatment can be loss of reduction, neurovascular injury, osteonecrosis and malunion, especially for four-part fractures (Greiwe 2015).

Intramedullary fixation can be used for 2-part fractures, less often for 3-, and 4-part fractures, but is rarely indicated because of its high failure rate (Domingue *et al.* 2021). Overall, hardware-related complications following intramedullary fixation are common, with 16-23% requiring further surgery and/or removal of the nail. Commonly, complications involve rotator cuff and articular cartilage injury and varus malunion (Greiwe 2015).

With regard to shoulder replacements, there are two different techniques. Hemiarthroplasty is performed for complex fractures of younger individuals, who would likely have complications with the placement of a fixation plate. Reverse total shoulder, on the other hand, is the procedure used for older patients with lower bone density (Domingue *et al.* 2021). In general, these two techniques are used for patients with poor bone quality and poor vascularity who undergo three- and four-part fractures of the proximal humerus (Greiwe 2015).

Lastly, fixation plates are used for the treatment of 2-, 3- and 4-part fractures. This type of device remains one of the principal articular preservation treatments for complex fracture patterns, (Murray *et al.* 2011) and has demonstrated better biomechanical outcomes than other devices (Solberg *et al.* 2009; Sproul *et al.* 2011). In this regard, the number of proximal humerus fractures treated with internal fixation has increased in recent years (Huttunen *et al.* 2012; McLean *et al.* 2019).
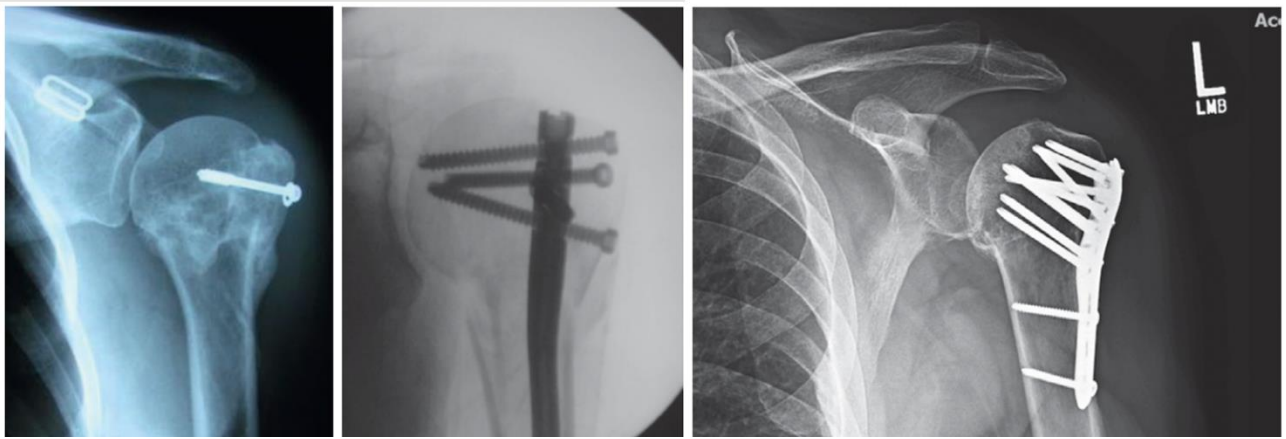
Figure 2.8 - Example of (left) percutaneous pinning, (centre) intramedullary nailing and (right) locking plate. Adapted from (Greiwe 2015)

2.3.3 Fracture Fixation Plates

Fracture plate fixation is performed through open shoulder surgery and the implant is inserted into the lateral portion of the proximal humerus. The plates and the screws act as an anchoring system for the fracture fragments, preventing major movements and providing the necessary stability during rehabilitation.

Locking plates and non-locking plates are two different types of implant designs used for fracture fixation, and locking plates have become increasingly popular due to their high success rate (Kubiak *et al.* 2006; Seide *et al.* 2007; Soileau *et al.* 2007; Walsh *et al.* 2006). Locking plates provide enhanced stability and support after the implant is fixed because of an anchoring mechanism between the threaded screw heads and the plate itself. This mechanism establishes a stable angular connection between the screws and the plate, which improves the overall stability and support of the implant (Cronier *et al.* 2010; Jabran *et al.* 2018).

Non-locking or compression plates were bulkier and showed a high failure rate, especially for osteoporotic patients and were also associated with a high rate of infection (Karataglis *et al.* 2011). However, the plates used now have a higher probability of success, allowing for early fracture mobilisation. Indeed locking plates provide greater angular stability, allowing the load to be distributed uniformly throughout the structure, rather than being concentrated on a single

screw (Karataglis *et al.* 2011; Maldonado *et al.* 2003; Seide *et al.* 2007). Furthermore, the stability of compression plates is dependent on the interaction between the plate and bone, as opposed to locking plates where it is dependent on the interaction between the bone and the screw. This results in a decreased level of friction, leading to a reduced biological impact (Egol et al. 2004). Moreover, locking plates have superior pull-out strength and stiffness, and failure requires simultaneous pull-out or failure of all screws.

There are different types of locking plates, developed by different manufacturers (Figure 2.9).

One type of commercial plate is Spatial Subchondral Support (S3; Zimmer Biomet, Warsaw, IN, USA)(Zimmer Biomet 2014). This design can have non-locking screws in addition to threaded ones, preventing damage to the articular surface (Le *et al.* 2019). The S3 plate is placed 3 cm distal to the greater tuberosity to achieve a 135° neck angle, improving the patient's postoperative response and reducing complications of subacromial impingement (Jabran *et al.* 2018).
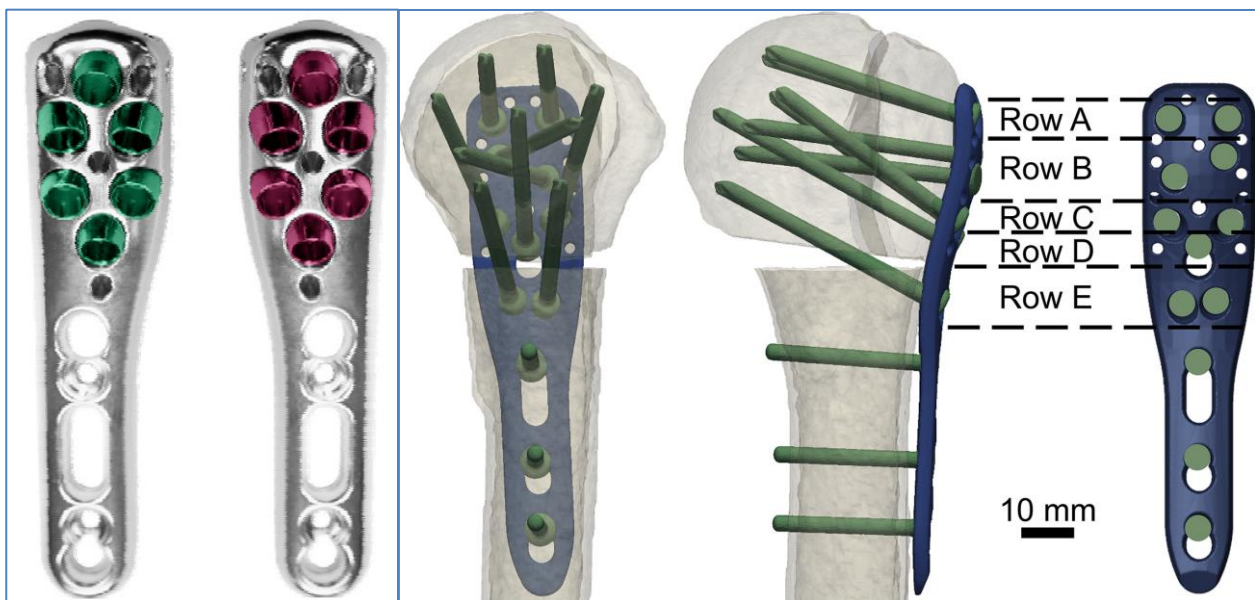
Figure 2.9 - Different designs of locking plates. (left) Spatial Subchondral Support (Gille et al. *2008). (right) PHILOS (Fletcher* et al. *2019c).*

The successor to this plate is the A.L.P.S. Proximal Humerus Plating System (Zimmer Biomet 2015a), which similarly takes full benefit of the principle of Spatial Subchondral Support plate. This plate is designed to reduce the risk of complications associated with treating proximal humerus fractures such as varus collapse, screw penetration into the articular surface and subacromial impingement.

Another popular product is PHILOS (Proximal Humerus Internal Locking System; Synthes, Paoli, PA, USA). It provides 9 screw holes for proximal fragments and at least 3 screw holes for distal shaft fixation, but the recommended procedure is to apply at least 4 screws in the head of the humerus, up to 9 if the bone quality is very poor (Synthes 2018).

It has been reported that locking plates have led to complications, mainly involving screw perforation in the bone. Therefore, a second-generation locking technology has been developed with variable-angle screws (Cronier *et al.* 2010). This type of plate allows the screw direction in the humeral head to be adjusted before locking, in contrast to the traditional locking technologies where screw angles are predefined (Jabran *et al.* 2018).

One example of variable angle plates is the Non-Contact Bridging plate (NCB, Zimmer, Warsaw, IN, USA) (Zimmer Biomet 2015b). This design allows polyaxial positioning of the screw (30°) in the head of the humerus and subsequent locking of the screw head for greater stability. Furthermore, these plates decrease the risk of periosteal blood impairment since they act as an internal fixator without contact between the plate and the bone surface (Figure 2.10). Extra care should be taken with the orientation of the screws, and additional technical guidelines on how to orientate the screw in the space are essential to ensure actual benefits and avoid the risk of interacting screws (Königshausen *et al.* 2012).
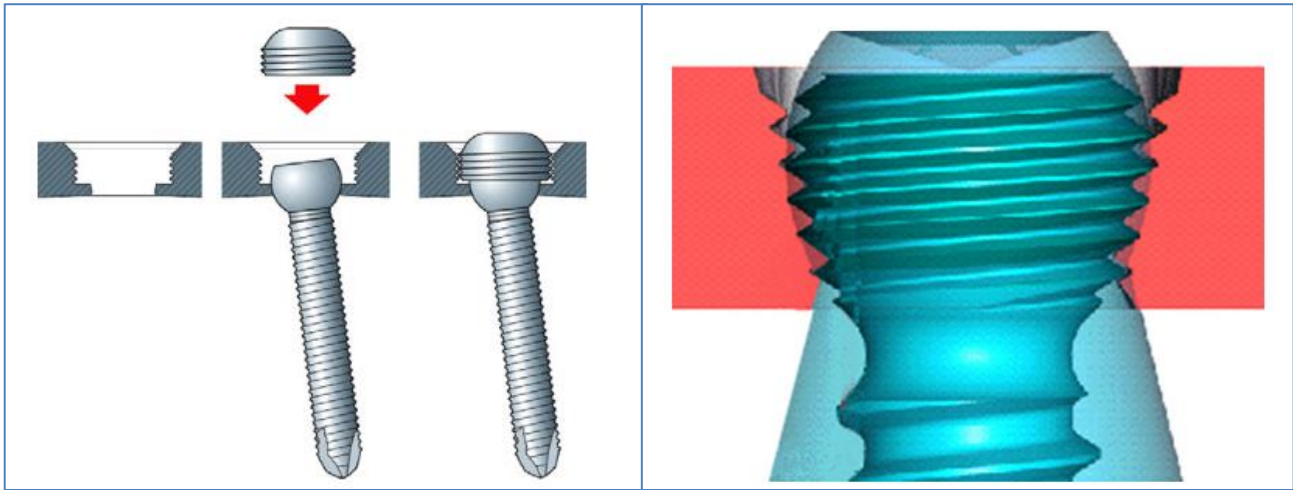
Figure 2.10 - Example of variable angle locking system. (left) Design from Zimmer. (right) design from AO. (Cronier et al. *2010)*

2.3.4 Clinical Complications

In comparison with compression fixation plates for proximal humerus fracture, the use of locking plates significantly increased the probability of success (Kubiak *et al.* 2006; Seide *et al.* 2007; Soileau *et al.* 2007; Walsh *et al.* 2006). However, especially for 3- and 4-part fractures, recovery of the injury remains difficult to achieve and post-operative failure rates of up to 36% have been reported (Kralinger *et al.* 2014). This is more frequent in osteoporotic patients over 60 years of age, for which an unexpectedly high rate of screw cut-out and revision surgery is observed (Kralinger *et al.* 2014; Owsley and Gorczyca 2008; Solberg *et al.* 2009; Sproul *et al.* 2011).

The majority of the complications are screw perforation (9%), varus collapse due to a malunion (7-16%), subacromial impingement (5%), avascular necrosis (4%), and bone resorption (2-10%). Overall, there is a reoperation rate of around 14% (Kavuri *et al.* 2018; Owsley and Gorczyca 2008; Sproul *et al.* 2011).

Excessive screw penetration into the bone is referred to as screw cut-out or perforation and is the most common reason for early revision surgery (Sproul *et al.* 2011). There are two different types, primary and secondary. Primary screw penetration describes the intraoperative placement of the screws into the glenohumeral joint. When the screws penetrate the articular

surface due to the collapse of the humeral head caused by varus collapse, avascular necrosis, or failure of fixation, it is referred to as secondary screw penetration (Kavuri *et al.* 2018). Primary screw cut-out is preventable since it is a technical error made by the surgeon. Secondary screw cut-out is more difficult to predict and may be the result of technical errors in fixation or due to the type of fracture and subsequent development of avascular necrosis (Sproul *et al.* 2011). A significant risk factor to consider in relation to this type of failure could be the age of the patient and the quality of the bone, as shown in the study conducted by Owsley et al. (Owsley and Gorczyca 2008).

Varus collapse of the humeral head is one of the major complications of locking plate fixation of proximal humerus fractures. Moreover, this failure can lead to secondary subacromial impingement and penetration of the screw into the articular surface of the glenohumeral joint. Risk factors related to this failure may be poor alignment of the implant and bone segments, lack of medial cortical support and bone quality (Kavuri *et al.* 2018; Sproul *et al.* 2011).

Impingement is another complication observed, especially in varus-angular fractures where impingement of the greater tuberosity is more likely. This problem is often symptomatic and may require removal of the plate. Subacromial impingement may be the result of intraoperative poor plate placement or the consequences of humeral head collapse. This type of failure can be anticipated and can be avoided if the implant placement procedure is performed correctly. Moreover, an increased risk of hardware impingement has been noted for a high surgical neck fracture followed by more proximal plate fixation (Greiwe 2015; Sproul *et al.* 2011).

Avascular necrosis (AVN) affects 11% of patients and can occur up to five years after injury (Sproul *et al.* 2011). There are many complications subsequently caused by AVN, including pain, poor functional performance, decreased range of motion and arthritis of the glenohumeral joint. (Kavuri *et al.* 2018). This kind of failure is associated with the type of fracture and the stability of the fixation, and the risk of incidence increases for three- and four-part fractures (Huri *et al.* 2020). Furthermore, AVN can cause bone resorption of the tuberosity, with a reported incidence

between 3% and 10% (Miyamura *et al.* 2021). The size of major tuberosity fragments may influence their resorption, increasing the risk when they are smaller than 15% of the humeral head, and associated with low bone quality. The study conducted by Miyamura found a higher rate of bone resorption for an unreduced greater tuberosity, and the risk of resorption is reduced when an optimal reduction is achieved (Miyamura *et al.* 2021).

Screw cut out and varus collapse can be related to hardware and bone failures, and their causes will be explored in this thesis. With regard to impingement and avascular necrosis, these are attributed to kinematic mechanisms and vascularisation within the bone, so these are aspects that will not be investigated as part of this thesis.

2.4 In-Vitro Testing

Several studies have been found in the literature, which analyse the biomechanical behaviour of different plate designs and assess their performance through in vitro experiments. The use of in vitro biomechanical studies is relevant when it is necessary to test the safety of a device, generally by subjecting it to a maximum load or breakage scenario. Moreover, in vitro testing with a bone sample provides a realistic model of clinical conditions, and cadaver models are considered the gold standard for evaluating bone failure, including screw cut-out under fatigue loading. This approach inherently incorporates factors such as contact mechanics, friction, and material properties, along with providing precise control over experimental conditions. In contrast, in vivo testing presents certain limitations, as it remains a model. For instance, it does not account for biological processes like bone remodelling, and there can be significant variability in the quality of cadaveric materials. However, as pointed out in the systematic literature review conducted by Jabran et al (Jabran *et al.* 2018), there is no defined protocol for testing fixation plates for humerus fracture and each study considers several conditions in a different way, such as loading conditions, load application methods, failure criteria and parameters determined to indicate implant performance.

The presented assessment of the biomechanical studies is based on the review conducted by Jabran et al. (Jabran *et al*. 2018). One of the main criteria for classifying biomechanical studies in the literature is the type of load applied to the humerus, which has been defined as axial, torsion, bending and flexion combined with an axial load, which is found to be the most popular methodology (Jabran *et al*. 2018)(Figure 2.11).

The simplest loading configuration was compression of the bone along its shaft axis, whereby the most common setup was to fix the distal portion of the humerus at the shaft and load the humeral head. In the majority of cases, the system was considered to have failed when complete or irreversible closure of the fracture gap occurred and when a non-linearity in the load-displacement curve was evident, where the point of failure was also considered to be a point of a large load drop.

The least common type of loading observed was the application of a pure torsion along the shaft axis, in which generally the distal humerus was fixed and a moment was imposed on the head of the humerus. The most common criterion used to define failure was to measure the angular displacement of the bone portion or the recording of interfragmentary motion using 3D motion analysis systems.

Another configuration used was to apply a load along one of the two axes perpendicular to the shaft axis resulting in extension/flexion or varus/valgus moment, to simulate a bending condition. Different scenarios for load application were presented by Jabran et al., such as direct shaft loading, direct head loading, eccentric loading without a rod, eccentric loading with a horizontal rod, and eccentric loading with a vertical rod. Generally, the stiffness was calculated from the force-displacement curve and failure load of the bone-plate construct.

The type of loading most frequently reproduced was flexion and axial loading, and almost all studies loaded the humeral head. Several of these angled the humerus by approximately 20° of abduction to simulate mainly shear loading. In this configuration, some studies based their

conditions on the clinical studies conducted by Bergmann et al., in which the glenohumeral contact forces were measured in vivo during activities of daily life (Bergmann *et al.* 2007). Typically, 3D motion analysis systems were used to measure humeral and inter-fragmentary motion.

In conclusion, despite the type of loading scenario examined, most studies analyse peak force output, stiffness, and fracture gap displacement to investigate construct stability. The
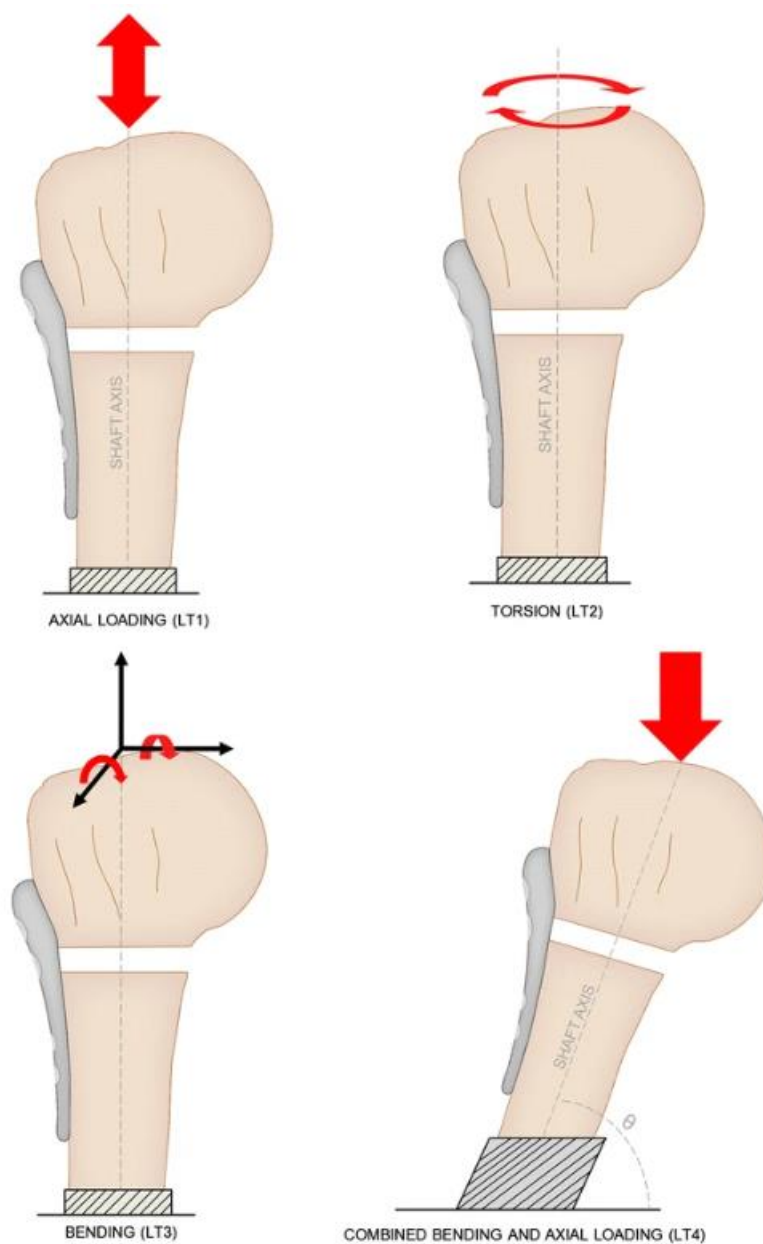
Figure 2.11 - Four types of loading conditions performed in the literature (Jabran et al., 2018)

limitations concerning these biomechanical studies are that they do not analyse the distribution of stresses and strains, information that is useful for a deeper understanding of the functionality of the fracture fixation plate. Moreover, the tests conducted involve the use of cadaveric specimens, so there is a problem of great variability between the specimens, only one implant design per specimen can be evaluated and the accessibility of cadaveric specimens is very limited.

2.5 Finite Element Modelling

An alternative to experimental studies is computational modelling, such as Finite Element (FE) analysis. This technique aims to virtually replicate an in vivo or in vitro scenario, allowing a potentially fast analysis of the problem and doesn't require the use of technical equipment and samples, which can be expensive and not always available. A powerful advantage of this technique is that it allows the measurement of results that cannot be detected experimentally, such as the stress and strain distribution in the bone and the implant. Furthermore, when developed correctly, FE analysis helps to improve implant design and gives the possibility to develop a preoperative plan, with the aim of improving fracture healing and reducing device failure (Lewis *et al.* 2021, 2022). In addition, FE analysis can contribute to faster research and development towards better treatment options and strategies (Castro-Franco *et al.* 2020; Lewis *et al.* 2021). On the other hand, FE modelling has some disadvantages. One significant limitation is its dependence on computer processing power, which can restrict the complexity of the models that can be analysed. Additionally, FE modelling requires approximations to represent the mechanical properties of bone materials, as well as the boundary and loading conditions. These approximations may not accurately reflect the complexities of real-world scenarios, potentially impacting the accuracy of the analysis.

This technique could be advantageous for studying the biomechanics of fixation plates for proximal humerus fracture since it would allow not only to estimate the deformations in the

bone and the stresses in the implant, but also to easily and quickly modify several parameters to analyse different aspects and identify the worst case leading to failure of the device.

The FE technique consists of discretising a virtual geometry into a mesh and imposing boundary and loading conditions. The elements are deformable, unlike rigid-body musculoskeletal models, allowing stresses and strains to be calculated.

The literature does not offer a large number of FE studies on the fixation plate for proximal humerus fracture and a total of 32 studies were found. However an increase in the use of this technique has been noticed in recent years. This rise in interest could be explained by the growing incidence of this type of fracture and the high failure rate of the device. This section describes all the works found in the literature through a non-systematic review, that was also recently conducted by another research group (Lewis et al. 2021). The table below summarises the main features of these studies (Table 2.2)

In general, these models are designed in the same way, i.e. the humerus is segmented from a CT image, then the fracture and implant placement are modelled, the geometry meshed, the materials defined and the loading and boundary conditions applied (Figure 2.12)
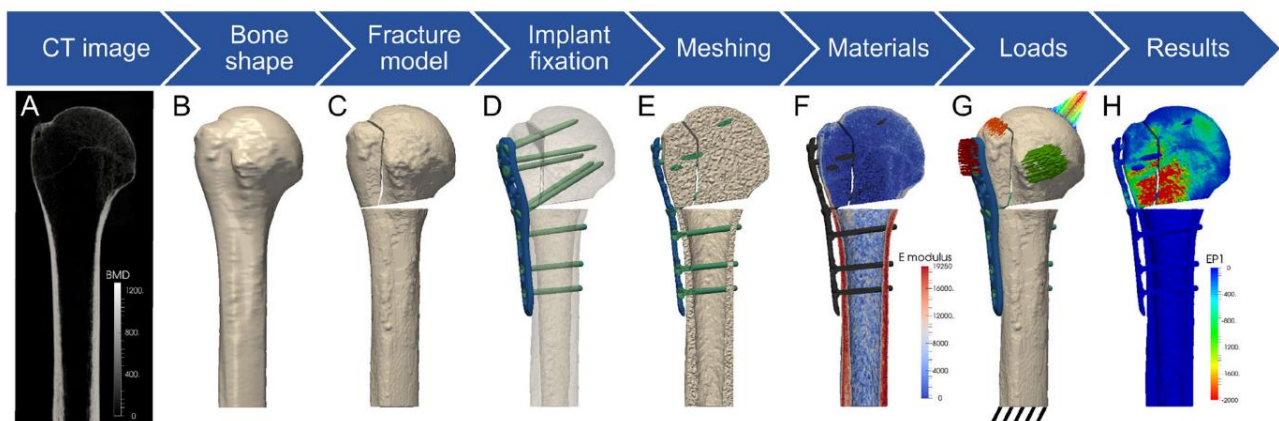
Figure 2.12 - Overview of finite element analysis workflow (Lewis et al., 2021)

First, the reported studies were conducted with the aim of comparing different systems with each other, for example, a locking plate and an intramedullary nail, optimising the use of

43

existing locking systems or optimising their design (Chen et al. 2017; Feerick et al. 2013; Li et al. 2022; Mischler et al. 2020a). These studies were developed with a subject-specific geometry, fourteen studies conducted the analysis on one patient (Chen *et al.* 2020; Jabran *et al.* 2019b) and the others on multiple patients (Schader *et al.* 2021; Steiner *et al.* 2018).

Continuum FE (cFE) modelling of bone was used in most studies and only one used microFE (µFE), which allowed to consider the microarchitecture of trabecular bone in the humerus head, obtaining accurate stiffness values and estimating bone surface displacement (Steiner *et al.* 2018). The lack of this type of study in the literature is explained mainly by the fact that the µFE modelling of the fixation of the entire proximal humerus requires large computational resources even for linear elastic simulations. Indeed, studies have been found in the literature that have considered an µFE analysis only for a reduced portion of the bone, analysing the behaviour at the screw-bone interface. These studies were not included in the literature search for this study. Regarding the material properties used in the cFE models, they were mostly considered heterogeneous based on Bone Mineral Density (BMD) derived from CT scans (Kennedy *et al.* 2013; Varga *et al.* 2018), alternatively, they were also simplified as homogeneous by assigning different values for cortical and trabecular bone (Jabran *et al.* 2019b). In general, the assignment of material properties based on BMD makes the model more complex, although the correct assumption of the equation between BMD and Young's modulus of bone remains problematic, considering that there is no accurate study conducted specifically for the humerus, and the relationships used have been derived mostly from those obtained for the femur (Helgason *et al.* 2008). For the µFE model, homogeneous properties at the tissue level were assumed.

Regarding the definition of the constraints, focusing on the bone-screw interface, most studies, whether µFE or cFE, bonded the two structures together, in order to make the simulation easier and faster (Maldonado *et al.* 2003; Steiner *et al.* 2018; Varga *et al.* 2018). Only one study imposed a friction coefficient between bone and screws, justifying this choice as mimicking the implant

inserted without osseointegration (Chen *et al.* 2017). Two studies analysed the difference between the interface of bonded and pseudo-threaded screws, while the first study found no difference (Inzana *et al.* 2016), the second noted that a longer thread increases the bending stiffness in varus (Le *et al.* 2019). However, it has to be considered that the use of threading in the simulation increases the computation time significantly, from 2 to about 290 minutes (Inzana *et al.* 2016). Furthermore, a study conducted by MacLeod *et al.* on the femoral shaft concludes that the use of a bonded condition is sufficient if the focus is not on analysing the bone-screw interface but on understanding the general behaviour of the system (MacLeod *et al.* 2012).

With respect to the loading conditions, several studies reproduced simplified conditions based on experimental setups conducted in previous publications, such as varus bending, torsional loading and, the most frequently used condition, bending combined with compression. Other studies have mimicked physiological activities. Most of these referred to the clinical study conducted by Bergmann to determine the magnitude and direction of the reaction force at the glenohumeral joint (Bergmann *et al.* 2011). Indeed, the most commonly reproduced condition consisted of the application of a force on the humeral head with an inclination of approximately 20° deviation from the axis of the humerus in order to represent a typical force scenario during common shoulder activities. None of these studies refers to the worst case, but the magnitude of the applied force refers to a functional load.

Regarding the outputs analysed, the most frequently reported outcome from the analysed results was the fracture's gap motion, which aimed to achieve maximum stiffness of the bone-implant system (Chen *et al.* 2020; Feerick *et al.* 2013). Additionally, some studies performed experimental analyses to validate the FE models by comparing bone displacement (Jabran *et al.* 2019b; Tilton *et al.* 2020a).

Other parameters mentioned in some studies are implant pressure and stress, with the aim of minimising them in order to avoid damage to the device. Furthermore, it was observed that high

pressures at the implant tips and high shear stresses along the implant-bone interfaces indicate a possible pull-out or push-out of the screws for each device (Feerick *et al.* 2013; Yang *et al.* 2015).

Finally, some studies have measured stress or strain across the entire bone or in bone fraction around the tips of the screws (Schader *et al.* 2021; Tilton *et al.* 2020b). Reduced stress values have been recorded in the presence of the calcareous screw (Yang *et al.* 2015), with longer screws (Kennedy *et al.* 2013) by inserting cement between the implant and the (Feerick *et al.* 2013), changing the screw angles (Jabran *et al.* 2018) and selecting more elastic implant materials (Feerick *et al.* 2013). However, these studies have not been experimentally validated since the stress within the bone cannot be measured. With regard to strain in bone, the general pattern of strain distribution appeared to be independent of individual bone quality and musculoskeletal activity, but its magnitudes were mainly influenced by bone quality (Maldonado *et al.* 2003). In a study conducted by Varga et al., the strain measurement around the screw tips was found to be an experimentally validated surrogate for the cyclic screw cut-out failure (Varga *et al.* 2017), consequently the resulting model was used to explore the behaviour of the system under different conditions. For example, it has been noted that the strain value in the bone can be reduced by appropriately selecting the number of screws in the head of the humerus (Inzana *et al.* 2016; Tilton *et al.* 2020b), by using longer screws (Fletcher *et al.* 2019a), maximising the spread of the screws (Fletcher *et al.* 2019a) and directing the screws in the humerus head in a more proximal direction (Mischler *et al.* 2020b). Furthermore, it has been observed that the correct position of the plate influences the value of strain, reducing its magnitude with a more proximal implant position (Fletcher *et al.* 2019b).

Furthermore, it can be seen that different parameters have a significant influence on the success of the fixation plate for proximal humerus fracture. In this regard, a few research groups have conducted several parametric studies to explore the variation of stress and deformation in the implant and the bone due to the change of several parameters, such as screw length (Fletcher

et al. 2019a), screw configuration and space positioning (Fletcher *et al.* 2019b; Jabran *et al.* 2019b; Mischler *et al.* 2020b; Tilton *et al.* 2020b), plate positioning (Fletcher *et al.* 2019b), fracture gap size and cement augmentation configurations (Tilton *et al.* 2020b; Varga *et al.* 2018, 2020). It is crucial to acknowledge that the parametric studies mentioned used a relatively small number of subjects for the analysis, which limited the understanding of subject variabilities. Indeed, some studies conducted the analysis on just one subject (Jabran *et al.* 2019b; Tilton *et al.* 2020b), while others included a small number of subjects, ranging from 24 to 47 (Fletcher *et al.* 2019a, 2019c, 2019b; Mischler *et al.* 2020b; Varga *et al.* 2018, 2020). Despite the fact that conducting a parametric study is useful to explore different scenarios, this type of study is merely informative and does not consider the interaction between parameters, nor the probability of observing a specific level of a parameter (Taylor and Prendergast 2015).

Table 2.2 - Summary of Finite Element studies that investigated humeral proximal fractures with fracture fixation plates

PAPER	LOADING COND	PARAMETERS	N SUBJECTS	OUTPUT	RESULTS
Maldonado *et al.* 2003	Neutral arm position 90° abduction 90° forward flexion	Bone quality Cement augmentation	N = 2	Bone strain Shear stresses at the bone-cement interface	Strain magnitudes affected by bone quality Strain distribution is independent of bone quality and musculoskeletal activity
Feerick *et al.* 2013	90° abduction	Cement augmentation The behaviour of four different implants	N = 1	Displacement between the fracture fragments Shear stress distribution in cortical bone	Cement reinforcement improves the plate stability
Kennedy *et al.* 2013	90° abduction	Cement augmentation	N = 1	Maximum pressure on bone-implant interface Implant stress	Improvement in initial stability and decrease in the implant-bone interface stress are achieved through augmentation
Čukelj *et al.* 2014	Abduction Adduction Axial compression Flexion	Fracture gap angle The behaviour of two different plates	N = 1	Bone displacement Fracture gap displacement	The angle of the fracture gap does not influence the overall displacement of the model
He *et al.* 2015	Abduction Adduction Flexion Extension Axial compression Int/Ext rotation	Bone quality The behaviour of two different implants	N = 1	Construct stiffness Fracture gap displacement Von Mises stress distribution on the implants	Enhanced stability with an additional medial buttress plate
Yang *et al.* 2015	90° abduction	Fracture gap length Calcar screw	N = 1	Max shear stresses of screw-bone interfaces Von Mises implant stresses	Calcar screws and medial cortical contact improve stability
Inzana *et al.* 2016	Varus bending	Screw configuration Bone-screw interface condition	N = 10	Averaged maximum principal strain in peri-screw bone region	Calcar screws reduce peri-implant bone strain

PAPER	LOADING COND	PARAMETERS	N SUBJECTS	OUTPUT	RESULTS
					Bone-screw interface modelling does not affect statistical findings
Zhang *et al.* 2016	Axial bending	Hole style	N = 1	Von Mises stress around screw holes and in the implant Stiffness of the construct Directional displacement within the gap	It might be more effective to use a configuration with separate locking and dynamic holes to prevent plate fracture
Chen *et al.* 2017	Compression	Fracture impaction Bone quality The behaviour of two different implants	N = 1	Bone and implant stress Gap displacement	Fracture impaction enhances stability and decreases peak stress on the metallic implant and bone in cases of both normal and poor bone quality
He *et al.* 2017	Compression Abduction Int/Ext rotation	The behaviour of four different implants Quality of the bone	N = 1	Construct stiffness Fracture gap Stress distribution on the implant Neck-shaft angle variation	Additional medial plating increases construct stability and decreases implant stress via buttressing
Varga *et al.* 2017	Compression with bending	NA	N = 20	Averaged minimum principal strain in peri-screw bone region	Bones strain around screw tips is highly correlated with experimental cycles to cut-out failure
Acklin *et al.* 2018	Adduction	Inclination of gliding screws	N = 3	Peak principal compressive strains at the tip of the four proximal screws	Lower maximum values for the gliding plate when comparing a 30° screw angle to a 20° screw angle and to the PHILOS plate.
Mendoza-Muñoz *et al.* 2018	Compression	Loading conditions Screws configuration Implant material	N = 1	Construct stiffness Von Mises stress in the implant	The use of medial support and steel plates decreases implant stress and

PAPER	LOADING COND	PARAMETERS	N SUBJECTS	OUTPUT	RESULTS
		Bone quality		Bone deformation	bone deformation and increases stiffness
Steiner *et al.* 2018	Displacement at the humeral head	Bone quality	N = 8	Construct stiffness Strain distribution in bone	Validated model, good correlation with experiments in terms of construct stiffness
Varga *et al.* 2018	Three experimental and three physiological loading modes	Effect of cement augmentation of screw tips	N = 47	Averaged compressive strain in peri-screw bone region	The biomechanical effect of augmentation strongly depends on the BMD, more osteoporotic samples experience greater benefit
Fletcher *et al.* 2019a	Three experimental and three physiological loading modes	Length of the screws	N = 42	Averaged compressive strain in peri-screw bone region	Longer screws reduce the risk of cut-out failure
Fletcher *et al.* 2019c	Three experimental and three physiological loading modes	Proximal-distal plate positioning Calcar screws	N = 26	Averaged compressive strain in peri-screw bone region	Plate proximalization and calcar screws reduce the predicted cut-out failure risk
Fletcher *et al.* 2019b	Three experimental and three physiological loading modes	Screw configuration	N = 26	Averaged compressive strain in peri-screw bone region	Besides screw number, screw spread and calcar screws are the most influential on predicted cut-out failure risk
Jabran *et al.* 2019c	Bending	Orientation of infero-medial screws Screws length	N = 1	Load Maximum von Mises implant stress Bone stress around screws Change of fracture gap closure	The bending stiffness of the bone-plate construct is primarily reliant on the infero-medial screws more than any other screws
Le *et al.* 2019	Adduction	Length of the threaded head screw	N = 1	Construct stiffness von Mises stress of screws and screw holes	Threading the head screws caused an increase in the mean von Mises stress within the corresponding screw holes. Longer threading increased varus bending stiffness

PAPER	LOADING COND	PARAMETERS	N SUBJECTS	OUTPUT	RESULTS
Chen *et al.* 2020	Compression 20° abduction Rotation	The behaviour of four different implants Quality of the bone	N = 1	Construct stiffness Fracture gap displacements Von Mises stress in the implants	Intramedullary strut increases medial support and augments the stability of lateral plating
Kim *et al.* 2020	Abduction	Inferomedial supporting screws group Bone quality	N = 1	Stress distribution in the humeral head Local maximum peri-implant stress in the bone around proximal screws	In osteoporotic individuals, it is advisable to use inferomedial supporting screws to replicate the stress distribution seen in a non osteoporotic bone
Mischler *et al.* 2020a	Three experimental and three physiological loading modes	Plate type Screw configuration	N = 26	Averaged compressive strain in peri-screw bone region	PHILOS provides in general better stability compared to the periarticular plate Screw configuration highly influential on the predicted cut-out failure risk
Mischler *et al.* 2020b	Three experimental and three physiological loading modes	Screw orientation	N = 20	Averaged compressive strain in peri-screw bone region	Optimized locking screw angles significantly decrease predicted cut-out failure risk
Tilton *et al.* 2020c	Compression 20° abduction 20° adduction	Additive manufacturing	N = 1	Strain around the screws Construct stiffness Relative displacement of fragments Maximum von Mises stress in the plate	The medial strut attached to the plate reduced the predicted cut-out risk
Tilton *et al.* 2020b	20° abduction	Fracture gap Screw configuration	N = 1	Compressive strain around the screws Construct stiffness Relative displacement of fragments	Certain configurations with a lower number of screws can be effective Using one calcar screw may be enough

PAPER	LOADING COND	PARAMETERS	N SUBJECTS	OUTPUT	RESULTS
				Maximum von Mises stress in the plate	Negligible effect of screw configuration for well-reduced fractures
Varga *et al.* 2020	Three experimental and three physiological loading modes	Cement augmentation	N = 24	Averaged compressive strain in peri-screw bone region	Configuration of screw augmentation strongly determines the mechanical benefit Augmentation of calcar screws is most beneficial to decrease predicted cut-out failure risk
Schader *et al.* 2021	Three experimental and three physiological loading modes	Screw orientation	N = 19	Averaged compressive strain in peri-screw bone region	The orientations of the locking screws could be optimized for specific subjects to reduce the risk of cut-out and enhance PHF fixation
Xu *et al.* 2022	Torsional testing 0° axial compression 20° abduction 20° adduction	4 different calcar screw fixation groups	N = 1	Averaged compressive strain in peri-screw bone region The average stress in the cylindrical bone regions around the tips of the calcar screws	The medial support is important
Muthusamy *et al.* 2022	15°, 40°, 65° and 90° abduction and flexion 35°, 50°, 70°, and 90° horizontal flexion	NA	N = 1, synthetic bone	Maximum deformation at the fracture site Maximum bone stress Maximum plate stress Mean screw stress	The stiffness significantly increased when the arm was lifted at high angles and moved with minimal horizontal flexion. At the fracture location, the shoulder joint experienced greater displacement and increased stress on the plate, screws, and bones with horizontal flexion compared to abduction and flexion movements.

PAPER	LOADING COND	PARAMETERS	N SUBJECTS	OUTPUT	RESULTS
Li *et al.* 2022	Abduction Adduction, Flexion Extension Axial compression Int/Ext rotation	The behaviour of three different implants	N = 1, synthetic bone	Fracture stability Rotation stability Implant stress	Lateral support reduced implant stress
Putzeys *et al.* 2024	Axial bending, Torsion bending Combined Compression and Bending	Wire pulling mechanism Screw orientation	N = 1, synthetic bone	Bone and implant stress	The push-pull mechanism results in a stress distribution that is more uniform

2.6 Surrogate models

As observed in the previous section, several FE models have been developed in the literature to analyse the behaviour of the fixation plate for proximal humerus fracture under different conditions. Moreover, it is clear that the success of this device is influenced by a large number of parameters, such as the positioning of the screws, the type of fracture, and the number of screws implanted in the head, etc. In order to consider the interaction and variation of all these parameters, the simple use of an FE analysis through a parametric study would be too expensive, as the number of simulations required would be too high. For this reason, it would be necessary to use more advanced computational techniques (Taylor and Prendergast 2015).

A more sophisticated approach compared to a parametric study is to develop a Design of experiments (DoE), which consists of exploring a fixed number of levels (N) for a large number of parameters (P) (Taylor and Prendergast 2015). Mendoza-Muñoz et al. developed 24 models from a full-factorial design to evaluate the effects of implant material properties, type of force applied, bone density and screw configuration. This is the only Design of Experiment found regarding the fixation plate for a proximal humerus fracture, and it did not explore the problem in significant detail (Mendoza-Muñoz *et al.* 2018). Instead, in a more general study conducted for femur fracture, the Taguchi method was used, in which only a few dimensions of the plate and screw design were taken as parameters. The number of experiments is significantly reduced by the Taguchi method, as it controls the level of design factors through the use of an orthogonal matrix. Additionally, it statistically assesses the significance of the design factors (Kim *et al.* 2011). DOE offers the advantage of considering the interaction of parameters, unlike parametric studies. It allows for the effective exploration of multiple parameters and the assessment of their individual impact on the expected outcome. However, it does not have a probability associated with fixed levels for each parameter (Taylor and Prendergast 2015).
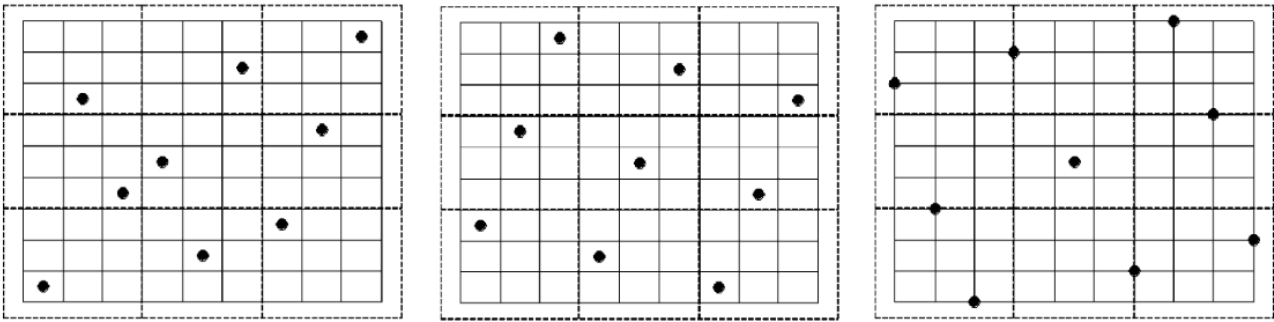
Figure 2.13 - Examples of Latin Hypercube sampling method. (Leary et al. *2003) Reprinted by permission of Informa UK Limited, trading as Taylor & Taylor & Francis Group.*

Probabilistic analysis is the best method to simultaneously explore multiple variables and assess the risk/probability for a given outcome. The simplest method that can be used to conduct this kind of analysis is to use Monte Carlo-based sampling, however, it requires the execution of a large number of analyses, which makes it a very computationally expensive solution (Taylor and Prendergast 2015). One of the techniques that have been used in the analysis of various orthopaedic devices to minimise the computational cost associated with probabilistic analyses are surrogate models, which are mainly used when the number of parameters is extremely high (Taylor and Prendergast 2015). Surrogate models mimic the behaviour of complex systems, and they can be either analytical models or black box models. It has been shown that surrogate models are an efficient approach for the design of computationally expensive models such as those found in aerospace systems and orthopaedic medical devices (Queipo *et al.* 2005).

The basic principle of the surrogate model is to approximate through the input/output data the fitness function that describes a given problem, and is generally extremely complex (Jin 2005). The development of the surrogate model consists of three stages. Firstly, selecting an appropriate sampling technique to provide a training dataset that covers the range of input data, then the choice of the most appropriate algorithm for the analysis of the problem and the optimisation or training of the model parameters. Finally, the accuracy of the surrogate model is evaluated (Queipo *et al.* 2005). Some applications in the field of computational orthopaedics are presented below, in which finite element models have been developed to create input and

output data in order to build surrogate models. Several sampling methods and surrogate models are listed.

The sampling methods used are different, as shown in the table below (Table 2.3), and the aim is to create a dataset that allows the surrogate model to be developed correctly and efficiently. One sampling method used is the fractional factorial design, which consists of selecting a subset of all possible combinations. Indeed, the size of all potential combinations increases exponentially with the complexity of the problem and is not suitable for computational orthopaedic biomechanics due to the significant computational time, so it is preferable to use a fractional method (O'Rourke *et al.* 2016). The Taguchi method, a form of fractional factorial design, was used to design a fixation plate for a femur fracture and, as previously mentioned, this method is based on the orthogonal matrix, and the number of samples can be decreased efficiently (Fang *et al.* 2020). The most frequently found method used for biomechanical orthopaedic problems is Latin Hypercube sampling. This method is found to be efficient, providing uniform coverage of the sampling space by partitioning it so that the probability of each interval is equal (Fitzpatrick *et al.* 2014; Queipo *et al.* 2005). It also has the ability to provide different sampling plans with very different performances in terms of uniformity, as is shown in the example in the figure (Leary *et al.* 2003) (Figure 2.13).

The second step consists of choosing the most appropriate surrogate model. Models can be parametric (e.g. polynomial regression, Kriging) and non-parametric (e.g. neural networks, random forest), the first assumes that the global relationship between inputs and outputs is known, while non-parametric models use different types of simple models at different regions of the data in order to build a global model (Queipo *et al.* 2005). The polynomial regression model is a methodology that examines the quantitative relationship between a function of interest and a fixed number of basis functions (Queipo *et al.* 2005). A quadratic model was used in a study of a fixation device for femur fracture and simplified regression models were implemented afterwards since the full quadratic models are complex and can be challenging to interpret (Wee

et al. 2017). A simple linear regression approach, in which the output of the surrogate model is a linear combination of inputs, was used in a study for a knee implant. This study showed promising results regarding the use of a surrogate model in this type of application (Fitzpatrick *et al.* 2014). Another model that was found to be able to capture the data trend efficiently is the Gaussian process model, which was used in two different studies on hip implants (Al-Dirini *et al.* 2020; Bah *et al.* 2011). Furthermore, the Kriging method evaluates the output in an unsampled region by calculating a weighted average of the known values of the function in the region surrounding the point (Queipo *et al.* 2005). In a study of an acetabular cup, this model gave accurate predictions while significantly reducing the computational time (O'Rourke *et al.* 2016). With regard to non-parametric models, the Random Forests method was used for another study on Total Hip Arthroplasty and gave satisfactory results. This model consists of a set of decision trees, each trained using a bootstrap technique, and computes predictions as the average result of all the decision trees. This technique was shown to be reasonably fast to define and quick to train (Donaldson *et al.* 2015). Artificial neural networks are a technique that uses a network of functions to model complex, non-linear relationships between input and output variables. An example of this was conducted in a study to define the trend in femoral neck strain (Taylor *et al.* 2017).

Table 2.3 - Example of biomechanical studies that used surrogate models

Paper	Investigation	Sampling method	Surrogate model
Bah *et al.* 2011	Investigate the statistical effects of implant positioning on the initial stability of a cementless total hip replacement (THR).	Latin Hypercube Sampling (LHS)	Gaussian model
Fitzpatrick *et al.* 2014	Develop a surrogate model that can predict micromotion across the entire bone-implant interface for a knee replacement	Latin Hypercube Sampling (LHS)	Linear regression
Donaldson *et al.* 2015	Develop a surrogate model of Total Hip Arthroplasty that can predict contact mechanics and permanent deformations	Latin Hypercube Sampling (LHS)	Random Forest
O'Rourke *et al.* 2016	Develop a time-efficient method for quantifying FE model sensitivity to input parameters and their interactions by using a surrogate model.	Random from Full factorial design	Linear and Kriging models
Taylor *et al.* 2017	Develop an adaptive neural network (ANN)-based surrogate model to predict femoral neck strains and fracture loads obtained from a previously developed population-based FE model	Latin Hypercube Sampling (LHS)	Adaptive Neural Network
Wee *et al.* 2017	Create mathematical models that describe the interaction between the parameters of fracture fixation constructs and resulting 3D biomechanics	Design of Experiments	Polynomial regression model Simplified regression model
Chatterjee *et al.* 2019	Generate ANN models from FE data to explore the microstrain in femur implants, analysing various implant configurations and several bone conditions	NA	Adaptive Neural Network
(Ziaeipoor *et al.* 2019) Ziaeipoor *et al.* 2019	Develop Multivariate Linear Regression-based (MLR) surrogate models to lower the computational cost of predicting femoral strains during regular activity in comparison when compared to finite element analysis	Latin Hypercube Sampling (LHS)	Multivariate Linear Regression
Al-Dirini *et al.* 2020	Create a computationally efficient, cohort-specific surrogate model for predicting the distribution of micromotion along the contact surface for various implant positions	Latin Hypercube Sampling (LHS)	Gaussian model

Paper	Investigation	Sampling method	Surrogate model
Fang *et al.* 2020	An orthogonal regression approach is used to find the relationships between the internal fixation parameters and the biomechanical indices with the lowest number of runs, in order to establish surrogate models of different biomechanical indices	Taguchi	Quadratic regression orthogonal method
Stowers *et al.* 2021	Use of an FE and Gaussian process approach to determine the stress and strain for skin sutures	Latin Hypercube Sampling (LHS)	Gaussian model
Dhason *et al.* 2023	Generate ANN models from FE data to analyse implant movements, exploring various material compositions of a fracture fixation plate	NA	Adaptive Neural Network

2.6.1 Deep Learning

Deep Learning models are a more complex type of surrogate model. Deep Learning is an advanced subset of Machine Learning inspired by the human brain's structure and function. These models consist of complex data-driven algorithm structures and have been successfully applied in various fields such as aerospace engineering, business, and science applications. Deep Learning algorithms have been developed to perform complex tasks such as image recognition, natural language processing, decision-making, and drug molecule prediction (Lecun *et al.* 2015).

The foundation of Deep Learning is the Neural Network, which consists of connections of basic units called neurons, connected to each other in several numbers of layers (López-Monroy and García-Salinas 2021). Each neuron receives an input, elaborates the information through linear transformations and is sent as output. The information received from multiple inputs is elaborated as follows: $z = (\sum_{i=1}^{n} w_i x_i) + b$, in which z is the output, x is the input, w is the weight value and b is the bias value. After the information is elaborated, usually an activation function is applied to the output z, depending on the type of problem. Throughout the training process, the loss function is used to indicate whether the model is effectively learning. Different types of loss functions are available, and all of them compare the predicted output of the model with the actual output. During the training process, the weights and biases of the network are updated through the backpropagation process to minimise the error between the predicted and actual output (López-Monroy and García-Salinas 2021). The term "Deep" in Deep Learning refers to the number of layers of neurons, including an input layer, multiple hidden layers, and an output layer.

There are different types of deep learning models, which vary based on their complexity and applications, such as feedforward Adaptive Neural Networks (ANN), Convolutional Neural Networks (CNN), and Graph Neural Networks (GNN).
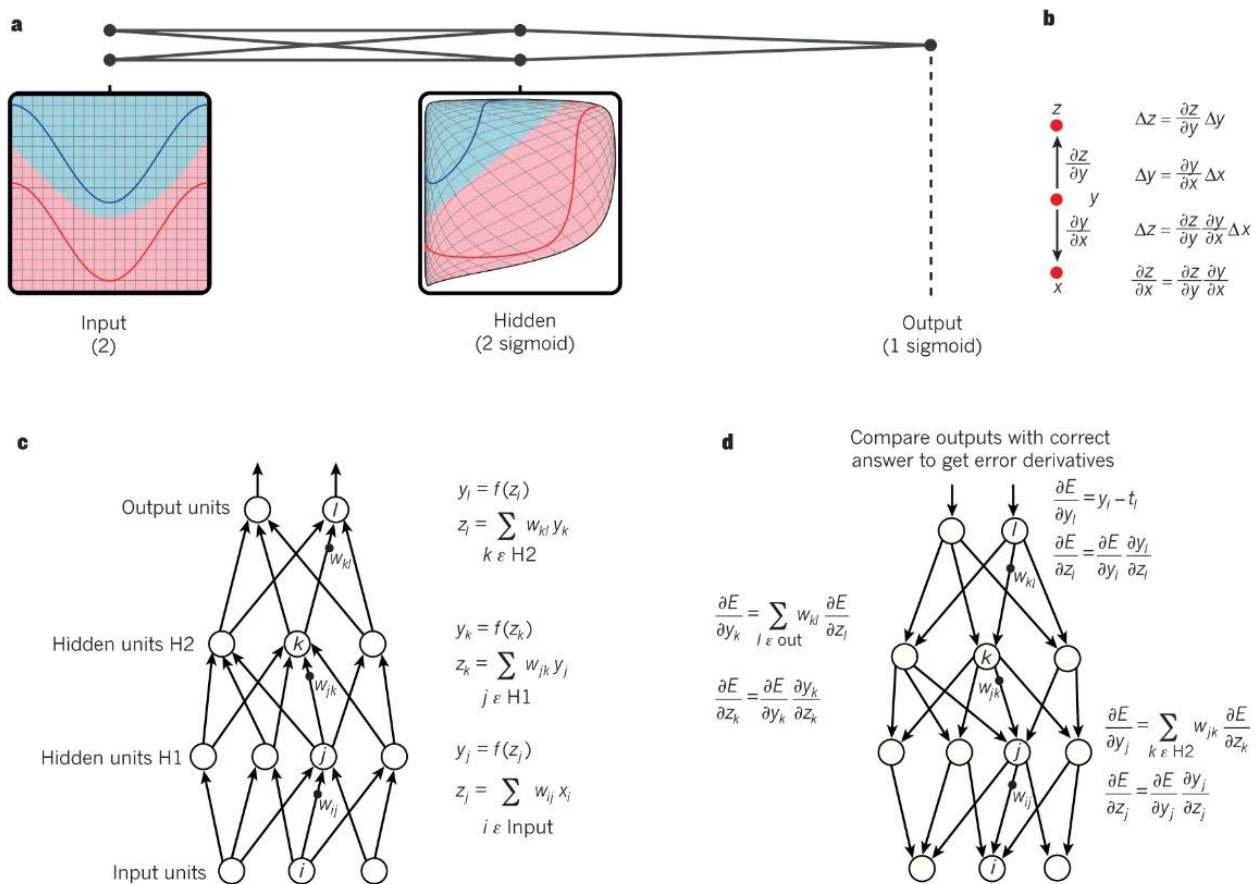
Figure 2.14 - Example of a multilayer neural networks and backpropagation (Lecun et al. 2015)

Adaptive Neural Networks (ANNs) are one of the simplest types of Deep Learning models, that often use a multilayer approach. ANNs usually adopt a feedforward approach, which entails mapping a fixed-size input to a fixed-size output. To move from one layer to the next, the inputs are sum weighted by a set of neurons from the previous layer, and compute the result through a non-linear function (Lecun et al. 2015). These models are usually trained using backpropagation functions (Sarker 2021)(Figure 2.14). Several ANN models have been developed for combining FE analysis with Deep Learning. These models have been used to estimate the accumulation of apparent fatigue damage in bone structures (Hambli 2011), predict femoral neck strains and fracture loads (Taylor et al. 2017), and estimate tissue adaptation loads based on trabecular bone density distribution (Zadpoor et al. 2013).

A Convolutional Neural Network (CNN) is a type of Deep Learning algorithm that is commonly used for image analysis. CNNs are a more advanced form of Artificial Neural Networks (ANNs), and they use convolutional and pooling layers to process 1D and 2D structured data such as signals, languages, and images (Lecun *et al.* 2015). CNNs are frequently used to perform tasks such as natural language processing, object detection, image segmentation and medical image analysis (Sarker 2021).

As an example, a CNN model had been developed to classify proximal humeral fractures from medical images (Chung *et al.* 2018). A CNN approach is also used, where the actual data from the FE mesh is utilised for training the models. One of the earliest applications of this approach was conducted on a 2D aortic structure (Liang et al. 2018). Another relevant work used a CNN approach to estimate the stress in an inhomogeneous structure (Krokos et al. 2022b).

2.6.1.1 Graph Neural Network

When working with 3D structures, it can be challenging and restrictive to use ANN and CNN approaches. Indeed, ANN models typically operate on parameters or features, while CNN models require data with a grid-like structure (Jiang and Chen 2023)(Figure 2.15). In FE
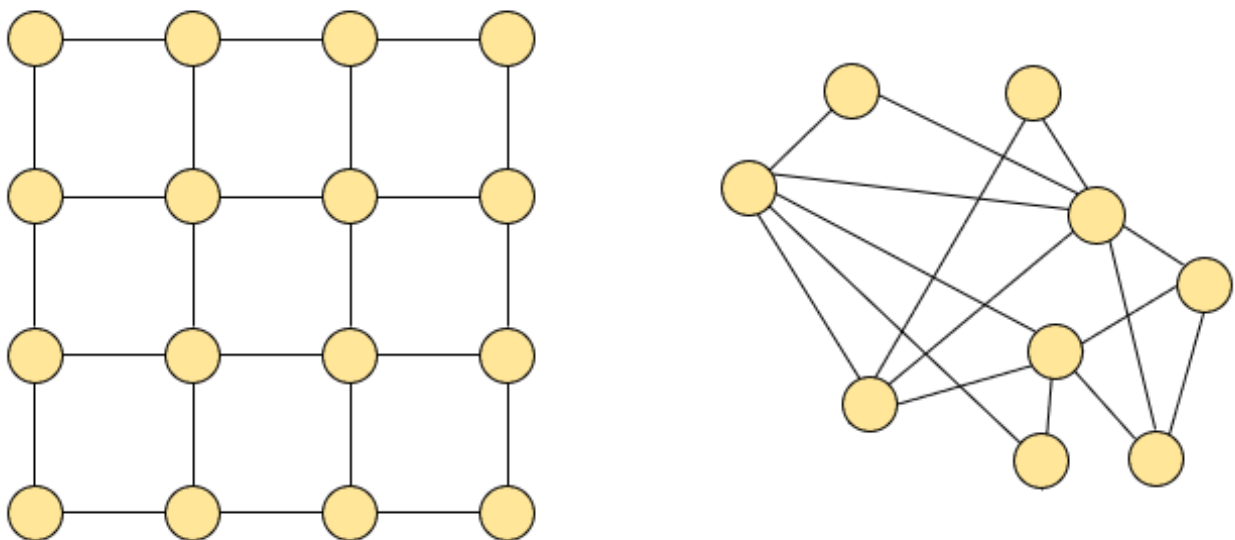
Figure 2.15 - Example of regular grid structure (left) and graph structure (right) (image crafted by the author).

applications, these methods are particularly challenging due to the mesh data, which often lacks an organised grid structure, therefore, a more flexible methodology is required. Graph Neural Networks (GNNs) are deep learning techniques that operate on graph structure.

This methodology has a wide range of applications in various fields such as social network analysis, biology, and computer vision (Zhou et al. 2020). Unlike CNN methods, GNN works on non-Euclidean structures and focuses on tasks such as node classification, link prediction, and clustering.
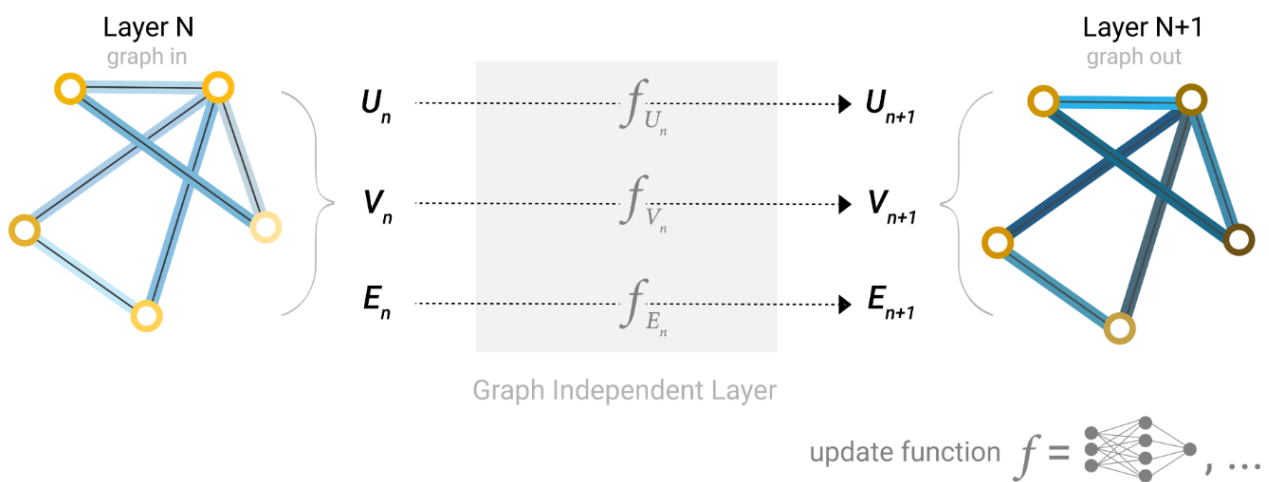
Figure 2.16 - A single layer of a simple GNN. A graph is the input, and each component (V, E, U) gets updated by a multilayer perceptron (MLP) to produce a new graph. Each function subscript indicates a separate function for a different graph attribute at the n-the layer of a GNN model. (Sanchez-Lengeling et al. 2021)

GNNs are a type of neural network that can process and analyse graphs. A graph G = (V; E) is defined by a set of nodes V and a set of edges E connecting these nodes (Georgousis et al. 2021). GNNs have the unique ability to train on different structures, making them very powerful.

GNNs apply separate neural network layers to each component of a graph. These layers are known as GNN layers. For every node, edge, and global information, the GNN layer is applied to elaborate the input data. The output graph of a GNN can be described with updated embeddings because the GNN updates each of the node, edge, and global-context representations, but does not modify the connectivity of the input graph (Sanchez-Lengeling et

al. 2021)(Figure 2.16). Overall, GNNs are a very powerful tool for analysing graphs and extracting useful information from them.

According to the literature, only one study was found that utilised a computational model on a 3D structure using FE analysis combined with a GNN approach. The study was conducted on a porous material with the goal of developing a model that could predict the stress in the structure
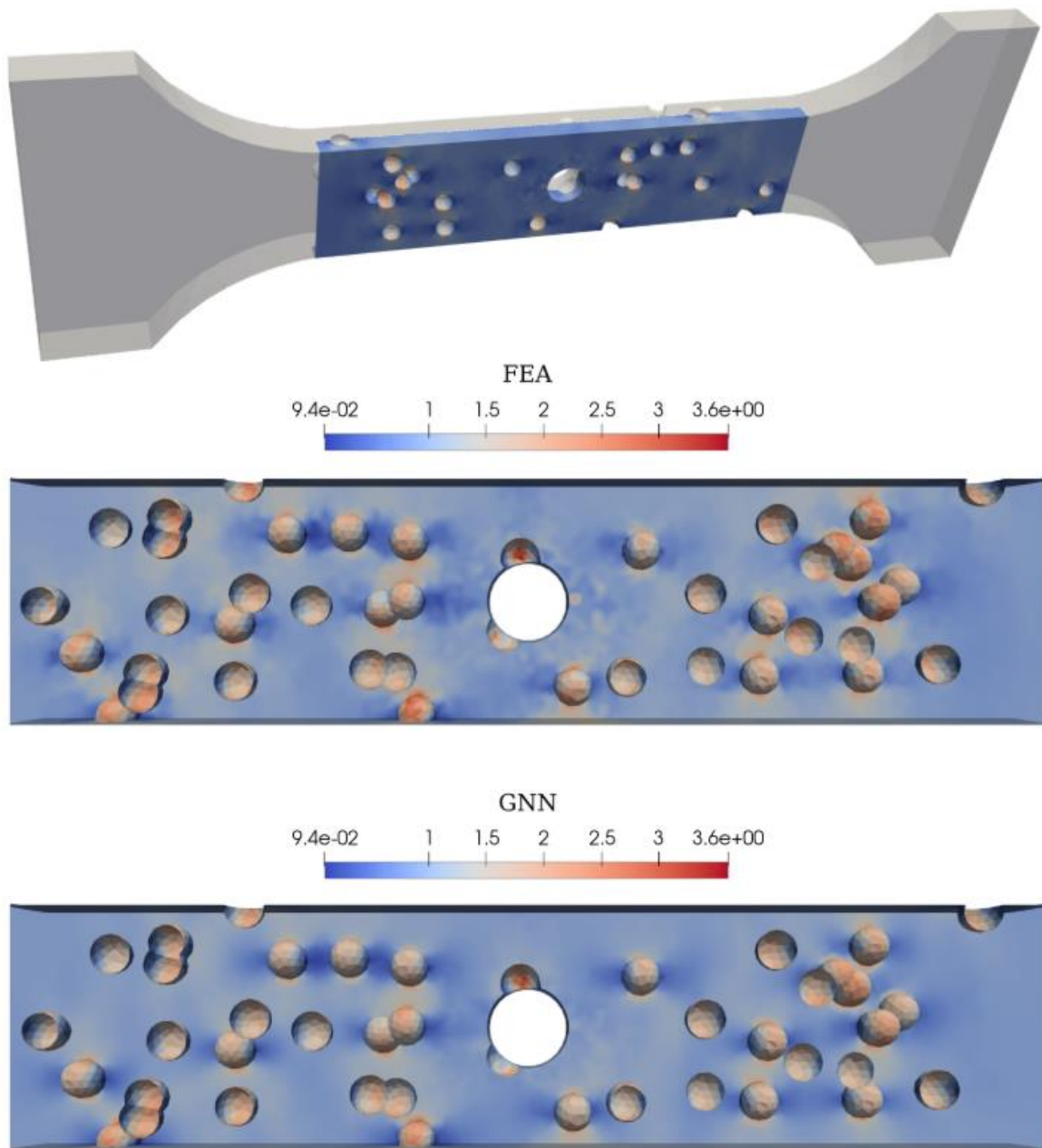
Figure 2.17 - Comparison between the Von Mises stress distribution as calculated by FEA (centre) and the Von Mises stress distribution as predicted by the GNN (bottom) (Krokos et al. 2024)

by altering the porous positions and dimensions. The study yielded positive results, indicating that a GNN methodology can be effectively carried out using data collected from FE analysis without compromising accuracy (Krokos et al. 2022a)(Figure 2.17).

2.7 Summary

Proximal humerus fracture is a common type of fracture in elderly patients. As the geriatric population is growing, the incidence of this type of injury is also on the rise. Surgical treatment for this fracture typically involves the use of fracture fixation plates and over the years, these plates have undergone several design changes to improve their success rate. The most significant change was the introduction of locked plates which have threads between the screw heads and the holes in the plates. More recently, second-generation fixation plates have been developed, which allow for the orientation of the screw to be varied before insertion into the bone. Despite these design improvements, the failure rate of fracture fixation plates can be as high as 35% which are clinically shown as screw perforation, malunion, subacromial impingement and avascular necrosis. The causes of these failures are not yet fully understood. Indeed, the bone-implant system is complex and involves multiple variables such as the position of the plate, the number of screws in the head of the humerus, and their orientation in space. FE analysis is a commonly used computational approach to better understand the behaviour of the plate-bone system. However, the studies conducted so far have been limited to a parametric analysis with a restricted number of variables. In fact, conducting a parametric analysis with a larger number of variables would be too computationally expensive. To overcome this limitation, advanced computational methods such as Deep Learning can be used. Deep Learning models have been successfully applied to various complex problems such as natural language processing, image recognition, and market trend predictions. In biomedical applications, these models have been used to predict femoral neck strains, and fracture loads, and classify humeral fractures from medical images.

Combining Deep Learning with FE analysis has the potential for a more efficient investigation of the failure of fracture fixation devices by varying the parameters of interest, implant and patient-related. Advanced computational methods are necessary to assess the influence of patient and implant-related parameters on bone strain in a more efficient and time-efficient manner, in order to understand the cause of failure and eventually improve fracture fixation devices.

Chapter 3

Assessing screw length impact on bone strain in proximal humerus fracture fixation via surrogate modelling[1]

3.1 Introduction

Proximal humerus fractures are one of the most frequent fractures in older patients (Sporer *et al.* 2006) and the incidence of this type of injury is expected to increase in the coming decades with the increase in the geriatric population (Palvanen *et al.* 2006). About 15% of proximal humeral fractures require surgery (Neer 1987), which usually involves the use of a fixation device like a fracture fixation plate. Fracture healing is not always achieved, and a failure rate of up to 35% has been reported in the literature (Kralinger *et al.* 2014). There is yet no clear understanding of the causes of these failures. Indeed, the complexity of the bone-implant system

---

[1]The study presented in this chapter is the subject of the following paper:

is determined by the involvement of a high number of variables, some related to the patient, i.e. bone quality and type of fracture, and some related to the surgery, i.e. the position of the plate, the number and the lengths of the screws in the head of the humerus and their orientation in space (Lewis *et al.* 2021). Several Finite Element (FE) studies have been developed with the aim of better understanding the behaviour of the plate-bone system, and the most detailed studies have been limited to a parametric analysis with a restricted number of parameters being assessed, performing from less than one hundred to up to a few thousand simulations, but not covering all the possible configurations for each study (Fletcher *et al.* 2019a, 2019c, 2019b; Jabran *et al.* 2019b; Mischler *et al.* 2020b; Tilton *et al.* 2020b; Varga *et al.* 2018, 2020). The parameters that have been evaluated in these studies are screw length (Fletcher *et al.* 2019a), screw orientation (Jabran *et al.* 2019b; Mischler *et al.* 2020b; Schader *et al.* 2021) and configuration (Fletcher *et al.* 2019c; Mischler *et al.* 2020a; Tilton *et al.* 2020b), plate positioning (Fletcher *et al.* 2019b), cement augmentation at the screw tips (Varga *et al.* 2018, 2020) and fracture configuration (Tilton *et al.* 2020b). Studies have been conducted on a single subject (Jabran *et al.* 2019b; Tilton *et al.* 2020b) and others on a group of subjects (Fletcher *et al.* 2019a, 2019c, 2019b; Mischler *et al.* 2020b; Varga *et al.* 2018, 2020). Indeed, from a computational point of view, it would be too expensive to develop a full parametric analysis with a large number of parameters, considering the time taken to develop and run each FE model.

Surrogate models are a potential solution as they are used to reduce the computational time in complex problems (Jin 2011). The principle of surrogate models is to approximate through input parameter values and the corresponding output responses a function that describes a given problem (Jin 2005). Kriging, Adaptive Neural Network (ANN) and Gaussian process (GP) based models are a few examples of surrogate models that have been used in the past to describe a biomechanical problem trained using finite element data, in particular for hip and knee implants (Al-Dirini *et al.* 2020; O'Rourke *et al.* 2016; Taylor *et al.* 2017). The ANN technique simply consists of a connection of neurons that process and send information to each other,

where the mechanism is inspired by the neurons in the brain. It has been demonstrated to be accurate for the prediction of femoral strain (Taylor *et al.* 2017) but hasn't yet been used to explore the impact on the bone of a fracture fixation device. The aim of this study is to develop an ANN based surrogate model to analyse the effect on bone strain of the length of the screws of a fracture fixation plate implanted in a humeral head of a single subject. We hypothesize that the ANN models can predict the variation of bone strain with the variation of a single implant parameter in a more cost-effective manner compared to a pure FE analysis. This could be particularly useful in more complex analyses in the future.

3.2 Method

3.2.1   FE model

A CT scan of a cadaver of a 61-year-old female from the New Mexico Decedent Image Database (NMDID) was used to generate a model of a proximal humerus (Edgar *et al.* 2020). The geometry of the humerus was segmented using Simpleware™ software (Version U-2022.12; Synopsys, Inc., Mountain View, USA), and the proximal region of the bone was extracted obtaining a segment of 160 mm length. A single fracture was virtually simulated by creating a 5 mm horizontal gap at the surgical neck of the humerus. This fracture was categorized as 11-A2.1 in the AO classification (Müller *et al.* 1990).

A fracture fixation plate (Austofix, Adelaide, Australia) was virtually implanted in the bone and secured with seven proximal screws and three distal screws each with a diameter of 3 mm. Screws of different lengths were used, based on the distance between the tip of the screw and the glenohumeral joint (TJD, Tip to Joint Distance), which was varied between 4, 8, 12 and 16 mm (Fletcher *et al.* 2019a), resulting in a total of 4^7 possible configurations. The plate was positioned 4mm posterior to the bicipital groove and 7mm distal to the top of the greater tubercle (Figure 3.1).
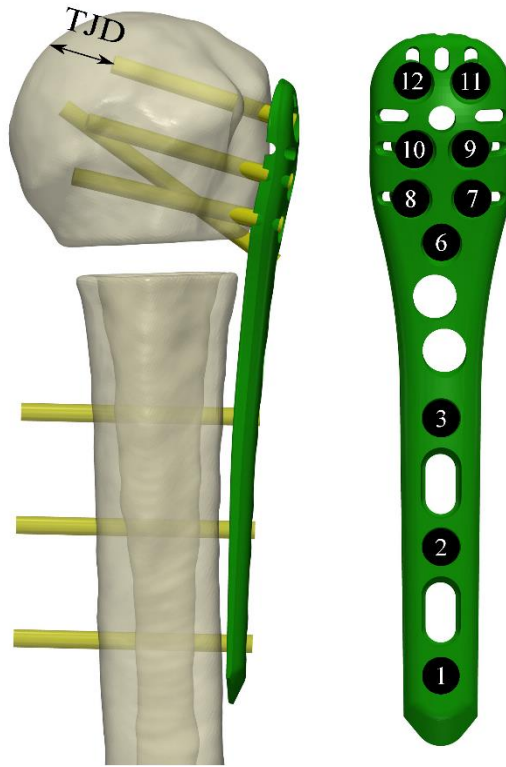
Figure 3.1 - Coronal view of the proximal humerus with the plate, showing the TJD (left). Numerations of the screws used for this study (right)

Bone mineral density (BMD) was calibrated within the CT image using the generic Eggermont approach (Eggermont *et al.* 2019). Elastic properties were determined by converting the local BMD value, ρ_{app}, of the CT image to elastic modulus, E, using the conversion equation provided by Morgan et al (Morgan *et al.* 2003).

$$E(MPa) = 6850\rho_{app}^{1.49}$$

The mean value of BMD in the humeral head had a value of 66.39 mgHA/cm³ (Kamer *et al.* 2016; Krappinger *et al.* 2012), corresponding to a subject with low density, as has been used in other studies (Fletcher *et al.* 2019c; Mischler *et al.* 2020b). The plate and the screws were titanium with a Young's modulus of 105 GPa and a Poisson's ratio of 0.3. A mesh of linear tetrahedral elements (C3D4) was generated with Synopsys' Simpleware™ FE module. A convergence study was performed to assess the appropriate size of the elements, resulting in the use of an element edge length between 0.5mm and 1.0 mm. Details of the convergence study

70

can be found in the Supplementary Material (Appendix A: Supplementary material for Chapter 3). Tied contacts were defined at the bone-screw and screw-plate interface. Loading conditions were defined according to Röderer's experimental study, reproducing an axial bending scenario (Röderer *et al.* 2013). Specifically, the bone and implant geometry were rotated by 25° around the anterior-posterior axis of the humerus, in agreement with Bergmann's clinical study (Bergmann *et al.* 2007). The nodes of the distal portion of the humerus were linked to a node positioned external to the geometry, corresponding to the midpoint between the condyles of the distal humerus, which was fixed in all directions. On the humeral head, the nodes of a circular region with a diameter of 20 mm were linked to a point externally located at 1 mm distance from the surface of the humerus, along the axis connecting the centre of the humeral head to the centre of the circular region. A vertically oriented force of 100N was applied to this external point, which was constrained only to move in the vertical direction. No rotations were allowed
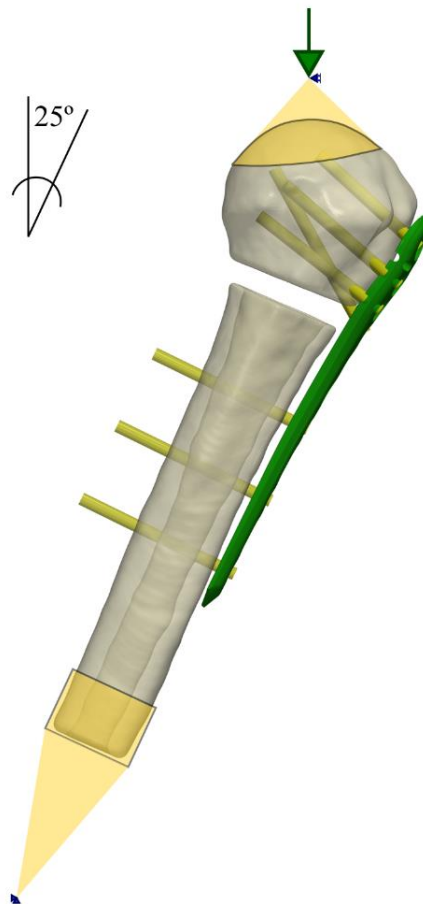
Figure 3.2 - Loading and boundary conditions

(Figure 3.2). These loading and boundary conditions were similarly defined from several studies from the AO foundation team, which were considered high quality and indeed a good comparable source.

For the analysis of the finite element models, ABAQUS (Version 6.14-3, Dassault Systèmes, Vélizy-Villacoublay, France) standard implicit static solver was used. The FE analysis was run on a total number of 1980 simulations, each of which had different screw length configurations, as explained in more detail in section 3.2.2. The minimum principal strain at the contact nodes around the proximal screws was computed from the elements data and used as the output data. Indeed, it has been shown that the average principal strain around the proximal locking screws can be a surrogate parameter of implant failure and estimate construct stability (Varga *et al.* 2017).

In order to speed up the process of the FE analysis, a systematic generation of the FE models was developed using Matlab (Mathworks, Natick, MA, USA). Through a customised script, the software is able to automatically select the desired parameters, generate the models through Simpleware, run the FE analysis on Abaqus and post-process the FE simulation data (Figure 3.3).

3.2.2 ANN Model

The Neural Network toolbox in Matlab (Mathworks, Natick, MA, USA) was used for the generation of the surrogate models. Different architectures were analysed, and the one with the best performance was selected; a three layers cascade forward network, with two hidden layers with 10 and 5 neurons and one linear output layer. The Levenberg Marquart backpropagation algorithm was used to train the ANNs (Figure 3.4).

The ANN was trained using 7 input variables, the TJD of each screw in the humeral head (screws 6-12 in Figure 3.1), resulting in a maximum number of possible combinations of 4^7, for a total of 16,384 configurations. Training of the ANN was performed using reduced sample sizes.

Latin Hypercube sampling method was used to generate training sets of 50, 100, 200, 500, and

1000 models, in order to assess the influence of training set size on the ANN performance. An
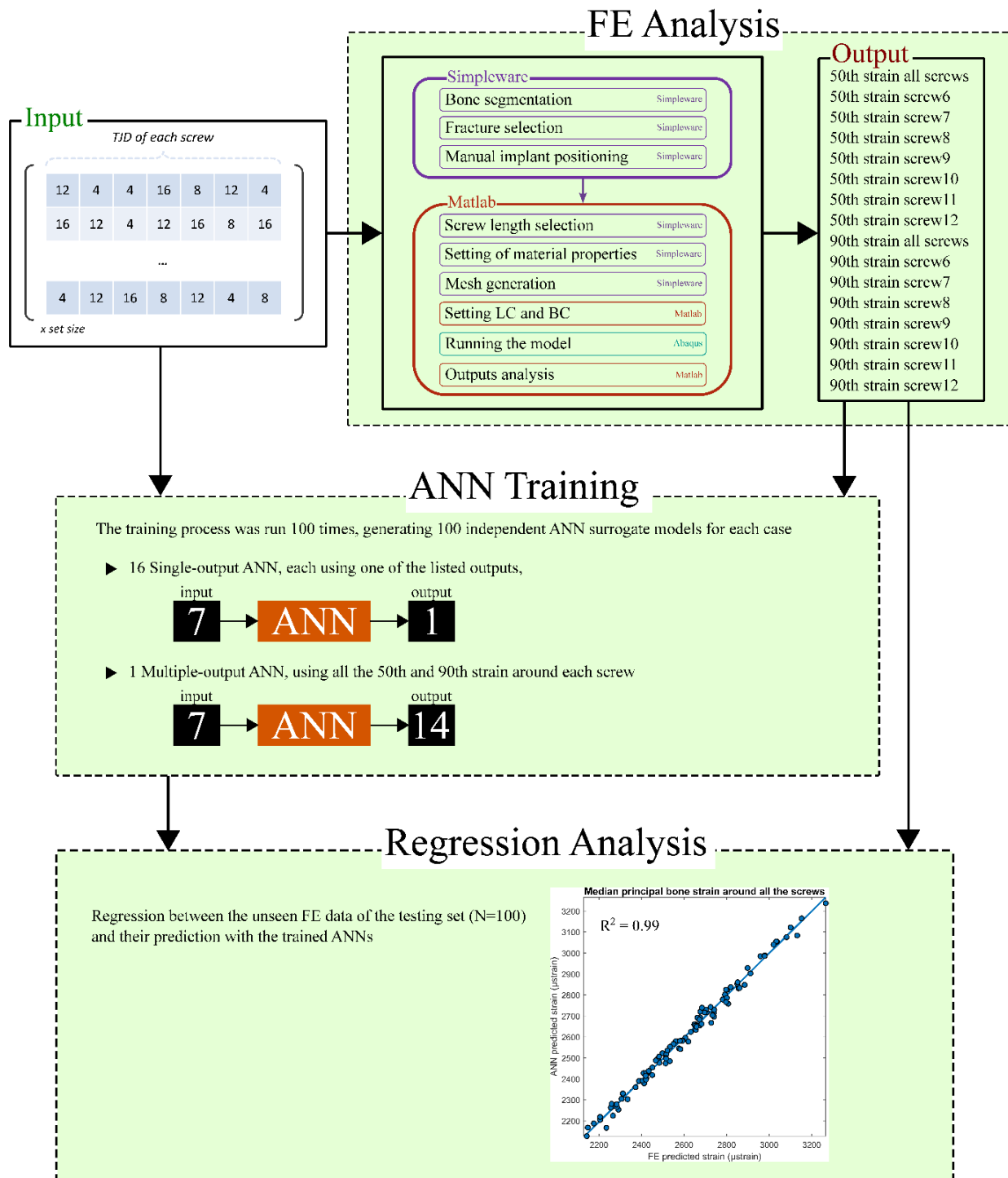
Figure 3.3 - Flow chart showing steps for the surrogate modelling, from the FE modelling, ANN training and regression analysis.

additional testing set of 100 samples was generated to test the quality of the predictions generated by the ANNs.

16 single-output ANNs were generated, to predict the 50th and the 90th percentiles of the principal strain around each screw and around all the screws. One multiple-output ANN was also generated, to predict the 50th and the 90th percentiles of the principal strain around each screw (Figure 3.3).
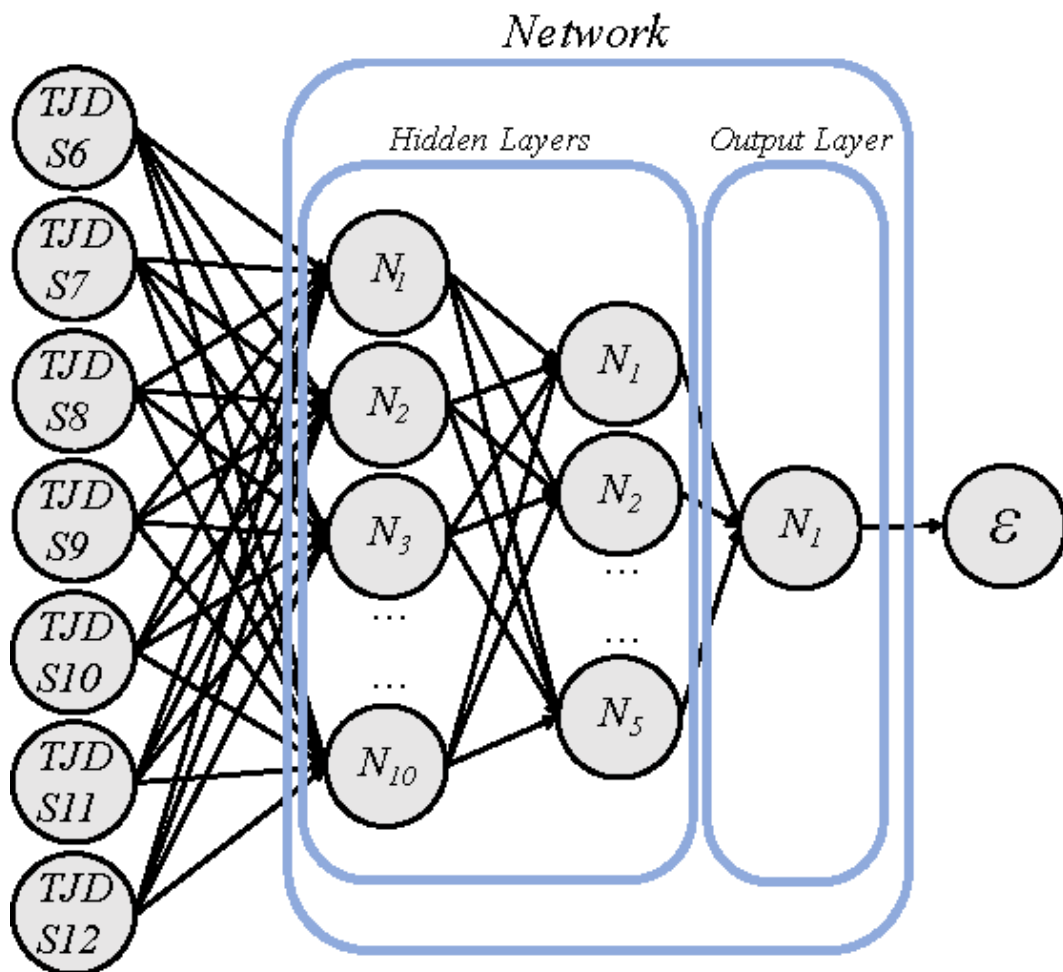
Figure 3.4 - Illustration of the Adaptive Neural Network

3.2.3   Assessment of surrogate model

During the generation of each ANN, 80% of the inputs were randomly selected for the training of the network, 10% for the verification of the training process and the other 10% were used

independently to assess the performance of the network. In order to assess the robustness of the networks, this process was performed 100 times, so 100 ANN models were trained and then used to replicate the results of the testing sets of 100 samples previously generated (Taylor *et al.* 2017). In order to assess the quality of the networks, the coefficient of determination (R^2), regression slope and root mean square error (RMSE) were reported, showing the result for the best ANN model and a mean of all the 100 ANN models. This process was done for all the ANNs trained using the training set of different sizes, and their results were compared to each other to define the best sample size (Figure 3.3).

For the multiple-output ANN, 100 networks were trained, and their outputs were averaged and then compared to the results from the testing set (Hashem 1997).

A further approach was used to test the accuracy of the ANN. An additional testing set of 30 FE models was developed, using tip-to-joint distance (TJD) of 6, 10 and 14 mm. The aim of this testing was to verify that the ANN can make predictions of configurations with values of inputs that are not being used in the original training process. The predictions of the ANN were compared with the output values of these FE models with a regression analysis.

3.2.4   Analysis of ANN predictions

As the last step, the single-output ANN models with the lowest error were used to make a prediction of all the 4^7 possible configurations, in order to identify the influence of the variation of screw length in the humeral head.

Moreover, once the single-output ANN trained to predict the principal bone strain around all the screws was developed, it was used to replicate a study from the literature (Fletcher *et al.* 2019a). Specifically, predictions of case studies where all the screws had the same TJD were produced, respectively with measures of 4, 8, 12 and 16 mm. Although a different implant was used, the aim was to capture the negative impact of shorter proximal screws on the bone minimal principal strain around the screws.

Analysis of variance (ANOVA) of the ANN model predictions was conducted in order to analyse the influence of each input on the output and identify any possible interactions between them. The ANOVA test was conducted on the ANN model predictions with a single output, for a total of 16 tests. In particular, the ANN models with minimum RMSE value were used. The ANOVA test generates the sum of squares for each input and each interaction term, which was then expressed as a percentage of the total sum of squares (%TSS) to determine the influence of each term.

3.3 Results

A total of 1,980 FE simulations were run and each one required a time of 15 minutes, from the generation of the mesh to post-processing the results. In Figure 3.5, several configurations are displayed as examples illustrating the results from the FE analysis (Figure 3.5).
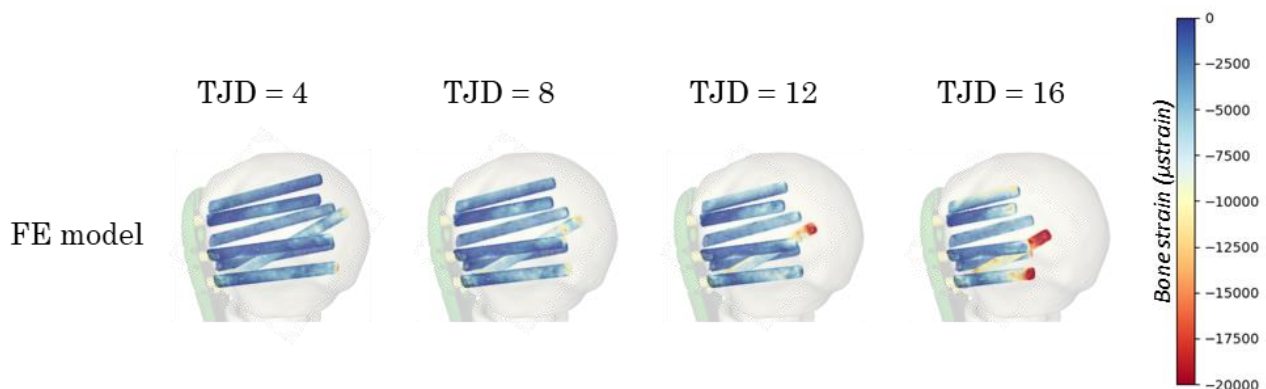
Figure 3.5 - Minimal Principal Strain distribution obtained from the FE analysis for the configurations having TJD of 4, 8, 12 and 16 mm

The accuracy of the ANN depends on the number of data sets used during its development. The ANNs were trained using a training set of 50, 100, 200, 500, and 1000 samples. These models were then utilized to predict the bone minimal principal strain of 100 configurations that were not included in the training set. The influence of training set size was performed only on the two single-output ANNs, one trained for the prediction of the 50th percentile of the minimal principal bone strain around all the screws and the second one trained for the prediction of the

90th percentile of minimal principal bone strain around all the screws. As the size of the dataset increased, the R^2 increased from 0.82 to 0.99 and RMSE dropped from 114.31 to 21.14 µstrain for the first ANN when increasing from 50 to 1000 training sets. The same behaviour was recorded for the second ANN, as the R^2 rose from 0.92 to 0.99 and the RMSE decreased from 466.14 to 62.78 µstrain (Table 3.1).

Table 3.1 - Performance of the single-output ANN on the testing set of 100 simulations for the prediction of minimal principal strain of the bone around all the screws. The influence of the training set size is shown. Results are displayed for the model with the lowest error, while the average of 100 models is shown in brackets.

Training set	50th percentile min principal strain				90th percentile min principal strain			
	R^2	Slope	RMSE, µ strain	RMSE, % strain	R^2	Slope	RMSE, µ strain	RMSE, % strain
50	0.827	0.985	114.321	3.501	0.923	0.924	466.149	4.208
	(0.612)	(0.912)	(191.280)	(5.858)	(0.772)	(0.852)	(811.366)	(7.323)
100	0.902	0.925	81.595	2.499	0.968	0.988	298.314	2.693
	(0.783)	(0.956)	(129.859)	(3.997)	(0.895)	(0.947)	(551.441)	(4.997)
200	0.944	0.954	61.053	1.870	0.992	0.970	149.858	1.353
	(0.866)	(0.901)	(93.040)	(2.850)	(0.977)	(0.973)	(245.956)	(2.220)
500	0.991	1.006	24.242	0.742	0.998	0.991	76.358	0.689
	(0.974)	(0.981)	(39.243)	(1.202)	(0.996)	(0.992)	(107.761)	(0.973)
1000	0.993	1.005	21.146	0.648	0.999	0.999	62.786	0.567
	(0.986)	(0.955)	(29.201)	(0.894)	(0.996)	(0.997)	(98.562)	(0.890)

Since the RMSE and the R^2 values of the ANN trained with 500 and 1000 data were comparable in values, all the subsequent presented ANN results were trained with the dataset of 500. An additional seven single-output ANNs were developed using the single value of strain around each screw as output. A strong correlation is shown in Table 3.2, and the error varied from 24.63 µstrain to 148.11 µstrain (Table 3.2).

A multiple-output ANN was trained 100 times, the results were averaged and compared with the results from the testing set. A high correlation was shown, with values of error higher but similar to the results from the single-output ANNs (Table 3.3). For simplicity, all the further analysis were made on single-output ANNs. In order to capture if the ANN was accurate in the prediction of the strain of models with additional values of TJD, once trained, the best ANN was used to predict the strain values from the 30 simulations with additional TJD values of 6, 10

and 14 mm. Comparing the output from the ANN with the FE models, a high correlation was detected (R^2 = 0.96-0.99) but with a significantly higher level of error (RMSE = 28.75 - 1190.72 μstrain) (Table 3.4, Table 3.5).

Table 3.2 - Performance of single-output ANNs on the testing set of 100 simulations for the prediction of bone strain around each single screw. Results are displayed for the model with the lowest error, while the average of 100 models is shown in brackets.

	50th min principal strain				*90th min principal strain*			
	R^2	Slope	RMSE, μ strain	RMSE, % strain	R^2	Slope	RMSE, μ strain	RMSE, % strain
Screw6	0.999	0.996	56.428	0.512	0.998	0.995	148.110	0.746
	(0.999)	(0.998)	(68.536)	(0.622)	(0.997)	(0.996)	(196.140)	(0.998)
Screw7	0.995	0.999	41.178	0.972	0.996	0.992	134.834	0.951
	(0.992)	(0.995)	(50.752)	(1.198)	(0.990)	(0.993)	(194.662)	(1.373)
Screw8	0.991	0.999	30.691	1.202	0.996	0.994	61.120	0.973
	(0.976)	(1.003)	(46.777)	(1.832)	(0.989)	(0.979)	(94.161)	(1.500)
Screw9	0.989	0.980	24.636	1.099	0.991	1.002	79.765	1.194
	(0.982)	(0.986)	(31.443)	(1.402)	(0.980)	(0.989)	(118.369)	(1.772)
Screw10	0.994	0.998	33.908	0.936	0.995	0.976	67.733	0.958
	(0.987)	(0.990)	(46.623)	(1.288)	(0.990)	(0.972)	(94.312)	(1.334)
Screw11	0.989	0.987	26.398	1.270	0.995	0.998	63.335	1.029
	(0.981)	(0.975)	(34.430)	(1.656)	(0.989)	(1.003)	(89.686)	(1.458)
Screw12	0.996	0.995	49.860	1.100	0.998	0.984	77.188	0.880
	(0.995)	(0.989)	(58.065)	(1.327)	(0.996)	(0.981)	(97.685)	(1.113)

Table 3.3 - Performance of the multiple-outputs ANN on the testing set of 100 simulations for the prediction of bone strain around each single screw

	50th min principal strain				*90th min principal strain*			
	R^2	Slope	RMSE, μ strain	RMSE, % strain	R^2	Slope	RMSE, μ strain	RMSE, % strain
Screw6	0.999	0.997	58.548	0.532	0.996	0.991	122.889	0.619
Screw7	0.994	0.997	42.014	0.992	0.995	0.971	72.072	0.508
Screw8	0.964	0.991	63.079	2.471	0.987	0.985	94.519	1.505
Screw9	0.986	0.981	28.172	1.257	0.993	0.968	82.310	1.232
Screw10	0.973	0.962	72.233	1.995	0.987	0.992	97.854	1.384
Screw11	0.981	0.894	40.536	1.950	0.997	0.980	87.996	1.430
Screw12	0.996	0.988	54.814	1.236	0.996	0.991	122.889	1.400

Table 3.4 - Performance of single-output ANNs on the testing set of 30 simulations with intermediate values of TJD for the prediction of bone strain around each single screw. Results are displayed for the model with the lowest error, while the average of 100 models is shown in brackets

	50th min principal strain				**90th min principal strain**			
	R²	Slope	RMSE, µ strain	RMSE, % strain	R²	Slope	RMSE, µ strain	RMSE, % strain
Screw6	0.986 (0.961)	0.936 (0.888)	289.135 (467.249)	2.625 (4.242)	0.926 (0.909)	0.818 (0.794)	1190.72 0 (1306.21)	5.996 (6.577)
Screw7	0.951 (0.943)	0.838 (0.831)	156.839 (169.441)	3.703 (4.000)	0.950 (0.907)	0.952 (0.902)	360.700 (500.819)	2.544 (3.533)
Screw8	0.988 (0.871)	1.029 (0.925)	35.817 (113.457)	1.403 (4.444)	0.952 (0.920)	0.953 (0.901)	176.880 (237.332)	2.817 (3.780)
Screw9	0.971 (0.942)	0.930 (0.870)	38.488 (55.388)	1.717 (2.470)	0.970 (0.931)	0.927 (0.909)	133.690 (197.194)	2.001 (2.952)
Screw10	0.994 (0.978)	0.983 (0.986)	28.757 (52.358)	0.794 (1.446)	0.984 (0.976)	0.969 (0.960)	103.655 (125.971)	1.466 (1.782)
Screw11	0.971 (0.923)	0.981 (0.944)	39.292 (64.391)	1.890 (3.097)	0.982 (0.970)	0.994 (0.973)	107.441 (139.388)	1.746 (2.265)
Screw12	0.990 (0.985)	1.016 (1.019)	71.948 (89.305)	1.622 (2.013)	0.994 (0.992)	0.995 (1.009)	104.948 (130.335)	1.196 (1.485)

Table 3.5 - Performance of the single-output ANN on the testing set of 30 simulations with intermediate values of TJD for the prediction of bone strain around all the screws. Results are displayed for the model with the lowest error, while the average of 100 models is shown in brackets

	50th min principal strain				**90th min principal strain**			
	R²	Slope	RMSE, µ strain	RMSE, % strain	R²	Slope	RMSE, µ strain	RMSE, % strain
All screws	0.959 (0.929)	0.961 (0.943)	49.815 (66.923)	1.526 (2.050)	0.959 (0.945)	0.868 (0.858)	364.228 (413.789)	3.288 (3.735)

The trained single-output ANNs were used to make predictions of principal bone strain around all the screws and each screw for all the possible 4^7 configurations. The ANN models took less than one minute of running time for the generations of the outputs of these configurations, significantly less than what an FE model would require. For the ANOVA of each single-output ANN model, the %TSS of each input was used to understand how the different screw lengths affect the screw strains. The interaction values between the input parameters were not reported since their %TSS was approximately zero. Variation in the length of screw 6 was found to be

influential, for both the 50th and 90th bone strain around all screws, as well as strain around

screws 6, screw 7, and screw 9 (Table 3.6, Table 3.7).

Table 3.6 - % of Total Sum Square – models of 50th percentile principal strain around the screws

50th percentile principal strain	All screws	Screw 6	Screw 7	Screw 8	Screw 9	Screw 10	Screw 11	Screw 12
Screw6	35.240	99.545	93.692	32.780	93.925	0.084	1.211	0.012
Screw7	0.378	0.231	4.769	0.568	0.054	0.014	0.430	0.158
Screw8	11.562	0.014	0.100	50.841	0.024	0.030	0.025	0.002
Screw9	0.019	0.004	0.138	0.064	3.572	0.026	1.770	0.026
Screw10	18.584	0.020	0.000	2.884	0.049	86.628	0.178	1.595
Screw11	6.953	0.017	0.005	0.015	0.463	0.004	83.163	0.019
Screw12	23.450	0.051	0.352	0.343	0.278	9.725	11.440	96.863

Table 3.7 - % of Total Sum Square – models of 90th percentile principal strain around the screws

90th percentile principal strain	All screws	Screw 6	Screw 7	Screw 8	Screw 9	Screw 10	Screw 11	Screw 12
Screw6	91.685	98.309	50.495	49.174	82.171	9.997	1.295	1.751
Screw7	2.351	0.482	42.951	1.167	0.218	0.009	0.951	0.375
Screw8	0.406	0.068	0.029	44.062	0.018	0.004	0.017	0.017
Screw9	0.070	0.044	0.276	0.046	12.317	0.019	1.377	0.048
Screw10	0.531	0.020	0.038	0.290	0.020	75.996	0.103	0.738
Screw11	0.433	0.001	0.001	0.009	0.018	0.005	91.321	0.099
Screw12	2.907	0.001	0.086	0.142	0.224	8.395	3.871	96.175

The results are shown by plotting the variation of the mean of the output values analysed with

the variation of screw length (Figure 3.7, Figure 3.6). The ANOVA test showed how screw 6 has

a major impact on the variation of the strain around the screws, especially on the 90th percentile

strain.

In order to give a graphical representation of the results with the variations of the TJD of each screw, the Sum of TJD was introduced and used to plot the variation of strain for each simulation (Figure 3.8, Figure 3.9).

Figure 3.8 shows the results of the 50th and 90th percentile principal bone strain around all screws for all 16,384 possible configurations. More specifically, the four simulations in which the TJD of each screw was fixed by the same size, respectively 4, 8, 12 and 16 mm, were highlighted. The results showed that the increase in TJD of all screws simultaneously was associated with a decrease in strain around all screws (Figure 3.8)(Fletcher *et al.* 2019a). Since the ANOVA test showed that screw 6 has the biggest impact on the variation of the 50th and 90th strain, a similar plot shows the variation of strain with the variation of the sum of TJD for each simulation, highlighting the value of TJD of screw 6 (Figure 3.9). That showed again how an increase of TJD for screw 6 caused an increase of bone strain around all the screws (Figure 3.9).
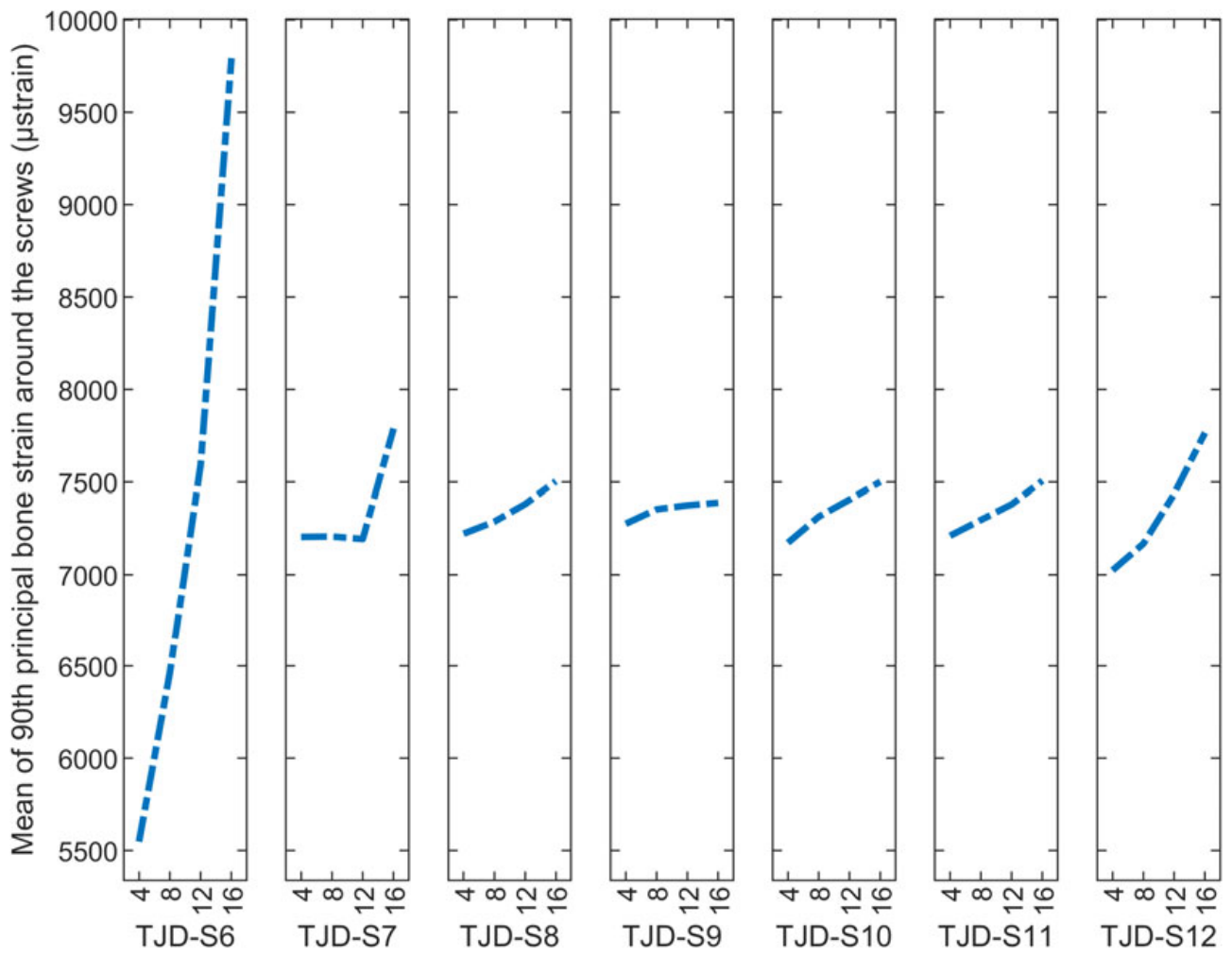
Figure 3.6 - Mean variation of the 90th principal strain around all the screws with the variation of the length of each screw
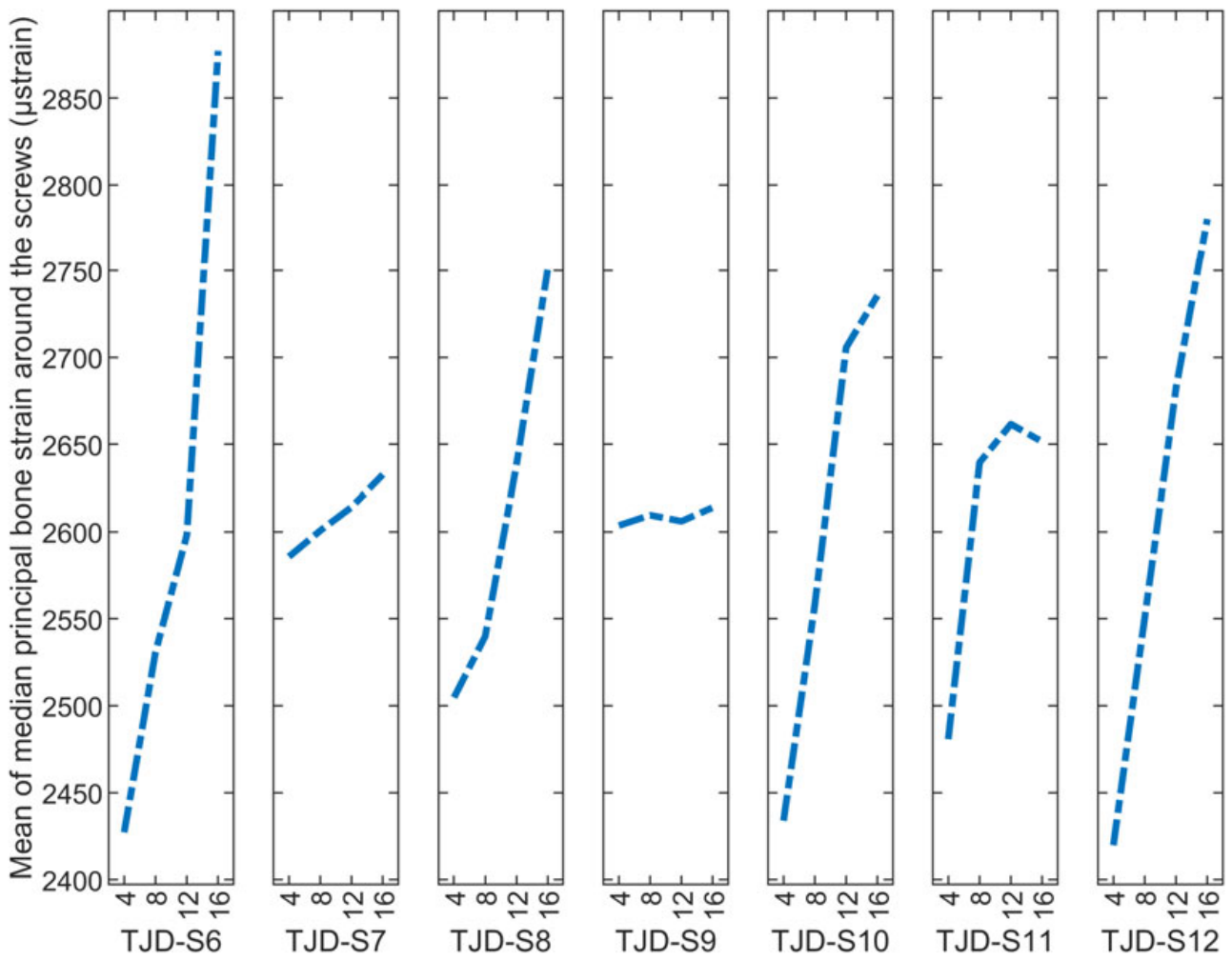
Figure 3.7 - Mean variation of the 50th percentile principal strain around all the screws with the variation of the length of each screw
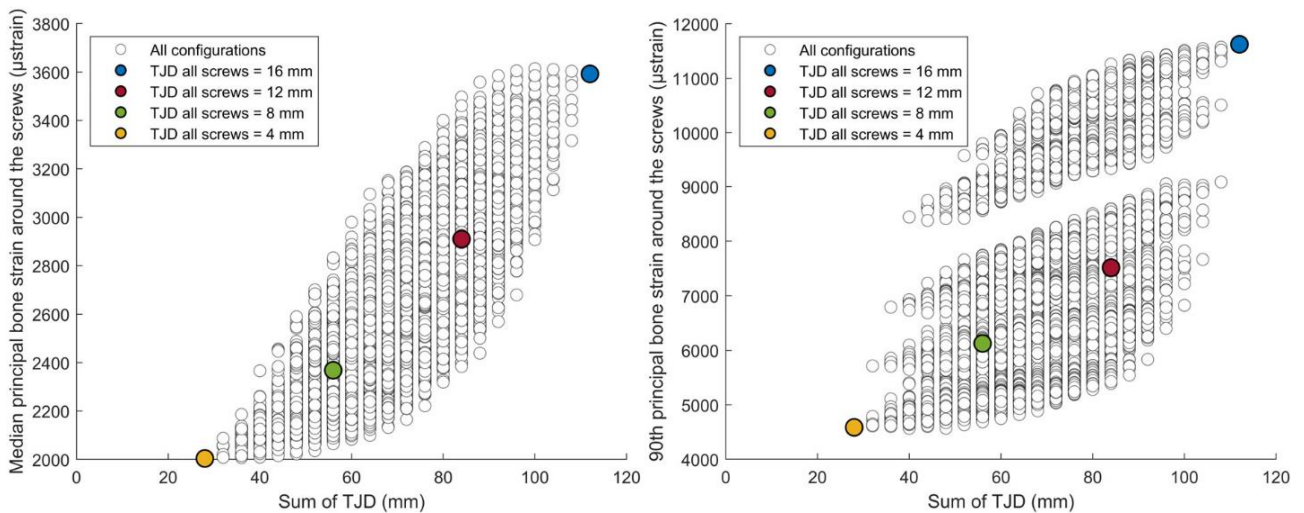
Figure 3.8 - Variation of 50th (left) and 90th percentile (right) of principal bone strain for all configurations. Focus on the configurations in which all the screws have the same TJD
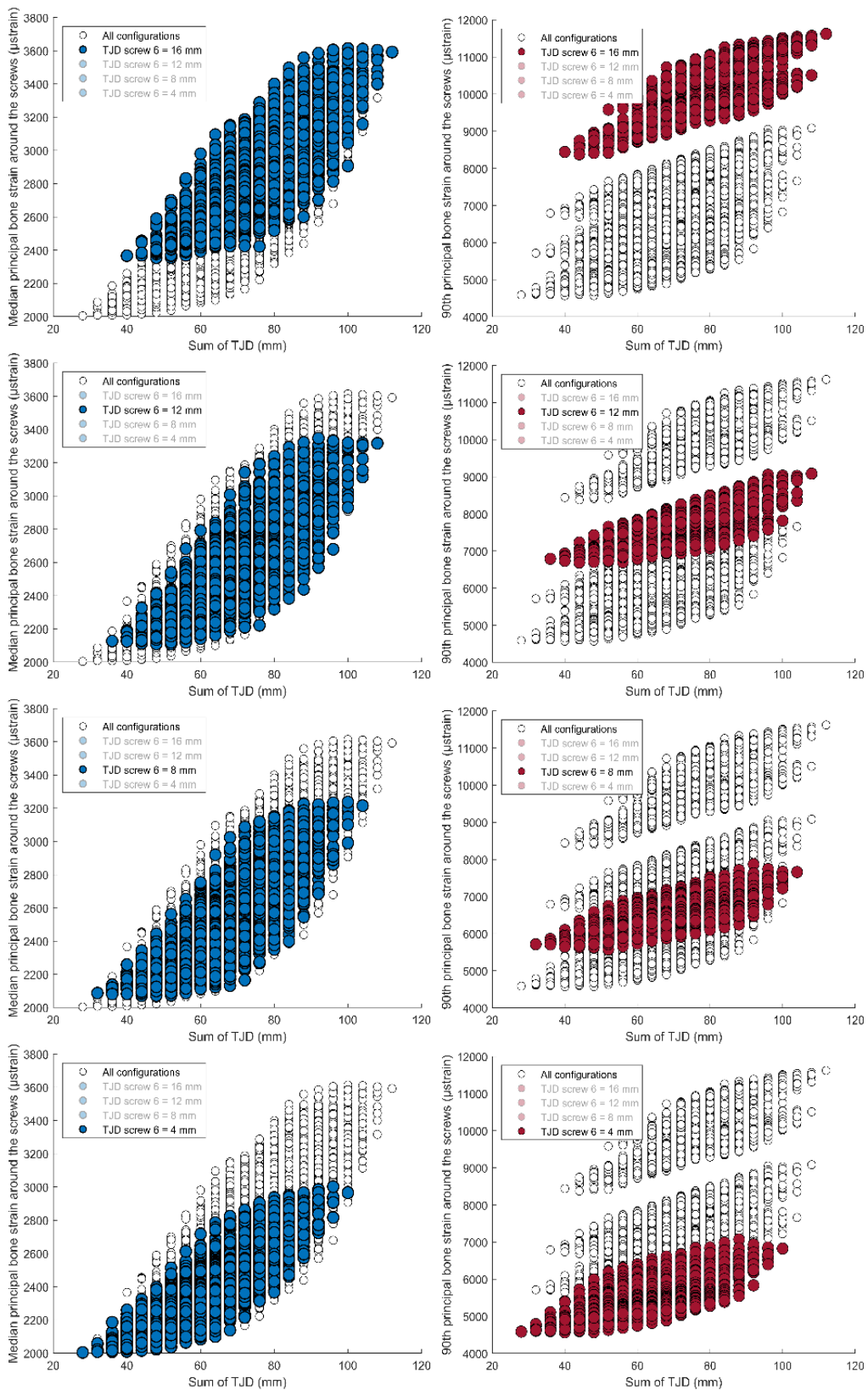
Figure 3.9 - Variation of the 50th percentile principal strain (blue) and the 90th percentile principal strain (red) of the bone for each length of screw 6 with the variation of the sum of the TJD of all the screws

3.4 Discussion

The motivation behind this study is that FE models can give an accurate prediction, but they are time consuming when the aim is to analyse the impact of a large number of parameters, making it necessary to run a high number of models. It has been shown that the problem of the fracture fixation device for a proximal humeral fracture is challenging because of the high number of variables and assessing the effect of all of them together through an FE analysis would be too computationally expensive. In this study, we developed a Neural Network model, which is a computationally cheap and reliable technique, to reproduce the strain around the screws with varying lengths, defined as the TJD of each proximal screw.

Firstly, the ANN model was shown to be sensitive to the training set size and acceptable results were obtained with a N=500 training set. Indeed, using a training set with 500 simulations resulted in high R^2 and low RMSE values when comparing the ANN predictions with FE results for 100 unseen cases (R^2=0.96-0.99, RMSE=24.6-148.11 µstrain). Moreover, the ANN was trained 100 times on the same data, and the comparison between the best and mean RMSE of the ANN predictions showed that the model was stable and not sensitive to the training data.

The ANN models with the lowest RMSE value were used to make predictions of bone strain around all the screws and bone strain around every single screw for 30 additional FE models developed having additional values of TJD (6, 10 and 14 mm). The results showed a high level of correlation (R^2=0.96-0.99) but a higher level of error (RMSE=28.7-1190.72 µstrain), suggesting that in order to make more accurate predictions, it is important to use the same input values used for the training process (Table 3.4, Table 3.5).

Examination of the strain values for individual screws, developing a single-output ANN and using the one over the 100 ANN models trained with minimum RMSE was shown to be equivalent to developing a multiple-output ANN and averaging their results over the ANN 100 models trained (Table 3.2, Table 3.3).

Once the accuracy of the networks was demonstrated, the network was used to make a prediction of all 4^7 configurations, with a running time of a few seconds. In terms of computational cost, making the predictions for all possible screw configurations with the ANN models was shown to be significantly faster than developing FE models for all the configurations. Indeed, it took around 20 minutes to develop, run and extrapolate the output from a single FE model. Furthermore, the training time for a single ANN was about one minute, and afterwards only took a few seconds to make new predictions. It would not be feasible to run 4^7 FE simulations, but this study showed it was possible to run a few FE models in order to train and develop an ANN model to use for the predictions of all the configurations.

Some of the 4^7 predictions obtained from the ANN model were compared to similar screw configurations reported in the literature. Firstly, analysing the values of strain around all the screws for the configurations with the same value of TJD for all the screws, the results showed that lower values of strain are obtained with longer screws, confirming the results of a previous FE study (Figure 3.8) (Fletcher *et al.* 2019a).

The advantage of running all the configurations using the ANN models made it possible to analyse the impact of each screw on the overall strain distribution and on the strain around each individual screw for a single subject. Overall, the results showed that the calcar screw (screw 6) had the greatest influence, confirming the importance of carefully choosing the appropriate length for this screw (Fletcher *et al.* 2019a; Gardner *et al.* 2007). In particular, the increase of TJD for screw 6 impacts significantly the strain around all the screws, and the same behaviour was reported in Fletcher's study for the calcar screws (Fletcher *et al.* 2019a). There was a marked increase in the strain when the TJD increased to 12 and 16 mm. Moreover, the results from the ANOVA test demonstrated how the variation of the length of screw 6 influenced the strain around screws 7, 8 and 9 (Table 3.6, Table 3.7).

There are several limitations in this study. Firstly, we developed the ANN from the data of only one subject. Although other studies found in the literature investigated the variations of implant

parameters on a single subject (Jabran *et al.* 2019b; Tilton *et al.* 2020b), for completeness of the study this methodology should be expanded to a group of subjects. We used non-homogeneous material properties, however, the equation used to describe the relation between Young's Modulus and BMD was specific to the femoral head (Morgan *et al.* 2003), not the proximal humerus. The loading and boundary conditions of the FE model were simplified, reproducing a simple load case with bonded conditions at the interface between the bone and the implant. Previous studies have shown no differences in results between tied and threaded conditions in terms of bone strain values around the screws (Inzana *et al.* 2016). No experimental analysis was conducted for the validation of the FE models. However, the value of strain around the screw was demonstrated to be a surrogate parameter of cut-out failure (Varga *et al.* 2017). Moreover, the range of values of 50th percentile principal strain around all the screws has shown to have comparable values to other FE studies conducted for the fracture fixation plates for a proximal humeral fracture (Mischler *et al.* 2020b; Varga *et al.* 2017). Lastly, only a single variation for the implant was made, the orientation of the screws was assumed to be fixed and the fracture configuration was simplified.

Due to the complexity of the problem and the high computational cost required to conduct an FE Analysis, the aim of this study was to demonstrate the use of advanced computational techniques for predicting the behaviour of a fracture fixation plate used in treating proximal humeral fractures in a simplified scenario. Giving the promising results of the ANN methodology and the computational efficiency demonstrated in this study, this gives us confidence that the complexity could be increased, by introducing further variables, such as the orientation or the configurations of the screws and including patient variability. Based on our experience, we would recommend a staged approach to increasing the complexity, in order to understand the impact this has performance of the surrogate model and the information required to give accurate predictions.

3.5 Conclusion

This study demonstrated that an ANN model can accurately reproduce the value of strain in the bone around the screws just with geometric information from the implant, using a multiple and a single output network. The advantage of this technique is that was possible to use the network to make a further and accurate analysis of all the possible configurations in a time efficient process. ANN models showed to be computationally inexpensive for new predictions, making it a possible tool for future applications, expanding it to assess the influence of more implant and bone parameters, such as screw orientation and patient variability.

Chapter 4

Assessing the influence of screw orientation on fracture fixation of the proximal humerus using finite element informed surrogate modelling[2]

4.1 Introduction

Fractures of the proximal humerus are one of the most frequent fractures in older subjects (Sporer *et al.* 2006), accounting for around 10% of all fractures, with a higher incidence for women >65 years old and predicted to increase with the increase of the older population (McLean *et al.* 2019; Palvanen *et al.* 2006). Up to 30% of these fractures are treated surgically

---

[2]The study presented in this chapter is the subject of the following paper:

Mini D, Reynolds KJ, Taylor M. (2024) *Assessing the influence of screw orientation on fracture fixation of the proximal humerus using finite element informed surrogate modelling*, Submitted to the International Journal for Numerical Methods in Biomedical Engineering

in Australia and the most common surgical procedure is the use of a fracture fixation plate (McLean *et al.* 2019). Fracture healing is not always achieved, with a reported failure rate of up to 35% (Kralinger *et al.* 2014). The design of fracture fixation plates has changed in the past decades intending to decrease the failure rate, and recently fracture fixation plates incorporating variable angle screws have been introduced (Cronier *et al.* 2010). In vitro testing is the gold standard for evaluating implant fixation biomechanics (Jabran *et al.* 2018), but it becomes impractical due to the high number of parameters to investigate. On the other hand, Finite Element (FE) techniques have been used to explore the impact of different variables on the bone strain and implant stress for a fracture fixation plate (Lewis *et al.* 2021), such as the position of the plate (Fletcher *et al.* 2019b), screw length (Fletcher *et al.* 2019a), bone quality (Fletcher *et al.* 2019c; Tilton *et al.* 2020b), the number of screws in the head of the humerus (Fletcher *et al.* 2019c; Tilton *et al.* 2020) and their orientation in space (Jabran *et al.* 2019b; Mischler *et al.* 2020b, 2022; Schader *et al.* 2021). However, these studies have been conducted on a limited number of configurations (Fletcher *et al.* 2019a, 2019c, 2019b; Jabran *et al.* 2019b; Mischler *et al.* 2020b; Schader *et al.* 2021; Tilton *et al.* 2020b; Varga *et al.* 2018, 2020), with simulation numbers ranging from a few hundred (Fletcher *et al.* 2019c, 2019b; Jabran *et al.* 2019b; Schader *et al.* 2021; Tilton *et al.* 2020b; Varga *et al.* 2018) to a few thousand (Fletcher *et al.* 2019a; Mischler *et al.* 2020b; Varga *et al.* 2020).

Only a few studies have investigated the impact of screw orientation on predicting mechanical failure of the locking plate fixation (Jabran *et al.* 2019b; Mischler *et al.* 2020b, 2022; Schader *et al.* 2021). Jabran study focused only on the variation of the orientation of two screws on a single subject, generating 538 configurations (Jabran *et al.* 2019b), and Mischler and Schader's studies were both conducted on a group of 19 subjects but varying the screw orientation of the 6 proximal screws only a one at a time, for a total of 88 configurations per sample (Mischler *et al.* 2020b; Schader *et al.* 2021). Additionally, in a more recent study, the Mischler group conducted experimental validation confirming that modifying the direction of specific screws led to

improvement in preventing cut-out failure when compared to using the standard fracture fixation plate (Mischler *et al.* 2022). The Jabran study only focused on the angle variations of two screws, while the first Mischler and Schader studies explored a considerably reduced number of screw angle combinations. Indeed, once a large number of simulations are needed to be investigated, FE analysis becomes too time-expensive and a more efficient computational technique is needed. Surrogate models can be used to make estimations of a larger number of configurations in a quicker way. Surrogate models aim to make a prediction of the fitness function of a complex problem with the use of input and output data (Jin 2005), reducing the computational effort. Specifically for biomechanical applications, Kriging (O'Rourke *et al.* 2016; Takian *et al.* 2021), Adaptive Neural Network (ANN) (Mini *et al.* 2024; Taylor *et al.* 2017) and Gaussian process (GP) (Al-Dirini *et al.* 2020; Bah *et al.* 2011) based models are some examples of surrogate models that have been used with a combination of FE data to describe problems with hip and knee implants. ANN, which consists of a network of interconnected neurons that exchange information with one another, has been shown to be a precise and effective technique (Taylor *et al.* 2017). Recently, our research group successfully developed an FE-informed ANN model to accurately predict bone strain with varying screw lengths for a proximal humeral fracture with a fracture fixation plate. The results of this study have demonstrated a high level of accuracy and have shown great promise for the application of ANN models in this field (Mini *et al.* 2024). The aim of the current study is to develop an FE-informed ANN model to analyse the effect of screw orientations on bone strain for a fracture fixation plate implanted in a humeral head, reducing computational time and exploring a wide range of possible configurations. A hypothesis of this study is that the variation of the angle of the calcar screw has a higher impact than the other screws, as has been shown in other studies (Mischler *et al.* 2020b).

4.2 Methods

In order to investigate how screw orientation affects bone strain, various FE models were generated by changing the direction of the screws in the proximal-distal and anterior-posterior directions. The input and output data from these models were then used to train different ANN models and their accuracy was analysed.

4.2.1 FE model

For the generation of the FE analysis, one CT image of a cadaver of a 61-year-old female donor was collected from the New Mexico Decedent Image Database (NMDID) (Edgar *et al.* 2020). The right humerus was manually segmented using Simpleware™ software (Version U-2022.12; Synopsys, Inc., Mountain View, USA) and the shaft was cut at 160 mm length from the humeral head. A single cut with a 5 mm gap at the surgical neck was virtually performed, representing
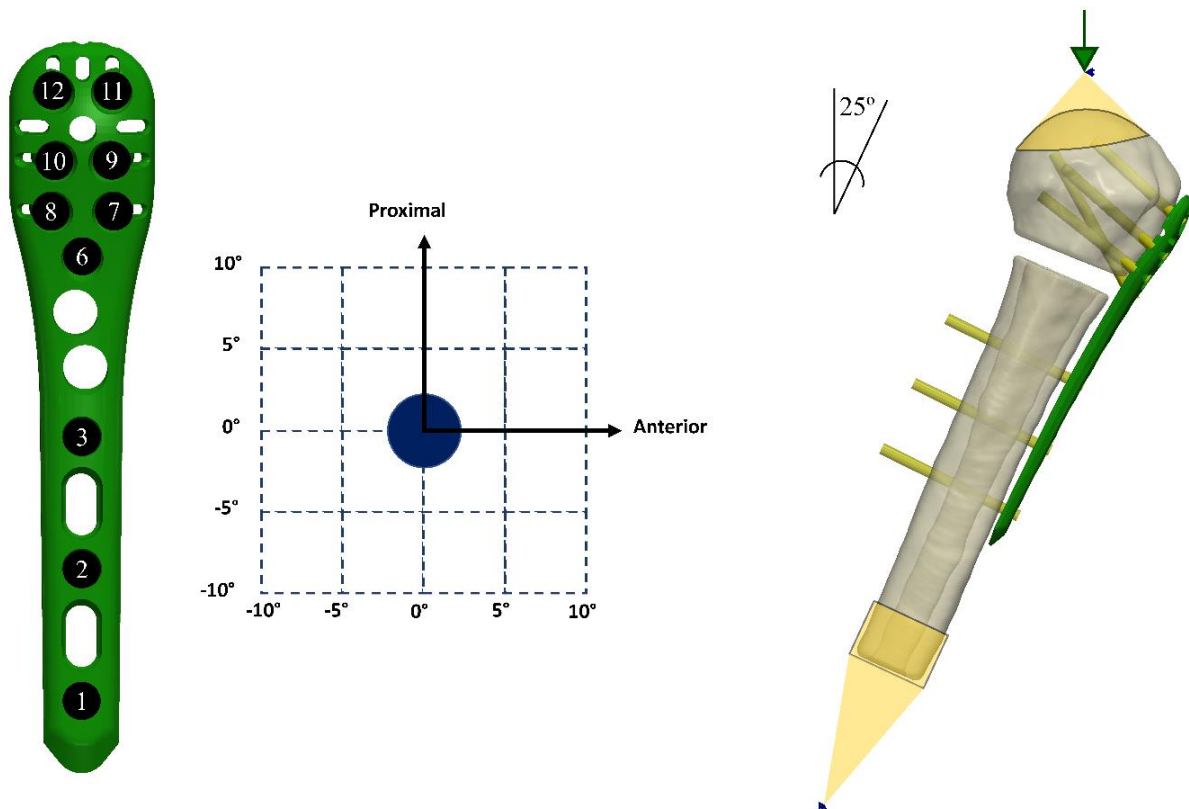
Figure 4.1 - Numerations of the screws used for this study (left). Possible orientation of the screw tips in distal-proximal and posterior-anterior direction (centre). The head of each screw was fixed in the neutral position (0°- 0°), and the tip of the screw was able to move from the neutral position with increments of 5° in the distal-proximal and posterior-anterior direction. Loading and boundary conditions (right) , image from (Mini et al. 2024)

a single two-part proximal humeral fracture AO/OTA 11-A2.1 (Foundation AO 2014). The bone fracture was virtually fixed using a fracture fixation plate (Austofix, Adelaide, Australia) (Figure 4.1). The plate was positioned 4 mm posterior to the bicipital groove and 7 mm distal to the top of the greater tubercle, and it was secured in place using seven proximal screws and three distal screws, each represented as a cylinder of 3 mm diameter. The orientation of the seven proximal screws was varied in the proximal-distal and anterior-posterior direction from the neutral position (0° - 0°) (Figure 4.1), which was defined as the standard configuration of the screws. The length of each of the screws was varied in order to have a fixed Tip to Joint (TJD) distance of 8 mm (Fletcher *et al.* 2019a), defined as the distance between the tip of the screws and the glenohumeral joint. The actual lengths of the screws varied based on the orientation of the screws. All materials were defined as linear elastic with a Poisson's ratio of 0.3. The material properties of the bone were defined as heterogenous, in which the bone mineral density (BMD) was derived from the CT image using the phantomless calibration methodology proposed by Eggermont et al. (Eggermont *et al.* 2019). The elastic modulus was converted from the local BMD using the Morgan et al. (Morgan *et al.* 2003) equation: $E(MPa) = 6850\rho_{app}^{1.49}$. The mean value of BMD in the humeral head had a value of 66.39 mgHA/cm^3, comparable to the range values found in the literature (Kamer *et al.* 2016; Krappinger *et al.* 2012). The screws and plate were defined as titanium alloy, with a Young's modulus of 105 GPa. Synopsys' Simpleware™ FE module was used to generate a mesh of the model with linear tetrahedral elements (C3D4), having an element edge length between 1 and 0.5 mm. Bone-screw and screw-plate interfaces were set as tied contact. The models were subjected to a vertical force, reproducing the axial bending scenario according to Röderer et al experimental study (Röderer *et al.* 2013). In agreement with Bergmann's clinical study (Bergmann *et al.* 2007), the model geometry was rotated by 25° around the anterior-posterior axis of the humerus. An external point that represented the midpoint between the condyles of the distal humerus was fixed in all directions and connected to the nodes of the distal portion of the humerus. A second external

point, located at 1 mm distance from the surface of the humeral head along the axis connecting the centre of the humeral head to the centre of the circular region, was linked to a circular region with a diameter of 20 mm located on the humeral head. A vertical force of 100N was applied on the second external point, free to move only in a vertical direction (Figure 4.1) and no rotations were allowed. Moreover, these loading and boundary conditions were similarly defined from several studies from the AO foundation team, which were considered high quality and indeed a good comparable source. The minimum principal strain of the bone was evaluated around the surface proximal screws, as this has been shown to be a surrogate parameter of failure (Varga *et al.* 2017). Indeed, from Varga's study it had been shown that the minimal principal strain around the screws can be used as a surrogate parameter to predict cut-out failure of the implant. A standard implicit static analysis of the models was run in ABAQUS (Version 6.14-3, Dassault Systèmes, Vélizy-Villacoublay, France). An automated workflow was developed using Matlab, as a way of speeding up the FE model generation, solution and post-processing pipeline.

The process included the generation of the models with the meshing process in Simpleware, setting of the boundary and loading conditions in Matlab (Mathworks, Natick, MA, USA), the running and output analysis using Abaqus and the post-processing of the data. Moreover, an automated process of selecting the screw length imposing a TJD of 8mm (Fletcher *et al.* 2019a) and detecting any screw collision in the model was developed and integrated into the workflow (Figure 4.2).

FE Analysis

Input

Orientation of each screw

10	5	10	0	0	5	0	0	10	5	0	10	5	10
0	0	10	5	0	10	5	5	0	0	10	5	0	0
						...							
5	10	0	0	5	10	0	10	5	10	10	0	0	5

x set size

Simpleware
- Bone segmentation — Simpleware
- Fracture selection — Simpleware
- Manual implant positioning — Simpleware

Matlab
- Screw length selection — Simpleware
- Setting of material properties — Simpleware
- Mesh generation — Simpleware
- Setting LC and BC — Matlab
- Running the model — Abaqus
- Outputs analysis — Matlab

Output
50th and 90th strain all screws
50th and 90th strain screw6
50th and 90th strain screw7
50th and 90th strain screw8
50th and 90th strain screw9
50th and 90th strain screw10
50th and 90th strain screw11
50th and 90th strain screw12

Screw collision

ANN Training

input 14 → ANN_collision → output 1

input 14 → ANN_strain14 → output 2

input 16 → ANN_strain16 → output 2

Regression Analysis

Accuracy of real collision prediction from ANN_collision
Regression FE predictions vs ANN_strain14 prediction
Regression FE predictions vs ANN_strain16 prediction

$R^2 = 0.99$

Full Factorial analysis

Prediction of 9^7 configurations using:

ANN_collision + ANN_strain16

ANN_strain14

Figure 4.2 - Flow chart showing steps for the surrogate modelling, from the FE modelling, ANN training and regression analysis.

4.2.2 Dataset Development

The proximal-distal and anterior-posterior orientation of the seven proximal screws was varied from the standard position, which corresponds to the original position of screws in the plate configuration (0° - 0°). In particular, the location of the head of each screw was fixed and the position of their tips was changed from the conventional position. For both directions, the tip was positioned from a range of -10° and 10°, varying with an increment of 5°. Therefore, the total number of possible configurations was 25^7. To reduce this number, three training sets of different sizes were developed with a Latin Hypercube sampling technique, generating training sets of 500, 2000, and 5000 FE models. Some simulations failed due to screws colliding when changing direction, resulting in 92, 370, and 879 successful simulations for each group, respectively. In addition, a final training set size of 7500 was considered as the sum of all the training sets, with a total of successful simulations of 1341. To test the surrogate models once developed, a testing set size of 500 simulations was created, which resulted in 91 successful simulations after removing invalid models due to screw collisions. The training datasets were used to develop and train three different ANNs, one for the detection of screw collision and the other two for the prediction of the principal minimum strain of the bone at the strain surface. All the ANNs were developed using the Neural Network toolbox in Matlab (Mathworks, Natick, MA, USA)(Figure 4.2).

4.2.2.1 ANN for detection of screw collision

The first ANN developed was a classification type network, developed in order to detect if any screw of a model was colliding with any others. The network used was structured with a single hidden layer composed of 25 neurons and one linear output layer, and the training algorithm used was a Levenberg-Marquard backpropagation function. The network was fed with the information of the anterior-posterior and proximal-distal position of each screw, for a total of 14 inputs. The output information was a binary representing the interaction or non-interaction of the screws. This ANN will be addressed as ANN$_{\text{collision}}$.

4.2.2.2 ANN for prediction of bone strain

The second and the third ANNs were composed of two hidden layers, with 10 and 5 nodes respectively, and one linear output layer. The input used to train the second model were the orientation of the screws and the information of screw collision, for a total of 16 inputs. The outputs used were the medial (50th) and 90th percentile of the minimal principal strain of the bone at the surface of the screws. In particular, ANN models were developed to predict the minimal principal strain of the bone around each single screw and around all the screws. A Bayesian regularization backpropagation function was used to train the ANNs. This ANN will be addressed as ANN_{strain_16}.

The third ANN was developed for the prediction of bone strain around the screws. The properties were the same as the ANN_{strain_16}, with the difference in the number of input data used to feed the model. Only information of screw orientation was used, without the information of collision of the screws, for a total of 14 inputs. This ANN will be addressed as ANN_{strain_14}.

4.2.3 Assessment of ANN

During the training process of all the ANNs, the training sets data were divided into 80% for the training, 10% for the validation and 10% for the test. Therefore, as the training samples were chosen randomly at the beginning of the training process, the training process of each case was conducted 100 times, generating 100 unique $ANN_{collision}$, ANN_{strain_14} and ANN_{strain_16} models, in order to assess the robustness of the networks. Those models were then used to make predictions of the 91 unseen cases previously generated for the testing set, and the accuracy of their prediction was evaluated. In particular, for the $ANN_{collision}$ the accuracy was defined by analysing the percentage of true prediction of the collision. For the ANN_{strain_16} and ANN_{strain_14}, a regression analysis was conducted reporting a coefficient of determination (R^2), regression slope and root mean square error (RMSE), to assess the quality of the predictions of the ANNs of the minimal principal strain of the bone around all the screws and the single screws. (Figure 4.2)

4.2.4 Analysis of ANN predictions

After the influence of the training set size was assessed, the best trained $ANN_{collision}$, ANN_{strain_14} and ANN_{strain_16} were used to make predictions of a reduced full factorial scenario, in which screws were able to vary in proximal-distal and anterior-posterior direction of only +10° and -10° from the neutral position. For this setup, the total number of possible configurations was 9^7. (Figure 4.2) An analysis was conducted to compare the accuracy of predictions between the use of $ANN_{collision}$ combined with ANN_{strain_16} and ANN_{strain_14}. To do so, an additional testing set of 500 simulations was defined through Latin Hypercube sampling method in which the TJD of the seven proximal screws had randomly distributed values of -10°, 0° and 10°, having a successful 96 simulations without collision. A two-sample t-test was conducted between the FE results and their predictions using $ANN_{collision}$ with ANN_{strain_16}, and their predictions using only ANN_{strain_14}.

Lastly, a comparison was made between the 50th and 90th percentile principal bone strain around the calcar screw (Screw 6) using the results of the full factorial generated with the $ANN_{collision}$ with ANN_{strain_16}, and the 50th and 90th percentile principal bone strain around the calcar screw (Screw 6) obtained from the training sets of 500, 2000, 5000, and 7500 FE models. This comparison was focused on the calcar screw as it has been shown to be highly influential (Mischler *et al.* 2020b). The purpose of this comparison was to determine whether there was a difference in the range of strain distribution between the predictions made on the full factorial space and the training sets. To test this, a t-test was conducted with a statistical significance level of $p < 0.001$.

4.3 Results

A total of 1,528 FE simulations were run and each one took between 15-20 minutes, from the generation of the mesh to post-processing the results. The training time of each ANN was a few minutes and, once trained, the ANNs prediction time of new configurations was just a few

seconds. In Figure 4.3, several configurations are displayed as examples illustrating the results from the FE analysis (Figure 4.3).
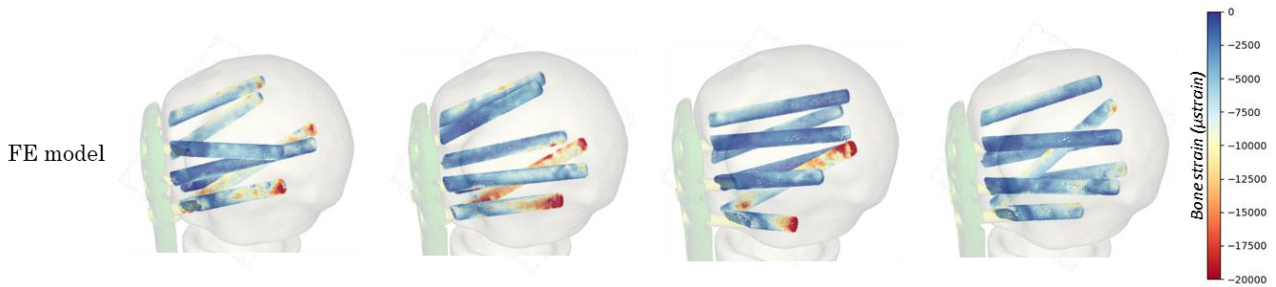
Figure 4.3 – An example of Minimal Principal Strain distribution obtained from the FE analysis for a few configurations

Firstly, the influence of the training set size was assessed for ANN$_{collision}$ and ANN$_{strain\_16}$.

For the prediction of collision, a percentage of true and false prediction has been reported, showing a minor improvement from the smallest to the largest training set, with a true best prediction from 82.6% to 84.4% (Table 4.1).

Table 4.1 - Performance of the ANN$_{collision}$ on the testing set of 91 successful simulations for the prediction of collision of the screws. The influence of the training set size is shown. Results are displayed for the model with the best accuracy, while the average of 100 ANN models is shown in brackets

Training set size (Reduced size)	Prediction of screw collision	
	% Prediction	% Error
500 (92)	82.60 (79.80)	17.40 (20.19)
2000 (370)	83.00 (80.75)	17.00 (19.25)
5000 (879)	83.40 (81.19)	16.60 (18.80)
7500 (1,341)	84.40 (81.75)	15.60 (18.24)

Regarding the prediction of principal bone strain around the screws, the influence of the training set size was assessed on the ANN$_{strain\_16}$ predicting the 50th and the 90th percentile principal bone strain around the screw surface. The ANN$_{strain\_16}$ demonstrated an improvement in accuracy using the larger training set size of 7500, reaching an R^2 value of 0.98 and a RMSE of 129.30 µstrain for the prediction of 50th percentile principal strain, and an R^2 value of 0.99 and a RMSE of 129.26 µstrain for the prediction of 90th percentile principal strain (Table 4.2).

Results of 50th and the 90th percentile principal bone strain predictions from the ANN$_{strain\_14}$ were reported as well, showing respectively a best R^2 of 0.91 and 0.92 and lowest RMSE of 95.40 and 306.35 µstrain (Table 4.3).

Table 4.2 - Performance of the ANN$_{strain\_16}$ on the testing set of 91 successful simulations for the prediction of bone principal strain around the screws. The influence of the training set size is shown. Results are displayed for the model with the best accuracy, while the average of 100 ANN models is shown in brackets

Training set size (Reduced size)	50th percentile min principal strain				90th percentile min principal strain			
	R^2	Slope	RMSE, µstrain	RMSE, % strain	R^2	Slope	RMSE, µstrain	RMSE, % strain
500 (92)	0.97 (0.97)	0.96 (0.95)	161.82 (176.75)	4.72 (5.15)	0.97 (0.97)	0.94 (0.93)	449.99 (459.19)	8.06 (8.23)
2000 (370)	0.98 (0.97)	0.98 (0.98)	140.92 (159.50)	4.11 (4.65)	0.99 (0.98)	0.99 (1.00)	209.71 (374.50)	3.76 (6.71)
5000 (879)	0.98 (0.98)	0.98 (0.98)	129.49 (138.18)	3.77 (4.03)	0.99 (0.99)	0.99 (0.99)	160.07 (211.98)	2.87 (3.80)
7500 (1,341)	0.98 (0.98)	0.99 (0.99)	129.30 (135.03)	3.77 (3.93)	0.99 (0.99)	1.00 (0.99)	129.26 (168.66)	2.32 (3.02)

Table 4.3 - Performance of the ANN$_{strain\_14}$ on the testing set of 91 successful simulations for the prediction of bone principal strain around the screws. Results are displayed for the model with the best accuracy, while the average of 100 ANN models is shown in brackets

Training set size (Reduced size)	50th percentile min principal strain				90th percentile min principal strain			
	R^2	Slope	RMSE, µ strain	RMSE, % strain	R^2	Slope	RMSE, µstrain	RMSE, % strain
7500 (1,341)	0.91 (0.84)	0.90 (0.85)	95.40 (124.116)	2.75 (3.62)	0.92 (0.87)	0.95 (0.92)	306.35 (392.85)	5.49 (7.04)

After having assessed the influence of training set size, additional ANNs were generated and trained using the input and output data from the training set with a size of 7500 to predict the 50th and 90th percentile principal strain around each single screw. The trained ANN$_{strain\_14}$ and ANN$_{strain\_16}$ were used to make predictions of the unseen testing dataset, resulting in a high value of R^2 and a low level of RMSE for all of them (Table 4.4, Table 4.5). In particular, the model ANN$_{strain\_16}$ showed a high level of R^2 ($R^2 > 0.99$) and low RMSE, ranging between 30.81 and 66.70 µstrain for the 50th percentile of strain and between 62.17 and 144.30 µstrain for the

90th percentile of strain (Table 4.4). On the other hand, the model ANN$_{strain\_14}$ had a lower accuracy than the ANN$_{strain\_16}$. For the prediction of the 50th percentile of strain, the R² range was 0.94-0.99 and the RMSE ranged between 69.66 and 163.52 µstrain. The regression with the 90th percentile of strain had a R² value between 0.84 and 0.98, and a RMSE value between 157.45 and 461.81 µstrain (Table 4.5).

Table 4.4 - Performance of the ANN$_{strain\_16}$ on the testing set of 91 simulations for the prediction of bone strain around each single screw. Results are displayed for the model with the best accuracy, while the average of 100 ANN models is shown in brackets

	50th percentile min principal strain				**90th percentile min principal strain**			
	R²	Slope	RMSE, µ strain	RMSE, % strain	R²	Slope	RMSE, µ strain	RMSE, % strain
Screw6	0.999 (0.999)	1.00 (1.00)	66.70 (89.25)	0.65 (0.88)	0.999 (0.988)	1.00 (1.01)	144.30 (372.83)	0.78 (2.03)
Screw7	0.998 (0.997)	1.00 (1.00)	57.24 (67.08)	1.12 (1.31)	0.997 (0.99)	0.98 (0.99)	205.28 (251.17)	1.31 (1.60)
Screw8	0.996 (0.992)	0.99 (0.99)	41.59 (55.62)	0.99 (1.32)	0.996 (0.994)	1.00 (1.00)	108.09 (143.62)	0.77 (1.02)
Screw9	0.997 (0.996)	1.00 (1.00)	40.09 (48.57)	1.37 (1.65)	0.992 (0.985)	0.99 (0.99)	184.04 (239.23)	2.39 (3.11)
Screw10	0.999 (0.998)	1.00 (1.00)	32.47 (42.06)	0.69 (0.89)	0.998 (0.994)	1.00 (1.00)	77.58 (121.77)	1.01 (1.58)
Screw11	0.999 (0.998)	1.00 (1.01)	24.02 (31.04)	0.83 (1.07)	0.997 (0.993)	1.00 (1.00)	84.16 (115.75)	1.30 (1.78)
Screw12	0.999 (0.999)	1.00 (1.00)	30.81 (37.04)	0.74 (0.89)	0.999 (0.999)	1.00 (1.00)	62.17 (81.87)	0.76 (1.01)

Table 4.5 -Performance of the ANN$_{strain\_14}$ on the testing set of 91 simulations for the prediction of bone strain around each single screw. Results are displayed for the model with the best accuracy, while the average of 100 ANN models is shown in brackets.

	50th percentile min principal strain				**90th percentile min principal strain**			
	R^2	Slope	RMSE, μ strain	RMSE, % strain	R^2	Slope	RMSE, μ strain	RMSE, % strain
Screw6	0.989 (0.973)	1.00 (0.98)	163.52 (237.73)	1.62 (2.35)	0.979 (0.946)	0.97 (0.96)	371.37 (583.09)	2.02 (3.17)
Screw7	0.954 (0.872)	0.93 (0.88)	123.57 (188.60)	2.42 (3.69)	0.970 (0.874)	0.92 (0.86)	461.81 (799.31)	2.95 (5.10)
Screw8	0.969 (0.915)	0.99 (0.90)	102.98 (149.99)	2.44 (3.56)	0.981 (0.946)	0.98 (0.96)	264.68 (388.92)	1.87 (2.75)
Screw9	0.946 (0.904)	0.95 (0.91)	91.422 (118.89)	3.11 (4.05)	0.846 (0.694)	0.88 (0.75)	423.15 (590.49)	5.50 (7.67)
Screw10	0.987 (0.965)	0.97 (0.97)	72.036 (109.05)	1.52 (2.30)	0.973 (0.923)	1.00 (0.95)	174.80 (284.37)	2.28 (3.70)
Screw11	0.979 (0.954)	0.97 (0.96)	63.485 (81.61)	2.20 (2.82)	0.970 (0.943)	0.97 (0.95)	182.96 (241.93)	2.82 (3.37)
Screw12	0.992 (0.981)	1.01 (1.01)	69.669 (94.53)	1.68 (2.28)	0.985 (0.936)	1.02 (0.98)	157.45 (250.67)	1.93 (3.08)

Afterwards, the best ANNs trained with 7500 FE data were used to make a prediction of a full factorial scenario, of 9^7 possible configurations, in which the screw could vary in proximal-distal and anterior-posterior direction of ±10° from the neutral position. The ANNs were used to make predictions of 50th and 90th percentile principal bone strain around all the screws and around every single screw of a new testing set.

A t-test was conducted between the FE predictions of an unseen dataset, that was set as a control group, their predictions using ANN$_{collision}$ with ANN$_{strain\_16}$ and their predictions using only ANN$_{strain\_14}$. No significant difference was found between the FE data and the predictions using ANN$_{collision}$ with ANN$_{strain\_16}$ ($p > 0.05$). Significant differences ($p < 0.05$) were found between the FE data and the ANN$_{strain\_14}$ predictions of 50th percentile principal bone strain around screw

6, screw7, and screw 8, and 90th percentile bone strain around all the screws, screw6, screw8, screw 9 and screw 10.
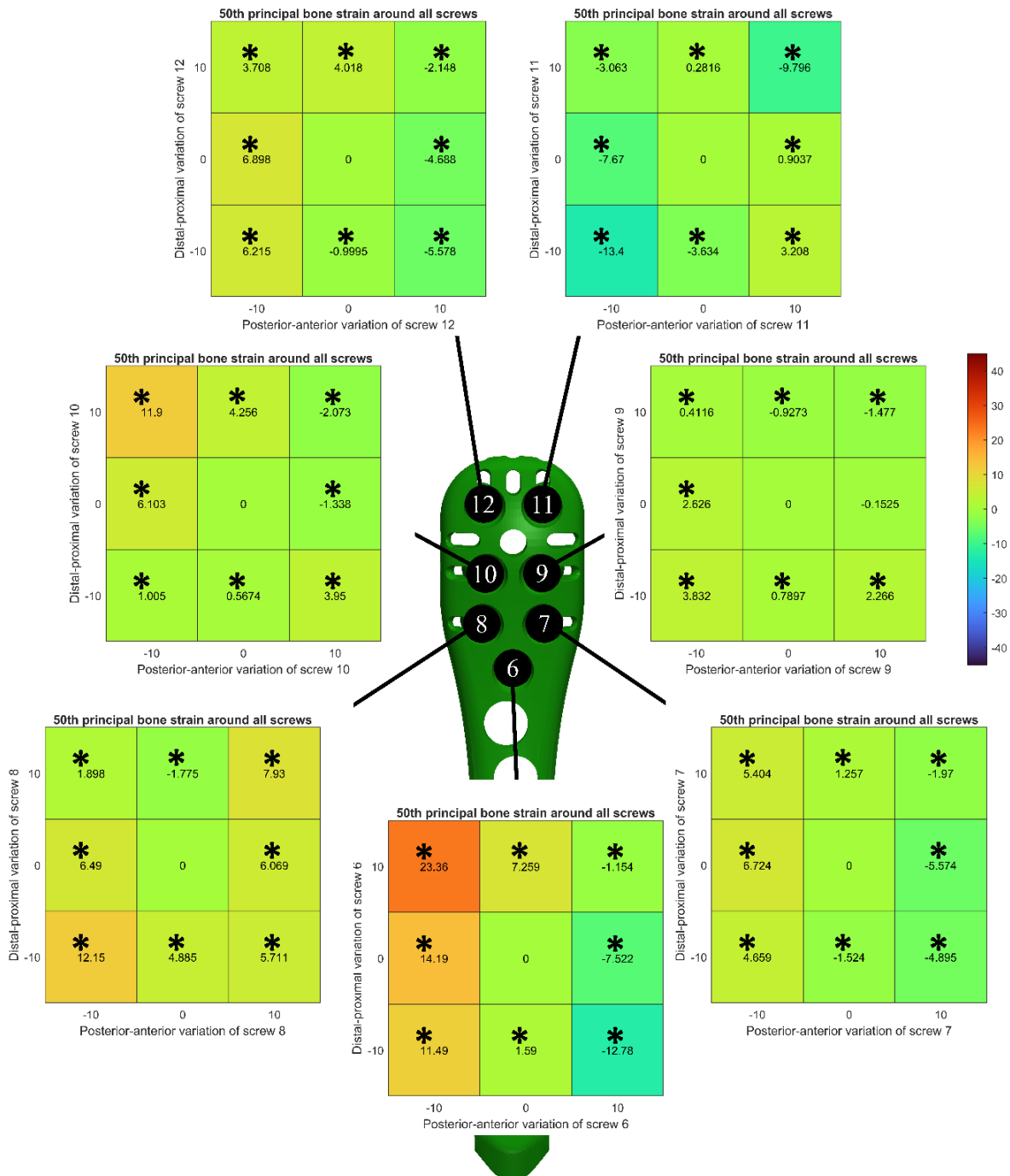
Figure 4.4 - Prediction of 50th percentile principal bone strain around all the screws. Each heatmap represent the percentage of variation of strain for each screw from its neutral position (* indicates p < 0.001).

Therefore, ANN$_{collision}$ with ANN$_{strain\_16}$ was used to make predictions of principal bone strain around the screws of all the 9^7 possible configurations since there was no statistical difference with the FE results. The influence of the orientation of each screw was assessed, showing
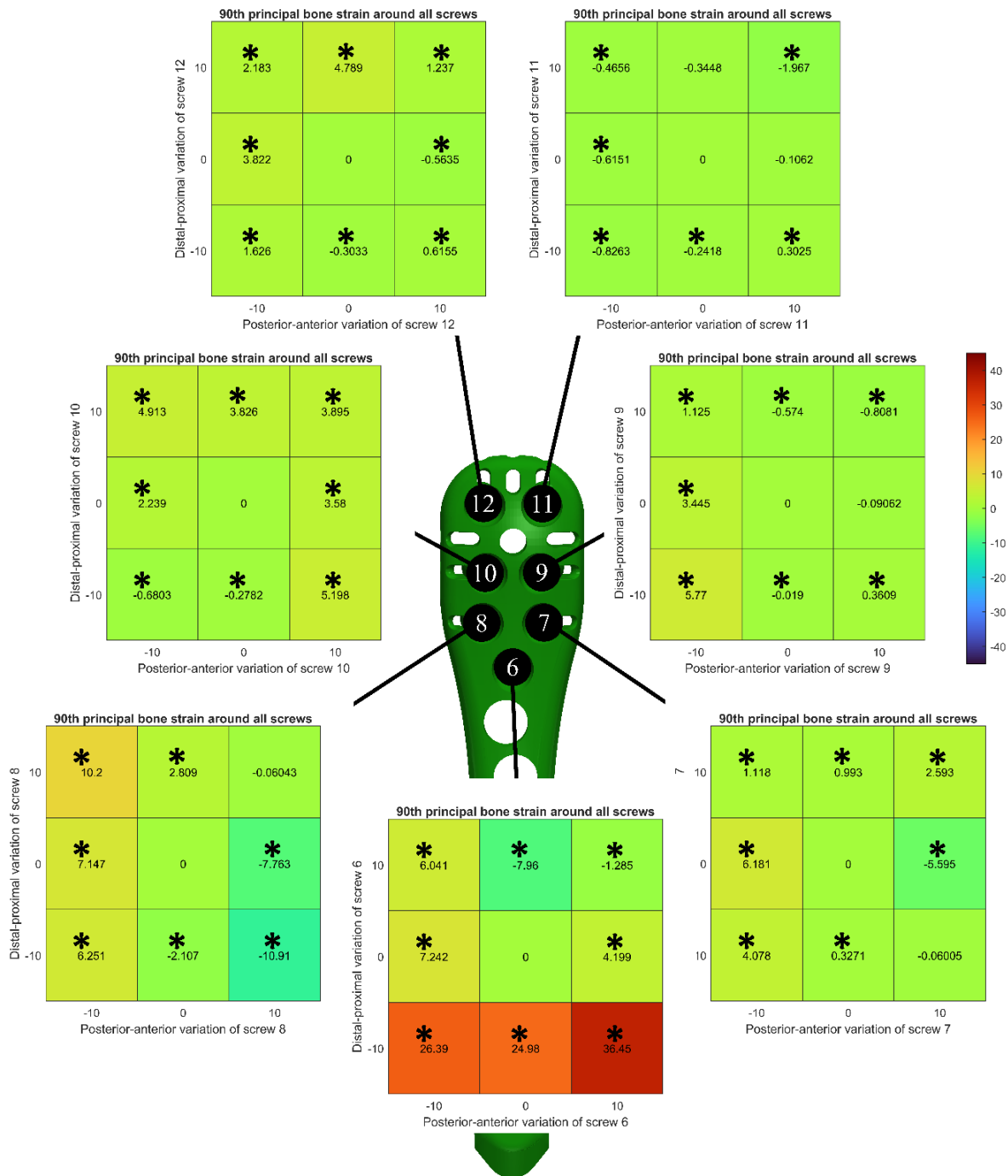
Figure 4.5 - Prediction of 90th percentile principal bone strain around all the screws. Each heatmap represent the percentage of variation of strain for each screw from its neutral position (indicates p < 0.001).*

significant difference (p<0.001) in most of the changed positions of their tips from the neutral position (Figure 4.4, Figure 4.5). In particular, screw 6, the one considered the calcar screw, had the biggest impact on variation of 90th and 50th percentile of bone strain around all the screws showing variations between -12,78% and +36.45% (Figure 4.4, Figure 4.5). Influence of variation of orientation of screw 6 on the strain of the bone around itself has been reported, showing variations between -22.49% and +45.06% (Figure 4.6).

As the calcar screw has a significant impact on the biomechanical outcomes of fracture fixation plates, a final comparison between the variation of the 50th and 90th percentile principal bone strain around the calcar screw (Screw 6) predicted with the use of the $ANN_{collision}$ with ANN_{strain_16}, and the variation of the same output predicted by the FE models for the training sets of 500, 2000, 5000 was conducted. There was a statistical difference (p<0.05) between the results from each of the training sets and the full factorial results (Figure 4.7), which show a higher range of bone principal strain.
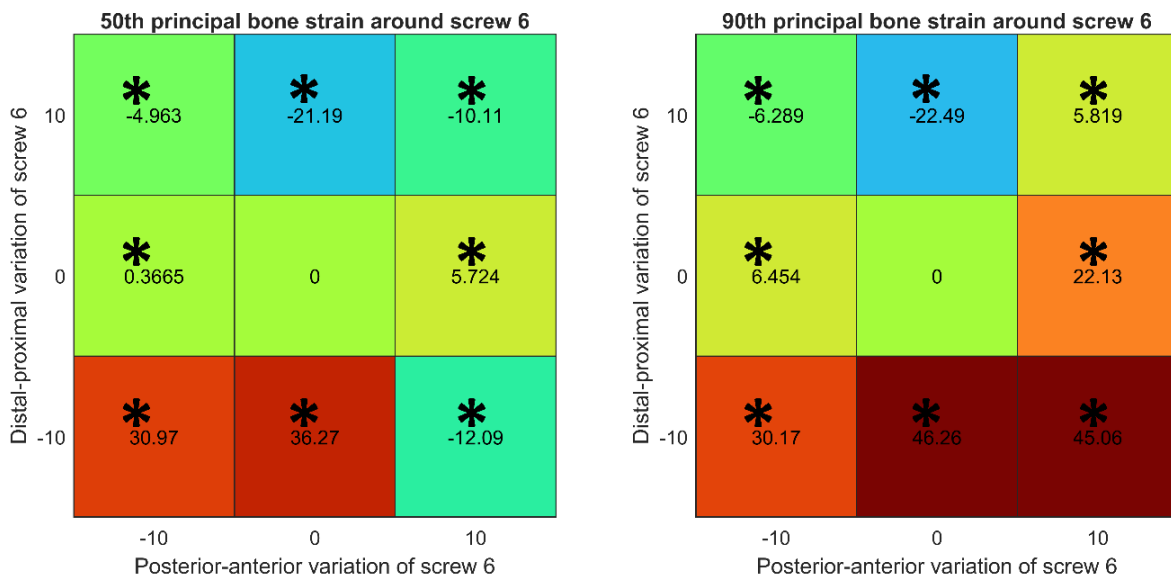
Figure 4.6 - Prediction of 50th and 90th percentile principal bone strain around screw 6 with the variation of orientation of screw 6. Each heatmap represent the percentage of variation of strain for each screw from its neutral position (indicates p < 0.001).*
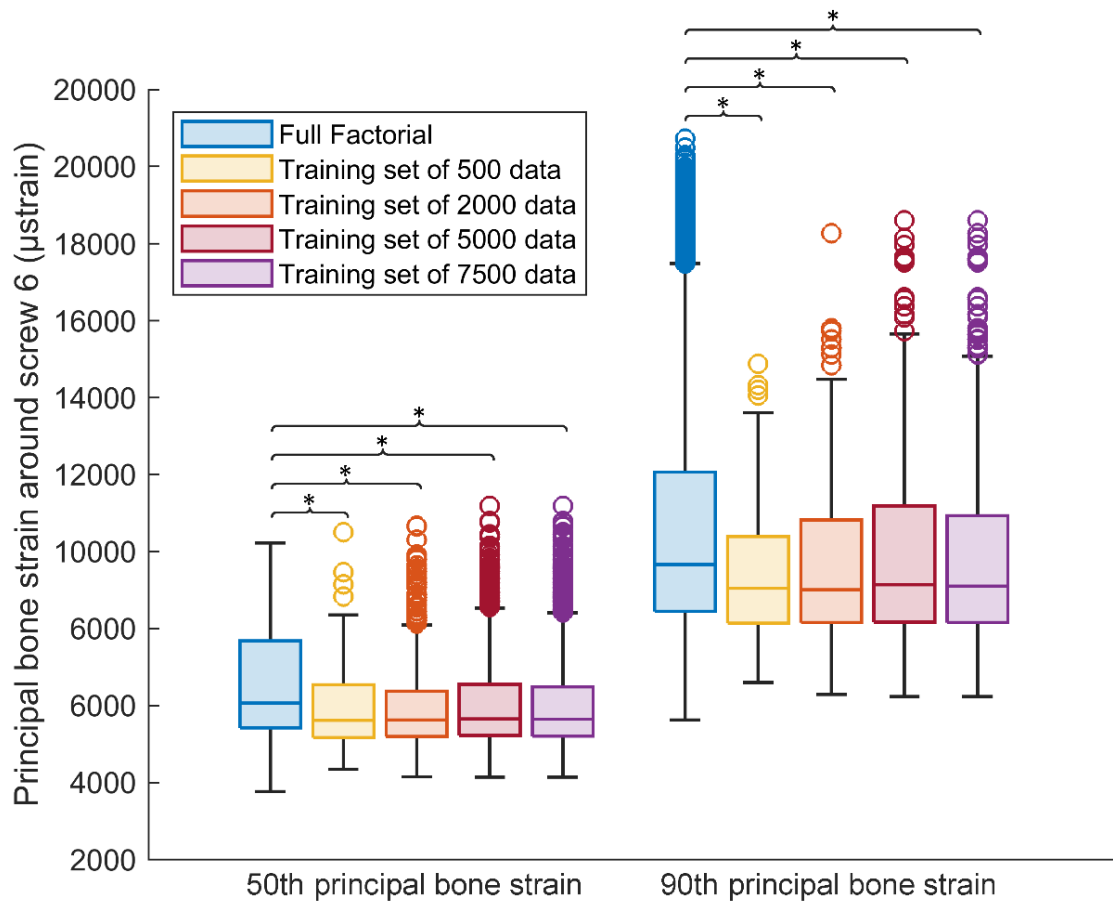
Figure 4.7 - Variation of 50th and 90th percentile bone principal strain for the ANN predictions of the full factorial scenario and the FE predictions of the training set of 500, 2000, 5000, and 7500 data. Statistical significance (p<0.05) is shown between

4.4 Discussion

The aim of this study was to develop a surrogate model, using FE data, in order to make efficient predictions of bone strain of a fractured humerus with a fracture fixation plate as a result of varying the orientation of the proximal screws. A new generation of fracture fixation plates with variable angle screws has been introduced in order to improve the outcome of fracture healing but a worst or optimal configuration hasn't yet been identified when changing the orientation of the proximal screws. Several FE studies have been conducted to investigate the impact of variable angle screws on the biomechanical performance of the fracture fixation plate, but they explored only a reduced number of configurations (Jabran *et al.* 2019b; Mischler *et al.*

2020b; Schader *et al.* 2021), or they investigate the impact of variable angle of only two proximal screws. Indeed, conducting an FE analysis on all possible configurations is not feasible due to the high number of possible combinations. A methodology combining FE data with ANN has been developed, with the aim of making accurate and faster predictions of bone strain around the screws, which had been shown to be a screw cut-out failure predictor (Varga *et al.* 2017). In a previous study, we developed a similar methodology to investigate the impact of variations in screw length on the humeral strain, and our findings confirmed the efficacy of using an ANN approach trained with FE data (Mini *et al.* 2024). For this particular biomechanical problem, two ANN have been developed, one for the prediction of screw collision and another one for the prediction of screw strain. The $ANN_{collision}$ had a maximum accuracy of 84.4%, which did not improve significantly with the increase in training set size. Two different training networks were developed for the prediction of principal bone strain, one having 14 input data (information of screw orientation) and the other having 16 input data (information of screw orientation and information of screw collision). ANN_{strain_16} showed a higher level of accuracy for the prediction of the unseen data from the testing set, especially on the prediction of the minimal principal strain of the bone around each single screw (Table 4.2, Table 4.3, Table 4.4), showing that the information of screw collision used as input improves the prediction of bone principal strain while the screw orientation is varied. Overall, the ANN_{strain_16} showed to be an accurate model for the prediction of bone principal strain around the screws.

Regarding the prediction of the full factorial scenario, a purely FE approach would have taken an impractical 66,430 CPU days, with each single simulation lasting a maximum of 20 minutes. However, combining FE analysis and ANN considerably reduced the running time. Specifically, to produce the full factorial analysis the ANN was trained with 7500 FE simulations, which only took 18.7 CPU days. Moreover, training the ANN models and using them thereafter only took a few minutes.

Predictions of a full factorial scenario of 9^7 simulations were made using ANN$_{collision}$ with ANN$_{strain\_16}$, in order to assess the influence of screw orientation on bone principal strain on the full factorial spectrum. The principal bone strain predicted by the ANN$_{collision}$ with ANN$_{strain\_16}$ was not significantly different from the FE data of the unseen testing set, unlike the case with ANN$_{strain\_14}$, which underestimated the same data. This indicates that even though the ANN$_{collision}$ alone did not demonstrate strong accuracy, when combined with ANN$_{strain\_16}$, they collectively made similar predictions of bone principal strain compared to the unseen FE results.

The variation in the orientation of the calcar screw, screw 6, showed to have the strongest influence on the 50th and 90th percentile of principal bone strain around all the screws (Figure 4.4, Figure 4.5). Analysing the 90th percentile of strain around all the screws, on average the distal position showed to be the least safe. When the screw 6 has a distal angle of 10°, the strain value increases up to 36% from the neutral position. In terms of the safest configuration, our findings agreed with those of Fletcher et al and Jabran et al, which identified the safest configuration was when the calcar screw is orientated in the proximal direction (Jabran *et al.* 2019b; Mischler *et al.* 2020b). Moreover, the impact of variations of screw 6 on the strain of the bone around itself has been reported, showing on average a decrease of more than 20% of the variation of 50th and 90th percentile principal bone strain in the proximal direction (Figure 4.6).

In this study we also compared the bone principal strain predictions obtained from the full factorial scenario generated with the ANNs with those of the FE simulations on the training sets of 500, 2000, 5000, and 7500. The analysis revealed a significant statistical difference. In particular, the ANNs models were able to make predictions on the full factorial scenario with higher principal bone strain variation than the ones predicted by the FE analysis on all the training sets (Figure 4.7). This showed how running a full factorial analysis with an ANN approach would help to identify potentially dangerous configurations, that would not have been detected through a small sample analysis.

This study has demonstrated that an FE-informed Neural Network approach can be used to explore the impact of variation of screw orientation on bone deformation for a fracture fixation plate for a proximal humeral fracture, showing the high impact that the calcar screw orientation has on the prediction of bone principal strain. This technique has the primary advantage of being cost-effective in terms of time, both during the training phase and when making predictions for unseen scenarios. The computational cost is strongly dependent on the generation of data used for the training process. Moreover, the process of generating new predictions was made easy and fast by utilizing only screw orientation and collision information as input data, eliminating the need for generating a new mesh for the prediction of new data. Indeed, the most cost demanding stage of this methodology is the generation of FE data, but this cost is significantly reduced once the surrogate model has been trained.

There are some limitations in this study. Firstly, the boundary and loading conditions of the FE analysis were simplified, and no muscular forces were taken into account. Regarding the material properties of the bone, the Morgan's equation referred to the femoral head was used to define the relation between Young's Modulus and BMD, as no equation referred to the humeral head was found in the literature (Morgan *et al.* 2003). The model was not experimentally validated, however, the methodology developed for the generation of the FE models reproduced in the Varga et al study and Mischler et al, that were experimentally validated (Mischler *et al.* 2022; Varga *et al.* 2017), which was similar to the methodology conducted in this paper for the generation of FE data. This study did not explore subject variabilities; indeed, the model was developed on a single subject, as the Jabran and Tilton studies, not taking into account the possible variations due to different anatomies (Jabran *et al.* 2018; Tilton *et al.* 2020b). Since Schader et al study (Schader *et al.* 2021) suggested that subject-specific optimization of the orientation of the screws could improve the biomechanical performance of the fracture fixation plate, in the future ANN models assessing patients' variabilities should be developed. Moreover, the model only included one fracture pattern in order to simplify the study. To improve the

model's complexity and applicability, future versions need to take into account different fracture configurations. Given that the accuracy of the ANN$_{collision}$ only reached 84.4%, additional techniques should be implemented and evaluated to predict screw collisions more accurately. Since the cause of the failure of these implants is unclear, the influence of other parameters should be investigated with this methodology, fracture pattern, screw configuration and also potential implant malposition. Additionally, the ANN model should include other outputs of interest to better understand the biomechanics of the fracture fixation plate, such as bone micromotion, fracture gap movements, and implant stress (Lewis *et al.* 2021).

Despite these limitations, this study demonstrated the potential of using a finite element informed Neural Network technique to develop an advanced computational model for investigating the variation of strain with the variation of implant screws orientation. Moreover, the technique proved to be a more efficient and less time-consuming approach than traditional methods. As also demonstrated in our previous research, using an ANN methodology is advantageous when studying problems with a wide solution space, and this approach allows for the identification of potentially dangerous configurations in a more computationally efficient manner. Our study's findings provide assurance regarding the use of this more efficient computational technique, and in the future, we intend to explore even more complex techniques to implement and potentially provide more informed decisions in a surgical setting.

4.5 Conclusion

In conclusion, a computational approach using FE and ANNs to predict bone deformation of the humerus with the variation of screw orientation was successfully developed. This methodology showed good accuracy for the prediction of deformation of the bone with the variation of screw orientation in the proximal-distal and anterior-posterior direction. The trained ANNs demonstrated the impact of the orientation of the calcar screws on the biomechanical

performance of the fracture fixation plate, in agreement with what was found in the literature, showing that more beneficial configurations can be reached with variable angle locking screws.

Chapter 5

Developing a Graph Neural Network model to predict a field of strains in fracture fixation of the proximal humerus varying implant parameters

5.1 Introduction

When conducting FE analysis to explore the effects of multiple parameters, the computational demands can be quite significant. In Chapters 3 and 4, an ANN methodology was developed to assist the FE analysis. The models demonstrated a high degree of accuracy in representing the 50th and 90th percentiles of bone strain around the screws. However, it is important to acknowledge that the ANN has a significant limitation of being incapable of replicating a distribution of strain on a geometry, thereby it doesn't allow the visualisation of the entire strain field. Other Deep Learning methods can be combined with the FE process to reduce computational costs and predict a distribution of strain. While some CNN methods have been developed for FE studies, they are limited to working with structured Euclidean data (Heidari

112

et al. 2022; Zhou *et al.* 2020). On the other hand, GNN is a powerful DL technique that operates on graph structures. When properly trained, GNN can predict features of graphs and 3D structures. GNNs are commonly used to analyse graph data such as social interactions, networks, and chemical structures. However, only a few studies have begun to explore the application of GNN in combination with FE analysis to investigate the strain and stress of 3D solid objects. For example, Krokos et al. utilised a GNN approach to study the variation of internal stress in an object by altering its porous structure, achieving high prediction accuracy from the model (Krokos *et al.* 2022a). This study, although not conducted in a complex scenario, demonstrated the potential of using a GNN approach with FE data to predict the behaviour of a material as its 3D structure changes.

To investigate the failure of fracture fixation devices for humerus fractures, FE methods are computationally demanding because of the high number of parameters and the wide solution space. The GNN approach could be used with FE data to predict the internal distribution of strain and stress when investigating the effect of multiple configurations for an orthopaedic implant. This approach can reduce the computational cost without sacrificing accuracy.

In this chapter, a GNN approach was developed using the FE data generated from Chapters 3 and 4. The goal was to predict the distribution of strain on the bone surface around the screws, with the variation in screw length and screw direction respectively. The accuracy of the model's predictions was analysed, and a comparison of the GNN models with the ANN model's prediction was conducted.

5.2 Methods

The study was based on the FE data developed in Chapters 3 and 4, which consisted of a single fractured humerus with a fracture fixation plate. In particular, in Chapter 3 the FE analysis was developed by varying the length of the proximal screws in the humeral head. In Chapter 4 the FE analysis was developed by varying the orientation in space of the proximal humeral

screws. Two distinct GNN models were developed on each dataset. The GNN for the first dataset will be addressed as GNN_{ScrewLength} and the GNN for the second dataset will be addressed as GNN_{ScrewDirection}.

5.2.1 Variation of screw length

5.2.1.1 Dataset development

The dataset used for this section was the FE dataset developed in Chapter 3, which consisted of a subject-specific analysis of a single humerus in which the screw length of the proximal screws was varied. As already described in section 3.2.1, an FE dataset was generated on a single image of a cadaver collected from the New Mexico Decedent Image Database (Edgar *et al.* 2020). The screw length of the seven proximal screws was the variable of interest, which was adjusted through the TJD between the cortical surface of the humeral head and the tip of the screws. In particular, the TJD measurements used were 4, 8, 12 and 16 mm. For this study, the five training sets of size 50, 100, 200, 500 and 1000 data and the testing set of 100 data were used, which were all already developed in Chapter 3. The data from each FE simulation was converted into graph structures consisting of a 3D connection of nodes. The features of the nodes of the bone around the seven proximal screws were elaborated and divided into input and output data, necessary for the training of the GNN_{ScrewLength}. In particular, the input data were defined as follows:

- x, y, z coordinates of the nodes of the bone surface around the proximal screws, each of them expressed in mm.
- Young Modulus of the nodes of the bone surface around the proximal screws, expressed in MPa. The data were extracted from the adjacent elements.
- Edge connectivity matrix. The nodes' connectivity was determined by connecting each node to the nodes positioned with a maximum distance of 2 mm.
- Edge length distance of the connected nodes.

The output value used was the Minimum Principal Strain of the bone of the surface around the proximal screws. The data were defined at the nodal level after they were extracted from the adjacent elements.

5.2.1.2 GNN_{ScrewLength} development

The study was conducted using Pytorch Geometric (Fey and Lenssen 2019), in particular, the GNN structure TAGConv (Du *et al.* 2017).

The GNN_{ScrewLength} consisted of an encoder, a processor, and a decoder (Figure 5.1)

- The encoder was made up of a single block containing a TAGConv and a ReLU layer. The TAGConv had an input size of 4, corresponding to the Young modulus at the nodes and the x, y, and z coordinates of the nodes. The output size of the TAGConv layer was set to 100.

- The processor consisted of six distinct blocks, each one composed of a TAGConv, and ReLu layer. The input and output size of each TAGConv layer was set to 100.

- The decoder consisted of a final distinct block of a TAGConv, and ReLu layer, and a final Linear layer. The input size of the TAGConv was 100, and the output was 1, corresponding to the minimum principal strain at the nodes.

For each TAGConv the number of loop K was set to 6. Each block elaborates also the information on edge connectivity and edge length.

This structure was used to train five different GNN_{ScrewLength} models using the five training sets of sizes 50, 100, 200, 500 and 1000, in order to investigate the amount of data needed for an accurate model.
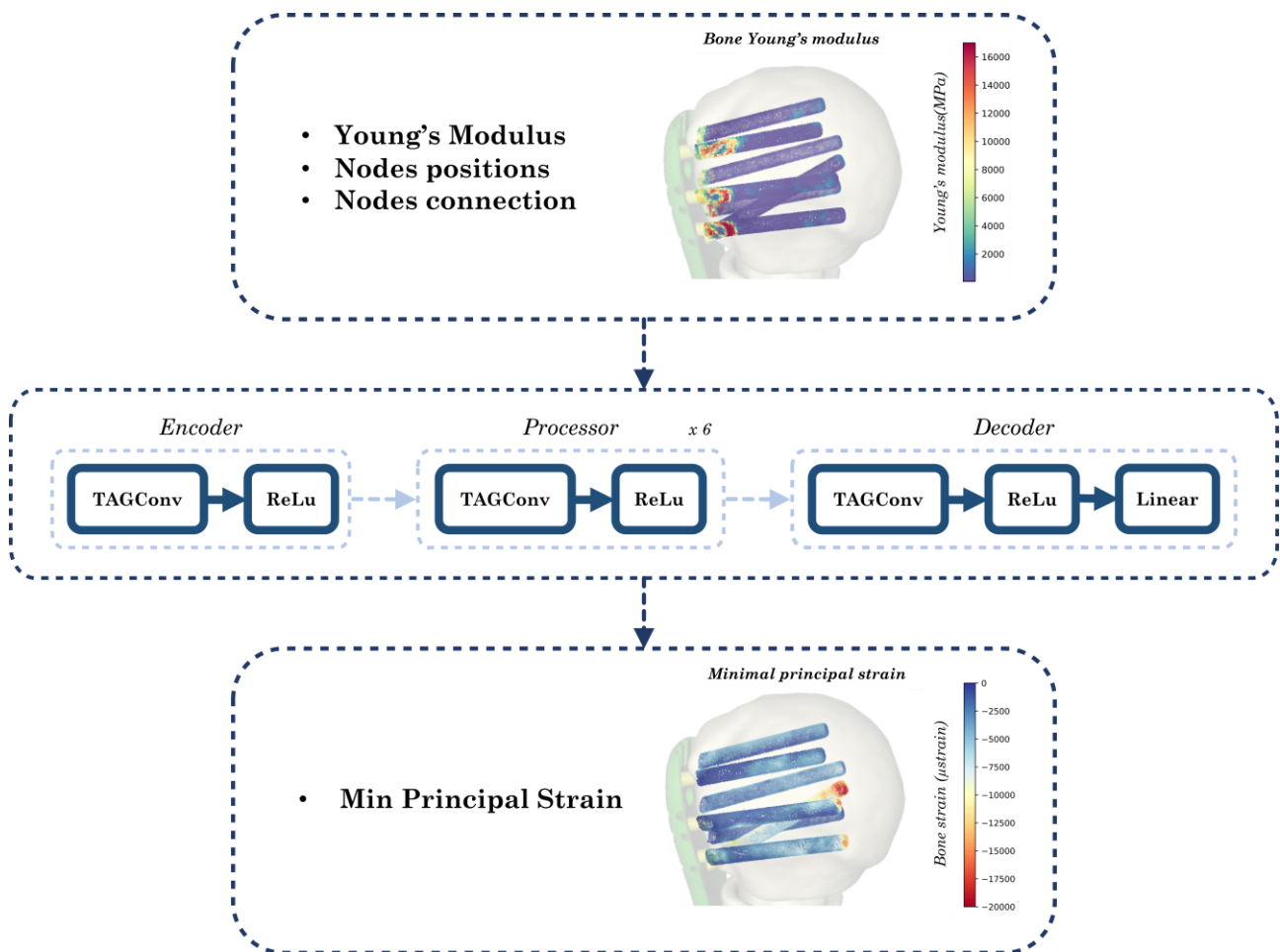
Figure 5.1 - Workflow of the GNN$_{ScrewLength}$ to investigate the influence of variation of screw length

5.2.1.3 Analysis of GNN$_{ScrewLength}$ predictions

To assess the robustness of the GNN$_{ScrewLength}$ model, the training process was separately performed 10 times using the entire dataset for each training session, resulting in the creation of 10 distinct trained models each time. Subsequently, all the configurations of the testing set of 100 models were used to assess the accuracy of the models. A regression analysis between those FE data and their prediction with the trained GNN$_{ScrewLength}$ models was conducted, looking at the R^2 values, the slope and the RMSE (µstrain and %). The best GNN$_{ScrewLength}$ model was considered to have the lowest %RMSE. The GNN models were trained using a NVIDIA RTX A400 GPU with 16 GB of RAM.

5.2.1.4 Reproduction of ANN case study

An additional FE dataset was created based on the conditions outlined in section 3.2.1. This dataset replicated configurations where all the screws had a TJD length of 4, 8, 12, and 16 mm, and where the TJD of all the screws was set to 8, except for screw 6, which was varied to 4, 8, 12, and 16 mm. In total, there were 7 different configurations. These specific configurations were chosen based on findings from Chapter 3 and previous literature, which demonstrated the impact of changing the screws' TJD equally and the effect of varying the length of the screw inserted in the calcar region, which in this configuration was represented by Screw 6. The bone strain of the new dataset was predicted using the best $GNN_{ScrewLength}$ model of each training set. The differences between the predictions from $GNN_{ScrewLength}$ with the FE real data and $GNN_{ScrewLength}$ vs the ANN developed in Chapter 3 were reported.

5.2.2 Variation of screw direction

5.2.2.1 Dataset development

The dataset used for this section was the FE dataset developed in Chapter 4, which consisted of a subject-specific analysis of a single humerus in which the screw direction of the proximal screws was varied. As already described in section 4.2.1, an FE dataset was generated on a single image of a cadaver collected from the New Mexico Decedent Image Database (Edgar *et al.* 2020). The variable of interest was the screw orientation of the seven proximal screws in the humeral head. In particular, proximal-distal and anterior-posterior were varied by ±5° and ±10° from the neutral position, set as 0°-0°. For this study, the four training sets of size 500, 2000, 5000 and 7500 data and the testing set of 500 data were used for the analysis, which were all already developed in Chapter 4. As the collision of the screws had to be taken into account, the training and testing sets had respectively a reduced number of 92, 370, 879, 1341, and 91 successful simulations. The data from each FE simulation was converted into graph structures consisting of a 3D connection of nodes. The features of the nodes of the bone around the seven

proximal screws were elaborated and divided into input and output data, necessary for the training of the GNN$_{\text{ScrewDirection}}$. In particular, the input data were defined as follows:

- x, y, z coordinates of the nodes of the bone surface around the proximal screws, each of them expressed in mm.

- Young Modulus of the nodes of the bone surface around the proximal screws, expressed in MPa. The data were extracted from the adjacent elements.

- Edge connectivity matrix. The connectivity of the nodes was determined by connecting each node to the nodes positioned with a maximum distance of 1.5 mm.

- Edge length distance of the connected nodes.

The output value used was the Minimum Principal Strain of the bone of the surface around the proximal screws. The data were defined at the nodal level after they were extracted from the element adjacent.

5.2.2.2 GNN$_{\text{ScrewDirection}}$ model development

The study was conducted using Pytorch Geometric (Fey and Lenssen 2019), in particular a combination of different GNN structures. In contrast to the approach outlined in Section 5.2.1.2, a different GNN methodology was used based on the observation that the performance of the GNN in Section 5.2.1.2 was almost identical. Therefore, the aim was to construct a model using an alternative GNN structure (Appendix B: Supplementary material for Chapter 5).

The GNN$_{\text{ScrewDirection}}$ consisted of an encoder, a processor, and a decoder (Figure 5.2).

- The encoder consisted of a single block composed of a GraphNorm (Cai *et al.* 2021), XConv (Li *et al.* 2018), ChebConv (Defferrard *et al.* 2016), Dropout and a ReLu layer. The input size of the GraphNorm was 1, corresponding to the young modulus at the nodes.

- The processor consisted of two distinct blocks, each one composed of an XConv, ChebConv, Dropout and a ReLu layer. The input and output size of each TAGConv layer was set to 100.

- The decoder consisted of a final distinct block of an XConv, ChebConv, Dropout, ReLu layer, and a final Linear layer. The input size of the linear layer was 100 and its output size was 1, corresponding to the minimum principal strain at the nodes.

For each XConv the hidden channels were 100, the dimension was set to 3 and the Kernel size was 5. Each XConv elaborates on the information of Young modulus and nodal positions. For each ChebConv the Chebyshev filter size K was 9. Each ChebConv elaborates on information of Young modulus edge connectivity and edge length distance. The four training sets were utilized to train four distinct GNN$_{ScrewDirection}$ models with the previously explained structure. The accuracy of the models was assessed based on the training set size.
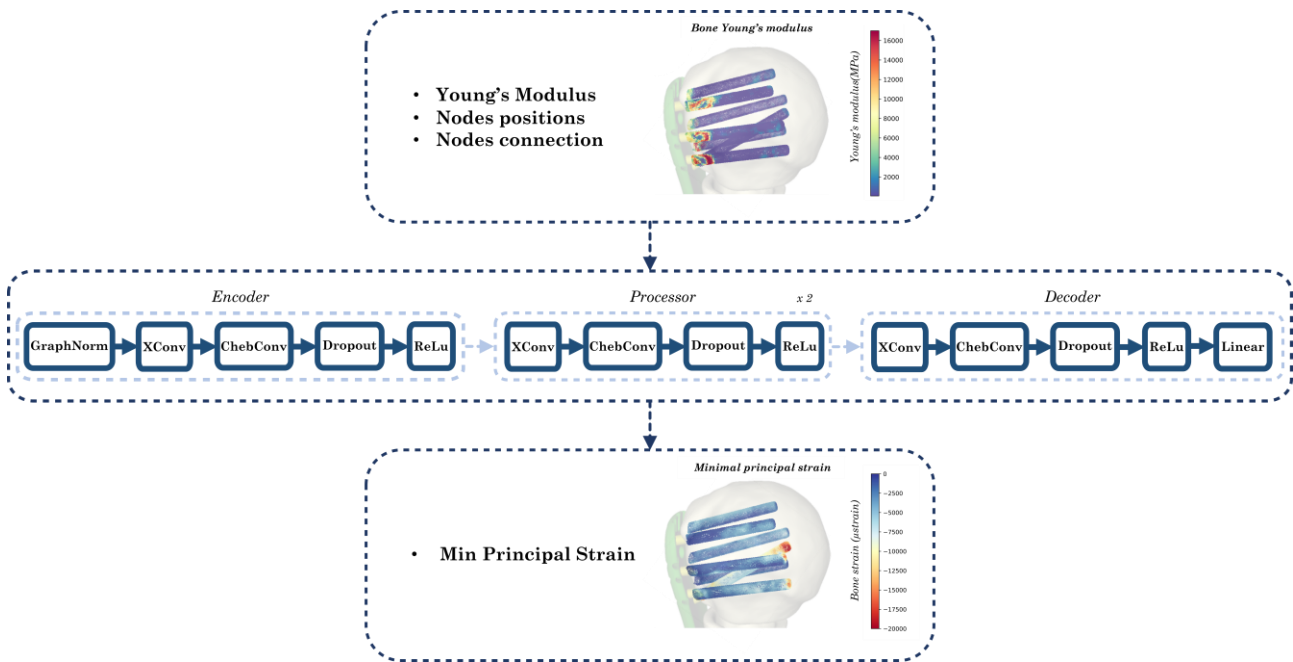
Figure 5.2 - Workflow of the GNN to investigate the influence of variation of screw orientation

5.2.2.3   Analysis of GNN$_{ScrewDirection}$ prediction

To assess the robustness of the GNN$_{ScrewDirection}$ model, the training process was separately performed 10 times using the entire dataset for each training session, resulting in the creation of 10 distinct trained models each time. Subsequently, all the configurations of the testing set of 500 models (91 successful simulations) were used to assess the accuracy of the models. A

regression analysis between those FE data and their prediction with the trained $GNN_{ScrewDirection}$ was conducted, looking at the R^2 values, the slope and the RMSE (µstrain and %). The GNN models were trained using a NVIDIA RTX A400 GPU with 16 GB of RAM.

5.2.2.4 Reproduction of ANN case study

A comparison between the predictions generated by the ANN developed in Chapter 4 and the predictions from the $GNN_{ScrewDirection}$ was conducted. The testing set of 500 simulations (91 successful) was considered, in particular the FE configurations with the maximum and minimum 90th percentile principal strain. The difference between the 50th and 90th percentile of bone principal strain between the predictions from the ANN and those from the $GNN_{ScrewDirection}$ was calculated.

5.3 Results

5.3.1 Variation of screw length

The overall FE analysis running time for each configuration, from the selection of the screw length to the post-processing of the output, took approximately 15-20 minutes. During the training process for the $GNN_{ScrewLength}$ models, it took 0.5, 1.1, 2.2, 6, and 12 hours to train them using training sets of 50, 100, 200, 500, and 1000 data, respectively. Once trained, each $GNN_{ScrewLength}$ model only took 15 seconds to predict the 100 FE configuration of the testing set (Table 5.1). An analysis of the influence of the training set size revealed that for regression between the FE data and their prediction with the $GNN_{ScrewLength}$ models, the R^2 and slope increased from 0.92 to 0.95, and the RMSE decreased from 889.75 to 681.88 µstrain (a reduction of 23%) when transitioning from a training set of 50 data points to one of 1000 (Figure 5.3, Table 5.1).

For each training set, the best $GNN_{ScrewLength}$ model was used to replicate the bone principal strain of the testing set of 7 configurations introduced in section 5.2.1.4. Among these

configurations, four had the same TJD measurement for all the screws, while the remaining three had a TJD of 8mm for all the screws, except for screw 6, which had a varied TJD. The models were able to accurately capture the increase in strain with the increase in TJD (Figure 5.4, Figure 5.5). The model trained with more data showed higher accuracy, particularly in demonstrating the increase in strain for the configuration where all screws had a TJD of 16 mm, resulting in a significant increase in strain at the tip of the cortical screw. A similar behaviour was observed for the model where screw 6 had a TJD of 16 mm while the rest of the screws had a TJD of 8 mm. Depending on the configuration, the model trained with more data tended to produce results closer to the FE predictions, with the R^2 increasing from 0.86-0.95 for the model trained with a smaller training set to 0.91-0.98 for those trained with a larger training set. Looking at the RMSE, this study reported a reduction from 5.29-2.50% μstrain for the model trained with 50 data to a %RMSE of 4.29-1.57% μstrain for the model trained with 1000 data (Table 5.2, Table 5.4).

For the comparison of the 7 configurations with the $GNN_{ScrewLength}$ and ANN models developed in Chapter 3, the absolute percentage difference was reported for both the 50th and the 90th minimal principal strain. Overall, a difference variation of 0.63-15.50 was reported, showing no influence on the training set size. The ANN model is assumed to provide better predictions than the GNN model, given that the error tends to be higher for the $GNN_{ScrewLength}$ predictions. When using a GNN model, accuracy is sacrificed in order to visualize the strain field (Table 5.3, Table 5.5).

Table 5.1 - Performance of the GNN$_{ScrewLength}$ on the testing set of 100 simulations for the prediction of the minimal principal strain of the bone around all the screws with the variation of screw length. The influence of the training set size is shown. Results are displayed for the model with the lowest error, while the average of 100 models is shown in brackets.

Training set size (n)	FE Running time (h)	GNN training time (h)	R^2	Slope	RMSE (µstrain)	RMSE (%)
50	16.6	0.5	0.92 (0.91)	0.92 (0.92)	889.75 (931.08)	3.76 (3.92)
100	33.3	1.1	0.93 (0.93)	0.94 (0.89)	816.73 (871.99)	3.32 (3.52)
200	66.6	2.2	0.94 (0.94)	1.00 (1.00)	744.23 (762.69)	3.11 (3.24)
500	166.6	6	0.95 (0.95)	0.95 (0.95)	700.74 (710.01)	2.81 (2.88)
1000	333.2	12	0.95 (0.95)	0.96 (0.95)	681.88 (693.29)	2.71 (2.76)
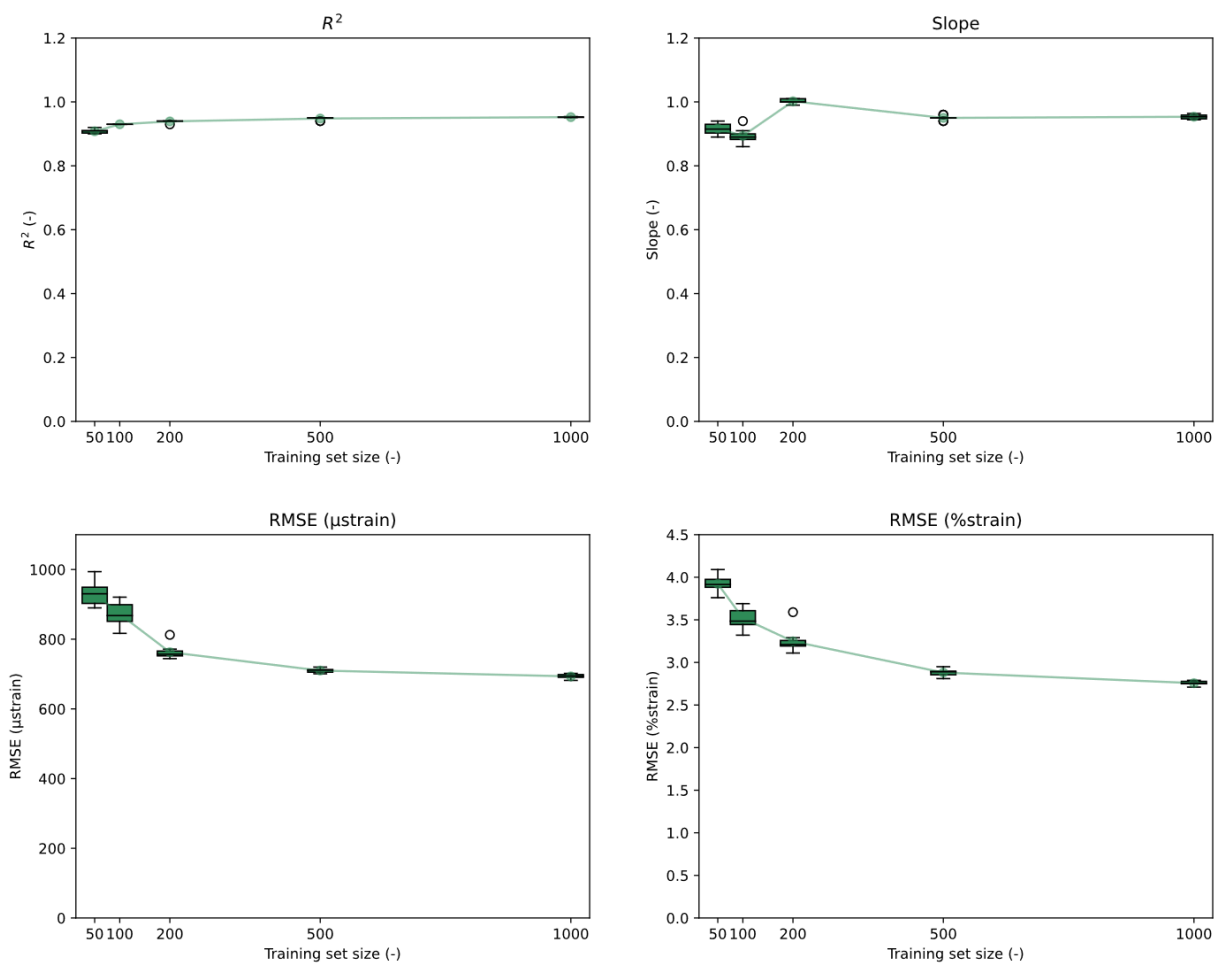
Figure 5.3 – Performance variation of the 10 GNN$_{ScrewLength}$ developed for each training set.

Table 5.2 - Regression of the minimal principal strain at the nodes between the FE data of the configurations having the same value of TJD for every seven proximal screws, and their prediction using the best $GNN_{ScrewLength}$ for each training set size

		Nodal minimal principal strain																				
		R^2					Slope					RMSE (µstrain)					RMSE (%)				
Training set size		50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000
TJD	4	0.89	0.92	0.90	0.94	0.95	1.06	1.04	1.12	1.02	1.04	722.5	569.8	774.6	492.0	452.6	4.71	3.71	5.05	3.20	2.95
	8	0.92	0.95	0.96	0.97	0.95	0.93	0.95	1.06	0.96	1.01	613.3	462.4	510.0	411.1	518.1	4.27	3.22	3.55	2.86	3.61
	12	0.95	0.97	0.97	0.98	0.98	0.97	0.93	1.01	0.97	0.94	699.2	613.2	466.9	424.8	459.2	2.50	2.28	1.74	1.58	1.71
	16	0.86	0.89	0.87	0.90	0.91	0.73	0.77	0.76	0.80	0.80	2339.7	2096.7	2164.4	2016.7	1948.3	5.15	4.62	4.77	4.44	4.29

Table 5.3 - Comparison of the configurations with the same value of TJD for each seven proximal screws predicted with the $GNN_{ScrewLength}$ model and the ANN model developed in Chapter 3. The absolute difference between the 50th and the 90th percentile minimal principal strain values obtained from the $GNN_{ScrewLength}$ and ANN models is reported.

		% of error of 50th percentile minimal principal strain between the ANN predictions and the $GNN_{ScrewLength}$ prediction					% of error of 90th percentile minimal principal strain between the ANN predictions and the $GNN_{ScrewLength}$ prediction				
Training set size		50	100	200	500	1000	50	100	200	500	1000
TJD	4	13.01	6.87	13.40	11.20	8.63	14.91	8.57	19.04	10.22	10.62
	8	6.65	1.67	8.93	7.78	5.69	0.21	0.38	9.40	0.64	2.75
	12	1.23	6.64	0.37	1.43	5.35	2.24	4.70	6.11	1.90	1.63
	16	12.02	15.08	10.58	12.49	15.51	11.55	11.51	4.47	8.72	8.83
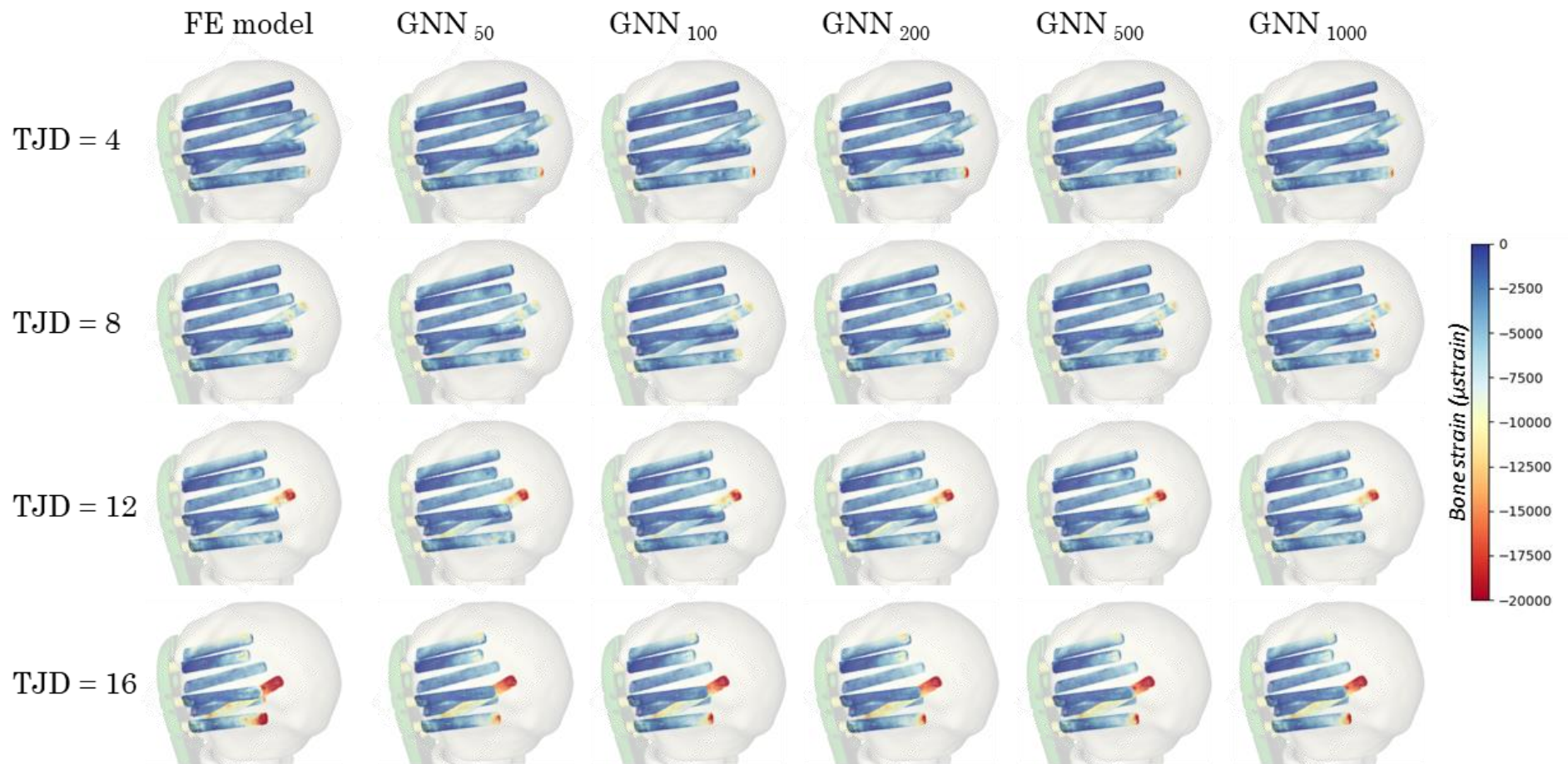
Figure 5.4 - Bone principal strain distribution of the configurations having same TJD for all the screws. Results from the FE model and the best GNN$_{ScrewLength}$ trained with each training set size.

Table 5.4 – Regression of the minimal principal strain at the nodes between the FE data of the configurations having TJD of all the screws set to 8 mm and having varied TJD for screw 6, and their prediction using the best GNN$_{ScrewLength}$ for each training set size

		R^2					**Slope**					**RMSE (μstrain)**					**RMSE (%)**				
	Training set size	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000
TJD	4	0.88	0.94	0.90	0.94	0.95	1.01	0.99	1.07	0.99	1.00	729.16	492.3	718.42	495.85	438.3	5.29	3.57	5.22	3.60	3.18
	8	0.92	0.95	0.96	0.97	0.95	0.93	0.95	1.06	0.96	1.01	613.3	462.4	510.0	411.1	518.1	4.27	3.22	3.55	2.86	3.61
	12	0.94	0.95	0.96	0.98	0.98	0.98	0.90	1.05	0.97	0.98	667.7	635.5	570.6	447.3	375.3	2.79	2.66	2.38	1.82	1.57
	16	0.92	0.92	0.94	0.94	0.95	0.87	0.92	0.94	0.94	0.97	1247.8	1176.9	1082.3	1075.7	974.5	2.91	2.74	2.52	2.50	2.27

Table 5.5 - Comparison of the configurations varying only the TJD of screw 6 predicted with the GNN$_{ScrewLength}$ model and the ANN model developed in Chapter 3. The absolute difference between the 50th and the 90th percentile minimal principal strain values obtained from the GNN$_{ScrewLength}$ and ANN models is reported

		% of error of 50th percentile minimal principal strain between the ANN predictions and the GNN$_{ScrewLength}$ prediction					**% of error of 90th percentile minimal principal strain between the ANN predictions and the GNN$_{ScrewLength}$ prediction**				
Training set size		50	100	200	500	1000	50	100	200	500	1000
TJD	4	9.12	5.26	10.71	8.08	4.48	8.95	4.59	12.65	5.72	4.47
	8	6.65	1.67	8.93	7.78	5.69	0.21	0.38	9.40	0.64	2.75
	12	1.60	1.97	2.08	2.42	0.88	1.99	5.57	7.18	0.92	0.63
	16	10.47	12.34	10.86	9.26	12.40	5.35	4.19	2.23	2.79	0.22
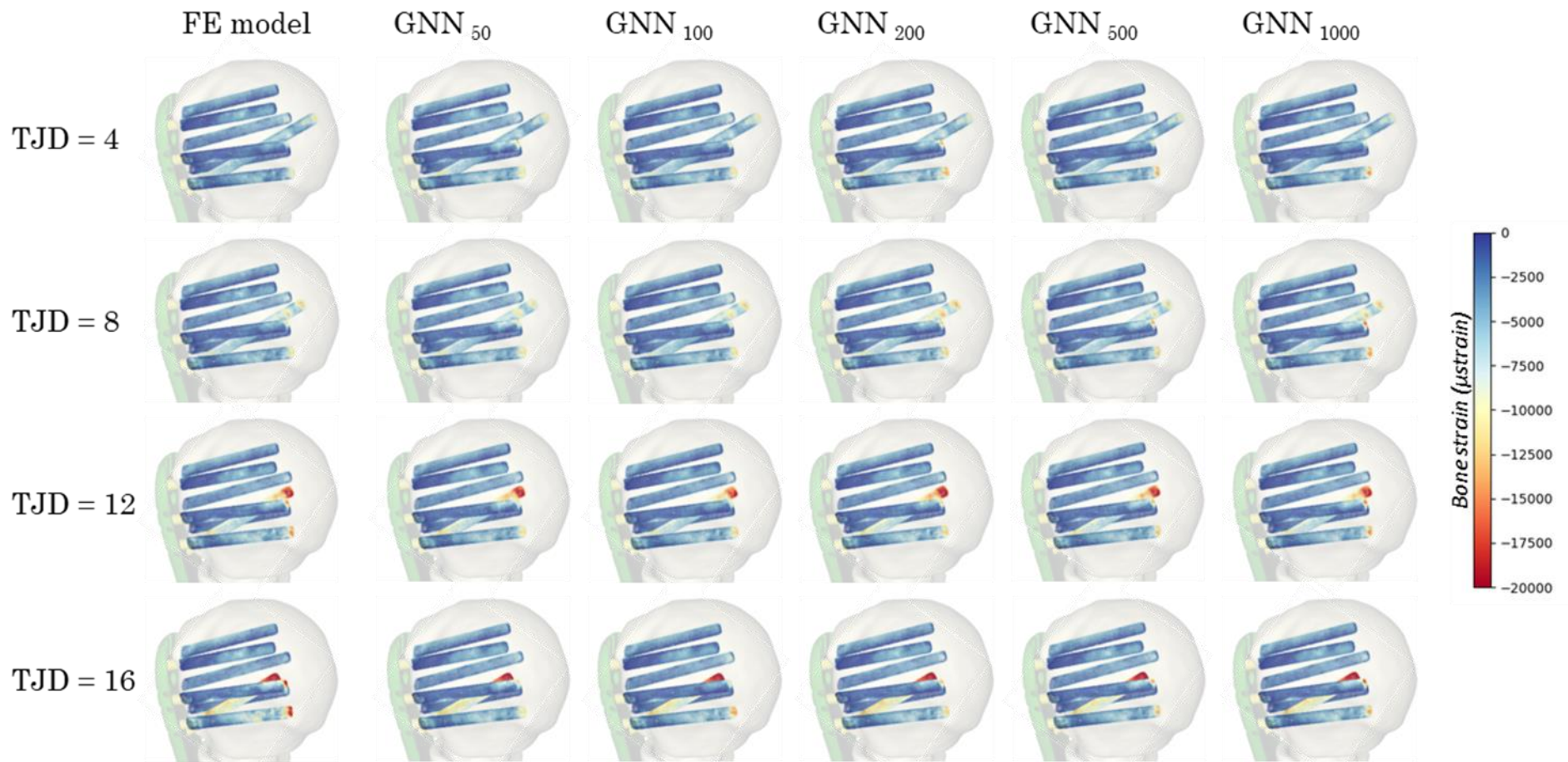
Figure 5.5 - *Bone principal strain distribution of the configurations having same TJD for screw 6. Results from the FE model and the best GNN*$_{ScrewLength}$ *trained with each training set size.*

5.3.2 Variation of screw orientation

The entire process for the FE analysis for each configuration, from selecting the screw angle to processing the outputs, took approximately 15-20 minutes. For the training process, it took 1, 4, 10, and 15 hours to train the GNN_ScrewDirection using training sets of 500, 2000, 5000, and 7500 data, respectively. Once trained, each GNN_ScrewDirection only took 15 seconds to predict the 500 FE configurations of the testing set (Table 5.6).
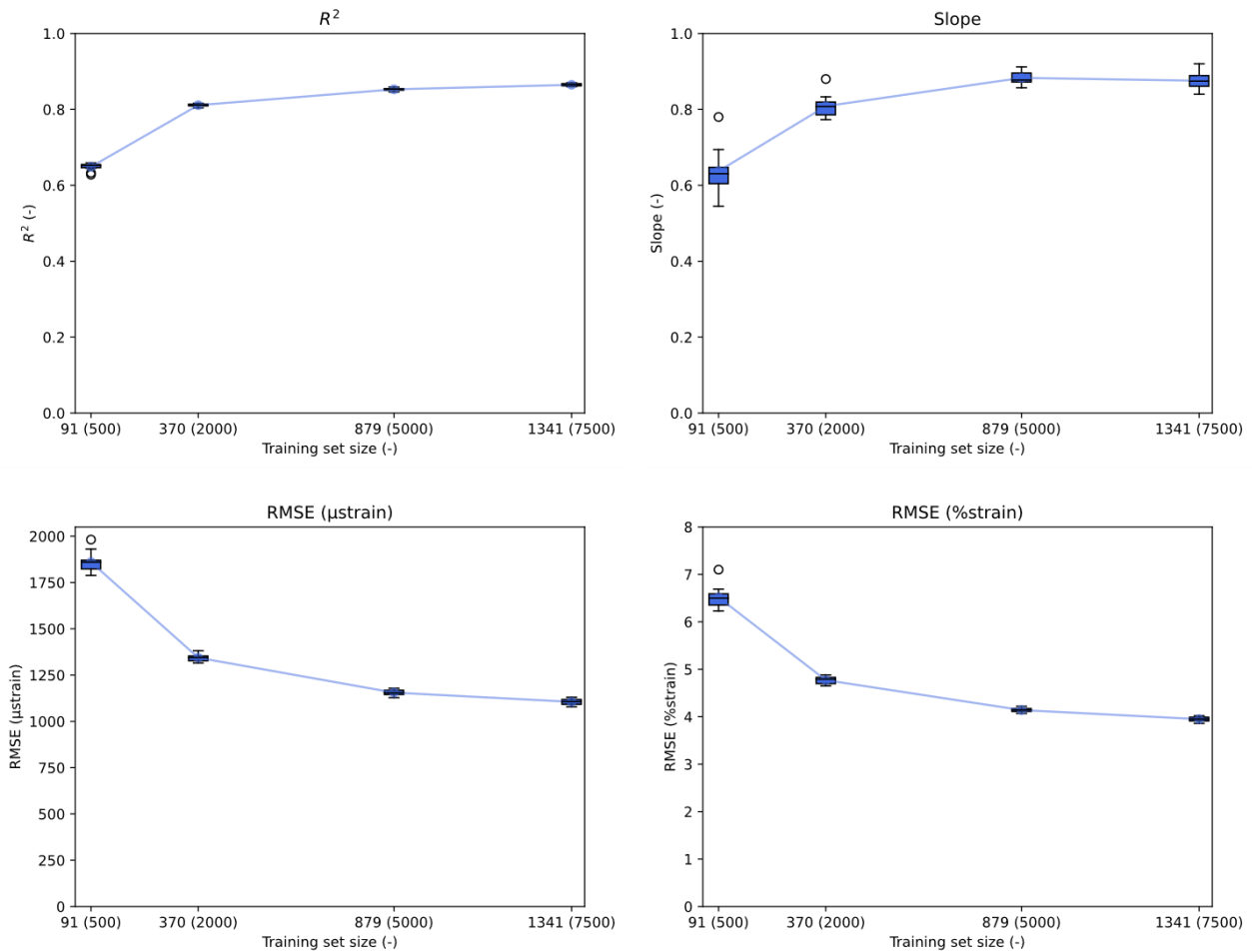
Figure 5.6 - Performance variation of the 10 GNN_ScrewDirection developed for each training set. The original size of the training set is shown with the reduced one in brackets.

When analysing the impact of the training set size, looking at the regression between the FE data and their prediction with the GNN_ScrewDirection models, it was observed that as the training set increased from 500 data to 7500, the R^2 increased from 0.66 to 0.87 and also the increased

from 0.69 to 0.89 Additionally, the RMSE decreased from 1788.21 to 1078.77 µstrain, indicating a reduction of 39.6% (Table 5.6, Figure 5.6).

In the analysis of the predictions of the configurations with the highest and lowest 90th percentile minimal principal strain, the impact of training set size was reported. The study found that increasing the training data led to improved accuracy, with the models better capturing high levels of strain along the cortical screw and the screws in the lower part of the humeral head (Figure 5.7). The results showed that models trained with more data produced outputs closer to the FE analysis, with the R^2 value increasing from 0.56-0.59 for the model trained with a smaller training set to 0.82-0.83 for those trained with a larger set. In terms of the %RMSE, the study observed a decrease from 10.67-9.3% of µstrain for the model trained with 500 data points to 6.66-5.68% of µstrain for the model trained with 7500 data points (Table 5.7).

A comparison was made between the configurations with the highest and lowest 90th percentile minimal principal strain predicted by the $GNN_{ScrewDirection}$ model and the ANN model developed in Chapter 4. The absolute percentage difference for both the 50th and 90th percentile minimal principal strain was calculated. The report showed a difference range of 0.33% to 19.99%, indicating that the training set size only influenced the 90th percentile minimal principal strain (Table 5.8).

Table 5.6 - Performance of the GNN$_{ScrewDirection}$ on the testing set of 91 simulations for the prediction of the minimal principal strain of the bone around all the screws with the variation of screw orientation. The influence of the training set size is shown. Results are displayed for the model with the lowest error, while the average of 100 models is shown in brackets.

Training set size (n)	FE Running time (h)	GNN training time (h)	R^2	Slope	RMSE (μstrain)	RMSE (%)
500 (92)	30.48	1	0.66 (0.65)	0.69 (0.63)	1788.21 (1859.24)	6.23 (6.51)
2000 (370)	123.12	4	0.81 (0.81)	0.82 (0.81)	1315.42 (1343.55)	4.65 (4.77)
5000 (879)	292.80	10	0.86 (0.85)	0.88 (0.88)	1127.99 (1153.56)	4.07 (4.14)
7500 (1,341)	448.80	15	0.87 (0.86)	0.89 (0.88)	1078.77 (1105.22)	3.86 (3.95)

Table 5.7 - *Regression of the minimal principal strain at the nodes between the FE data of the configurations with the highest and lowest 90th percentile minimal principal strain, and their prediction using the best GNN$_{ScrewDirection}$ for each training set size*

	Nodal minimal principal strain															
	R2				Slope				RMSE (µstrain)				RMSE (%)			
Training set size	500 (92)	2000 (370)	5000 (879)	7500 (1341)	500 (92)	2000 (370)	5000 (879)	7500 (1341)	500 (92)	2000 (370)	5000 (879)	7500 (1341)	500 (92)	2000 (370)	5000 (879)	7500 (1341)
Max 90th percentile strain	0.59	0.80	0.86	0.83	0.53	0.75	0.88	0.84	2478.0	1716.1	1355.5	1511.1	9.3	6.44	5.09	5.68
Min 90th percentile strain	0.56	0.72	0.77	0.82	0.73	0.77	0.83	0.86	1262.3	969.0	871.56	787.6	10.67	8.19	7.37	6.66

Table 5.8 - *Comparison of the configurations with the highest and lowest 90th percentile minimal principal strain predicted with the GNN$_{ScrewDirection}$ model and the ANN model developed in Chapter 4. The absolute difference between the 50th and the 90th percentile minimal principal strain values obtained from the GNN$_{ScrewDirection}$ and ANN models is reported.*

	% of error of 50th percentile minimal principal strain between the ANN predictions and the GNN$_{ScrewDirection}$ prediction				**% of error of 90th percentile minimal principal strain between the ANN predictions and the GNN$_{ScrewDirection}$ prediction**			
Training set size	500 (92)	2000 (370)	5000 (879)	7500 (1341)	500 (92)	2000 (370)	5000 (879)	7500 (1341)
Max 90th percentile strain	15.08	11.14	15.91	19.99	21.14	9.56	1.01	0.33
Min 90th percentile strain	9.01	10.19	10.67	7.98	7.34	10.67	5.23	2.61
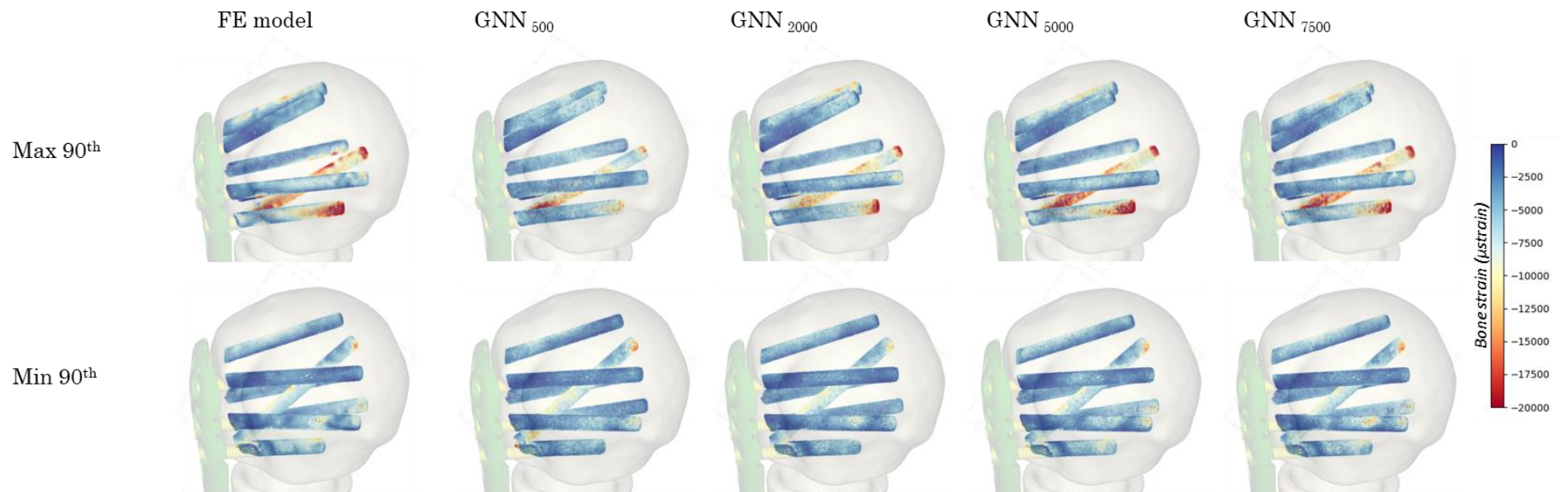
Figure 5.7 -Bone principal strain distribution of the configurations with max and min and 90th percentile principal bone strain. Results from the FE model and the best GNNScrewDirection trained with each training set size
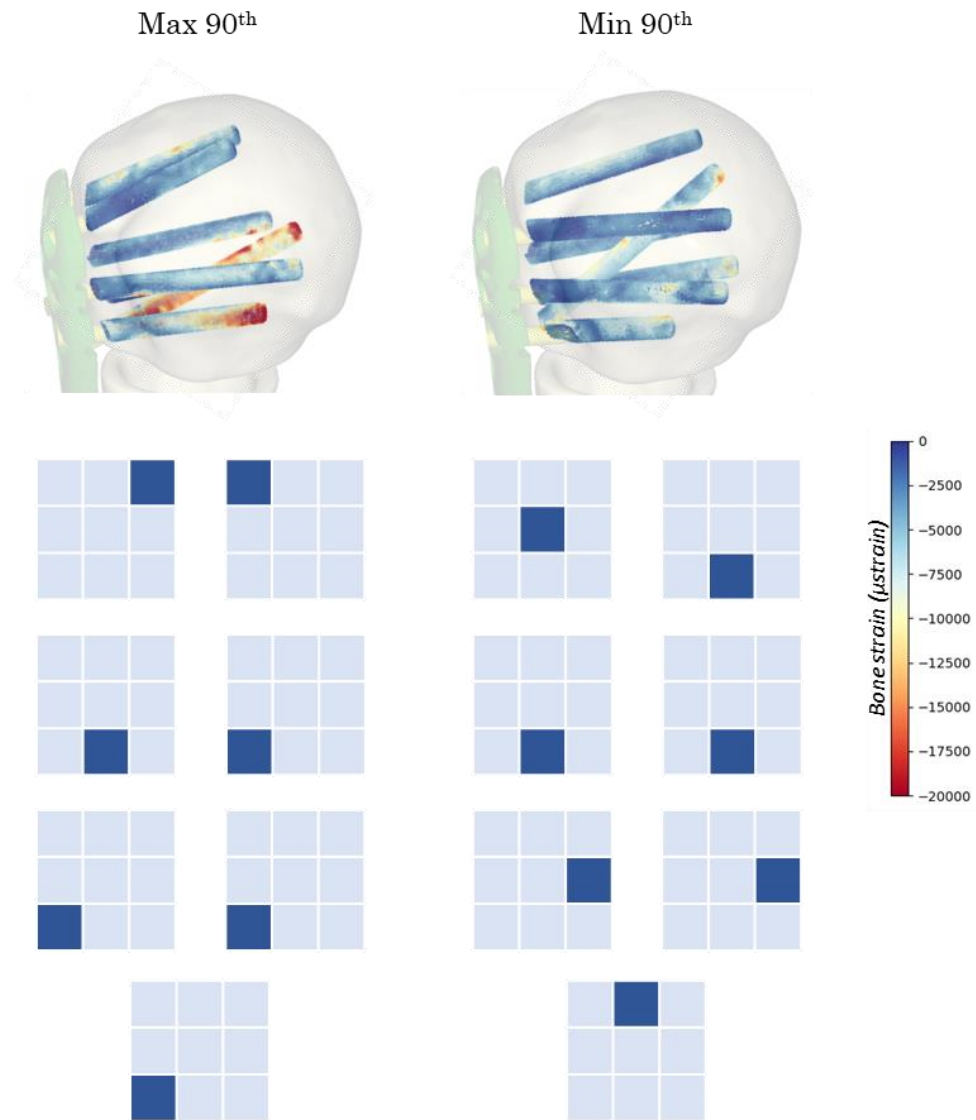
Figure 5.8 - Bone principal strain distribution of the configurations with max and min 90th percentile principal bone strain. Focus on the screw anterio-posterio and proximal-distal position.

5.4 Discussion

The aim of this study was to develop an advanced DL model to assist the FE analysis for quickly and accurately predicting bone strain varying implant parameters. Additionally, the aim was also to develop a technique capable of reproducing the field of strain inside the bone without losing accuracy. To achieve this, two different GNN models were developed to predict strain for each dataset developed in Chapters 3 and 4, which consisted in varying the screw length and screw orientation of the implant. Overall, the GNNs model efficiently predicted FE data and accurately estimated the distribution of bone strain, in terms of RMSE of the predictions. Focusing on the model's accuracy, the GNN models were able to make accurate predictions of strain distribution, for both datasets, showing a best RMSE of 2.71% of strain for the $GNN_{ScrewLength}$, and a best RMSE of 3.86% of strain for the $GNN_{ScrewDirection}$. However, the $GNN_{ScrewDirection}$ had lower accuracy compared to the $GNN_{ScrewLength}$. Despite the use of different structures for the two datasets, it is unlikely that this difference can solely account for the lower performance of $GNN_{ScrewDirection}$. Indeed, the Appendix already showed that the same structure used for $GNN_{ScrewLength}$ performed equally with the dataset that involved variation in screw direction (Appendix B: Supplementary material for Chapter 5). However, the lower performance of $GNN_{ScrewDirection}$ may be attributable to the higher variability of node positions in space, given that the screws were moved in proximal-distal and anterior-posterior positions. Moreover, due to the higher computational cost of the $GNN_{ScrewDirection}$ model and due to the limits of the available computing power, the maximum ledge length connectivity was set 0.5 mm lower than the one specified for the $GNN_{ScrewLength}$ problem. This adjustment may have contributed to a reduction in the accuracy of the results. Given the limited computing resources, with only a GPU equipped with 16GB of RAM, it would be beneficial to consider and incorporate more powerful computing options such as compute clusters or cloud-based GPUs in future works.

Overall, the GNN$_{ScrewLength}$ predictions for the cases shown in Figure 5.4 and Figure 5.5 were visually similar. The model effectively captured the strain distribution around the surface of the screws, demonstrating higher accuracy for the models trained with more data and showing strong regression. Similarly, the case shown in Figure 5.7 illustrated how the GNN$_{ScrewDirection}$ trained with less data was unable to capture the higher increase in strain.

When comparing the predictions from the ANN models and the GNN models for GNN$_{ScrewLength}$ and GNN$_{ScrewDirection}$, it was noticed that there was little difference in the predictions of the 90th percentile of strain for both datasets. However, a more noticeable variance was observed in the predictions of the 50th percentile of strain. This suggests that while the ANN and GNN approaches provided similar results in detecting higher strain values within the bone, there was a discrepancy in predicting the 50th percentile of bone strain. When comparing the computational requirements of the two methodologies, it is evident that the GNN approach demands more computational power and training time. However, it excels in predicting and displaying the actual strain distribution on the surface of interest, in the same amount of time required by the ANN method to give a prediction for the same amount of data.

The studies of this chapter were conducted on an NVIDIA RTX A400 GPU with 16 GB of RAM. The models could not be developed on the CPU and CUDA had to be installed on the computer in order to utilize the PyTorch Geometric library. Focusing on the computational time, for both types of datasets, the training time of the GNN models was strictly dependent on the training set size. In particular, the more data we used to train the model, the more accuracy we gained, but the more training time was necessary. Notably, the difference in model accuracy was less significant when comparing the two largest training sets than when comparing the largest set with the others. This indicates that the best predictions were achieved with the largest dataset, and further improvements would not be gained by increasing the amount of data. However, the training time of the GNNs was significantly lower compared to generating the same amount of FE data, demonstrating the computational efficiency of this DL approach. For a more

comprehensive analysis, the volume of the bone around the screws should be considered rather than just focusing on the surface area. This will allow for a more thorough examination of the effects of the implant's impact on the bone surrounding the proximal screws of the plate. The main current limitation is related to the constraints on computer resources. The analysis currently utilizes all of the GPU memory available, and to expand the analysis to a specific volume of interest, additional resources would be required.

This study is the first to aim to predict bone strain using an advanced computational technique for proximal humeral fractures with a fracture fixation device. Some studies in the literature have conducted parametric analyses to understand the behaviour of fracture fixation devices by varying different variables, implant and patient-related (Fletcher *et al.* 2019c, 2019a, 2019b; Mischler *et al.* 2020b; Varga *et al.* 2017, 2018, 2020). However, these studies have primarily focused on the mean strain around the screws and have not investigated the variation of strain distribution with the variation of the parameters of interest. Examining the internal distribution could be beneficial for gaining a better understanding of the problem. Analysing a problem with a very large solution space using only FE analysis is challenging due to the computational costs. DL models could potentially assist the FE analysis, improving computational power. Although GNNs have shown computational efficiency, they have not yet been fully explored for addressing a biomechanical problem using FE analysis.

A series of limitations need to be addressed. Firstly, the development of $GNN_{ScrewLength}$ and $GNN_{ScrewDirection}$ only involved the investigation of a small number of GNN structures. It is important to test additional algorithms, particularly a GNN algorithm specifically tailored to this type of problem. Furthermore, given the diverse methods within the DL branch, further investigations are necessary to determine if there are alternative methods that could effectively address this problem. Only the variation of a single parameter was considered. There are a high number of parameters related to the implant, such as screw length, screw orientation, and implant position, as well as patient-related factors like bone quality and fracture configurations,

which all play a role in the outcome of the treatment of proximal humeral fractures with a fracture plate (Lewis *et al.* 2021). Unfortunately, it is not clear which parameter is the most influential and how to reduce the failure rate. Therefore, a more complex model should be developed to address all these variables. The model was developed based on a single subject, so it was not capable of capturing and predicting the variations among different subjects. Indeed, to enhance the reliability of the model, it should be trained on a larger dataset that includes multiple subjects. The model was developed using only the nodal data on the bone surface around the screws. To effectively capture the entire field of strain in the bone, the model should be trained on volumetric data, potentially encompassing the entire bone segments. However, the increase of data for the training process would drastically increase the computational time and would also demand more GPU resources in terms of memory. Lastly, a computational limitation is that the model requires information at the nodal level to make predictions for new configurations. This requires transforming the surface of interest into nodal data, assigning material properties and coordinates in 3D space, and defining the connections between the nodes. While this process adds some computational cost, it reduces the data analysis to a smaller surface compared to the entire mesh of the fractured bone with the implant. Therefore, data extraction during the model settings adds some computational cost to the process, but they are very limited. Moreover, this cost is still relatively minor compared to running an FE analysis, and no mesh morphing is required. Overall, the GNN proved to be a very efficient and time-saving technique. The training time was shorter than the FE data generation, and the GNN model was able to make new predictions in just a few seconds. The study demonstrated how GNN can be used to improve the FE computational power for the analysis of a fracture fixation plate for proximal humeral fractures and potentially could be explored for other orthopaedic problems.

In this chapter, it was demonstrated that well-trained GNNs are capable of making accurate predictions. A high R^2 value was achieved even for the smallest training set when varying the

screw length. However, when varying the screw orientation, the R^2 was lower, likely due to the challenge of having nodes located in different positions due to the change in screw orientation. Nevertheless, the model was able to capture the strain distribution. Furthermore, the GNN technique proved to be highly time-efficient. The training time was much lower than the FE running time for the same amount of data, and once trained, it only took a few seconds to make predictions for a new scenario.

Chapter 6

Developing a Graph Neural Network model for a group of subjects to predict bone strain distributions in fracture fixation of the proximal humerus

6.1 Introduction

The current approach to analysing patient variability for a specific medical implant involves using FE, which can be computationally expensive, especially in cases with complex loading and boundary conditions, or when a high number of parameters and different implant configurations are involved. While simple surrogate models like ANN models could help reduce computational costs, they are not able to predict the distribution of strain or stress for a 3D geometry. Additionally, using various geometries and 3D meshes for subject-specific analysis can be a challenging and time-consuming process. Therefore, it is crucial to develop a more effective methodology.

Graph Neural Networks (GNN) are a type of Deep Learning (DL) method that operates on graph structures. GNNs are mesh-independent and can be used to predict different 3D geometries, therefore taking into account subject variability. Krokos et al used GNN methodology to predict stress within 3D porous structures while varying certain geometric features such as the size and positions of the pores (Krokos *et al*. 2022a). However, GNN techniques have not yet been explored in orthopaedic scenarios, especially for subject-specific analysis to predict bone strain with the use of a fracture fixation device.

In this chapter, the aim was to assess the ability of a GNN model to predict bone strains across patients, using subject-specific FE data of the humerus with a fracture fixation plate for a group of subjects. A semi-automated workflow was developed to perform FE analysis on CT scans of 434 subjects, each with a fracture fixation plate for a proximal humeral fracture. The parameters of the implant remained unchanged, and the FE data were used to train a GNN model. The main focus of this study was to assess the time efficiency and accuracy of the model.

6.2 Methods

The study was based on patient-specific humeri with fracture fixation plate models of 434 subjects. A semi-automated FE process was developed, and their data were used to train a GNN model.

6.2.1 FE model

For the generation of the FE data, CT images of 434 cadavers of > 60-year-old donors were collected from the New Mexico Decedent Image Database (NMDID) (Edgar *et al*. 2020). The cohort consisted of 203 females and 231 males. Synopsys' Simpleware™ AS Ortho module was used for the segmentation of 81 left and 353 right humeri (Version U-2022.12; Synopsys, Inc., Mountain View, USA). All the humeri were cut at 160mm in length from the top of the humeral head and a virtual osteotomy of a 5 mm gap was performed at the surgical neck, simulating a single two-part proximal humeral fractures AO/OTA 11-A2.1 (Foundation AO 2014). A fracture

fixation plate (Austofix, Adelaide, Australia) with twelve screws was virtually inserted in the bone, positioned 4mm posterior to the bicipital groove and 7mm distal to the top of the greater tubercle (Austofix, Adelaide, Australia). The proximal screw length was automatically set for each subject to maintain a constant Tip To Joint (TJD) distance of 8 mm (Fletcher *et al.* 2019a) and their orientation in the space was fixed. The diameter of all the screws was set to 3 mm. All materials were modelled as linear elastic, with a Poisson's ratio of 0.3. The screws and plate were made of titanium alloy, having a Young's modulus of 105 GPa. The Eggermont methodology was used to obtain each subject's Bone Mineral Density (BMD) (Eggermont *et al.* 2019), which was then used to convert to Young Modulus using the Morgan et al. equation (Morgan *et al.* 2003): $E(MPa) = 6850\rho_{app}^{1.49}$. An FE mesh with linear tetrahedral elements (C3D4), with an element edge length between 1 and 0.5 mm was generated using the Synopsys' Simpleware™ FE module. Refer to Appendix A for the mesh convergence. The contact between the bone and the screws and between the plate and the screws were set as tied. According to Röderer's experimental work (Röderer et al. 2013), the models were exposed to a vertical force to simulate an axial bending scenario. The model's geometry was then rotated by 25° around the humerus's anterior-posterior axis, in line with Bergmann's clinical study (Bergmann et al. 2007). A point external to the humerus, representing the midpoint between the condyles of the lower end of the upper arm bone, was fixed in all directions and connected to the nodes of the lower portion of the humerus (Figure 3.2). In addition, another external point was connected to a 20 mm-diameter circular area on the surface of the humeral head. This point was positioned 1 mm from the surface of the humeral head along the axis connecting the centre of the humeral head to the centre of the circular area. The second external point, which could only move vertically, was then subjected to a vertical force of 100N. Similar to the studies carried in chapter 3 and 4, these loading and boundary conditions choice were made based on the high quality studies developed from the AO foundation team. The outputs of interest were the principal components of the strain at the surface of the bone around the proximal screws, as it

had been shown to be a parameter of failure for fracture fixation plates used for proximal humeral fractures (Varga *et al.* 2017). The models were run in ABAQUS (Version 6.14-3, Dassault Systèmes, Vélizy-Villacoublay, France) as a standard implicit static analysis.

In order to be time efficient, the process of generating the FE data for each subject was partially automated using a chain of Matlab (Mathworks, Natick, MA, USA) scripts. The processes included the segmentation of the humeri using the autosegmentor tool, the translation of the reference system based on the anatomical references of each subject, the virtual osteotomy, the selection of screw length, the generation of the mesh, the selection of material property based oh CT image, the setting of the boundary and loading conditions on and then the FE simulation running and output analysis (Figure 6.1).
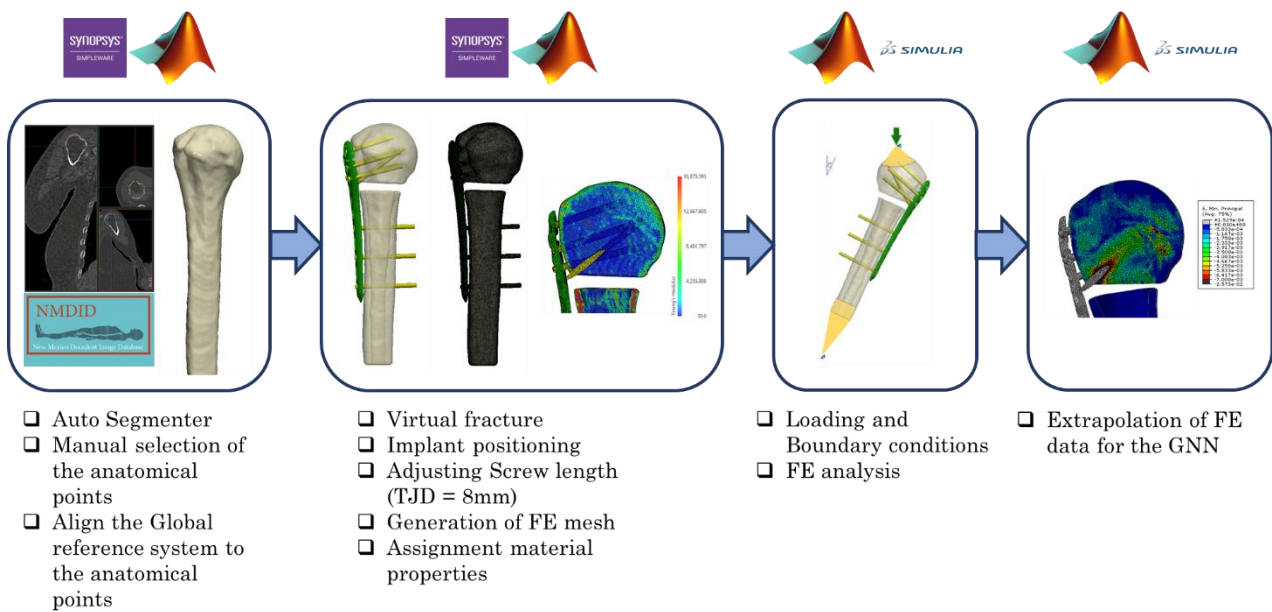
Figure 6.1 - Workflow developed for the generation of FE data

6.2.2   Dataset development

After running FE simulations for the 434 subjects, the data from each analysis was converted into graph structures consisting of a 3D network of nodes. The input and output data for

generating the graphs were extracted from the nodal information of the bone surface around the screws for each subject. These graph structures were then used to train the GNN models.

In particular, the input data were defined as follows:

- x, y, z coordinates of the nodes of the bone surface around the proximal screws, which were normalised. For each subject, the coordinate data were normalised based on the coordinates data of their cortical bone screw (screw 6), to reduce the variability of the data since the coordinate system was subject-specific, based on the anatomical reference points of the humerus.

- Young Modulus of the nodes of the bone surface around the proximal screws, which were normalised. For each node, this data was derived from the adjacent elements. The data of all subjects were normalised based on the Young Modulus data of a subject which was picked randomly.

- Edge connectivity matrix. The nodes' connectivity was determined by connecting each node to the nodes positioned with a maximum distance of 2 mm.

- Edge length distance of the connected nodes.

The output values used were the minimum, middle and maximum principal strain of the bone of the surface around the proximal screws. The data were defined at the nodal level after being extracted from the adjacent element.

The graphs obtained from the cohort of 434 subjects were divided into four groups of 100, 200, 300, and 400 subjects for the training set and one group of 34 subjects for the testing set. The training sets were used to train the GNN models, with the focus of investigating the amount of data needed to build an accurate model. The testing set was used to test the GNNs once they were trained.

6.2.3 GNN Development

The GNN structure used for this study was developed using PyTorch Geometric (Fey and Lenssen 2019), using mainly the structure ChebConv (Defferrard *et al.* 2016). The GNN consisted of an encoder, a processor and a decoder (Figure 6.2).

- The encoder consisted of a single block composed of a ChebConv, Dropout and ReLu layer. The input size of the ChebConv was 4, corresponding to the young modulus at the nodes and the x, y and z coordinates of the nodes, and the output size was set to 100.

- The processor consisted of two distinct blocks, each one composed of a ChebConv, Dropout and ReLu layer. The input and output size of each ChebConv layer was set to 100.

- The decoder consisted of a final distinct block of a ChebConv, Dropout and ReLu layer, and a final Linear layer. The input size of the ChebConv was 100, and the output was 3, corresponding to the principal component of strain at the nodes.
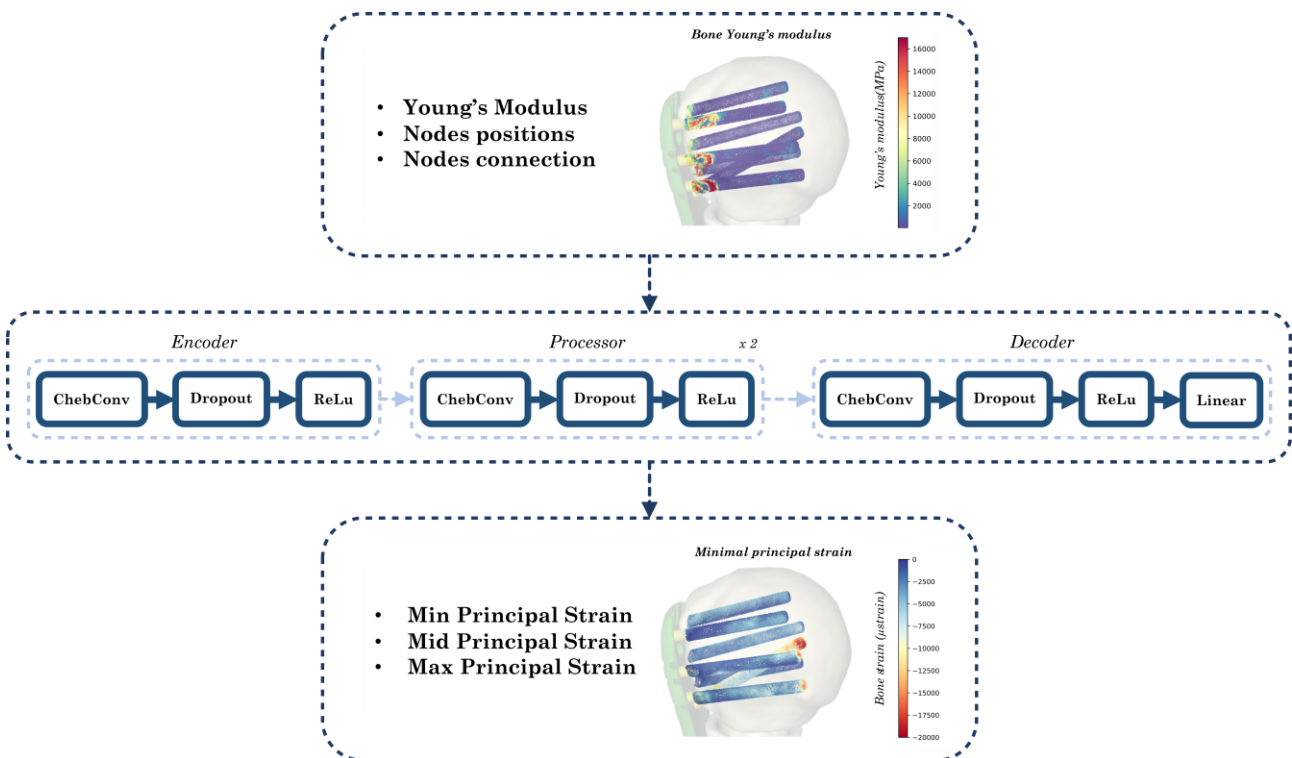
Figure 6.2 - Workflow of the GNN to investigate the subject variabilities

For each ChebConv the Chebyshev filter size K was 9. Each ChebConv elaborates information about young modulus, nodal coordinates, edge connectivity and edge length distance.

6.2.4 Analysis of GNN predictions

In order to evaluate the robustness of the GNN models, the training process was conducted 10 times for each training set, resulting in 10 unique GNN models for each training set. The data for the training sets were selected randomly from the total 434 graphs available. The accuracy of the GNN models was assessed using the remaining set of 34 FE models. Specifically, a regression analysis was performed comparing the nodal prediction of the strains around the screws from the FE real data to their predictions from the trained GNN model. In this analysis, the R^2 values, the slope, and the RMSE (in µstrain and %strain) were reported. The best GNN was considered to have the lowest % of RMSE. The GNN models were trained using a NVIDIA RTX A400 GPU with 16 GB of RAM.

Additionally, a categorical analysis was performed. Specifically, the 34 subjects from the testing sets were divided into three different groups based on the mean value of the BMD of the bone around the seven proximal screws. The groups were arbitrarily defined as subjects with a mean BMD value lower than 92 mg/cm^3, subjects with a value between 92 and 119 mg/cm^3, and subjects with a value higher than 119 mg/cm^3. A Kruskal-Wallis H-test was conducted between the GNN groups, and a Dunn-Bonferroni pairwise comparison was carried out if statistical difference was achieved (p*<0.05) to detect which groups are statistically different, with a p-value lower than 0.05. The same process was conducted for the FE groups. Finally, a Mann-Whitney U test was conducted between each FE and GNN for each division group within the same BMD range to determine any statistical differences, with a set p*<0.05.

6.3 Results

The FE analysis took approximately 37-42 minutes per subject, from the segmentation process of the humeri to extracting the outputs of interest after running the FE simulation. The training

of the GNNs with the training set of 100, 200, 300, and 400 data sets took 1, 2, 3, and 4 hours respectively. After the training process, each GNN took only 15 seconds to make predictions for the 34 unseen subject cases of the testing set.

Prior to selecting the model outlined in section 6.2.3, several GNN methods with different structures were analysed (Table 6.1). A single GNN for each structure was trained using increasing training sets ranging from 50 to 400. This helped to evaluate the accuracy of each model progressively. The final model was chosen based on having the highest R^2 and slope, as well as the lowest RMSE. Additionally, it demonstrated a shorter training time compared to the other models when using a training set of 400 subjects (Figure 6.3).

The following results come from the model explained in section 6.2.3. Focusing on the influence of training set size on the GNN accuracy, increasing the number of training data sets improved the prediction accuracy of the models (Figure 6.4). Looking at the regression results between the Minimal Principal Strain of the nodes around the screws of the FE models of the testing set and their predictions with the GNNs, the R^2 value increased from 0.69 for the training set of 100 data to 0.75 for the training set of 400 data. The GNN models under predict the strains, as the slope is always lower than 1. In particular, the slope increased from 0.78 for the training set of 100 data to 0.80 for the training set of 400 data. The RMSE had a reduction of 5.6% between the predictions from the GNN trained with 100 data and the one trained with 400, decreasing from a value of 759 µstrain to a value of 716 µstrain (Table 6.2).

The predictions of the Maximum Principal Strain had a similar trend to the prediction of Minimal Principal Strain. In particular, the R^2 value and the slope value increased respectively from a value of 0.67 and 0.76 to a value of 0.74 and 0.82, from the predictions made with the GNN trained with 100 data and the one trained with 400. With a similar decreasing rate, the RMSE value dropped from 750 to 665 µstrain (Table 6.2).

The GNN models were less efficient at predicting the Middle Principal strain values. In particular, a slight increase of R^2 was reported, respectively of 0.37 for the training set of 100 to 0.44 for the training set of 400. The absolute value of RMSE didn't change, but the % of RMSE slightly dropped, from 10.68% to 9.69% respectively for the training set of 100 and 400 data (Table 6.2).

In analysing the strain prediction based on the division of subjects into groups by their mean BMD, it was observed that the magnitude of both minimum and maximum strain increased as BMD decreased among the subject groups. Statistical significance was observed between the GNN group with the lowest BMD and the other two GNN groups in terms of predicting minimum and maximum strain. Additionally, there was a significant difference between the FE group with the lowest BMD and the FE group with the highest BMD for predicting strain. However, there was no statistically significant difference between every FE and GNN for each group, as the p-value was higher than 0.05 (Figure 6.5).

The GNN and FE predictions for strain in subjects with minimum, 25th percentile, 50th percentile, 75th percentile, and maximum BMD values over the testing set of 34 subjects have been documented. In the case with the highest BMD, no significant increase in strain was observed. However, for the subjects with the lowest BMD, an increase in strain was observed at the tip of the cortical screws, and the GNN predictions successfully captured this increase. It was observed that there was a gradual increase in BMD at the bottom of the more distal screws with the overall increase of BMD within subjects (Figure 6.6).

Table 6.1-Details of the GNN structures used and tested. X: nodes matrix. A: adjacency matrix. W: edge weight matrix. K: trainable filter, P$_i$: neighbouring point positions of x$_i$. γ$_Θ$, h$_Θ$: neural networks. α: parameters that learn how much information to keep in the mean.

GNN	Type	Input	Output	Algorithm
TAG (DU *ET AL.* 2017)	MessagePassing, topology adaptive graph convolutional networks operator	Node features Edge index Edge weights (opt)	Node features	$$X' = \sum_{k=0}^{K} (D^{-1/2} A D^{-1/2})^k X W_k$$ $$D_{ii} = \sum_{j=0} A_{ij}$$
XConv (LI *ET AL.* 2018)	Module, convolutional operator	Node features Positions Batch vector (opt)	Node features	$$x_i' = Conv(K, \gamma_\Theta(P_i - p_i) \times (h_\Theta(P_i - p_i) \parallel x_i))$$
ChebConv (DEFFERRARD *ET AL.* 2016)	MessagePassing, chebyshev spectral graph convolutional operator	Node features Edge indices Edge weights (opt) Batch vector (opt) Maximum lambda value (opt)	Node features	$$X' = \sum_{k=1}^{K} Z^{(k)} \cdot \Theta^{(k)}$$ $$Z^{(1)} = X$$ $$Z^{(2)} = \hat{L} \cdot X$$ $$Z^{(k)} = 2 \cdot \hat{L} \cdot Z^{(k-1)} - Z^{(k-2)}$$ $$\hat{L} = \frac{2L}{\lambda_{max}} - I$$
GraphNorm (CAI *ET AL.* 2021)	Module, graph normalization over individual graphs	Node features Batch vector (opt)	Node features	$$x_i' = \frac{x - \alpha \odot E[x]}{\sqrt{Var[x - \alpha \odot E[x]] + \epsilon}} \odot \gamma + \beta$$

147
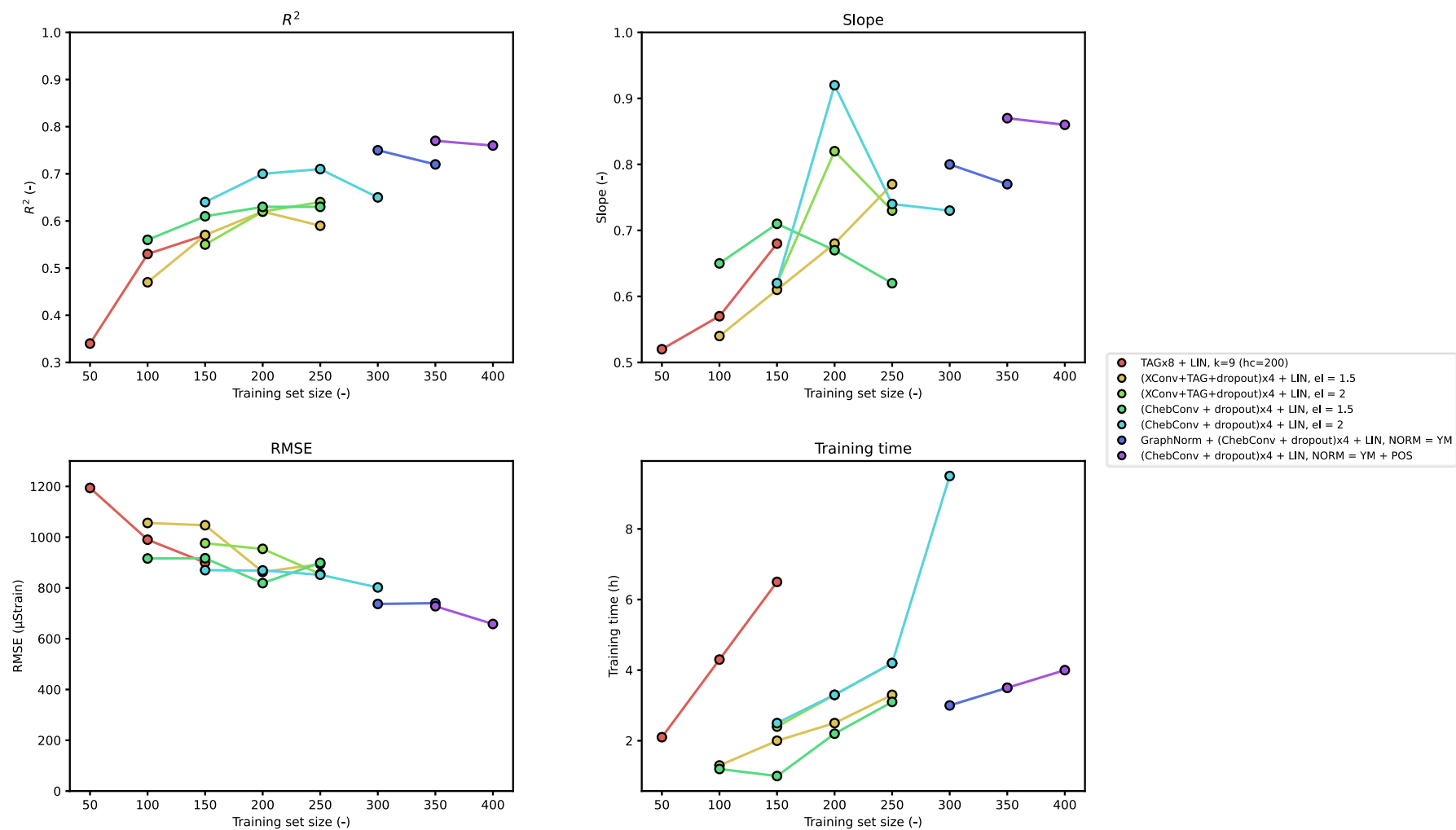
Figure 6.3 - Performance of different GNN models with the increase of training set size

Table 6.2 - Performance of the GNN on the testing set of 34 simulations for the prediction of minimal, middle and maximum principal strain of the bone around all the screws. The influence of the training set size is shown. Results are displayed for the model with the lowest error, while the average of 100 models is shown in brackets. The computational time difference between the FE and the GNN processes is also displayed

Training set size (n)	FE Running time (h)	GNN training time (h)	Minimal Principal strain				Mid Principal strain				Maximum principal strain			
			R^2	Slope	RMSE (µstrain)	RMSE (%)	R^2	Slope	RMSE (µstrain)	RMSE (%)	R^2	Slope	RMSE (µstrain)	RMSE (%)
100	61-70	1	0.71 (0.69)	0.71 (0.78)	374 (759)	5.50 (5.90)	0.37 (0.38)	0.39 (0.39)	104 (107)	10.68 (10.72)	0.70 (0.67)	0.77 (0.78)	702 (721)	6.54 (6.79)
200	122-140	2	0.73 (0.73)	0.81 (0.79)	696 (733)	5.03 (5.37)	0.41 (0.41)	0.41 (0.42)	105 (106)	9.54 (10.21)	0.72 (0.71)	0.80 (0.80)	656 (697)	6.04 (6.26)
300	183-210	3	0.75 (0.73)	0.79 (0.80)	718 (728)	5.13 (5.56)	0.44 (0.42)	0.42 (0.42)	102 (105)	9.44 (10.23)	0.76 (0.72)	0.80 (0.79)	669 (691)	5.96 (6.28)
400	244-280	4	0.76 (0.75)	0.74 (0.80)	751 (716)	4.93 (5.31)	0.44 (0.44)	0.44 (0.43)	102 (105)	9.69 (10.18)	0.74 (0.74)	0.85 (0.81)	660 (676)	5.63 (6.14)
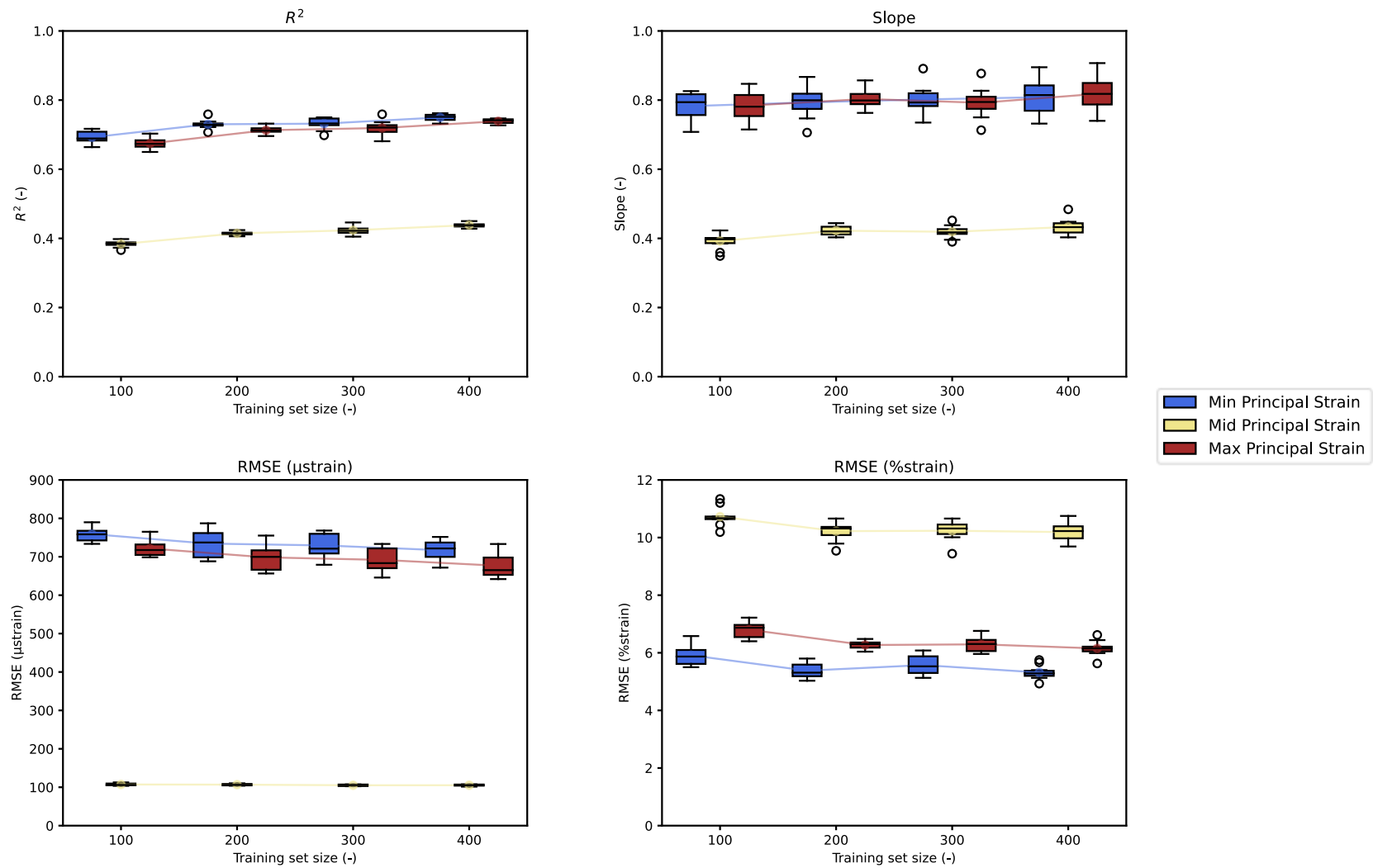
Figure 6.4 - Performance variation of the 10 GNNs developed for each training set. The variation for the training set prediction and the testing set prediction is shown
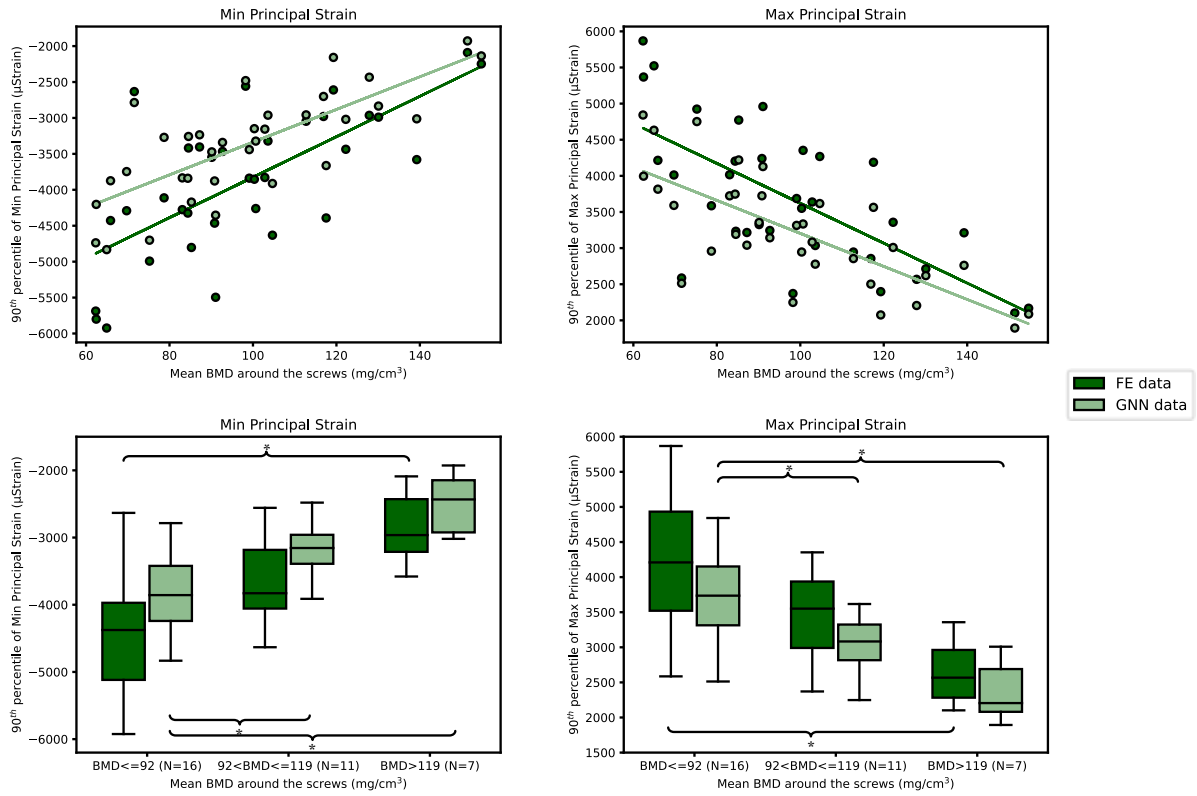
Figure 6.5 - Regression and categorical analysis for the FE and GNN data of the 34 subjects of the testing set.
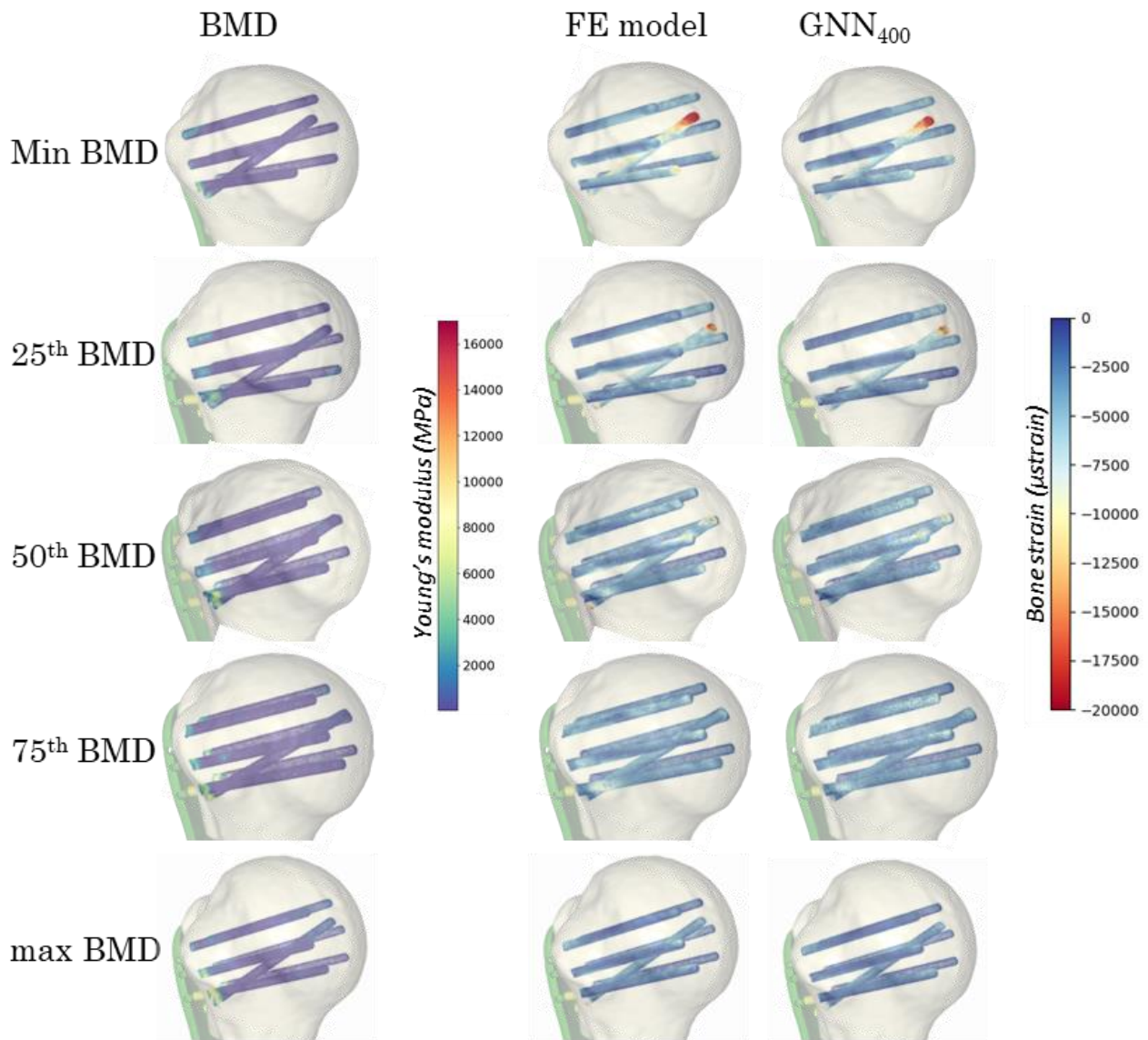
Figure 6.6 – BMD distribution and Min principal strain distribution of strain in the humerus for the subjects with minimum, 25th percentile, 50th percentile, 75th percentile and maximum value of BMD over the testing set of 34 subjects. Results from the FE model and the best GNN trained with the training set of size 400.

6.4 Discussion

Developing an efficient computational tool using an FE technique to predict the bone strain at a population level when applying an implant is a complex and time-consuming task. Indeed, FE analysis can be demanding, especially when conducting extensive analyses on multiple subjects. However, creating a more efficient tool has the potential to investigate the biomechanics of a specific implant in a more time-efficient manner, which could significantly improve surgical planning for implants and minimize the risk of failure for certain implant designs in the future.

The objective of this study was to develop an efficient computational model using a GNN technique and an FE analysis to predict the distribution of bone deformation in the humeral head among individuals with a proximal humeral fracture and a fracture fixation plate. To efficiently generate FE data for a cohort of 434 subjects, a semi-automated process was developed. Firstly, several architectures of GNN models were explored and ultimately the best one was selected based on its regression results between 35 FE data from the testing set and their predictions with the GNN model, as well as its computation time. The selected model was the most complex one, and it was demonstrated that increasing the complexity of the model led to an improvement in the predictions (Figure 6.3).

When analysing the impact of the training set size during the training process of the GNN models, it was observed that increasing the amount of training data led to improved accuracy in predicting Min, Mid, and Max bone strain, in terms of RMSE of the predictions. The predictions were more accurate for the Min and Max strain compared to the Mid strain, possibly due to the fact that the Mid Principal strain values are very close to zero, fluctuating between low positive and low negative values. This made it challenging for the model to identify consistent patterns (Figure 6.4, Table 6.2). The data presented in Figure 6.4 suggests that increasing the number of subjects in the training sets results in higher accuracy. This is evidenced by the decrease in RMSE from 5.50 to 4.93 % of µstrain for the Min Principal strain, from 10.68 to 9.69 % of µstrain for the Mid Principal strain, and from 6.54 to 5.63 % of µstrain for the Max Principal strain when transitioning from a training set of 100 data to a training set of 400 data (Figure 6.4). The values obtained for both the Min and Max principal strain appear to be within an acceptable range, as they are fairly close to 1 considering the absolute strain values. However, further improvements to the GNN model are needed to achieve a more substantial reduction in RMSE and to improve overall model performance.

In the analysis of the 90th percentile of the Min and Max principal strains around the screws for the 34 subjects in the testing set, there is a notable increasing trend in the Min principal strain

and a decreasing trend in the Max principal strain with higher values of BMD within subjects. In comparison to the FE analysis results, the GNN model underpredicts the magnitude of both the Min and Max principal strains. When the data were categorised into three groups based on their BMD values around the screws, a statistical difference was observed between the GNN group with the lowest BMD and the others. This demonstrates that the GNN model can effectively distinguish between subjects with low and high BMD values when presented with unseen data. Furthermore, GNN has the capability to accurately predict the distribution of strain on a three-dimensional surface, enabling accurate identification of areas with elevated strain (Figure 6.6).

After analysing the predictions of the GNN trained with each training set and the evaluation of the GNN's prediction of the 90th percentile of the Min and Max principal strain around the screws, it is evident that the models consistently underestimate the strain values (Figure 6.4, Figure 6.5). Interestingly, there is minimal variance between the results of the GNN from the training process and the GNN from the test analysis of unseen scenarios (Appendix C: Supplementary material for Chapter 6), making it unlikely that the model is overfitting. Therefore, it seems improbable that increasing the data for the training set would significantly enhance the GNN's performance. Rather, it is more likely that the model is underfitting, and it is probable that the GNN algorithms used for this specific problem may not be performing optimally, suggesting the exploration of alternative models.

This study is the first one to use an efficient computational technique to predict bone strain in proximal humeral fractures with a fracture fixation device for a cohort of subjects. Developing an efficient methodology for predicting bone deformation in multiple patients with a fracture fixation implant is challenging and time-consuming. Previous studies by the AO foundation group have developed a computational framework for generating FE data on a cohort of subjects, focusing on varying implant parameters (Fletcher *et al.* 2019b, 2019a, 2019c; Mischler *et al.* 2020b, 2022; Schader *et al.* 2021; Varga *et al.* 2018). Although advantageous in their parametric

approach, those models were developed on a limited number of subjects, between 20 and 47 subjects, and used only an FE approach. An efficient computational method is essential to reduce computational costs when dealing with a larger cohort of subjects. Ziaeipoor et al. conducted a comprehensive study using Principal Component Analysis (PCA) to predict the strain distribution in intact femurs across a large population. The study faced challenges related to the necessity use of mesh morphing for the model construction and predicting the strain for unseen subjects. Additionally, the model did not consider bone fractures, which is another limitation of the study (Ziaeipoor *et al.* 2020). This study innovatively used a combination of GNN and FE approaches to predict bone deformation in proximal humeral fractures treated with a fracture fixation plate, taking into account variations in subject geometries and material properties. The method was found to be computationally inexpensive, with the main computational cost being the generation of training data through the FE process. Indeed, the training duration was directly proportional to the amount of data in the training set, reaching a maximum of 4 hours. Furthermore, the model demonstrated the ability to accurately predict the 35 unseen cases.

A series of limitations should be addressed. Firstly, the loading and boundary conditions of the FE model were simplified and assumed a tied interface between the bone and the implant. Furthermore, the models were not experimentally validated. However, it is worth noting that similar conditions were employed in the majority of studies found in the literature (Fletcher *et al.* 2019b, 2019a, 2019c; Mischler *et al.* 2020b, 2022; Schader *et al.* 2021; Varga *et al.* 2017, 2018), some of which were experimentally validated (Varga *et al.* 2017, 2018). Another limitation is that only the surface bone data around the screws was analysed. Indeed, for a more comprehensive analysis, data from a volumetric region should be used. This approach, however, may result in a significant rise in the requirement for GPU resources and greater computational expenses for training the GNN models. It is essential to thoughtfully consider these factors, even though the use of volumetric data is likely to enhance the accuracy of the models by

increasing the amount of input data for each subject. This study has only focussed on the bone strain. Additional factors such as micromotions of the bone segments or implant stress could also be taken into account and, in principle, the GNN technique should be flexible enough to work with other type of variables, not limited to bone strain alone. The implant parameters remained unchanged, with only the subject variations being considered. In the future, it would be beneficial to first enhance the model's accuracy and then increase its complexity by introducing additional variables for analysis. These variables may include screw length, plate position, and screw orientation. Moreover, expanding the training data set beyond 400 subjects would also be beneficial. Lastly, for the development of this chapter, only a limited number of GNN models from the PyTorch library were explored. However, it is essential to expand the investigation by exploring a wider range of GNN and DL approaches with varying degrees of complexity. Additionally, it would be advantageous to explore other DL methods, such as advanced image-based Convolutional Neural Network (CNN) models or Generative Adversarial Networks (GANs) models. However, one challenge in exploring these techniques would involve converting the FE dataset into an image format, which could be time-consuming, especially considering the 3D geometry involved. The aim of this study was to create an efficient computational model using a GNN technique and an FE analysis to predict the distribution of bone deformation in the humeral head among individuals with a proximal humeral fracture and a fracture fixation plate. The methodology used in the study demonstrated that GNN and FE techniques can be effectively applied to this purpose. In addition to their efficiency and computational time advantages, using a GNN technique eliminates the need for mesh morphing processes, significantly reducing the computational time required compared to other simpler surrogate techniques or an extended FE analysis. Furthermore, GNN techniques provide enhanced flexibility in predicting unseen scenarios. Although the study was simplified, it suggests that by increasing complexity and incorporating additional parameters for investigation, a GNN approach could be developed to analyse biomechanical problems and enhance the traditional FE approach. Ideally, a computational approach such as the one

described in this chapter could, in the future, enhance surgical planning and assist surgeons in improving the outcome of medical applications.

Chapter 7

Discussion, Conclusion, and Future Work

7.1 Discussion

Understanding the biomechanical performance of a fracture fixation plate used to treat proximal humeral fractures is a complex task due to numerous influencing parameters. These parameters can be categorised as implant-related, including screw length, screw orientation, and plate position, as well as patient-related, such as bone quality, bone size, and fracture configuration. FE analysis is commonly used to study the biomechanical behaviour of these fixations' solutions and their worst-case scenario, but it can be computationally intensive when dealing with a large number of potential configurations. Previous studies that used FE analysis to investigate the performance of fracture fixation plates for proximal humeral fractures were limited to a small number of subjects, up to 47, and a restricted number of conditions, resulting in a few thousand simulations. This approach demonstrated that relying solely on FE analysis becomes extremely demanding when exploring a greater number of configurations and when aiming to explore the subject variabilities. The main objective of this thesis was to develop efficient computational models to assist the FE analysis and reduce computational time. Specifically, the focus was on creating computational models of varying complexity to examine the impact of different implant variables and subject variability.

Fracture fixation plates are considered the gold standard for treating proximal humeral fractures, particularly in elderly patients. Some FE analysis studies found in the literature have examined the impact of various implant parameters on bone deformation through parametric analysis. These parameters include screw length, screw orientation, plate position, and cement augmentation. However, these analyses are limited, as they only explore a few parameter combinations and they run a limited number of simulations (Fletcher *et al.* 2019b, 2019a, 2019c; Mischler *et al.* 2020b, 2022; Schader *et al.* 2021; Varga *et al.* 2018). Indeed, FE Analysis can be computationally demanding, especially when exploring multiple configurations and dealing with complex boundary and loading conditions. In this regard, DL methodologies could help and assist FE analysis, enhancing its computational capabilities. Training a DL model using FE data can lead to quick and accurate predictions for unforeseen scenarios, ultimately saving the computational time typically required for a purely FE-based process.

The first two studies carried out in this thesis consisted of combining an FE and an ANN approach to investigate the impact of two different implant characteristics on the humeral strain. Specifically, two distinct FE datasets were created by varying the length of the proximal screws of the implant and modifying the screw orientation within the humeral head. In the field of orthopaedic applications, only a limited number of studies in the literature have employed a combined FE and ANN model. For instance, these models have been used to predict femoral neck strains and fracture loads (Taylor et al. 2017), as well as femoral strain variations based on implant design features and bone quality (Chatterjee et al. 2019). Additionally, they have been used to detect bone displacement based on bone plate features for a femoral fracture (Dhason et al. 2023). However, none of these studies have investigated the influence of implant parameters of a fracture fixation plate on humeral strain. The studies in both Chapters 3 and 4 involved exploring up to a thousand simulations, resulting in a larger number of configurations compared to previous FE studies in the literature (Fletcher *et al.* 2019a; Jabran *et al.* 2019a; Mischler *et al.* 2020b; Schader *et al.* 2021). Additionally, training ANN models using FE data

offered several benefits, particularly in reducing the computational time required for predicting new, unseen scenarios. This approach provided a significant advantage, as the process of predicting new scenarios only took a few seconds and necessitated only the length or orientation data of the screws of the new configurations. Indeed, no mesh information is necessary to make predictions for new scenarios, which makes this method a quick and efficient tool for generating new predictions. However, one of the main limitations of the ANN approach is that, while it is computationally inexpensive, it does not provide a prediction of the local strain distribution, but rather only a generalized mean or maximum value of the strain occurring within the humerus.

To address this issue, the aim of the studies of Chapter 5 and Chapter 6 was to develop an efficient GNN model combined with the FE analysis, to make predictions of a distribution of strain with the variation of implant parameters and within subjects. Research in the literature has demonstrated that GNN models are suitable for analysing mesh structures and can effectively learn from structured data such as molecular structures and 3D geometries. This presents a significant advantage for FE analysis, particularly considering the computational cost involved, as the integration of GNN with FE analysis overcomes the need for mesh morphing when analysing multiple structures. An early notable study successfully combined FE and GNN approaches to predict stress within a porous structure while varying its porous characteristics. This study provided valuable insights into the current use of FE and GNN methodologies, although it was limited in that it did not account for variations in material properties inside the structure, focusing solely on the geometric variation of the pores (Krokos *et al.* 2022a). In the research outlined in Chapters 5 and 6, an innovative approach utilizing GNN was conducted to predict bone strain. Several GNN models were trained using the FE data developed in Chapters 3 and 4, as well as on a new dataset from a cohort of 434 subjects. The architecture of the GNN models was chosen from those available in the Pytorch Geometric library, after evaluating their performance. The GNN models demonstrated the ability to accurately predict new scenarios, effectively capturing the impact of implant and subject

variabilities. One significant advantage of this approach is the substantial reduction in computational time compared to FE analysis. While the model's training time could extend over several hours depending on the data volume, once trained, it takes only a few minutes to generate new predictions. This efficiency makes it highly suitable for real-time applications and surgical planning assistance.

The research presented in this thesis highlights the potential of using an FE dataset to develop an effective computational model for predicting unseen scenarios. The ANN model, while relatively simple, demonstrated its ability to provide a quantitative prediction of overall bone strain, rather than an actual distribution. This capability is valuable for assessing whether the strain levels are within a safe range or if the bone is approaching failure. The input data required to generate new predictions were straightforward, involving only implant features such as screw length and screw orientation. However, it is important to note a limitation of the model, as it does not incorporate patient-specific variables, given that the ANN models were trained on a single specimen. To incorporate patient-specific characteristics into the model, patient characteristics such as humerus length, diameter of the humeral head, and BMD need to be included. On the other hand, the GNN approach offers more advanced capabilities, as it is able to predict the distribution of strain, enabling real-time visual analysis of the impact of the parameter of interest.

The research outlined in this thesis offers significant promise for use in biomedical implant applications due to its high level of accuracy and computational efficiency, making it a valuable advancement in the field. Indeed, the methodologies used have the potential to support in vitro testing, provide guidelines for improved surgical practices and pre-operative planning, and even potentially enable real-time adjustments during surgeries. In the context of in vitro testing for regulatory approval, it is often necessary to test for worst-case scenarios. However, identifying such scenarios can be challenging in biomechanical problems that involve a wide solution space and a high number of variables. For instance, the biomechanical challenges of fracture fixation

plates for proximal humeral fractures, as described in the thesis introduction, exemplify these complexities. According to a study by Zdero et al, a review was performed to identify research studies that conducted experimental or computational analyses to explore the impact of plate and/or screw variables on fracture fixation performance. The review revealed that a wide range of variables were investigated, highlighting the potential challenges of comprehensively examining all variables using either experimental or finite element approaches (Zdero *et al.* 2024). More efficient computational techniques such as the ones described in this thesis hold the potential to identify worst-case scenarios for specific complex problems. Additional FE analysis can then be conducted to confirm the results from the surrogate models, followed by validation through mechanical testing. Moreover, the comprehensive analysis presented in this thesis has the potential to provide valuable guidance for surgeons in the application of fracture fixation plates and for other biomechanical scenarios. Chapters 3, 4, and 5 revealed that careful consideration should be given to the length and orientation of cortical screw (screw 6), as these factors greatly influence the strain in the bone and consequently the success of the implant, consistent with findings from previous studies (Fletcher *et al.* 2019b; Mischler *et al.* 2020b). For instance, the results suggest that a TJD ranging from 4 to 8 mm should be chosen to minimize the overall strain in the humerus, especially for the cortical screw, which also should be directed more proximally to reduce the risk associated with high strain. Therefore, it is important to carefully consider the length and orientation of the cortical screw based on this evidence. Additionally, Chapter 6 demonstrated that lower BMD results in increased bone strain, which could lead to failure of fixation (Mischler *et al.* 2020b). While previous research has examined the impact of screw length and poor bone quality, the use of a DL technique combined with an FE analysis could offer substantial benefits in examining scenarios where the effects of certain variables are not extensively documented and potentially assist surgeons in making more informed decisions. This approach may also facilitate the integration of multiple variables for the study of complex scenarios. Lastly, the techniques outlined in this thesis have the potential to serve as a highly efficient computational tool for surgical planning and for evaluating the

potential risk of failure for each individual patient. It is important to highlight that further substantial development is required to enhance the effectiveness of the tool. Nonetheless, the findings presented within this research offer valuable encouragement for the expansion and integration of the tool into practice.

7.2 Limitations

The ultimate goal of this project is to rapidly predict the strain for any patient with any fracture pattern while accounting for various implant variables. While this thesis has made significant progress, there is still some way to go in achieving this goal, and it is important to acknowledge the limitations associated with the FE and DL models developed in this thesis. In the FE analysis, the bone was modelled as a linear elastic material, and the equation used to derive the elastic modulus from the bone mineral density was based on the femoral head (Morgan *et al.* 2003). It is important to note that there is currently no literature addressing the humeral head in this context. Using a more complex material behaviour for the bone in the model would increase the FE running time, but it is unlikely to have a significant impact on the DL training and prediction time. Regarding the accuracy of the prediction, integrating a more complex bone behaviour could result in more realistic outcomes, although the impact may not be substantial (Zhou and Al-Qadi 2024). For all the FE studies, in order to simulate a simplified scenario a simplified fracture pattern was applied, resulting in an unrealistic scenario. For future applications and improvements of the model, it is essential to replicate a more realistic fracture patterns to achieve a more accurate scenario. The FE analysis used simplified loading and boundary conditions, focusing solely on an axial-bending scenario and excluding muscular forces from the model. Moreover, no sensitivity analysis was conducted to explore the impact of the variation in loading and boundary conditions. Increasing the complexity of the loading and boundary scenario, similar to the material properties, would lead to longer FE computational times and more realistic FE data. This would consequently improve the accuracy of DL models,

however, it would not necessarily increase the training and prediction time of the DL models. The FE and DL models were not experimentally validated. Nonetheless, the loading and boundary conditions used for the FE analysis reproduced those found in a group of works that have been experimentally validated. The technique explored in this thesis has the potential to be used to develop a computational tool capable of examining a broader spectrum of boundary and loading conditions, not limiting the investigation to a single scenario, resulting in a more realistic investigation. In this thesis, the analysis was focused specifically on the bone strain around the surface of the screws. This particular output was chosen because previous research has shown it to be a surrogate indicator of potential cut-out failure of the proximal humeral plate (Varga *et al.* 2017). Moreover, the AO Foundation's team conducted several studies to examine how variations in different parameters affect bone strain around screws. These studies have served as guidelines for the FE developed in this thesis (Fletcher *et al.* 2019b, 2019c, 2019a; Mischler *et al.* 2020a, 2020b, 2022; Schader *et al.* 2021; Varga *et al.* 2017, 2018, 2020). In order to carry experimental tests consistently with the literature, it is recommended to use a biaxial testing machine for conducting bending, axial rotation, and varus bending tests similar to the analysis conducted by Unger and Roderer (Röderer *et al.* 2013; Unger *et al.* 2012), and reproduced in the FE studies by the AO Foundation's group (Fletcher *et al.* 2019b, 2019c, 2019a; Mischler *et al.* 2020a, 2020b, 2022; Schader *et al.* 2021; Varga *et al.* 2017, 2018, 2020). However, these testing methods have the limitation of not being able to predict the strain inside the bone. Therefore, additional parameters such as micromotions and fracture gap distance should be used to validate the models developed in this thesis. For a more thorough analysis, Digital Volume Correlation (DVC) methods could be employed in conjunction with these testing methods to estimate the strain inside the bone and validate the models, although requiring a µCT scan machine (Roberts *et al.* 2014).

The ANN models developed in Chapters 3 and 4 were simplified. They were generated using data from a single subject, they did not account for patient-specific input and were not designed

to be transferrable to other individuals. Consequently, the model's applicability was limited to the specific subject on which it was trained.

The GNN models used in Chapters 5 and 6 were originally designed for other applications, and as a result, the algorithms were not originally customized for the specific biomechanical problem addressed in this thesis. Furthermore, only a limited number of GNN models have been used and tested. As a result, it would be highly beneficial to further investigate the efficacy of alternative GNN structures and to develop specific algorithms tailored to address the challenges of this particular biomechanical issue. A limitation of these GNN models is their dependence on mesh data from an unseen scenario to produce new predictions, which can lead to time-consuming processes, especially when dealing with complex scenario geometries.

Every study conducted in this thesis focused on addressing one variable at a time in order to simplify the problem and gradually increase complexity. For the studies conducted in Chapters 3, 4 and 5, the datasets developed were based on a single patient, and it was analysed the effect of a single implant variable. On the other hand, the dataset for Chapter 6 was generated on 434 subjects, but all the implant conditions were uniform. The studies conducted in every chapter were designed to analyse one variable at a time in order to test the capabilities of the FE-informed surrogate models in making predictions within a simplified scenario. The ultimate goal is to develop a complex model that can make predictions considering all variables influencing the problem. However, the reality and complexity of computational models suggest starting the analysis with a simplified approach and then progressing from there. The models developed in this thesis have demonstrated the ability to make accurate predictions, providing confidence to progress to more complex analyses.

Finally, to execute a surgical plan, a CT scan is often necessary, although it is not always a standard part of the care protocol. Moreover, a larger dataset of subjects is necessary for training the models presented in this thesis, which can be challenging since CT scans are typically less available in open sources and more costly to acquire. However, there is ongoing research focused

on developing models to convert 2D X-ray data, which are usually more available and less expensive, into 3D images, which could be advantageous in the future (Maken and Gupta 2023).

7.3 Future Work

The research conducted in this thesis represents the initial development of a more advanced computational tool to enhance the use of traditional FE analysis for the assessment of orthopaedic devices. This tool aims to evaluate the risk associated with fracture fixation plates for proximal humeral fractures, taking into account patient-specific factors and variability in implant design. By integrating FE with a DL approach, there is potential to significantly enhance the assessment of the surgical phase and improve the identification of risks associated with specific implant-related decisions. This computational tool holds promise for enhancing pre-clinical testing protocols and supporting surgeons in the planning of surgical procedures.

Looking ahead to future projects, the ANN methodology should be further explored by expanding it to a cohort of subjects in order to assess patient variability and by evaluating the impact of different fracture configurations. Additionally, it would be valuable to assess additional implant parameters, including screw configuration, plate position, and screw augmentation. Incorporating these parameters with the ones evaluated in Chapters 3 and 4 would enable to examine the collective impact of implant parameters and subject parameters. This comprehensive computational tool has the potential to assist surgeons in efficiently identifying high-risk scenarios associated with the implant procedure.

In addition to the GNN methodology developed in Chapters 5 and 6, there is a need to expand its scope to evaluate not only individual variables like screw length and subject variabilities but also to assess multiple implant variables across a cohort of subjects. Furthermore, similarly for the ANN, it is important to consider additional implant parameters such as plate position, and screw augmentation, and expand the subject's dataset to include different ethnicity groups, address gender biases, and enclose a wider age range. To enhance the accuracy of the DL

techniques, several adjustments and alternative explorations could be considered. For instance, by incorporating multiple variations, such as screw orientation and plate positioning, multiple GNN blocks trained on the individual variations could be used to assess these individual factors collectively, thereby improving overall accuracy. Furthermore, the region of interest could be expanded from the bone surface surrounding the screws to a volumetric region of the bone around the screws or even the entire humeral head. However, it's important to note that these modifications would considerably increase the computational requirements of the model during the training process. Although these implementations would increase the computational demands of the FE and DL models, the promising results presented in this thesis demonstrate that this methodology can function as a valuable computational tool for fracture fixation devices. Consequently, it is worthwhile to continue advancing these models and extend their application to more realistic scenarios.

In order to establish the reliability and applicability of the tool developed in this thesis, the methodology used should be replicated, which involves integrating FE data with a DL approach to investigate further orthopaedic issues outside the current scope. Similar to this thesis, an expanded dataset should be developed to generate data and evaluate the risk factors associated with various biomedical devices, such as joint implants and prosthetics. This approach will validate the effectiveness of the computational tool and expand its potential applications in the field of orthopaedics. The encouraging results outlined in this thesis provide confidence that the DL methods could be used to explore the impacts of various fracture fixation scenarios on different bone segments. In the context of other orthopaedic challenges, such as joint replacement, DL methods should be adapted to address specific issues, and they are likely to yield positive outcomes, provided that a comprehensive dataset is used during the model training process. For example, the integration of FE models with DL techniques can be used to investigate how various designs of knee prostheses affect the success of the implants across diverse populations. While FE models are well-established and widely applied in knee

replacement research, ongoing studies aim to enhance their design to improve their success rates. By training DL techniques on FE data, researchers can assess how different designs may influence implant loosening and, ultimately, failure. It is essential to train the model on an extensive dataset, particularly one that includes a wide range of subjects, to accurately capture the variability among individuals. This approach, following thorough experimental testing, could have some potential in the surgical planning of knee replacement. By providing personalised data, it can assist surgeons in selecting the most advantageous implant tailored to the individual needs of each patient.

The research conducted in this thesis demonstrated how leveraging DL techniques could be advantageous for the development and application of orthopaedic devices. To begin with, the Verification and Validation process of in silico trials for biomedical devices typically involves assessing the worst-case scenario, necessitating the analysis of various configurations. This process can be time-consuming when numerous variables need to be explored and a high number of potential configurations exist. In this context, DL techniques could facilitate the development of configurations that require analysis, thereby reducing the associated costs. It is important to establish standards for the development of DL models, as this could potentially enhance the in silico testing process for medical devices in the future. Furthermore, the implementation of efficient computational approaches, as explored in this thesis, could significantly aid in surgical planning for orthopaedic devices. Ideally, the creation of a database would allow surgeons to visualize the real-life impact of different implant choices, enabling them to make more informed decisions during surgical procedures.

7.4 Conclusion

The research presented in this thesis forms the groundwork for the creation of an effective computational tool to assist in the evaluation of fracture fixation plates for proximal humeral fractures. By integrating FE analysis with a DL approach, this work has enabled the exploration

of implant conditions and subject variability in a computationally efficient manner, demonstrating both time efficiency and reliability. Integrating such a tool into the design and development process of new fracture fixation designs as a standard practice could enhance the production of orthopaedic devices reducing the risk of failure. Additionally, this methodology could be integrated into a comprehensive tool for surgical planning, assisting surgeons to make informed decisions quickly while visualizing the impact of different variables on patients.

References

Al-Dirini, RMA, Martelli, S, and Taylor, M (2020) Computational efficient method for assessing the influence of surgical variability on primary stability of a contemporary femoral stem in a cohort of subjects. *Biomechanics and Modeling in Mechanobiology*, **19**(4), 1283–1295. doi:10.1007/s10237-019-01235-0.

Bah, MT, Nair, PB, Taylor, M, and Browne, M (2011) Efficient computational method for assessing the effects of implant positioning in cementless total hip replacements. *Journal of Biomechanics*, **44**(7), 1417–1422. doi:10.1016/j.jbiomech.2010.12.027.

Bergmann, G, Graichen, F, Bender, A, Kääb, M, Rohlmann, A, and Westerhoff, P (2007) In vivo glenohumeral contact forces-Measurements in the first patient 7 months postoperatively. *Journal of Biomechanics*, **40**(10). doi:10.1016/j.jbiomech.2006.10.037.

Bergmann, G, Graichen, F, Bender, A, … Westerhoff, P (2011) In vivo gleno-humeral joint loads during forward flexion and abduction. *Journal of Biomechanics*, **44**(8), 1543–1552. doi:10.1016/j.jbiomech.2011.02.142.

Boileau, P, and Walch, G (1997) The three-dimensional geometry of the proximal humerus. *Journal of Bone and Joint Surgery - Series B*, **79**(5), 857–865. doi:10.1302/0301-620X.79B5.7579.

Cai, T, Luo, S, Xu, K, He, D, Liu, TY, and Wang, L (2021) GraphNorm: A Principled Approach to Accelerating Graph Neural Network Training. *Proceedings of Machine Learning Research*, **139**, 1204–1215.

Cardoso, FA, and Gasparik, A (2023) *Activity and the shoulder: From soft tissues to bare bones Behaviour in our Bones: How Human Behaviour Influences Skeletal Morphology*.

doi:10.1016/B978-0-12-821383-4.00003-6.

Carofino, BC, and Leopold, SS (2013) Classifications in brief: The neer classification for proximal humerus fractures. *Clinical Orthopaedics and Related Research*, **471**(1), 39–43. doi:10.1007/s11999-012-2454-9.

Castro-Franco, AD, Mendoza-Muñoz, I, González-Ángeles, Á, Cruz-Sotelo, SE, Castañeda, AM, and Siqueiros-Hernández, M (2020) Trends in the characterization of the proximal humerus in biomechanical studies: A review. *Applied Sciences (Switzerland)*, **10**(18). doi:10.3390/APP10186514.

Charlton, IW, and Johnson, GR (2006) A model for the prediction of the forces at the glenohumeral joint. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, **220**(8), 801–812. doi:10.1243/09544119JEIM147.

Chatterjee, S, Dey, S, Majumder, S, RoyChowdhury, A, and Datta, S (2019) Computational intelligence based design of implant for varying bone conditions. *International Journal for Numerical Methods in Biomedical Engineering*, **35**(6), 1–17. doi:10.1002/cnm.3191.

Chen, H, Zhu, Z guo, Li, J tao, Chang, Z hao, and Tang, P fu (2020) Finite element analysis of an intramedulary anatomical strut for proximal humeral fractures with disrupted medial column instability: A cohort study. *International Journal of Surgery*, **73**(September 2019), 50–56. doi:10.1016/j.ijsu.2019.11.026.

Chen, YN, Chang, CW, Lin, CW, … Li, CT (2017) Numerical investigation of fracture impaction in proximal humeral fracture fixation with locking plate and intramedullary nail. *International Orthopaedics*, **41**(7), 1471–1480. doi:10.1007/s00264-017-3401-x.

Chung, SW, Han, SS, Lee, JW, … Kim, Y (2018) Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta Orthopaedica*, **89**(4), 468–473. doi:10.1080/17453674.2018.1453714.

Cronier, P, Pietu, G, Dujardin, C, Bigorre, N, Ducellier, F, and Gerard, R (2010) The concept of locking plates. *Orthopaedics and Traumatology: Surgery and Research*, **96**(4 SUPPL.). doi:10.1016/j.otsr.2010.03.008.

Dahan, G, Trabelsi, N, Safran, O, and Yosibash, Z (2019) Finite element analyses for predicting anatomical neck fractures in the proximal humerus. *Clinical Biomechanics*, **68**(May), 114–121. doi:10.1016/j.clinbiomech.2019.05.028.

Defferrard, M, Bresson, X, and Vandergheynst, P (2016) Convolutional neural networks on

graphs with fast localized spectral filtering. *Advances in Neural Information Processing Systems*, (Nips), 3844–3852.

Dhason, R, Roy, S, and Datta, S (2023) Surrogate model assisted design optimization of composite bone plates to achieve selective stress shielding. *Materials and Manufacturing Processes*, **38**(16), 2034–2043. doi:10.1080/10426914.2023.2195914.

Dickerson, CR, Chaffin, DB, and Hughes, RE (2007) A mathematical musculoskeletal shoulder model for proactive ergonomic analysis. *Computer Methods in Biomechanics and Biomedical Engineering*, **10**(6), 389–400. doi:10.1080/10255840701592727.

Domingue, G, Garrison, I, Williams, R, Riehl, JT, and Trauma, CO (2021) Management of proximal humeral fractures: a review. *Current Orthopaedic Practice*, **32**(4), 339–348.

Donaldson, FE, Jr, EN, and Coburn, JC (2015) Prediction of contact mechanics in metal-on-metal Total Hip Replace- ment for parametrically comprehensive designs and loads. *Journal of Biomechanics*, **48**, 1828–1835. doi:10.1016/j.jbiomech.2015.04.037.

Drake, RL, Vogl, W, Mitchell, AWM, and Gray, H (2010) *Gray's Anatomy for Students*.

Du, J, Zhang, S, Wu, G, Moura, JMF, and Kar, S (2017) Topology Adaptive Graph Convolutional Networks., 1–13. Retrieved from http://arxiv.org/abs/1710.10370

Edgar, H, Daneshvari Berry, S, Moes, E, Adolphi, N, Bridges, P, and Nolte, K (2020) New Mexico Decedent Image Database., Office of the Medical Investigator, University of New Mexico. Retrieved from doi.org/10.25827/5s8c-n515

Eggermont, F, Verdonschot, N, van der Linden, Y, and Tanck, E (2019) Calibration with or without phantom for fracture risk prediction in cancer patients with femoral bone metastases using CT-based finite element models. *PLoS ONE*, **14**(7), 1–13. doi:10.1371/journal.pone.0220564.

Egol, KA, Kubiak, EN, Fulkerson, E, Kummer, FJ, and Koval, KJ (2004) Biomechanics of locked plates and screws. *Journal of Orthopaedic Trauma*, **18**(8), 488–493. doi:10.1097/00005131-200409000-00003.

Fang, R, Ji, A, Zhao, Z, Long, D, and Chen, C (2020) A regression orthogonal biomechanical analysis of internal fixation for femoral shaft fracture. *Biocybernetics and Biomedical Engineering*, **40**(3), 1277–1290. doi:10.1016/j.bbe.2020.07.006.

Feerick, EM, Kennedy, J, Mullett, H, FitzPatrick, D, and McGarry, P (2013) Investigation of metallic and carbon fibre PEEK fracture fixation devices for three-part proximal humeral

fractures. *Medical Engineering and Physics*, **35**(6), 712–722. doi:10.1016/j.medengphy.2012.07.016.

Fey, M, and Lenssen, JE (2019) Fast Graph Representation Learning with PyTorch Geometric., (1), 1–9. Retrieved from http://arxiv.org/abs/1903.02428

Fitzpatrick, CK, Hemelaar, P, and Taylor, M (2014) Computationally efficient prediction of bone-implant interface micromotion of a cementless tibial tray during gait. *Journal of Biomechanics*, **47**(7). doi:10.1016/j.jbiomech.2014.02.018.

Fleischhacker, E, Siebenbürger, G, Helfen, T, Gleich, J, Böcker, W, and Ockert, B (2021) Varus malposition relates to functional outcomes following open reduction and internal fixation for proximal humeral fractures: A retrospective comparative cohort study with minimum 2 years follow-up. *Injury*, **52**(3), 506–510. doi:10.1016/j.injury.2020.09.003.

Fletcher, JWA, Windolf, M, Grünwald, L, Richards, RG, Gueorguiev, B, and Varga, P (2019a) The influence of screw length on predicted cut-out failures for proximal humeral fracture fixations predicted by finite element simulations. *Archives of Orthopaedic and Trauma Surgery*, **139**(8), 1069–1074. doi:10.1007/s00402-019-03175-x.

Fletcher, JWA, Windolf, M, Richards, RG, Gueorguiev, B, Buschbaum, J, and Varga, P (2019b) Importance of locking plate positioning in proximal humeral fractures as predicted by computer simulations. *Journal of Orthopaedic Research*, **37**(4). doi:10.1002/jor.24235.

Fletcher, JWA, Windolf, M, Richards, RG, Gueorguiev, B, and Varga, P (2019c) Screw configuration in proximal humerus plating has a significant impact on fixation failure risk predicted by finite element models. *Journal of Shoulder and Elbow Surgery*, **28**(9), 1816–1823. doi:10.1016/j.jse.2019.02.013.

Foundation AO (2014) Müller AO Classification of Fractures—Long Bones.

Fukunaga, T, Miyatani, M, Tachi, M, Kouzaki, M, and Kawakami, Y (2001) Muscle volume is a major determinant of joint torque in humans. *Acta Physiologica Scandinavica*, **174**(4), 239–300.

Gardner, MJ, Weil, Y, Barker, JU, Kelly, BT, Helfet, DL, and Lorich, DG (2007) The importance of medial support in locked plating of proximal humerus fractures. *Journal of Orthopaedic Trauma*, **21**(3), 185–191. doi:10.1097/BOT.0b013e3180333094.

Garner, BA, and Pandy, MG (2001) Musculoskeletal model of the upper limb based on the visible human male dataset. *Computer Methods in Biomechanics and Biomedical Engineering*,

4(2), 93–126. doi:10.1080/10255840008908000.

Georgousis, S, Kenning, MP, and Xie, X (2021) Graph Deep Learning: State of the Art and Challenges. *IEEE Access*, **9**, 22106–22140. doi:10.1109/ACCESS.2021.3055280.

Gille, J, Schulz, AP, Queitsch, C, Paech, A, and Jürgens, C (2008) Initial Results of the S3-Humerus Plate., 133–136.

Greiwe, RM (2015) Proximal humerus fractures: Percutaneous fixation, proximal humeral nailing, and open reduction and internal fixation. In *Shoulder and elbow trauma and its complications*, Elsevier Ltd., , 83–112. doi:10.1016/B978-1-78242-449-9.00005-4.

Hambli, R (2011) Apparent damage accumulation in cancellous bone using neural networks. *Journal of the Mechanical Behavior of Biomedical Materials*, **4**(6), 868–878. doi:10.1016/j.jmbbm.2011.03.002.

Hashem, S (1997) Optimal linear combinations of neural networks. *Neural Networks*, **10**(4), 599–614. doi:10.1016/S0893-6080(96)00098-6.

Heidari, N, Hedegaard, L, and Iosifidis, A (2022) *Graph convolutional networks Deep Learning for Robot Perception and Cognition*, Elsevier Inc. doi:10.1016/B978-0-32-385787-1.00009-9.

Helgason, B, Perilli, E, Schileo, E, Taddei, F, Brynjólfsson, S, and Viceconti, M (2008) Mathematical relationships between bone density and mechanical properties: A literature review. *Clinical Biomechanics*, **23**(2), 135–146. doi:10.1016/j.clinbiomech.2007.08.024.

Högfors, C, Karlsson, D, and Peterson, B (1995) Structure and internal consistency of a shoulder model. *Journal of Biomechanics*, **28**(7), 767–777. doi:10.1016/0021-9290(94)00134-P.

Holloway, KL, Bucki-Smith, G, Morse, AG, … Pasco, JA (2015) Humeral Fractures in South-Eastern Australia: Epidemiology and Risk Factors. *Calcified Tissue International*, **97**(5), 453–465. doi:10.1007/s00223-015-0039-9.

Holzbaur, KRS, Murray, WM, and Delp, SL (2005) A model of the upper extremity for simulating musculoskeletal surgery and analyzing neuromuscular control. *Annals of Biomedical Engineering*, **33**(6), 829–840. doi:10.1007/s10439-005-3320-7.

Holzbaur, KRS, Murray, WM, Gold, GE, and Delp, SL (2007) Upper limb muscle volumes in adult subjects. *Journal of Biomechanics*, **40**(4), 742–749. doi:10.1016/j.jbiomech.2006.11.011.

Huri, G, Familiari, F, Moon, YL, Doral, MN, and Muccioli, GMM (2020) *Shoulder arthroplasty:*

the shoulder club guide Springer International Publishing, Springer International Publishing.

Huri, G, and Paschos, NK (2017) *The Shoulder*, London: Springer International Publishing. Retrieved from https://link.springer.com/10.1007/978-1-4471-4462-5_2

Huttunen, TT, Launonen, AP, Pihlajamäki, H, Kannus, P, and Mattila, VM (2012) Trends in the surgical treatment of proximal humeral fractures – a nationwide 23-year study in Finland. *BMC Musculoskeletal Disorders*, **13**(261).

Iglesias-Rodríguez, S, Domínguez-Prado, DM, García-Reza, A, … Castro-Menéndez, M (2021) Epidemiology of proximal humerus fractures. *Journal of Orthopaedic Surgery and Research*, **16**(1), 1–11. doi:10.1186/s13018-021-02551-x.

Inman, VT, and Abbott, LC (1991) Observations of the Function of the Shoulder Joint. *Clinical Orthopaedics and Related Research*, 3–12.

Inzana, JA, Varga, P, and Windolf, M (2016) Implicit modeling of screw threads for efficient finite element analysis of complex bone-implant systems. *Journal of Biomechanics*, **49**(9), 1836–1844. doi:10.1016/j.jbiomech.2016.04.021.

Jabran, A, Peach, C, and Ren, L (2018) Biomechanical analysis of plate systems for proximal humerus fractures: A systematic literature review. *BioMedical Engineering Online*, **17**(1), 1–30. doi:10.1186/s12938-018-0479-3.

Jabran, A, Peach, C, Zou, Z, and Ren, L (2019a) Biomechanical comparison of screw-based zones of a spatial subchondral support plate for proximal humerus fractures. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, **233**(3), 372–382. doi:10.1177/0954411919827984.

Jabran, A, Peach, C, Zou, Z, and Ren, L (2019b) Parametric Design Optimisation of Proximal Humerus Plates Based on Finite Element Method. *Annals of Biomedical Engineering*, **47**(2), 601–614. doi:10.1007/s10439-018-02160-6.

Jiang, C, and Chen, NZ (2023) Graph Neural Networks (GNNs) based accelerated numerical simulation. *Engineering Applications of Artificial Intelligence*, **123**(April), 106370. doi:10.1016/j.engappai.2023.106370.

Jin, Y (2005) A comprehensive survey of fitness approximation in evolutionary computation. *Soft Computing*, **9**(1), 3–12. doi:10.1007/s00500-003-0328-5.

Jin, Y (2011) Surrogate-assisted evolutionary computation: Recent advances and future

challenges. *Swarm and Evolutionary Computation*, **1**(2), 61–70. doi:10.1016/j.swevo.2011.05.001.

Kamer, L, Noser, H, Popp, AW, Lenz, M, and Blauth, M (2016) Computational anatomy of the proximal humerus: An ex vivo high-resolution peripheral quantitative computed tomography study. *Journal of Orthopaedic Translation*, **4**, 46–56. doi:10.1016/j.jot.2015.09.006.

Karataglis, D, Stavridis, SI, Petsatodis, G, Papadopoulos, P, and Christodoulou, A (2011) New trends in fixation of proximal humeral fractures: A review. *Injury*, **42**(4), 330–338. doi:10.1016/j.injury.2010.10.016.

Kavuri, V, Bowden, B, Kumar, N, and Cerynik, D (2018) Complications Associated with Locking Plate of Proximal Humerus Fractures. *Indian Journal of Orthopaedics*, **52**(may), 161–169. doi:10.4103/ortho.IJOrtho.

Kennedy, J, Feerick, EM, McGarry, P, FitzPatrick, D, and Mullett, H (2013) Effect of calcium triphosphate cement on proximal humeral fracture osteosynthesis: a finite element analysis. *Journal of Orthopaedic Surgery (Hong Kong)*, **21**(2), 167–172. doi:10.1177/230949901302100210.

Kim, JD, Kim, NS, Hong, CS, and Oh, CY (2011) Design optimization of a xenogeneic bone plate and screws using the Taguchi and finite element methods. *International Journal of Precision Engineering and Manufacturing*, **12**(6), 1119–1124. doi:10.1007/s12541-011-0149-x.

Kim, SH, Szabo, RM, and Marder, RA (2012) Epidemiology of humerus fractures in the United States: Nationwide emergency department sample, 2008. *Arthritis Care and Research*, **64**(3), 407–414. doi:10.1002/acr.21563.

Königshausen, M, Kübler, L, Godry, H, Citak, M, Schildhauer, TA, and Seybold, D (2012) Clinical outcome and complications using a polyaxial locking plate in the treatment of displaced proximal humerus fractures. A reliable system? *Injury*, **43**(2), 223–231. doi:10.1016/j.injury.2011.09.024.

Kralinger, F, Blauth, M, Goldhahn, J, … Hanson, B (2014) The influence of local bone density on the outcome of one hundred and fifty proximal humeral fractures treated with a locking plate. *Journal of Bone and Joint Surgery - American Volume*, **96**(12), 1026–1032. doi:10.2106/JBJS.M.00028.

Krappinger, D, Roth, T, Gschwentner, M, … Kralinger, F (2012) Preoperative assessment of the

cancellous bone mineral density of the proximal humerus using CT data. *Skeletal Radiology*, **41**(3). doi:10.1007/s00256-011-1174-7.

Krokos, V, Bordas, SPA, and Kerfriden, P (2022a) A Graph-based probabilistic geometric deep learning framework with online physics-based corrections to predict the criticality of defects in porous materials. *ArXiv*, 1–48. Retrieved from http://arxiv.org/abs/2205.06562

Krokos, V, Bordas, SPA, and Kerfriden, P (2024) A graph-based probabilistic geometric deep learning framework with online enforcement of physical constraints to predict the criticality of defects in porous materials. *International Journal of Solids and Structures*, **286–287**(March 2023), 112545. doi:10.1016/j.ijsolstr.2023.112545.

Krokos, V, Bui Xuan, V, Bordas, SPA, Young, P, and Kerfriden, P (2022b) A Bayesian multiscale CNN framework to predict local stress fields in structures with microscale features. *Computational Mechanics*, **69**(3), 733–766. doi:10.1007/s00466-021-02112-3.

Kubiak, EN, Fulkerson, E, Strauss, E, and Egol, KA (2006) The evolution of locked plates. *Journal of Bone and Joint Surgery*, **88**(SUPPL. 4), 189–200. doi:10.2106/00004623-200612001-00020.

Lasanianos, NG, Kanakaris, NK, and Giannoudis, P V. (2015) Proximal Humeral Fractures. In *Trauma and Orthopaedic Classifications: A Comprehensive Overview*, Springer, London, , 51–55. doi:10.1007/978-1-4471-6572-9_12.

Le, L, Jabran, A, Peach, C, and Ren, L (2019) Effect of screw thread length on stiffness of proximal humerus locking plate constructs: A finite element study. *Medical Engineering and Physics*, **63**, 79–87. doi:10.1016/j.medengphy.2018.12.004.

Leary, S, Bhaskar, A, and Keane, A (2003) Optimal orthogonal-array-based latin hypercubes. *Journal of Applied Statistics*, **30**(5), 585–598. doi:10.1080/0266476032000053691.

Lecun, Y, Bengio, Y, and Hinton, G (2015) Deep learning. *Nature*, **521**(7553), 436–444. doi:10.1038/nature14539.

Lewis, GS, Mischler, D, Wee, H, Reid, JS, and Varga, P (2021) Finite Element Analysis of Fracture Fixation. *Current Osteoporosis Reports*, **19**(4), 403–416. doi:10.1007/s11914-021-00690-y.

Lewis, GS, Wee, H, Vicory, J, Armstrong, AD, and Reid, JS (2022) Virtual Simulation for Interactive Visualization of 3D Fracture Fixation Biomechanics. *The Journal of the American Academy of Orthopaedic Surgeons*, **30**(1), e51–e58. doi:10.5435/JAAOS-D-20-

01322.

Li, D, Lv, WX, Chen, WM, … Yu, B (2022) Application of a lateral intertubercular sulcus plate in the treatment of proximal humeral fractures: a finite element analysis. *BMC Surgery*, **22**(1), 1–8. doi:10.1186/s12893-022-01557-4.

Li, Y, Bu, R, Sun, M, Wu, W, Di, X, and Chen, B (2018) PointCNN: Convolution on X-transformed points. *Advances in Neural Information Processing Systems*, **2018**-**Decem**, 820–830.

Liang, L, Liu, M, Martin, C, and Sun, W (2018) A deep learning approach to estimate stress distribution : a fast and accurate surrogate of finite-element analysis. *The Royal Society*, (ML).

Lippert, LS (2006) *Clinical Kinesiology and Anatomy*, 4th Editio, Company, A. Davis, Philadelphia.

López-Monroy, AP, and García-Salinas, JS (2021) *Neural networks and deep learning Biosignal Processing and Classification Using Computational Learning and Intelligence: Principles, Algorithms, and Applications*, Elsevier Inc. doi:10.1016/B978-0-12-820125-1.00021-X.

Lucas, DB, and Francisco, S (1973) Biomechanics of the Shoulder Joint. *Archives of Surgery*.

MacLeod, AR, Pankaj, P, and Simpson, AHRW (2012) Does screw-bone interface modelling matter in finite element analyses? *Journal of Biomechanics*, **45**(9), 1712–1716. doi:10.1016/j.jbiomech.2012.04.008.

Maken, P, and Gupta, A (2023) 2D-to-3D: A Review for Computational 3D Image Reconstruction from X-ray Images. *Archives of Computational Methods in Engineering*, **30**(1), 85–114. doi:10.1007/s11831-022-09790-z.

Maldonado, ZM, Seebeck, J, Heller, MOW, … Duda, GN (2003) Straining of the intact and fractured proximal humerus under physiological-like loading. *Journal of Biomechanics*, **36**(12), 1865–1873. doi:10.1016/S0021-9290(03)00212-4.

Maravic, M, Briot, K, and Roux, C (2014) Burden of proximal humerus fractures in the French National Hospital Database. *Orthopaedics and Traumatology: Surgery and Research*, **100**(8), 931–934. doi:10.1016/j.otsr.2014.09.017.

McLean, AS, Price, N, Graves, S, Hatton, A, and Taylor, FJ (2019) Nationwide trends in management of proximal humeral fractures: an analysis of 77,966 cases from 2008 to 2017. *Journal of Shoulder and Elbow Surgery*, **28**(11), 2072–2078. doi:10.1016/j.jse.2019.03.034.

Mendoza-Muñoz, I, González-ángeles, Á, Jacobo-Galicia, G, Castañeda, AM, and Valenzuela-Gutiérrez, J (2018) Evaluation of significant effects on locking plates design for a 2-part fracture of the surgical neck of the humerus using finite element and statistical analysis. *Revista Materia*, **23**(4), 2122–2133. doi:10.1590/s1517-707620180004.0556.

Mini, D, Reynolds, KJ, and Taylor, M (2024) Assessing screw length impact on bone strain in proximal humerus fracture fixation via surrogate modelling. *International Journal for Numerical Methods in Biomedical Engineering*, (May), 1–17. doi:10.1002/cnm.3840.

Mischler, D, Babu, S, Osterhoff, G, … Varga, P (2020a) Comparison of optimal screw configurations in two locking plate systems for proximal humerus fixation - a finite element analysis study. *Clinical Biomechanics*, **78**(February). doi:10.1016/j.clinbiomech.2020.105097.

Mischler, D, Schader, JF, Dauwe, J, … Varga, P (2022) Locking Plates With Computationally Enhanced Screw Trajectories Provide Superior Biomechanical Fixation Stability of Complex Proximal Humerus Fractures. *Frontiers in Bioengineering and Biotechnology*, **10**(June), 1–11. doi:10.3389/fbioe.2022.919721.

Mischler, D, Windolf, M, Gueorguiev, B, Nijs, S, and Varga, P (2020b) Computational optimisation of screw orientations for improved locking plate fixation of proximal humerus fractures. *Journal of Orthopaedic Translation*, **25**(October 2019), 96–104. doi:10.1016/j.jot.2020.02.007.

Miyamura, S, Lans, J, Min, KS, Waryasz, GR, Murase, T, and Chen, NC (2021) Bone resorption of the greater tuberosity after open reduction and internal fixation of complex proximal humeral fractures: fragment characteristics and intraoperative risk factors. *Journal of Shoulder and Elbow Surgery*. doi:10.1016/j.jse.2020.09.014.

Morgan, EF, Bayraktar, HH, and Keaveny, TM (2003) Trabecular bone modulus-density relationships depend on anatomic site. *Journal of Biomechanics*, **36**(7), 897–904. doi:10.1016/S0021-9290(03)00071-X.

Müller, ME, Koch, P, Nazarian, S, and Schatzker, J (1990) *The Comprehensive Classification of Fractures of Long Bones The Comprehensive Classification of Fractures of Long Bones*, Springer Berlin, Heidelberg. doi:10.1007/978-3-642-61261-9.

Murray, IR, Amin, AK, White, TO, and Robinson, CM (2011) Proximal humeral fractures: Current concepts in classification, treatment and outcomes. *Journal of Bone and Joint Surgery - Series B*, **93 B**(1), 1–11. doi:10.1302/0301-620X.93B1.25702.

Neer, CS (1987) Displaced proximal humeral fractures: part I. Classification and evaluation. 1970. *Clinical Orthopaedics and Related Research*, **442**, 77–82. doi:10.1097/01.blo.0000198718.91223.ca.

Nikooyan, AA, Veeger, HEJ, Chadwick, EKJ, Praagman, M, and Van Der Helm, FCT (2011) Development of a comprehensive musculoskeletal model of the shoulder and elbow. *Medical and Biological Engineering and Computing*, **49**(12), 1425–1435. doi:10.1007/s11517-011-0839-7.

O'Rourke, D, Martelli, S, Bottema, M, and Taylor, M (2016) A Computational Efficient Method to Assess the Sensitivity of Finite-Element Models: An Illustration with the Hemipelvis. *Journal of Biomechanical Engineering*, **138**(12). doi:10.1115/1.4034831.

Owsley, KC, and Gorczyca, JT (2008) Displacement/screw cutout after open reduction and locked plate fixation of humeral fractures. *Journal of Bone and Joint Surgery - Series A*, **90**(2), 233–240. doi:10.2106/JBJS.F.01351.

Palvanen, M, Kannus, P, Niemi, S, and Parkkari, J (2006) Update in the epidemiology of proximal humeral fractures. *Clinical Orthopaedics and Related Research*, (442), 87–92. doi:10.1097/01.blo.0000194672.79634.78.

Patel, AH, Wilder, JH, Ofa, SA, … Sherman, WF (2022) Trending a decade of proximal humerus fracture management in older adults. *JSES International*, **6**(1), 137–143. doi:10.1016/j.jseint.2021.08.006.

Poppen, NK, and Walker, PS (1978) Forces at the glenohumeral joint in abduction. *Clinical Orthopaedics and Related Research*, **NO. 135**. doi:10.1097/00003086-197809000-00035.

Prinold, JAI, Masjedi, M, Johnson, GR, and Bull, AMJ (2013) Musculoskeletal shoulder models: A technical review and proposals for research foci. *Journal of Engineering in Medicine*, **227**(10), 1041–1057. doi:10.1177/0954411913492303.

Queipo, N V., Haftka, RT, Shyy, W, Goel, T, Vaidyanathan, R, and Kevin Tucker, P (2005) Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*. doi:10.1016/j.paerosci.2005.02.001.

Roberts, BC, Perilli, E, and Reynolds, KJ (2014) Application of the digital volume correlation technique for the measurement of displacement and strain fields in bone: A literature review. *Journal of Biomechanics*, **47**(5), 923–934. doi:10.1016/j.jbiomech.2014.01.001.

Röderer, G, Brianza, S, Schiuma, D, … Tami, A (2013) Mechanical assessment of local bone

quality to predict failure of locked plating in a proximal humerus fracture model. *Orthopedics*, **36**(9), 1134–1140. doi:10.3928/01477447-20130821-14.

Roux, A, Decroocq, L, El Batti, S, … de Peretti, F (2012) Epidemiology of proximal humerus fractures managed in a trauma center. *Orthopaedics and Traumatology: Surgery and Research*, , 715–719. doi:10.1016/j.otsr.2012.05.013.

Sanchez-Lengeling, B, Reif, E, Pearce, A, and Wiltschko, AB (2021) A Gentle Introduction to Graph Neural Networks. *Distill*. doi:10.23915/distill.00033.

Sarker, IH (2021) Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, **2**(6), 1–20. doi:10.1007/s42979-021-00815-1.

Schader, JF, Mischler, D, Dauwe, J, Richards, RG, Gueorguiev, B, and Varga, P (2021) One size may not fit all: patient-specific computational optimization of locking plates for improved proximal humerus fracture fixation. *Journal of Shoulder and Elbow Surgery*. doi:10.1016/j.jse.2021.06.012.

Schlegel, TF, Hawkins, RJ, and Frcs, C (1994) Displaced Proximal Humeral Fractures : Evaluation and Treatment. *Journal of the American Academy of Orthopaedic Surgeons*, (C), 54–66.

Seide, K, Triebe, J, Faschingbauer, M, … Jürgens, C (2007) Locked vs. unlocked plate osteosynthesis of the proximal humerus - A biomechanical study. *Clinical Biomechanics*, **22**(2), 176–182. doi:10.1016/j.clinbiomech.2006.08.009.

Soileau, R, Cartner, J, and Zheng, Y (2007) Locked versus conventional plate-screw fixation in osteoporotic bone: A review. *Techniques in Orthopaedics*, **22**(4), 247–252. doi:10.1097/BTO.0b013e31815dccdd.

Solberg, BD, Moon, CN, Franco, DP, and Paiement, GD (2009) Surgical treatment of three and four-part proximal humeral fractures. *Journal of Bone and Joint Surgery - Series A*, **91**(7), 1689–1697. doi:10.2106/JBJS.H.00133.

Sporer, SM, Weinstein, JN, and Koval, KJ (2006) The Geographic Incidence and Treatment Variation of Common Fractures of Elderly Patients. *Journal of the American Academy of Orthopaedic Surgeons*, **14**(4), 246–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16585366

Sproul, RC, Iyengar, JJ, Devcic, Z, and Feeley, BT (2011) A systematic review of locking plate

fixation of proximal humerus fractures. *Injury*, **42**(4), 408–413. doi:10.1016/j.injury.2010.11.058.

Steiner, JA, Hofmann, UAT, Christen, P, Favre, JM, Ferguson, SJ, and van Lenthe, GH (2018) Patient-specific in silico models can quantify primary implant stability in elderly human bone. *Journal of Orthopaedic Research*, **36**(3), 954–962. doi:10.1002/jor.23721.

Synthes (2018) PHILOS and PHILOS Long ®., , 3–4.

Takian, W, Rooppakhun, S, Ariyarit, A, and Sucharitpwatskul, S (2021) Optimal conformity design of tibial insert component based on ISO standard wear test using finite element analysis and surrogate model. *Symmetry*, **13**(12). doi:10.3390/sym13122377.

Taylor, M, Perilli, E, and Martelli, S (2017) Development of a surrogate model based on patient weight, bone mass and geometry to predict femoral neck strains and fracture loads. *Journal of Biomechanics*, **55**. doi:10.1016/j.jbiomech.2017.02.022.

Taylor, M, and Prendergast, PJ (2015) Four decades of finite element analysis of orthopaedic devices: Where are we now and what are the opportunities? *Journal of Biomechanics*, **48**(5). doi:10.1016/j.jbiomech.2014.12.019.

Tilton, M, Armstrong, A, Sanville, J, … Manogharan, GP (2020a) Biomechanical Testing of Additive Manufactured Proximal Humerus Fracture Fixation Plates. *Annals of Biomedical Engineering*, **48**(1), 463–476. doi:10.1007/s10439-019-02365-3.

Tilton, M, Armstrong, AD, Wee, H, Hast, MW, Manogharan, G, and Lewis, GS (2020b) Finite element-predicted effects of screw configuration in proximal humerus fracture fixation. *Journal of Biomechanical Engineering*, **142**(8), 1–7. doi:10.1115/1.4045907.

Unger, S, Erhart, S, Kralinger, F, Blauth, M, and Schmoelz, W (2012) The effect of in situ augmentation on implant anchorage in proximal humeral head fractures. *Injury*, **43**(10), 1759–1763. doi:10.1016/j.injury.2012.07.003.

Van Der Helm, FCT (1994) A finite element musculoskeletal model of the shoulder mechanism. *Journal of Biomechanics*, **27**(5). doi:10.1016/0021-9290(94)90065-5.

Varga, P, Grünwald, L, Inzana, JA, and Windolf, M (2017) Fatigue failure of plated osteoporotic proximal humerus fractures is predicted by the strain around the proximal screws. *Journal of the Mechanical Behavior of Biomedical Materials*, **75**(June), 68–74. doi:10.1016/j.jmbbm.2017.07.004.

Varga, P, Inzana, JA, Fletcher, JWA, … Windolf, M (2020) Cement augmentation of calcar

screws may provide the greatest reduction in predicted screw cut-out risk for proximal humerus plating based on validated parametric computational modelling. *Bone and Joint Research*, **9**(9), 534–542. doi:10.1302/2046-3758.99.BJR-2020-0053.R1.

Varga, P, Inzana, JA, Gueorguiev, B, Südkamp, NP, and Windolf, M (2018) Validated computational framework for efficient systematic evaluation of osteoporotic fracture fixation in the proximal humerus. *Medical Engineering and Physics*, **57**, 29–39. doi:10.1016/j.medengphy.2018.04.011.

Walsh, S, Reindl, R, Harvey, E, Berry, G, Beckman, L, and Steffen, T (2006) Biomechanical comparison of a unique locking plate versus a standard plate for internal fixation of proximal humerus fractures in a cadaveric model. *Clinical Biomechanics*, **21**(10), 1027–1031. doi:10.1016/j.clinbiomech.2006.06.005.

Wee, H, Reid, JS, Chinchilli, VM, and Lewis, GS (2017) Finite Element-Derived Surrogate Models of Locked Plate Fracture Fixation Biomechanics. *Annals of Biomedical Engineering*, **45**(3). doi:10.1007/s10439-016-1714-3.

Wu, G, Van Der Helm, FCT, Veeger, HEJ, … Buchholz, B (2005) ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion - Part II: Shoulder, elbow, wrist and hand. *Journal of Biomechanics*, **38**(5), 981–992. doi:10.1016/j.jbiomech.2004.05.042.

Yang, P, Zhang, Y, Liu, J, Xiao, J, Ma, LM, and Zhu, CR (2015) Biomechanical effect of medial cortical support and medial screw support on locking plate fixation in proximal humeral fractures with a medial gap: A finite element analysis. *Acta Orthopaedica et Traumatologica Turcica*, **49**(2), 203–209. doi:10.3944/AOTT.2015.14.0204.

Yılmaz, S, Vayısoğlu, T, and Çolak, MA (2020) Shoulder Anatomy. In Huri, G., F. Familiari, Y. L. Moon, M. N. Doral, and G. M. Marcheggiani Muccioli, eds., *Shoulder Arthroplasty*, Springer, , 3–44. doi:10.1007/978-3-319-98908-2_2.

Zadpoor, AA, Campoli, G, and Weinans, H (2013) Neural network prediction of load from the morphology of trabecular bone. *Applied Mathematical Modelling*, **37**(7), 5260–5276. doi:10.1016/j.apm.2012.10.049.

Zdero, R, Brzozowski, P, and Schemitsch, EH (2024) Biomechanical design optimization of proximal humerus locked plates: A review. *Injury*, **55**(2), 111247. doi:10.1016/j.injury.2023.111247.

Zhou, J, Cui, G, Hu, S, … Sun, M (2020) Graph neural networks: A review of methods and

applications. *AI Open*, **1**(December 2020), 57–81. doi:10.1016/j.aiopen.2021.01.001.

Zhou, Q, and Al-Qadi, IL (2024) Graph Neural Networks to Simulate Flexible Pavement Responses Using Three-Dimensional Finite Element Analysis Data. *Transportation Research Record*, (iv). doi:10.1177/03611981241242783.

Ziaeipoor, H, Martelli, S, Pandy, M, and Taylor, M (2019) Efficacy and efficiency of multivariate linear regression for rapid prediction of femoral strain fields during activity. *Medical Engineering and Physics*, **63**, 88–92. doi:10.1016/j.medengphy.2018.12.001.

Ziaeipoor, H, Taylor, M, and Martelli, S (2020) Population-Based Bone Strain During Physical Activity: A Novel Method Demonstrated for the Human Femur. *Annals of Biomedical Engineering*, **48**(6), 1694–1701. doi:10.1007/s10439-020-02483-3.

Zimmer Biomet (2014) S3 Proximal Humerus Plating System - Surgical Technique., Warsaw, IN, USA.

Zimmer Biomet (2015a) A.L.P.S.® Proximal Humerus Plating System - Surgical Technique., Warsaw, IN, USA, , 1–30.

Zimmer Biomet (2015b) NCB ® Proximal Humerus System.

Appendix A: Supplementary material for Chapter 3

Convergence Study

A convergence study was conducted in order to select the mesh size for all the FE simulations run in this study. Five different FE models were generated, varying the maximum and the minimum edge lengths of the elements of the mesh, as shown in the table. (Table A.1)

Table A.1 - Mesh dimension of each model and their results of bone principal strain around all the screws

Model number	Elements number	Min element edge length (mm)	Max element edge length (mm)	50th percentile principal bone strain around the screws (µstrain)	90th percentile principal bone strain around the screws (µstrain)	Running time (min)
1	120482	2.0	2.5	1798.18	4510.56	5
2	197715	1.5	2.0	2050.91	5147.50	5
3	435899	1.0	1.5	2155.45	5510.33	8
4	935165	0.5	1.0	2360.77	6129.38	15
5	2591842	0.4	0.7	2664.00	6872.86	42

The loading, boundary conditions and material properties were set the same as explained in the methods section. The principal bone strain around all the screws was analysed, and each model's results were compared.

As shown in the table below, between models 5 and 4 we obtained a deviation of bone strain up to 10%, which was assumed acceptable since the running time of the model with the finest mesh was significantly higher (Table A.1).

Assessing different surrogate models

Two more surrogate models were developed in this study, a Squared Exponential Gaussian Process Regression (GPR) model and a Linear regression model.

As for the development of the ANN models, the surrogate models were trained with the training set of 500 FE simulations and then tested with the set of 100 FE simulations. The regression between the results from the FE models and their prediction from the surrogate models has been reported.

Overall, for both predictions of bone principal strain made with the Squared Exponential GPR model and Linear Regression model, the R^2 values were high but the values of RMSE for the linear regression model were significantly higher. The results from the Squared Exponential GPR were comparable with the results from the ANN models, although they showed a slightly high level of RMSE (Table A.2, Table A.3, Table A.4).

Moreover, the Squared Exponential GPR model and Linear Regression model are not capable to make predictions of multiple outputs, differently from the ANN model.

Table A.2 - Prediction of bone principal strain around all the screws

Surrogate model	50th percentile min principal strain			90th percentile min principal strain		
	R^2	Slope	RMSE, μ strain	R^2	Slope	RMSE, μ strain
Squared Exponential GPR	0.99	0.99	24.575	0.99	0.99	98.84
Linear Regression	0.86	0.90	93.7	0.93	0.91	434.03

Table A.3 - Squared Exponential GPR model's predictions of bone principal strain around each screw

	50th percentile min principal strain			90th percentile min principal strain		
	R²	Slope	RMSE, μ strain	R²	Slope	RMSE, μ strain
Screw6	0.999	0.996	61.112	0.997	0.994	195.084
Screw7	0.992	0.993	49.206	0.987	0.978	227.871
Screw8	0.988	1.00	35.677	0.991	0.968	91.055
Screw9	0.976	0.978	36.115	0.973	0.986	136.542
Screw10	0.991	0.984	41.085	0.993	0.972	80.115
Screw11	0.985	0.991	29.794	0.991	1.006	78.650
Screw12	0.996	0.987	51.315	0.997	0.979	85.294

Table A.4 - Linear regression model's predictions of bone principal strain around each screw

	50th percentile min principal strain			90th percentile min principal strain		
	R²	Slope	RMSE, μ strain	R²	Slope	RMSE, μ strain
Screw6	0.979	0.977	315.43	0.939	0.944	875.413
Screw7	0.916	0.921	160.50	0.692	0.674	1107.041
Screw8	0.501	0.498	229.62	0.685	0.602	526.478
Screw9	0.908	0.915	71.388	0.899	0.872	267.430
Screw10	0.942	0.924	105.676	0.940	0.917	226.851
Screw11	0.114	0.140	235.037	0.970	0.985	150.473
Screw12	0.965	0.950	153.212	0.967	0.956	283.566

Appendix B: Supplementary material for Chapter 5

GNN development

Performance of the GNN described in section 5.2.1.2 on the dataset described in section 5.2.2.1, regarding the dataset on the variation of screw orientation (Figure B.1).
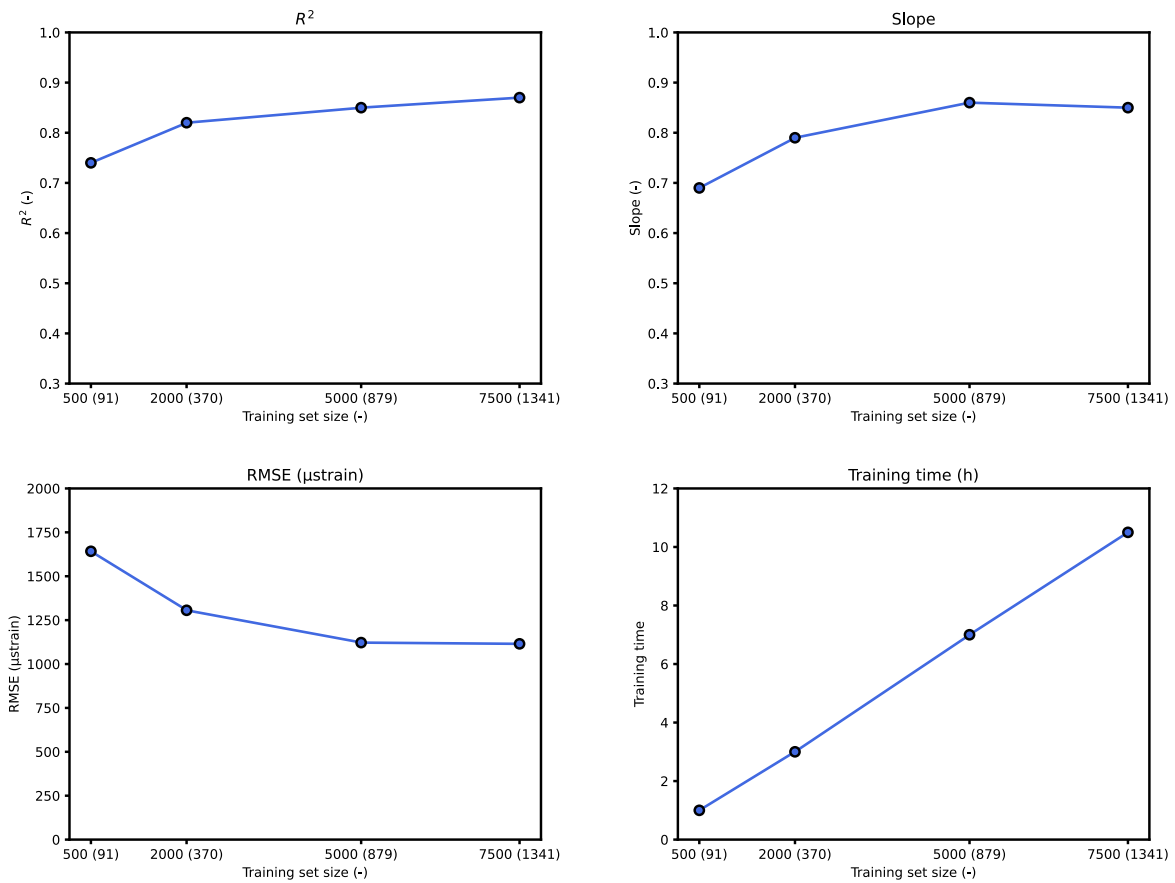

Figure B.1 - Performance of the GNNScrewDirection developed with the model described in section 5.2.1.2 for each training set. The original size of the training set is shown with the reduced one in brackets

Appendix C: Supplementary material for Chapter 6

GNN training results

Performance of the training process of the GNN described in section 6.2.3 (Figure C.1, Table C.1).
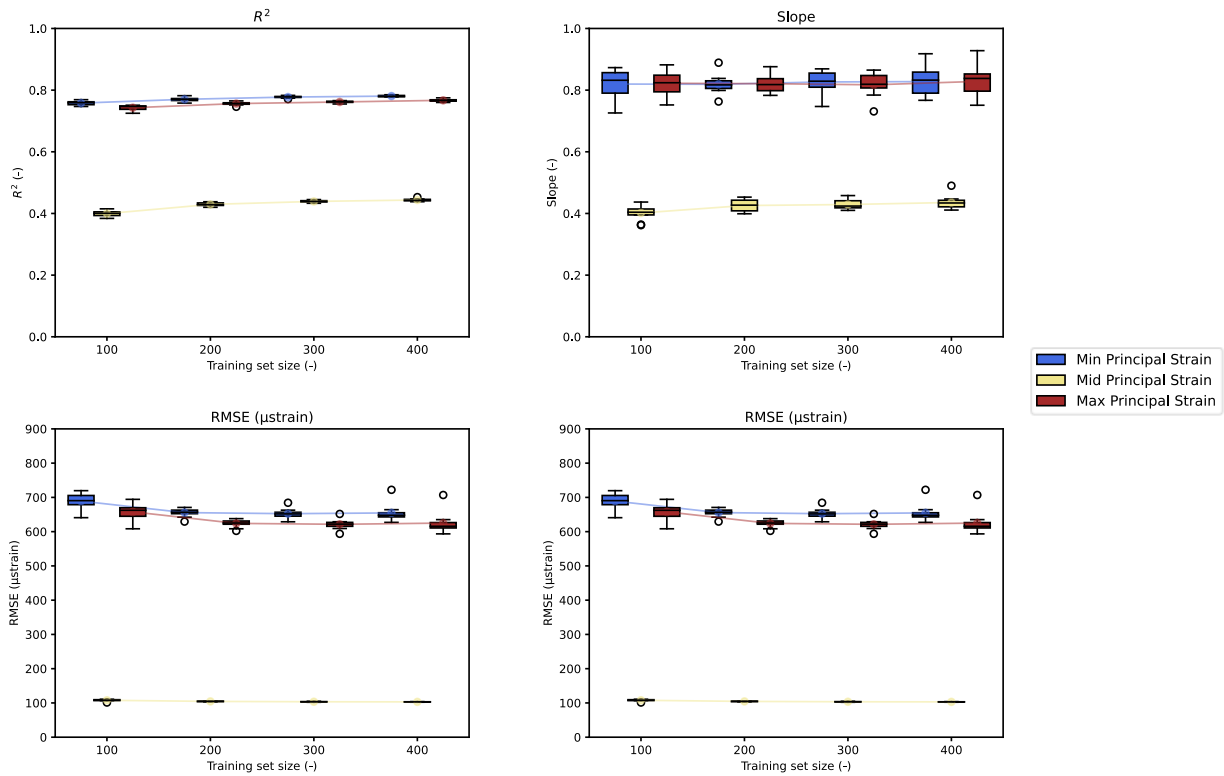

Figure C.1 - Performance variation of the training of 10 GNNs developed for each training set. The variation for the training set prediction and the testing set prediction is shown.

Table C.1 - *Performance of the training of the GNN on the testing set of 34 simulations for the prediction of minimal, middle and maximum principal strain of the bone around all the screws. The influence of the training set size is shown. Results are displayed for the model with the lowest error, while the average of 100 models is shown in brackets. The computational time difference between the FE and the GNN processes is also displayed*

Training set size (n)	FE Running time (h)	GNN training time (h)	Minimal Principal strain				Mid Principal strain				Maximum principal strain			
			R²	Slope	RMSE (µstrain)	RMSE (%)	R²	Slope	RMSE (µstrain)	RMSE (%)	R²	Slope	RMSE (µstrain)	RMSE (%)
100	61-70	1	0.75 (0.76)	0.80 (0.82)	641 (689)	5.02 (5.14)	0.41 (0.40)	0.43 (0.40)	107 (108)	10.2 (10.5)	0.75 (0.74)	0.84 (0.82)	651 (657)	5.65 (5.98)
200	122-140	2	0.78 (0.77)	0.81 (0.82)	651 (655)	4.77 (4.91)	0.44 (0.43)	0.45 (0.43)	103 (104)	9.97 (10.3)	0.76 (0.76)	0.80 (0.82)	625 (624)	5.59 (5.71)
300	183-210	3	0.78 (0.78)	0.84 (0.83)	629 (652)	4.78 (4.88)	0.44 (0.44)	0.43 (0.43)	102 (103)	10.1 (10.2)	0.77 (0.76)	0.82 (0.82)	594 (621)	5.55 (5.69)
400	244-280	4	0.79 (0.78)	0.83 (0.83)	627 (654)	4.72 (4.91)	0.45 (0.44)	0.43 (0.44)	102 (103)	9.98 (10.1)	0.78 (0.77)	0.84 (0.83)	593 (624)	5.43 (5.72)